Photo by J.R. Mitchell

Understanding Estimates of AI Sire Fertility

J.M. DeJarnette, R.L. Nebel, C.E. Marshall
Select Sires Inc., Plain City, Ohio

KEY POINTS

Sire management and quality control programs in an AI center have two goals, namely: (1) ensure production and marketing of straws containing highly fertile sperm in sufficient numbers to maximize the conception rate obtained with each sire; and (2) minimize differences in fertility among sires whose semen is released for sale. For 90% of the bulls marketed, their fertility is within a 6 percentage unit range. On a population basis, we will never be able to measure "fertility" more precisely than ±3 percentage units, because of the many factors beyond anyone's control including: binomial variation, herd environment, measurement errors, and bias in semen use. Failure to recognize limitations in any estimate of potential fertility leads to over interpretation of small differences among sires in apparent fertility. Available data show that major AI centers do an admirable job in achieving the goals laid out above.

INTRODUCTION

A primary goal of dairy and beef producers is to achieve high pregnancy rates with semen from genetically superior sires. The primary goal of every AI organization is to provide customers with a product that allows them to meet these goals. Recent trends toward reduced reproductive performance in dairy cattle have heightened awareness all aspects of bovine reproduction. Many producers seek to exploit every opportunity to improve conception rates in their herds, which has led to increased emphasis on fertility estimates in the sire selection process.

To make sound decisions regarding any selection trait, producers need to be cognizant of the magnitude of response that might be realized and the corresponding accuracy and (or) limitations that might be associated with calculation of the estimate for that trait. This brochure will: 1) define and describe measures of fertility; 2) review the relationship between conception rate and semen quality; 3) review how knowledge of the relationship between conception rate and semen quality is implemented in both semen processing and quality control at an AI center; 4) point out limitations in estimating or predicting future fertility of an AI sire and 5) review methodologies used for estimating recent fertility of an AI sire. Hence, the goal of this document is to provide producers with a scientifically sound summary of these topics in terms a layperson can understand, so that informed decisions can be made regarding the emphasis to place on estimated fertility in the sire selection process.

MEASURES OF FERTILITY

It is critical to list and define terminology sometimes used in conjunction with discussion of sire fertility and attempt to point out the strengths and weaknesses of each. It is also important to note that all measures are ***estimates***, with an associated error rate (precision) which is often difficult to define and will vary greatly between measures.

Fertility is defined as “the state or quality or being fertile”. For our purposes, this is a very general and vague term. Even severely sub-fertile animals would still meet the definition of “fertile”. Hence this term is of limited value because one does not know exactly what measure of fertile capacity is being referred to. The utility of the term fertility is that it is easy to use in sentence structure for writing purposes.

Fertilization rate is the rate of fusion of male and female gametes that result in the formation of a new individual, the zygote. In most cases, fertilization rate cannot be measured in cattle. Failure of cattle to become pregnant may be due to embryonic mortality following fertilization. Without harvested embryos/ovum within a few days of insemination, it is impossible to distinguish between fertilization failure and embryonic mortality. This will be a common limitation to all fertility definitions that follow. Fertilization rate does, however, set the upper limit for all measures of reproductive capacity that follow.

Non-return rate is an estimate of outcome based on the lack of a detected or recorded return to estrus following some specified interval (60 to 90 days) after AI. Non-return rate “assumes” the farmer is observing his animals each day (ideally 2X) to catch those returning to estrus and "assumes" that failure to return to estrus is due to pregnancy. Hence, there is an error rate associated with not looking for and (or) not seeing cows in heat. Additionally, there is an error rate associated with failure to record a repeat service. In some non-return systems there is no cross check to ensure the animal is even still in the herd. Thus, cows that are culled after AI might be assumed to be pregnant when they fail to receive another AI. It should go without saying that non-return calculations are prone to high error rates.

Conception rate - Using a "text book definition", conception rate is basically synonymous with fertilization rate and therefore cannot really be measured in the commercial setting. However, for the purposes of this manuscript and consistent with reproductive terminology used in academic circles, conception rate will be defined as the percentage of females diagnosed as pregnant by conventional means (rectal palpation, ultrasonography, hormonal assay) at some specified interval after AI (e.g., 25 or 60 days). In addition to the limitations cited for fertilization failure, conception rate cannot account for embryo or fetal death that might subsequently occur. However, for the purpose of estimating sire fertility potential, conception rate will, in most cases, be associated with the least amount of error.

Pregnancy rate - Historically, pregnancy rate and conception rate have often been used interchangeably. However recent definitions have evolved to include a component of time, which is usually defined as a 21 day estrous interval. In addition, pregnancy rate includes in the denominator females that were "eligible" to be inseminated irrespective of whether they were bred or not. Thus, pregnancy rate is calculated in one of two ways:

1) Number of females that conceived in a defined time period divided by the number of females eligible for breeding during the defined time period.
2) Estrous detection rate x conception rate during the defined time interval.

Although pregnancy rate is usually expressed as the function of a 21 estrous interval, the time interval can be altered to be expressed as: fixed time AI pregnancy rate (1 day), synchronized pregnancy rate (3 to 5 days), or breeding season pregnancy rate (60 to 90 days). Because it includes cows in the denominator that should have been bred but were not, pregnancy rate is an excellent barometer of herd management and cow fertility. However in most cases, pregnancy rate should NOT be used in discussions of sire fertility as bulls could be penalized for lack of pregnancy in cows that were never bred.

PRINCIPLES OF SEMEN QUALITY AND FERTILITY

Essential to understanding the relationship of semen quality to fertility are the concepts of compensable and uncompensable sperm quality traits and their interactions with numbers of sperm per dose (Salisbury and Vandemark, 1961; Saacke 1998). If the number of sperm per AI dose is increased based on measurement of a compensable semen trait, there is an increase in fertility up to a maximum level (termed "upper-limit value"; discussed later). It is important to note however that the increase in fertility is not proportionate over the entire range of sperm dosages. Compensable semen quality traits are generally believed to be associated with measures of sperm viability (i.e., motility, acrosomal integrity, cell membrane integrity, etc.).

Conversely, if the number of sperm per AI dose is increased based on measurement of an uncompensable semen trait, there is NO increase in fertility with additional sperm beyond that associated with a compensable trait that might be added concurrently. Uncompensable semen quality traits are closely associated with abnormal sperm morphology, DNA integrity, and sustenance of normal embryonic development following fertilization (Kidder et al., 1954; Bearden et al., 1956). In all cases, the female population sets an upper limit based on the herd's fertility, health, lactation stress, nutrition, diligence of management, etc.

Figure 1 illustrates the impact on fertility as the number of live sperm per AI dose is increased **and** how this differs depending on the proportion of sperm with a compensable vs. uncompensable defect in the insemination dose. Data for 5 hypothetical bulls are depicted, although real-world data have been published (Sullivan and Elliott, 1968; Filseth et al., 1992; den Daas et al., 1998). Bull A is one typically ejaculating semen containing few defective sperm of any type, and most defective sperm have a compensable problem (e.g., abnormal acrosome; almost all live sperm are motile). Hence, observed fertility increases dramatically as number of live sperm per AI dose is increased. Semen from bull B, and especially C, has a higher proportion of sperm with a compensable defect (e.g., abnormal acrosome; immotile sperm). Maximum fertility is still obtained with these bulls, but it takes perhaps twice as many live sperm per AI dose. Semen from Bull D, and especially E, is different from that of

Bulls A, B, or C. For semen from Bull D, within the population of defective sperm, the ratio of cells with an uncompensable defect to cells with a compensable defect is sufficiently great that it is detectable in terms of depressed fertility. For Bull E, this ratio is even more unfavorable, and sperm with uncompensable defects might total one-third of the proportion with a compensable defect; remember some sperm have no defects.

Due to the heterogeneous nature of sperm in any population, the rate at which individual bulls and (or) ejaculates approach the maximum level of fertility allowed within a given population of females is a function of the severity and ratio of compensable and uncompensable sperm defects within the sample. Bulls A, B and C differ in the numbers of sperm per dose required to achieve maximum fertility, but not in actual observed fertility. Irrespective of sperm numbers per AI, sires D and E will not obtain a level of fertility comparable to that of sires A, B, or C due to uncompensable semen quality characteristics.

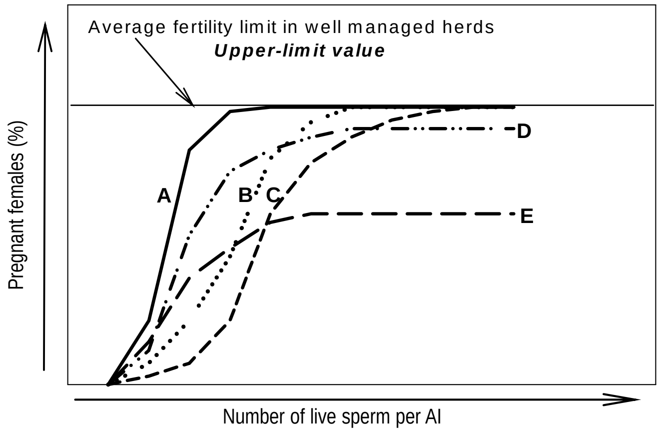

Figure 1. Relationship of number of live sperm number per AI dose and fertility for bulls whose semen differs in quality. Semen from bulls A, B or C displays essentially no sperm with uncompensable defects, whereas semen from bulls D or E includes a substantial proportion of sperm with uncompensable defects. Note the "upper-limit value", which might be higher or lower in a given herd than for the average of all well managed herds depicted. In herds where the upper limit for fertility is lower than depicted, pregnancy rates obtained with bulls A-E will be proportionally lower. (Adapted from Salisbury and VanDemark, 1961; den Daas et al., 1998).

Importantly, the depression in observed fertility due to uncompensable sperm for Bulls D and E, is intrinsic in that sperm produced by those bulls cannot be "fixed" in the laboratory by adding more sperm or by magical freezing potions added to the extender medium. Most sperm with uncompensable defects reflect a problem in spermatogenesis within the testis perhaps exacerbated by summer heat, health problems, or other factors. It is uncertain if the propensity for uncompensable defects in semen from a sire will show up in semen from his sons or grandsons. However, because such defects tend to result in early death of an embryo, a given defective spermatozoon typically does not result in a living son.

AI CENTER QUALITY CONTROL PROGRAMS

The primary goal of the AI center is to ensure every straw of every collection from every bull is a highly fertile product and equal in fertility potential to all other straws of all other collections from that bull or most other bulls. Although the goal is to make all products of comparable fertility, it goes without saying that the fertility potential of all bulls was NOT created equal. How then could an AI center presume to accomplish such a feat?

First, it is extremely important to understand the difference between research and commercial applications. **Substantial expression of differences in fertility is obligatory in research. Minimal expression of differences in fertility is obligatory for commercial success.**

Research projects are designed to measure the impact of various processing steps or semen quality characteristics on fertility potential. In either case, research projects ***intentionally introduce variation in straws of semen sent to the field in an attempt to better understand which processing changes or traits are important and which are not.***

A research experiment that compared semen samples all of which had similar quality characteristics would yield very little new information. However, well designed studies will contain semen samples with an anticipated difference in outcome or considerable range in measured value for the attribute(s) under investigation. Only in this way can the researcher determine: 1) Does the change in processing procedure really help? 2) Which semen quality attributes are important to fertility? 3) At what level or incidence do these attributes impact fertility? and 4) What is the magnitude of impact that these factors have on fertility? Through more than 60 years of research and experience, AI center personnel have greatly enhanced their knowledge and understanding of the effects of various components of semen quality on fertility potential.

Quality control programs at an AI center put to use knowledge gained from research to intentionally minimize variation in quality of semen distributed and maximize fertility potential of each straw released for sale. This is accomplished through implementation of safeguards at several levels, each embodying numerous carefully documented procedures. The first safeguard is before and during selection of bulls; not only must they be genetically superior, but also structurally sound, healthy, disease free, and physically capable of producing semen of acceptable quality for use in AI. The next safeguard is after arrival at the AI center and entrance into the quarantine barn, where disease-free status of bulls is confirmed. Pre- and post-thaw evaluations of initial collections provide evidence of the bull's ability to produce semen of acceptable quality to enter the salable inventory. Failure in any area results in sire removal. Assuming satisfactory initial test results, semen is stock-piled for distribution. At this point, the most important safeguard with respect to consistent maximization of fertility potential of each straw distributed comes into action: post-thaw quality control.

The primary safeguard to insure top quality semen is embodied in a comprehensive **post-thaw semen evaluation program.** The sole goal of this program is to identify and discard any ejaculate that is substandard on the basis of laboratory tests and, hence, might result in less than normal fertility if straws were distributed for use in AI.

At Select Sires, representative straws of each freeze code are randomly selected and evaluated by specially trained personnel. Results are systematically monitored by a supervisor who also insures that precision of each test is maintained. For each freeze code, these quality control tests include: number of doses produced; percentage of motile sperm shortly after thawing and also after incubation at 37° C for 3 hours; percentage of sperm with an intact and normal acrosome after incubation at 37° C for 3 hours. Every month, the percentage of morphologically normal sperm in semen from each bull is determined. For bulls with marginal levels of any semen quality characteristics, morphology is evaluated on each and every collection until acceptable normal levels are re-achieved. Any freeze code failing one of these quality criteria, is re-examined, and, assuming it fails again, all straws from the collection are discarded. Each quality test for each bull is tracked to facilitate detection of gradual trends or abrupt changes.

Bulls whose semen consistently fails to meet any quality control standard on multiple successive collections are taken off the collection schedule. Such is the case when summer heat and humidity may adversely affect semen quality. This ensures that their semen does not reach the saleable inventory. Such a bull will be monitored, and sometimes it might require greater than 6 months for quality to be restored. When semen quality characteristics improve, the bull will be returned to the collection schedule. Alternatively, these bulls may be eliminated from the program, as incapable of producing semen meeting the stringent criteria of our quality control program.

IMPORTANCE OF THE NUMBER OF SPERM PER AI STRAW

The number of sperm per insemination dose is important to the producer purchasing semen and to the AI organization processing semen. The **producer** wants assurance that the straws purchased will: (1) contain sufficient sperm of sufficient quality to maximize the probability of pregnant females (both cows and heifers), and (2) cost him the least possible amount of money for a straw of the genetic quality and line selected for purchase. The **AI organization** wants to: (1) sell only straws that contain a number of sperm anticipated to provide optimum probability of conception in a producer's herd, based on historic data for bulls in general, and specific semen quality and fertility data for the individual bull; and (2) receive a fair selling price consistent with the value of a given bull derived from established genetic merit, production of salable doses of semen per week, popularity, and other factors. Since costs of genetic testing (progeny proof), bull maintenance, and semen collection and processing are, for the most part, invariable expenses, the number of salable doses produced per week is a major factor in determining selling price. Further, as discussed previously, placing “more than enough sperm” in an AI straw does not improve the probability of a pregnancy. For these reasons, it is in the interest of both the producer and AI organization that the AI organization strives to insure that sufficient sperm are in each AI straw stored for distribution, but not 30 times the number needed to achieve optimum fertility under typical field conditions.

It is a common “perception” that many of the problems with bull fertility and AI conception rates are due to insufficient number of sperm in AI doses from a particular bull, or all bulls in general. There is no doubt that the number of sperm per AI dose can impact fertility. This concept was proposed by Salisbury and VanDemark in 1961, and has been confirmed in countless studies, including several that document the **minimum** number of sperm required per dose and the fact that this **minimum number differs for each bull** (Sullivan and Elliott, 1968; den Daas et al. 1998). Extensive research using Holstein AI sires indicates there is a **“threshold value”** above which adding more sperm to an AI dose **does not** increase fertility. The threshold value for sperm number per AI dose required to reach maximum fertility potential ranges from as low as 0.5 million sperm per dose to as high as 12 million per dose. Hence, for some bulls having more than 1-2 million sperm per AI dose does not improve fertility, whereas for other bulls, inclusion of 10-15 million sperm per AI dose would be required to reach the upper limit value. For the typical bull, the number of sperm required to reach the threshold value is between 2.5 and 5 million total sperm per AI dose (Filseth et al., 1992; van Giesson et al., 1992; den Daas et al., 1998). A recent world-wide survey of semen processing practices at major AI organizations (Vishwanath, 2003) found the average AI dose contained approximately 20 million total sperm (range 10 to 40 x 10^6). This confirms that AI organizations are responsibly placing considerably more sperm in each straw than are necessary to achieve maximum fertility potential by skilled inseminators in all but the most poorly managed herds. Remember AI of cows not in estrus will not provide pregnancy irrespective of sperm number per dose.

Although the AI center may increase numbers of sperm per dose for certain sires, in an attempt to compensate for marginal sperm quality and keep the bull available to producers, this effort often is futile. Sperm from many of these bulls often possess traits that are uncompensable, so that fertility remains somewhat below average due to sperm quality, despite the extra sperm numbers in the straw. This is particularly true for sires of many beef breeds that in general do not produce semen of comparable quality to sires of dairy breeds.

How many extra sperm are wasted in each AI dose? This is important to producers because it affects availability of semen and price per dose. Rather than placing 5 to 10 time mores sperm in each straw than are really needed, a conscientious AI organization might package semen from certain high demand sires with only 2 to 5 times more sperm per straw than are needed to maximize fertility in most herds. This will increase the supply of highly desired genetics. This practice is attempted ONLY if post-thaw evaluation of sperm quality (see above) remains high AND if close monitoring of fertility data suggests no decrease in fertility should be expected. Thus, semen from certain high demand sires will have fewer sperm per dose than semen packaged from other sires, but in all cases, each AI dose contains at least 2 to 5 times more sperm than needed to reach the upper-limit value (Figure 1; Salisbury and VanDemark, 1961). Remember, as discussed above, insemination of more sperm than necessary to reach the threshold value does not result in detectably higher conception rates. The producer should remember that the alternative to efficient utilization of sperm from high demands sires is to reduce number of straws processed, which in-turn adversely influences both availability and price for the customer.

The producer should remember that high demand sires are a minority in any AI organization (perhaps 10 or 15% of all bulls available). Most bulls produce far more sperm than required to meet consumer demand and their straws contain far more sperm (typically 15 to 40 million sperm per straw) than required to reach the threshold value.

MISLEADING ARTIFACTS IN SEMEN QUALITY AND FERTILITY DATA

The efforts of quality control programs used by major AI centers result in several interesting artifacts. The first is what often appears to be a negative correlation between sperm number per dose and fertility. This happens because: 1) bulls with below average semen quality often achieve below average fertility even though number of sperm per straw is intentionally and markedly increased, while 2) popular bulls that produce sperm with above average quality maintain above average fertility even when their semen is extended to a relatively low number per dose (but still 2 to 5 times the threshold value). Secondly, there is little relationship between measures of semen quality in the lab and estimated sire fertility when only those samples that passed quality control and reached the saleable inventory are considered. This latter should not surprise anyone. It simply confirms that the quality control program is working as desired, and that semen that might be of noticeably lower fertility was not released for distribution. Discarding semen of relatively poor quality minimizes variation in quality (and fertility) of semen that reaches the salable inventory.

Importantly, because of intense screening and culling of both bulls and ejaculates for the past 70 years, the range in fertility potential of dairy bulls available today from major AI organizations is greatly reduced compared to that of almost any other population of males. In fact, intense selection by AI organizations for semen quality and fertility over 70 years likely eliminated many undesirable genotypes with respect to fertilizing potential of sperm and early embryo survival. This is evidenced by the fact that regardless of which sire fertility estimating system is used, ~90% of dairy sires will rank within a range of -3 to +3 units of the average for all bulls. Further, only 2 to 3% of sires will have an estimated relative fertility less than or equal to 4 percentage points below average, and even these sires are not sterile but only slightly subfertile. Additionally, the measurement error for estimated fertility for a given sire is seldom (if ever) less than ±2 percentage units, and frequently will be >3 percentage units because precision is a function of number of observations and other factors.

It is also important to remember that any time an average is calculated, by definition, half the animals included in the dataset must be "below average". However, if we cull all animals that are below average today, tomorrow when we calculate a new average, there will simply be a new group of sires that are now defined as below average. This logic could be repeated until only a single sire is left and then, by definition, he would just be of average fertility. These facts tend to suggest that AI centers are doing an admirable job of achieving their goal of producing ***a consistent quality product and further implies that below average fertility bulls should NOT be equated to subfertile bulls.***

LIMITATIONS IN ESTIMATING OR PREDICTING FUTURE FERTILITY OF AN AI SIRE

Estimated sire fertility is the ultimate criterion to define the fertility potential of an individual sire and the adequacy of an AI center quality control program. The key word in this sentence is **"estimated"**. Far too much emphasis is placed on the magnitude of the deviation (i.e., +2 rather than -2 or -3) as though it had great precision and accuracy. To grasp an adequate understanding of both the value and limitations of estimated sire fertility, one first must have a clear grasp of the concept of binomial variation.

Binomial variation

Fertility potential is a trait that is statistically referred to as a binomial variable. A binomial variable is one for which there are only two possible outcomes. A given insemination can only result in one of two possible outcomes: the cow either becomes pregnant or she does not. Binomial variation is basically the influence of random chance on the observed outcome. Mathematical formulas will quite accurately predict the probability of any given outcome of a binomial variable based on the average of the population and the sample size being considered. These formulas have in turn been used as the foundation for a multibillion dollar industry that thrives quite profitably in Las Vegas. The gaming industry uses these formulas to convert numerous potential outcomes from numerous games of chance into simple odds of a binomial variable: you win or the house wins. These formulas are real and they work. If, after reading this document, you still doubt the laws of binomial variation, please visit one of these establishments. They will be more than happy to provide you with advanced level training.

The most common analogy used to describe a binomial variable is the coin toss. With each coin toss there is a 50:50 probability of obtaining a "heads" or a "tails". With repeated series of coin tosses, we expect to "average" a 50:50 ratio of heads to tails. However the result of any single toss is totally independent of the result of any previous or subsequent coin toss. When you toss a coin a given number of times and get an answer that is not a 50:50 ratio, you have just experienced the impact of binomial variation. For the sake of simplicity, let's focus on a single example of 4 coin tosses. With all binomial variables, the potential number of combinations of outcomes increases exponentially and is equal to the sample size squared. In the example of 4 coin tosses, there are 16 possible outcomes with each having a probability of occurrence of 6.25% (Table 1). Although there are 16 potential combinations of outcomes, the number of different possible results for a binomial variable is always equal to the sample size + 1. Thus, in our example there are really only 5 possible results (100, 75, 50, 25 or 0% heads) but multiple combinations that yield the same result. The "average" percentage of heads across all combinations is 50%; however, if we add-up the cumulative probabilities of all the possible combinations that result in "exactly" a 50:50 ratio, we find that there is only a 37.5% probability (6.25% x 6 possible combinations) that our answer from 4 coin flips will actually result in exactly a 50:50 ratio of heads to tails. Most important to note is the fact that there is a 62.5% probability that our result from 4 flips will NOT equal a 50:50 ratio of heads to tails, even though we know for a fact that the coin has a head on one side and a tail on the other.

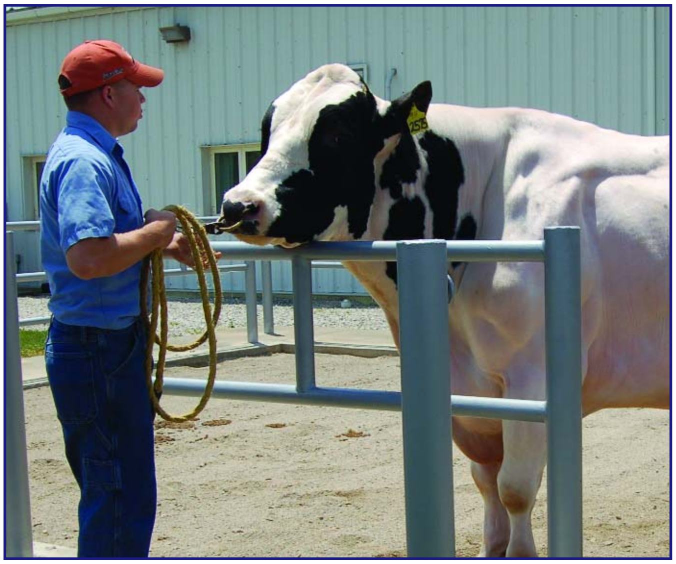

Table 1. Potential combinations of outcomes from a single series containing 4 coin tosses.

Toss 1	Toss 2	Toss 3	Toss 4	% Heads	Probability (%)
Heads	Heads	Heads	Heads	100	6.25
Heads	Heads	Heads	Tails	75	6.25
Heads	Heads	Tails	Heads	75	6.25
Heads	Heads	Tails	Tails	50	6.25
Heads	Tails	Heads	Heads	75	6.25
Heads	Tails	Heads	Tails	50	6.25
Heads	Tails	Tails	Heads	50	6.25
Heads	Tails	Tails	Tails	25	6.25
Tails	Heads	Heads	Heads	75	6.25
Tails	Heads	Heads	Tails	50	6.25
Tails	Heads	Tails	Heads	50	6.25
Tails	Heads	Tails	Tails	25	6.25
Tails	Tails	Heads	Heads	50	6.25
Tails	Tails	Heads	Tails	25	6.25
Tails	Tails	Tails	Heads	25	6.25
Tails	Tails	Tails	Tails	0	6.25

Another important nuance of binomial variation is the number of observations on the absolute possible outcomes. For example, using any “odd number” of coins tosses (1, 3, 5, etc.) it is mathematically impossible to achieve an answer that this is exactly “50:50” even though we know there are only 2 sides to the coin and each are different.

Conception is seldom a 50:50 probability

The coin toss example was used to initiate this discussion simply because it is the most commonly used analogy. However, the coin toss analogy is only relevant for herds that do in fact have a 50% conception rate. Most dairy herds have conception rates for lactating cows in the 30 to 40% range and therefore do NOT have an equal probability for pregnant or open outcomes from a given insemination. Additionally, we must remember that the observed fertility rate is:

Observed fertility = (male’s fertility) x (female’s fertility)

Although we recognize there is likely no such thing as a sire with a 100% conception rate, for the sake of illustration in this discussion, let’s assume there is and use binomial probabilities to predict what will happen if 10 units of semen are used in a herd where the female conception rate is fixed at 30%. In this example, there are 11 possible outcomes ranging from 0 to 100% pregnant, but there are now 100 possible combinations (10 x 10) to achieve these 11 results. The outcome of each insemination is independent of all others and is equivalent to taking your chances of picking a white marble from a bag containing 30 white marbles and 70 black marbles. With each selection, the odds are stacked 70:30 against a successful (white marble) outcome. The greatest probability of any single answer is 3 pregnancies (white marbles), however this is only expected to occur 26.7% of the time (Figure 2). This means that 73.3% of the time the result will be something other than a 30% conception rate. It could be more or could be less, but the observed answer has nothing to do with sire fertility which we know in this example to be 100%.

It is also very important to note that the probability that a 100% conception rate sire will achieve a value < 30% is greater than the probability that he would achieve a value > 30%. As illustrated in Figure 2, the cumulative probability of all combinations of < 3 pregnancies (0, 1, and 2 pregnancies) is 38.2%, whereas the 7 potential outcomes > 3 pregnancies (4 thru 10 pregnancies) only add up to a 35.1% probability. This illustrates why the coin flip is not the best example when discussing the effects of binomial variation on fertility as very few herds have “exactly” 50% conception rates and thereby equal probability of pregnant and open results from a given insemination. In the present example, it should not be surprising if we sometimes go through 10 inseminations with a 100% conception rate sire and achieve 0 or only 1 pregnancy as binomial probabilities predict this will occur 2.8% and 12.1% of the time, respectively.

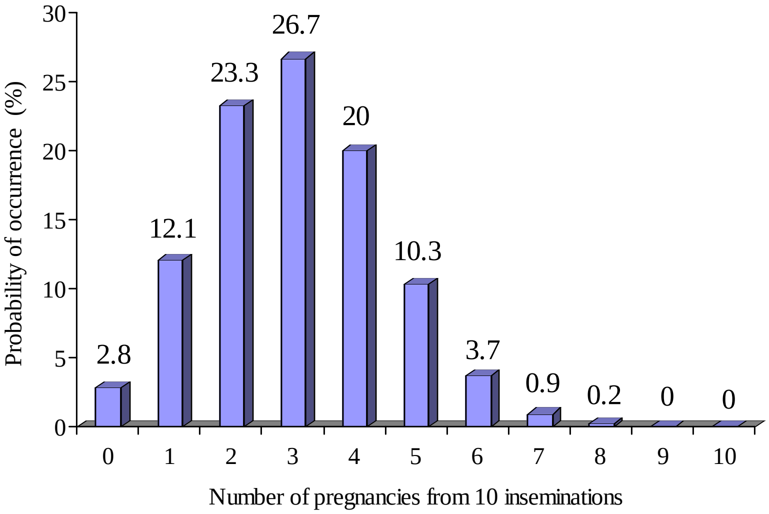

Figure 2. Probability of achieving various numbers of pregnancies as a result of use of 10 straws of semen from a 100% conception rate sire when used in a herd with a 30% true female fertility potential.

The only way to overcome the effects of extreme deviations from random chance is to take lots of chances. However with each successive attempt, another potential outcome is added to the matrix, which proportionately reduces the probability of occurrence of each individual outcome. If we expand the example to 100 services (marble selections), there are now 10,000 possible combinations of outcomes with 101 possible answers. The probability that the observed result for this 100% conception rate sire will be "close" to 30% increases with increasing numbers of services, but the probability that the answer will be "exactly" 30% decreases to only 8.7%. In this 100 service example, binomial probabilities predict that a 100% conception rate sire has 46.2% probability of achieving $\leq$ 29 pregnancies and a 45.1% probability of achieving $\geq$ 31 pregnancies.

The point of this extensive discussion of binomial variation is to emphasize that any given "observed value" for a binomial variable is most likely NOT the "real value". Even with very large numbers of services, we cannot always say with confidence that numeric differences among sires are "real" differences. The effects of numbers of observations on binomial confidence intervals in a herd with a 30% conception rate are presented in Table 2. With only 10 services the confidence interval is ±41% making it all but impossible to say that any bull is different than average no matter what the result. To confidently say that an observed 7% difference in conception rate is truly different from average, we must have 300 or more services to the sire in question. Even large herds seldom have a sufficient number of services to a given sire to detect meaningful differences in conception potential.

Table 2. Effect of sample size on minimum detectable difference and range in conception values that cannot be statistically be considered different than average conception (2-tail test, P = 0.05, 80% power) for a herd with a 30% theoretical average fertility rate.

Number of services per sire	Minimum detectable difference from average	Range in conception rate not different from "average" value
10	±41%	-11% to 71%
50	±18%	12% to 48%
100	±13%	17% to 43%
300	±7%	23% to 37%
500	±6%	24% to 36%
1000	±4%	26% to 34%

Assigning "blame"

Producers and AI center sales personnel often print some type of breeding summary sheet from the on-farm record keeping system or enter fertility discussions attempting to assign blame and credit in this direction or that. However as previously mentioned:

Observed fertility = (male's fertility) x (female's fertility)

In the previous section we "fixed" female fertility at 30% and sire fertility at 100% to illustrate the impact of binomial variation on the observed fertility outcome when all other things were equal. In those examples, we were confident that the ONLY thing being measured was the effect of binomial variation. In the real world there are no 100% conception rate sires and all other things are never equal. Thus, in the real world the observed outcome in a 30% conception rate herd is the product of the "average" sire fertility times the "average" female fertility. A 30% herd conception rate can be achieved by any of the combinations of male and female fertility listed below. It's also important to note that the lowest fertility group establishes the maximum for the herd. For instance, if female fertility drops to 25%, the maximum fertility potential of the herd will be 25% and this will only be achieved if the producer can magically find some 100% conception rate sires that really do not exist. Otherwise, sire conception would have to be >100% to achieve >25% herd conception.

Table 3. Combinations of male and female fertility that yield a 30% herd conception rate.

Male fertility (%)	Female fertility (%)	Herd fertility (%)
100	30	30
80	38	30
60	50	30
40	75	30

Table 3 above refers to “average” conception for males and females although we know that not all males or females have the same probability of conception. Within herd, the female probability for each insemination varies based on factors such as parity, milk production, days in milk, cow health, accuracy of estrus detection, etc. and the male component can vary based on semen quality, sperm numbers, semen handling, technician competence, timing of insemination, etc. With each insemination there are a plethora of factors influencing the probability of success, but only 2 possible outcomes: successful or unsuccessful. Consider two adjacent cows in a free-stall barn to be bred on the same day. Due to the factors above, the probability of success for cow A might be 0.45 for female fertility and 0.7 for male fertility (0.315 probability of success), whereas cow B is 0.3 for female fertility and 0.8 for male fertility (0.24 probability of success). For each service, how do we determine who deserves the most credit for the observed outcome: the male, the female, or binomial variation?

As discussed and illustrated elsewhere (Amann and Hammerstedt, 2002; Amann 2005), the inescapable fact is, female fertility has a profound impact on what is observed. This is readily apparent by use of the same frozen semen in heifers or lactating cows, or in well managed vs. poorly managed herds. It is common to assume that the “cattle” in different herds differ in potential fertility, given that summaries report great herd-to-herd differences in fertility. However is this logical or are the females convenient “scape-goats” to shift attention away from people problems?

Females in most herds have been bred for several generations using AI and 80% are progeny of the same 200-300 bulls and are mated to a common population of sires. Note that as average fertility of the US cow population declines due to production and other stresses or management, the estimated fertility of bulls also declines. Through no fault of the bulls, they might not look as good today as they did 20-30 years ago. Additionally as female fertility declines, the ability to accurately assign the proportion of conception potential due to sire fertility also declines resulting in smaller deviations in sire fertility estimates. Given the commonality in genetics of females and sires used, it seems that people and cow health, including stress from management, nutrition, lactation, housing, and disease, are the major causes for disparity in observed pregnancy rates among herds.

ESTIMATES OF SIRE FERTILITY

“Estimated sire fertility” is best defined as a value, based on mass population data, summarized to include numerous “corrections” for parity, milk production, season, etc. and usually expressed as a deviation from an overall population mean. Systems to estimate sire fertility typically involve large datasets obtained from dairy records processing centers. Through use of complex statistical models, these systems attempt to simultaneously distinguish the magnitude of the male and female components to observed fertility. The “units” of a deviation are often implied to approximate 1 percentage point change in the proportion of animals pregnant at approximately 60 days after AI. However, most estimates actually are based on non-return rates which typically contain considerably more error than actual conception rates.

There is no doubt that the true fertility of sires in an AI organization differs and, hence, average conception rates obtained with their semen will differ. Despite the best efforts of a quality control program and culling of semen of poor quality, differences in sire fertility will be detected in summaries of reproductive data. The important questions for a producer, and also the AI center, are: (1) is the observed difference real or does it reflect an error of measurement in the estimation; (2) does the detected difference have a biological impact in a producer's herd; and (3) does the detected difference have economic consequences on a producer's profit.

To precisely and accurately measure apparent fertility one must have: (1) a common unit of measure, such as how the females are judged pregnant and the interval after AI when the observation is made; (2) no errors in recorded data pertaining to the insemination, such as wrong date, bull, or female; (3) no errors in measurement, be it detection of females returning to estrus, palpation per rectum, or ultrasonic evaluation; (4) accurate and complete records, without biases from culling certain animals or undisclosed preferential treatment of individuals; and (5) a sufficient number of "complete records" (last calving date, date of AI, parity, milk production, sire ID, outcome status, etc.) to adjust for environmental effects and to overcome the effects of random chance and binomial variation. Common sense and random checks cause one to conclude that each of these sources of error permeate any data set. Hence, summaries of such data are approximations with a "measurement error" totally independent of any uncertainty associated with binomial variation or the model for statistical analysis. Much attention is often given to the statistical model, but the impact of measurement errors tends to be unrecognized.

Within herd, conception rates are also influenced by environmental factors such as: season of breeding, parity of cow, days in milk, service number, milk production level, cow health status, technician, basis for decision to inseminate (standing heat, rubbed tail paint, timed-AI), etc. Estimates of sire fertility are calculated using complex statistical models in attempt to account for as much of this environmental "noise" as possible. However, just because a sire fertility estimate includes a given parameter in the model statement, that does not mean 100% of the impact or bias that might be introduced by this parameter has been accounted for. For example, the impact of parity in one herd may be much greater than the impact in another herd. Most models will adjust this based on an average impact, which will over estimate the contribution in some herds and under estimate the contribution in others. This can sometimes be overcome by analyzing the effect on a within herd basis (nesting), however this creates another source of error when a given herd has an insufficient number of animals in a given sub-class to accurately measure the impact on fertility.

Unfortunately, there are an abundance of factors that impact fertility potential that are unknown or, if known, impossible to adjust for because data are not available to calculate correction factors. For example, we know there are good strong standing heats and we know that there are questionable heats. Does anyone doubt that some inseminators bias their sire selection among these insemination opportunities and opt for less expensive semen for questionable heats and reserve more valuable semen for insemination opportunities with a greater probability of success? Similarly, some cows are known to be less fertile than others due to postpartum health, which might cause a producer to bias his sire selection decision for more (or less) valuable semen when breeding such cows predicted to be more (or less) fertile than their contemporaries. The Canadian sire fertility estimates attempt to minimize such effects by including the price of the straw of semen used for each insemination in the model. Similarly, the microenvironment within a herd can change from week-to-week or even day-to-day. For example, one week the temperature humidity index (THI) might be 55% and a group of 25 Ovsynch cows that received all their shots are bred to Sire A. However, the next week the THI is 88% and Sire B is used on a group of 50 Ovsynch cows of which only 80% received all of their shots. All of these factors add noise to any estimate of sire fertility and make it impossible to account for all the factors affecting the binomial outcome: pregnant vs. open.

It would be nice if we could simply add up the number of pregnancies and divide it by the total number of services to calculate sire fertility potential. However, ignoring these environmental factors is destined to yield erroneous conclusions. As another extreme but perhaps common example, Sire A is a high-demand, proven sire priced at $50 per dose and is typically used for first or second service only of highly fertile first lactation cows showing standing heat. As he graduated in August, available fertility data are based on inseminations occurring in the months of Sept to April only. Sire B is an unproven young sire whose semen sells for $5 per dose and is typically used on late lactation cows or questionable heats and since he was released in May, most available fertility data are from summer inseminations. Sire B might actually be more fertile than sire A, but we would never be able to determine that from a simple arithmetic mean.

Table 4. Summary of factors contributing to variation in estimated sire fertility.

- Sire and semen fertility
- Binomial variation
- Female/Environmental Factors
    - Accountable
        - Herd
        - Parity
        - Breed
        - Milk production
        - Days in milk
        - Service number
        - Season/month
    - Not adjusted for because of lack of information
        - Quality of heat
        - Health status of female
        - Technician
        - Semen handling
        - Insemination timing
        - Errors in data recording
        - Errors in data measurement
        - Microenvironment
        - Bias in semen usage

From the above discussion, one might conclude: Why bother calculating sire fertility estimates at all? The fact is sire fertility estimates, if calculated properly, do provide a valuable service to both the AI center and the producer. Each system does the best job possible with the data provided to estimate the fertility contribution of each individual male. The unfortunate reality is that far too few industry personnel have been adequately educated on the accuracy limitations of these evaluation systems resulting in the common over interpretation of the biological and economic important of relatively minor numeric "observed differences" (not to be confused with the real differences which we can only approximate based on our observations). The bottom line is that effective and efficient use of sire fertility evaluations dictates a cautious and conservative approach when interpreting the relevance of minor differences reported in sire fertility estimates. This is especially true when the value for the bull of interest is based on fewer than 1000 observations.

METHODS OF EVALUATING SIRE FERTILITY

Technician non-return rates

Professional Technician non-return rates (NR) are the oldest method of evaluating sire fertility. This approach has been around since 1938, long before use of frozen semen or computers. Through the early 1970s, NR served the industry well to detect differences associated with improved procedures to process and distribute semen and to detect bulls with unusually low or high fertility potential. However as the name implies, technician non-return rates simply mean that an animal was inseminated and the professional technician was not called to re-service the animal within a given amount of time, usually 60 to 90 days. Usually, first service data were tabulated separately from subsequent breedings. The NR system assumes that the reason an animal did not receive a repeat AI was that she conceived and remained pregnant through day 60. However, she might have left the herd, been placed with a bull, serviced by another technician, or simply a victim of a poor heat detection. There was no guarantee a repeat service would be recorded even when it occurred. Thus, as often as not, the reason an animal was not re-bred may have nothing to do with conception.

Even when repeat inseminations were accurately recorded, we often know little more about the animal than herd, month of AI, and the technician that bred her. We then are at the mercy of random chance to assume each sire's usage is equally balanced across other factors known to affect conception such as: accuracy of heat detection, parity, days in milk, milk production, etc. It is generally accepted that NR rates are highly prone to bias and error. In fact, many NR systems maintain a 70% average even though we know the average conception rates in today's dairy herds are closer to the 30 to 40% range. Despite these limitations, technician NR rate is still a rapid (5 to 6 months) method to estimate fertility of a newly released sire or detect distinctly subfertile bulls. The outcome from Select Sires' technician NR system was termed Relative Breeding Efficiency I (RBE I).

Estimated Relative Conception Rates (ERCR)

Early computerized record systems were replaced in 1990 by the Dairy Records Management Systems (DRMS) in Raleigh, NC. They developed an estimate termed "ERCR" which is a 70-day, first service NR calculation corrected for numerous environmental factors. As an east coast processing center, most data obtained by DRMS are from east coast and mid-west herds with limited data from western states. Unlike the technician NR systems, ERCR uses on farm records which increases the likelihood of having all repeat services reported. Lack of returns to estrus due to cows being culled is usually easily detected. The ERCR statistical model adjusts for numerous environmental factors influencing conception including: herd, month, year, parity, days in milk, and milk yield. Further data edits exclude herds likely reporting erroneous data to produce a fairly realistic average NR rate. To ensure reasonably accurate estimates, ERCRs are only published on sires with $\geq$ 300 services. To accumulate a larger number of services and thereby enhance accuracy, ERCR calculations are based on a 3-year rolling database. Unfortunately, this inclusion of historic data makes it difficult to detect changes in sire fertility over time.

Recognizing the accuracy limitations in their calculations (to fully account for binomial variation or confounding environmental factors), DRMS suggests that all bulls within ±3 ERCR units of the average value (0) should be considered to be of average fertility, and primary selection emphasis should be placed on other economically important traits. In May of 2006, USDA-AIPL began calculating ERCR using the same model and data as used by DRMS. AIPL plans to continue calculating ERCR with no changes to the model or data sources until a new methodology, currently under development, is completed and validated. Eventually the new system will be implemented and ERCR will be discontinued.

Agritech Analytics

Agritech Analytics (ATA) of Visalia, CA began releasing sire fertility estimates in 2004. Their system uses on-farm data from large herds, most of which are located in California. Similar to ERCR, ATA adjusts estimates for numerous environmental factors. In contrast to ERCR, ATA relies entirely on on-farm pregnancy diagnoses to determine if an animal is pregnant or open. This should enhance accuracy of the measurements entered into the mathematical model. The ATA system also takes advantage of repeat services by including up to 5 services per cow per lactation. A sire will receive an evaluation if he has at least 10 services in each of 10 or more herds. Thus, a sire could receive an evaluation with as few as 100 services, in which case, the accuracy of such estimates should be viewed with extreme caution. The ATA system uses a perpetual database, meaning that once a data bit (1 AI, no matter when) is in the system, it is used in all subsequent calculations unless an editing step deletes the herd. This diminishes the system's ability to detect changes in sire fertility over time.

In addition to printing deviations from the current average, and the associated number of services, ATA assigns 1 to 5 stars to each bull to represent the quintile ranking of sire fertility. The top 20% of sires receive 5-stars and the lowest 20% receive only 1-star. However, producer interests are likely best served by focusing on the deviations and disregarding the quintile rankings as these are likely to give a "perception" of differences in sire fertility where they really do not exist. For example, in the Feb. 2007 ATA evaluation (Figure 3), the 5-star sires ranged in fertility deviations from +1.7 to +7.0, a spread of 5.3 points. Similarly, the 1-star group ranged from -1.8 to -8.4, a spread of 6.6 points. However, the fertility spread from the worst 2-star bull (-1.7) to the best 4-star bull (+1.6) was only 3.3 percentage points; approximately half of that within the 1- or 5-star groups. Exclusive use of the star system would give the perception that a 4-star bull is of higher fertility than a 2- or 3-star bull, which simply is not the case. The total range in estimated fertility is in essence only 10 points, and a range of 6 points encompassed ~90% of all bulls summarized (1367/1559). Based on the magnitude of the deviations and associated number of services, many bulls in the 1- or 5-star groups should be considered "average", as they would be in other ranking systems.

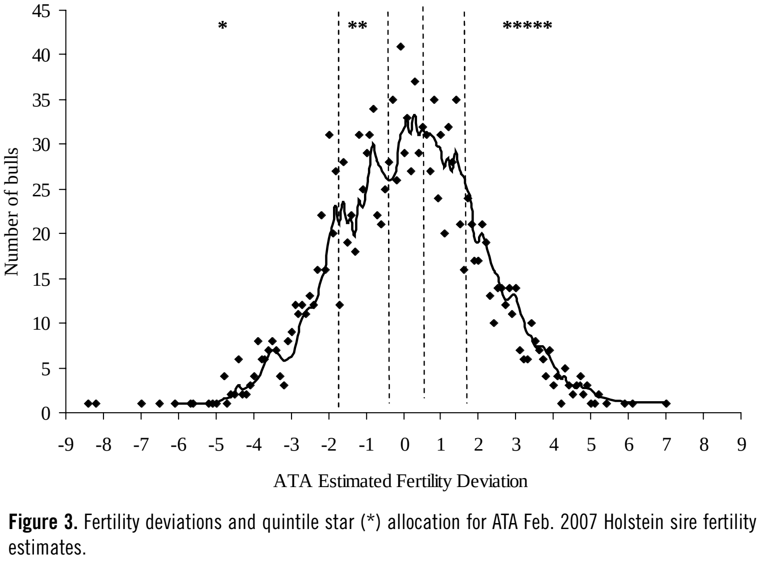

Figure 3. Fertility deviations and quintile star (*) allocation for ATA Feb. 2007 Holstein sire fertility estimates.

Relative Breeding Efficiency II

Select Sires has developed an exclusive, in-house, computerized method to calculate sire fertility estimates refer to **"Relative Breeding Efficiency II" (RBE II)**. Data used in this calculation are obtained from Select Sires' progeny test herds that process records at DRMS in Raleigh; hence, the calculation is biased towards east-coast and mid-west herds. RBE II uses all services between 40 and 300 days in milk to calculate NR estimates of sire fertility. RBE II uses similar data editing procedures and a similar model to both ATA and ERCR which results in a more realistic average NR (~40%). In addition, RBE II includes several parameters we have found to be very important that are not included in other systems. In particular, RBE II adjusts estimated sire fertility for the interval in days since the last service during the current lactation. This is because, a service occurring within 3 days of a prior service (double breeding) has an above average NR rate, though it's likely many of these cows actually conceived to the prior service. More importantly and more common, services occurring 4 to 17 days after a prior service (short heat interval) have NR rates that are significantly below average. By these approaches, the RBE II system presently is the only system that attempts to adjust for the accuracy of estrous detection when estimating sire fertility. To enhance our ability to detect changes in sire fertility over time, RBE II is restricted to the most recent 12 month's worth of data, which generally contains about 1,000,000 services from ~1900 bulls with each sire having ≥50 services retained after editing. Although sires with 50 or more services are included in the analysis for contemporary purposes, fertility estimates are only calculated for sires with ≥200 services, which typically ranges between 700 and 800 sires in each evaluation.

What to do when systems don't agree?

It is no surprise that estimated fertility for certain bulls will disagree across the 3 computer-based ranking systems. As described above, methodologies include/exclude different factors, retain/delete records older than 12 months, and are affected differently by both environmental and binomial variation. Indeed, comparing estimated sire fertility for individual males by across-system paired value correlations reveals that the r-values are on the order of 0.5 to 0.6. This means, that for a typical bull only 25 to 35% of the variation in one estimate of fertility is accounted for by the other. Is this because one system is no good, or that "noise" is dominating the estimates?

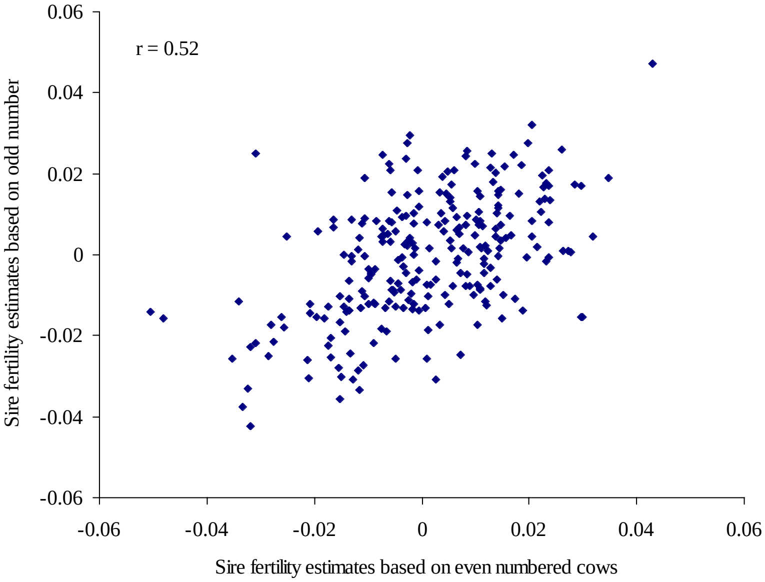

Figure 4. Scatter plot of estimated sire fertility deviations produced from a single dataset when even and odd numbered cows were analyzed separately. Each of the 399 sires represented had ≥200 services in each dataset (odd and even).

To demonstrate the impact of binomial variation, environment, and unexplained variation ("noise") on RBE II estimates, we used the March 2007 dataset and split cows into two groups based on "odd" and "even" last digit in the cow ID number. We then re-calculated individual estimates of sire fertility from the 2 datasets. Even though each dataset included the same herds and same time-frame, the correlation coefficient between the two datasets for 399 sires with ≥ 200 services in each dataset was only 0.52 (Figure 4). Increasing the minimum service number threshold to 500 services in each data set only improved correlations to 0.59 across 127 sires. This exercise demonstrates that computer based estimates of a bull's relative fertility potential have limited value. Approximately 50% of what we attribute to an individual bull, and hence differences among bulls, in reality should be associated with some unidentified factor(s) or simply noise. This does not mean that such estimates are useless and not worth looking at. It simply means **we must exercise a conservative approach when trying to extrapolate the biological and economic impact of sire fertility from minor numeric differences in sire fertility estimates.**

Thus, when estimated sire fertility disagrees between calculations, which one do you believe? **Answer: Believe them all!** Each system does the best job it can with the data included. However, each system has limitations to fully account for confounding variances, errors in measurement, and binomial variation. The real problems only occur when we try to "make too much" out of minor differences in estimated fertility. A bull with a "-2" deviation on one system and a "+1" deviation on another system is NOT really a disagreement as both systems are telling us that this is a sire with average fertility.

COMPOSITE FERTILITY INDEX

Although Select Sires is quite proud of our RBE II evaluation system, we recognize the value in other systems. From the discussion thus far, if you know nothing else, it should be readily apparent that when it comes to sire fertility estimates, ***there is strength in numbers***. In an effort to take advantage of this concept, Select Sire's developed the **Composite Fertility Index (CFI)** which combines the fertility deviations from all available sources and weights each based on the associated number of services. In CFI, a +1 deviation with 1000 services will carry twice the weight of a -1 with 500 services. As in RBE II, a final step in CFI is to convert this weighted deviation into a single index value. When you use CFI, you are using ALL available sire fertility data from all available systems (ERCR, ATA, RBE II).

The top, middle, and bottom 10 bulls for CFI index in the Feb 2007 evaluation are presented in Table 5 to provide an illustration of how the system works. From these examples it's clear to see there often is agreement and disagreement among individual sire fertility evaluation systems and that reliance on any single evaluation system could steer one in a wrong direction. However, combining results across evaluations systems gives greater confidence that estimates approximate the true fertility potential of the respective sires. Select Sires is extremely pleased to offer CFI to our customers to facilitate more accurate fertility estimates and more sound decisions in sire selection.

Table 5. An illustration of the top, middle, and bottom 10 sires for CFI in Feb. 2007.

	ATA		ERCR		RBE II		CFI		
Sire	Services	Dev	Services	Dev	Services	Dev	Services	Dev	Index
1	1260	3.3	1038	4	796	4.1	3094	3.7	4.7
2	1314	3.4	1084	2	1067	5.3	3465	3.5	4.7
3	381	1.2	855	4	901	1.9	2137	2.6	4.4
4	4156	2.8	5729	2	6578	2.5	16463	2.4	4.3
5	8318	3.0	13362	2	8770	2.2	30450	2.3	4.3
6	2147	2.9	1342	1	997	2.2	4486	2.2	4.2
7	18542	2.2	11761	2	5902	2.1	36205	2.1	4.2
8	6532	1.4	3351	3	1662	1.8	11545	1.9	4.1
9	2644	2.1	4326	2	1885	0.9	8855	1.8	4.1
10	2058	1.4	3208	2	2326	2.3	7592	1.9	4.1
11	5898	0.4	5497	0	2118	0.7	13513	0.3	3.6
12	4062	0.9	4537	0	5169	0.1	13768	0.3	3.6
13	1487	0.5	344	0	238	-3.1	2069	0.0	3.5
14	967	-0.5	573	-2	1198	1.1	2738	-0.1	3.5
15	4141	1.0	4979	0	2350	-1.1	11470	0.1	3.5
16	11561	-0.1	3270	1	2480	-0.4	17311	0.1	3.5
17	1003	1.1	1660	0	2594	-0.4	5257	0.0	3.5
18	2407	2.8	5060	0	3011	-2.3	10478	0.0	3.5
19	619	-1.5	746	1	609	-0.5	1974	-0.2	3.4
20	146	1.9	308	-2	716	-0.1	1170	-0.4	3.4
21	705	-3.1	2976	-2	2907	-1.0	6588	-1.7	2.9
22	2148	-0.5	4566	-3	4555	-1.0	11269	-1.7	2.9
23	3726	-0.3	12574	-3	6874	-0.5	23174	-1.8	2.9
24	4131	-1.7	6108	-1	8320	-2.2	18559	-1.7	2.9
25	1316	-2.5	1897	-3	2231	-1.6	5444	-2.3	2.7
26	2233	-1.0	1557	-4	1824	-3.4	5614	-2.6	2.6
27	4171	-2.5	2728	-4	3827	-1.9	10726	-2.7	2.6
28	4620	-2.1	8490	-3	9073	-2.9	22183	-2.8	2.6
29	1608	-2.5	1249	-4	1676	-2.6	4533	-3.0	2.5
30	879	-3.7	708	-3	282	-1.9	1869	-3.2	2.4

HOW TO USE SIRE FERTILITY ESTIMATES

The first step in learning to use sire fertility estimates is to recognize that they are all, "imprecise estimates", especially when based on <1000 services. It can be extremely common place for fertility estimates to change with large changes in the number of services. This is typical for new release sires as their initial estimates may only contain 200 to 500 services. However, as more data are added, the accuracy of the estimate increases which may or may not result in the same value that was obtained in the initial estimate. Many producers will take notice of the sires that start with large positive deviations and then subsequently regress to a lower value, but often make little note of the equal number of sires that increased ranking with increasing numbers of services. Many are tempted to associate these changes with changing semen quality, sire fertility, or altered sperm number dosages and often ignore the most plausible explanation of improved accuracy of the estimate with greater numbers of services.

Sire fertility estimates are of most value to the producer when they are used as a secondary selection criteria after a group of bulls that meet the genetic goals of the herd have been selected. With all other things being equal, the producer would be wise to select the sire with the higher fertility estimate. However, wherever fertility estimates with <1000 services are within ±3% of each other, they should all intents and purposes be considered identical. Estimates with <500 services should be considered "subject to change" with the most likely direction of change being toward "average". In reality, the most important question to ask of the sire fertility estimate is: "Should this sire be considered of normal, acceptable fertility (±3 of average)". If the answer is yes, then sire fertility should be of little consequence relative to the genetic merit of the sire. When sires deviate more than 3% below the mean of all bulls, the producer may then choose to ask: "Do the genetic merits of this bull warrant some limited use despite the potential reduction in conception rates?" **In all cases, avoid the temptation to over interpret the biological or economic significance of minor deviations in sire fertility.**

WHAT ARE "TIMED-AI SIRES"?

With the popularity of fixed time AI programs, many producers often ask: Do some sires perform better in Timed-AI programs than others? The short answer is yes, however the more important question is: Are these also the same sires that perform best when used after detected estrus? The answer is: *There are no scientific data that would give us reason to suspect that sires will somehow magically re-rank themselves in fixed time AI programs as opposed to an estrous detection based program.* In fact, there are data to suggest that if a producer really wants calves from a below average fertility sire, Timed-AI may be the best option because it minimizes differences in sire fertility as opposed to use in estrous detection based AI.

In a landmark study, Macmillan and Watson (1975; Figure 5) evaluated the interaction of sire fertility group and timing of AI on non-return rates. When cows were inseminated very early in estrus, the above average fertility sires were basically unaffected however the non-return rates of average and below average fertility sires dropped considerably. As insemination occurred closer to ovulation, the fertility of the average and below average fertility sires climbed to the point that they were basically indistinguishable from the above average fertility sires. These data suggest that sire fertility may be largely a function of sperm longevity in the female reproductive tract. Thus, if insemination is occurring at long intervals prior to ovulation, the above average fertility sire would be less likely to result in reduced conception due to a longer viable life span of sperm. In contrast, these data also imply that if AI is occurring near to the time of ovulation, sire fertility/sperm longevity is less of an issue. Thus sire fertility estimates based exclusively on Timed-AI will likely exhibit a narrow "range" in deviation estimates, but little to no evidence to suggest the sires will be "re-rank" themselves. The bottom-line is, these data suggest an above average fertility sire is above average fertility irrespective of insemination timing.

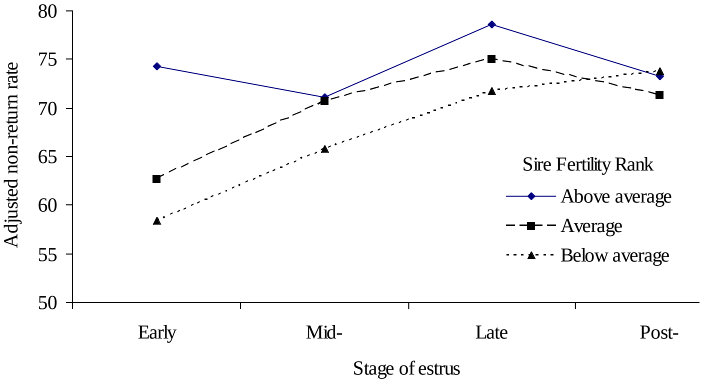

Figure 5. Effects of sire fertility group and stage of estrus at insemination on non-return rates. (Adapted from Macmillan and Watson, 1975)

To extrapolate the results of Macmillan and Watson to timed AI in today's Ovsynch programs, we must first ask the following question: When is the interval and (or) variance in time from AI to ovulation the greatest: estrous detection based programs or Ovsynch? Numerous studies indicate 80+% of Ovsynch treated cows will ovulate within a 6 to 8 hour window. That's a feat that not even the best of estrous detection based programs can "hold a candle to". In a twice a day estrous detection programs, we start our estimate of time to ovulation with ±6 hours of error. In the typical tail-painting program, we are starting with ±12 hours of error. Add to that the 25 to 30 hour interval from estrus to ovulation for those cows that "read the book". In addition, ~30% of the cows exhibiting a natural heat do not read the text books and will ovulate somewhere outside this 25 to 30 h window. We will be hard pressed to find the herd where estrus detection is a more accurate predictor of ovulation and insemination timing than is Ovsynch.

Regardless of how sires are to be used (estrus or TAI), when any 2 sire fertility evaluation systems are in disagreement, producers should be careful not to place too much credence in one system over another. This is the advantage of Select's Composite Fertility Index (CFI). Select's CFI is one stop shopping for sire fertility data from ERCR, ATA, and Select's in-house RBE II evaluation.

SUMMARY

Sire selection and quality control programs in the AI center have the goals of (1) ensure production and marketing of straws containing highly fertile sperm in sufficient numbers to maximize pregnancy rate obtained with each sire; and (2) minimize differences in fertility among sires whose semen is released for sale. Unfortunately, we probably will never be able to measure fertility more precisely than ±3 percentage units, because of the many factors beyond anyone's control including binomial variation, environment, measurement errors, and bias in semen use. Failure to recognize limitations in any estimate of fertility potential often results in over interpretation of the biological and economic significance of small differences among sires in apparent fertility. Available data show that major AI centers do an admirable job in achieving the goal laid out above. By any measure, ~90% of sires marketed achieve fertility deviations within ±3% of average. Select Sires' Composite Fertility Index (CFI) combines data from ATA, ERCR, RBE II, and other sources, to enhance the accuracy of sire fertility estimates by basing them on more services. The CFI system simplifies interpretation of estimated sire fertility by providing a single value representing the weighted deviation for all available systems. Use of CFI should allow our customers to make informed and sound decisions in respect to fertility in the sire selection process.

Acknowledgements

The authors greatly appreciates the review and constructive critique provided by each of the following reviewers, which greatly enhanced the value of this manuscript: Dr. Rupert P. Amann, Professor Emeritus, Colorado State University; Dr. Richard G. Saacke, Professor Emeritus, VA Tech; Dr. Duane Norman, USDA-AIPL; John Clay, Dairy Records Management Systems; and Chuck Sattler, Select Sires, Inc.

REFERENCES

Amann, R. P. 2005. Weaknesses in reports of "fertility" for horses and other species. Theriogenology 63:698-715.

Amann, R. P. and R. H. Hammerstedt. 2002. Detection of differences in fertility. J. Androl. 23:317-25.

Bearden HJ, Hansel WM, Bratton RW. 1956. Fertilization and embryonic mortality rates of bulls with histories of either low or high fertility in artificial breeding. J Dairy Sci 39:312-318.

den Daas JHG, de Jong G, Lansbergen LMTE,van Wagtendonk-de Leeuw AM. 1998. The relationship between the number of spermatozoa inseminated and the reproductive efficiency of individual dairy bulls. J Dairy Sci 81:1714-1723.

Filseth O, Komisrud K, Graffer T. 1992. Effect of dilution rate on fertility of frozen bovine semen. In: Proc XII Intl Cong Reprod and Artif Insem. Hague: Vol 3:1409-1411.

Kidder HE, Black WG, Wiltbank JN, Ulberg LC, Casida LE. 1954. Fertilization rates and embryonic death rates in cows bred to bulls of different levels of fertility. J Dairy Sci. 37:691-697.

Macmillan, K. L. and J. D. Watson. 1975. Fertility differences between groups of sires relative to the stage of oestrus at the time of insemination. Anim. Prod. 21:243-249.

Saacke RG. 1998. AI fertility: Are we getting the job done? In: Proc. 17th Tech. Conf Artif Insem and Reprod, Natl Assoc Animal Breeders, Columbia, MO, p. 6-13.

Salisbury GW, VanDemark NL. 1961. Significance of semen quality. In: Physiology of reproduction and artificial insemination in cattle. 1st ed. W. H. Freeman and Co. San Francisco, p. 359-379.

Sullivan JJ, Elliott FI. 1968. Bull fertility as affected by an interaction between motile spermatozoa concentration and fertility level in artificial insemination. In: Proc VI Inter Cong Anim Reprod and Artif Insem. Paris: Vol. 2:1307.

van Giessen RC, Zuidberg CA, Wilmink W, Veene W, den Daas N. 1992. Optimum use of a bull with high genetics. In: Proc XII Intl Cong Anim Reprod and Artif Insem. Hague: Vol 3:1493.

Vishwanath R. 2003. Artificial insemination: the state of the art. Theriogenology 59:571-584.