



Suite 6, 146 Marsh Street,
Armidale NSW AUSTRALIA 2350

email office@wagyu.org.au
phone 02 8880 7700

Wagyu Webinar #4: Carcase Data to EBVs. Professor Rob Banks, Dr Yuandan Zhang and Dr Matt McDonagh

PRESENTED BY PROFESSOR ROB BANKS, DR YUANDAN ZHANG AND DR MATT MCDONAGH

Matt McDonagh

Welcome everybody, today's webinar is about carcase data to EBVs and really how it is used in BREEDPLAN. We will have three speakers, President of the AWA Mike Buchanan to introduce the subject, and then Professor Rob Banks and Dr Yuandan Zhang of AGBU to talk us through a demonstration of carcase data to EBV calculations and then I will give a quick summary of data trends and takeaway messages.

This is the second webinar in the EBV series, we have been working with Neogen to bring this information, who are service provider partner with the AWA to bring members services in genetic testing.

[Promotional video of Neogen]

Mike Buchanan

Welcome everyone to the second session on Data to EBVs, I'm Mike Buchanan, President of the AWA. A special welcome to Professor Banks and Dr Zhang of AGBU, our science partners, and Dr Mat Wolcott. The carcase EBVs we are talking about were released back in 2014 and at the heart of most Wagyu breeding improvement in Australia and highly influential in Wagyu breeding around the world outside of Japan. At six years of age, we believe they are ripe for review, which may lead to more webinars in the next few months. In this particular introduction, I want to cover how and why we are reviewing, then Matt McDonagh will provide some analysis of our data and then hand over to Professor Banks and Dr Zhang to explain how it works.

To start with the obvious, we all know that the measurement and targeted improvement of a wide range of livestock traits is what EBVs are intended, to optimise breed development. We also know there is a parallel reality. Every year, millions of dollars are invested on the basis of Wagyu BREEDPLAN values, and nothing is more important in this market than the values for carcase traits in Wagyu BREEDPLAN. These EBVs and their downstream Indexes represent core foundation data for seedstock pricing. Someone has to take responsibility for Wagyu BREEDPLAN quality data. AWA is the publisher of Wagyu BREEDPLAN, and it therefore must take responsibility for making sure that this data is accurate. The final responsibility rests with the Board to ensure that this foundation to Wagyu industry values is as good as it can get. The buck stops with the Board, and you will see some of the faces of the Board onscreen now. You can see them all on the website, and all are volunteers for the advancement of the breed.

The Board commitment and mission statement to BREEDPLAN, 'the AWA has a duty to review and continuously improve Wagyu BREEDPLAN for ongoing breed development'. The single objective is to get it as right as possible, so that the money invested can be done so with confidence.

We have a duty to review - how do we do that? We don't have a bunch of PhD scientists on the Board for a complex

ABN 35 0037 00721



Suite 6, 146 Marsh Street,
Armidale NSW AUSTRALIA 2350

email office@wagyu.org.au
phone 02 8880 7700

field. An adversarial review would involve employing outside experts to review, and that's really not the way we do things. We work cooperatively with AGBU and grateful to the team for delivering what we need and that we have proper explanations for EBVs and Indexes.

We should also note confidentiality issues. At no stage does the Board have access to member data, anything submitted as confidential remains so. The Board does need data as part of the review, but is made available as anonymised extracts, or summarised generated by AWA or AGBU.

Does the Board have real skin in the Wagyu BREEDPLAN themselves? The answer is that yes, without a doubt, and we can demonstrate. Looking at the top 500 sires list, which is public data available from the website.

To be on this list, the animals require high accuracy EBVs, more than 10 registered progeny and so on. For any breeder, it takes long term commitment to capture and submit data, and then again with progeny data over a lengthy period. To check this, try searching for sires as distinct from males in the Wagyu BREEDPLAN database.

The dark green cell at the top indicates that the data has been sorted based on the Self-replacing Index, the right cells indicate three directors with the most top 500 published sires at the time of this search. GIN is Ginjo Wagyu, which is me, the result for me is based on more than 10 years of data submission into BREEDPLAN. To get a broad picture, you can see those three directors have bred and represent, around 8% of the top sires list. So yes, the AWA Board does have skin the game and commitment which goes beyond donating time and effort, and we definitely use Association tools to improve Wagyu. We have already added new tools this year with the Wagyu Breeder \$Index and the prefectural analysis and there are big plans for the year ahead. I will now pass you back to Matt McDonagh.

Matt McDonagh

Thanks Mike, to introduce Rob and Yuandan, who have been working with us for the past six months, all of which can't be produced for you today, we have had a lot of meetings with the Board, GIC and AGBU to strengthen our knowledge of how the data contributes to EBVs and it is a real pleasure to get to this point. Rob and Yuandan will be walking us through their demonstration of data to EBVs by way of carcass data examples.

Rob Banks

By way of introduction, I will do the introduction and summing up while Yuandan will go through the detail in the middle. I would also add, that as Matt said, we have been working with the AWA for the past six months to bring you this information, but AGBU has been working on Wagyu data for around six years through a number of projects including the development of Wagyu BREEDPLAN. The past six months has been indepth R&D topics that have been identified by the AWA, drilling into them with more detail and providing reports and interaction with the Board and the AWA. To reinforce Mike, today's session is about how carcass data is turned into EBVs and follows on from our session a couple of weeks ago, which is more general about data to EBVs. Wagyu is somewhat unique in the amount of carcass data that is collected in Australia and obviously it is extremely important for evaluation, and is unique in Australian breeds in that sense. Carcass traits are important to profitability value chains, that collection of data is strategically important for your breed and getting the maximum value out of that data collection, which we will touch on in our discussion.

Wagyu in Australia is doing a good job of collecting and using this data, and as a quick summary line, a very high proportion of it is useful to producing high quality EBVs. We will go through how that is generated and a few points

ABN 35 0037 00721



Suite 6, 146 Marsh Street,
Armidale NSW AUSTRALIA 2350

email office@wagyu.org.au
phone 02 8880 7700

that are worth thinking about afterward.

We are talking about taking data and generating EBVs and the reason is because we want to identify the value of the genes of each animal and the data that is collected on each carcass is one of the clues that we have to identify the genetic merit – what are the genes worth in the traits we are interested in. The ultimate goal is to come up with an estimate of how good those genes are – the EBVs.

The basic process of taking data to create EBVs is relatively simple, and summarised here. The first step is to remove the extreme values and by that we mean those figures that are extremely unlikely. An example may be if there is a record where the marbling score is shown as 29, we would know that there is likely to be an error and it will be checked. Checking those extreme values is to create 'clean' data.

The next process is to taking account of non-genetic factors that are affecting records, and we do that within contemporary groups. Contemporary groups are a group of animals that have been managed together in the same place and same time. We regard them as a fair group for comparison as they have basically been in the same environment, and within that group we will take into consideration any non-genetic factors that we have identified through research that has an impact on the trait, such as age and weight. Age is important, for some traits, the age of the mother is important.

By looking at how animals behave in a contemporary group, and then taking out the non-genetic factors, we try to get to as clean and fair as possible, a result that is an observation for EBVs. The EBVs are generated by combining the adjusted observations from different traits and related animals, and we know the relationship between animals through pedigree and genomics, and we know the relationship between traits because of the shared genes in common. We take account of those two factors to combine data with animals with its known relatives and traits and the end result is the estimated breeding value.

Some of the terms shown here, will be shared, so this is a glossary of terms we will use today. Yuandan and I will try to repeat what these terms mean. The most important two, which would be the least familiar, is the coefficient of variation, which is often abbreviated to CV, and measures the variation/diversity of a trait in a particular group of animals. The other is contemporary group described as CG or cg, which tells us about the group of animals running together.

In this example, we show how data is used for carcass traits to generate EBVs – we using actual data from the Wagyu database. We have looked at, starting with a raw phenotype or raw measure, then sequentially taking account contemporary group and other non-genetic factors and seeing what effect that particular record has on the ultimate EBVs produced for that animal. A significant part of that is to extract a subset of the data to remove a part of the animal's own records to see how the phenotype compares to their own EBV when the record is not included. Similar to checking the progeny to compare against EBVs. We will work through that process today.

Yuandan Zhang

Thank you Rob.

The data we will use for this demonstration is based on the carcass traits. From this slide, you can see we have four carcass traits – carcass weight, eye muscle area, P8 fat depth and marbling score (sometimes referred to as carcass intramuscular fat CIMF). With the data set, the first column is the total records for each of the traits, so for carcass

ABN 35 0037 00721



Suite 6, 146 Marsh Street,
Armidale NSW AUSTRALIA 2350

email office@wagyu.org.au
phone 02 8880 7700

weights we have 9,922 records and so on. The second column is a subset of the total, where the phenotype of the animals for the past 3 years has been removed to do the BREEDPLAN evaluation. The third column is the difference between the two, and is treated as the test animal because the phenotype is now removed, this information is then used to predict the EBV.

With the data adjustment, for carcase weight, it is relative simple, we adjust for two factors, the contemporary group which refers to the animal's environment, secondly adjust for age/carcase weight as it affects the other carcase traits.

Listing the adjustments, we start with the raw phenotype as supplied by the AWA database, then 2 or 3 levels of adjustment. The first is for the contemporary group, removing the non-genetic effects. The 2nd adjustment for contemporary and age or carcase weight. The 3rd adjustment is contemporary group, plus age, plus carcase traits.

These adjusted phenotype values are then submitted to BREEDPLAN evaluation. First, the data is phenotype values and treat it as the EBV estimate (EBVe). The different ways to do that evaluation is with genotype and phenotype is treated as a genomic EBV or alternatively the pedigree without the genomic information.

For the test animals we remove the phenotype which are treated as low, to predict the EBV. Again we can use the genotype information or the pedigree information. These become the test EBV tests, EBVt.

These terms are used for comparison. For progeny prediction we can also use the calculations from the parents' EBVs, known as the middle parents EBV (EBVm).

After calculating those values we do the comparison based on phenotypes and the EBV to find the correlations, to see if the different values are associated and compare them based on the different adjustments for the test animals compared to the full database.

We use the correlations to display some key points. The correlations table show the BREEDPLAN evaluations for EBVs. The table shows the levels of adjustments for the animals, and the predicted values for the test animals. The variance is also given to see what the spread is in the calculations. If the correlation between two measures is 0, then it is said to have no correlation, a value of 1 is perfect correlation. At >0.5 there is a high correlation.

The correlations chart with estimated EBVs for test animals, shows the carcase weight traits, from the left we have the raw phenotype, then the adjustments (contemporary, age, weight), then the last 2 bars shows the EBV with and without phenotype, and the mid-parent value.

The raw phenotype correlation shows a good relationship that increases with the adjustments. Adjusting for contemporary groups makes a change, but adjusting for age and carcase weight does not significantly change the results. The same graph for eye muscle area, the contemporary group adjustment made a significant difference. The correlation between the phenotype and the EBV in the test group is high, which means we can demonstrate the relationship with other relatives to predict the EBVs with higher accuracy.

The P8 fat depth, adjustment works in the contemporary group, but for marbling score the adjustments don't really change.

The charts looking at predicted EBVs for test animals correlation with and without the phenotype data, so it is based on their relatives information. Even without the phenotype we can get an EBV with relatively high correlation.

ABN 35 0037 00721



Suite 6, 146 Marsh Street,
Armidale NSW AUSTRALIA 2350

email office@wagyu.org.au
phone 02 8880 7700

Rob Banks

It is important to note that the Wagyu data set, but because there is quite a lot of carcass data in the data set, animals have plenty of relatives across different contemporary groups, which is why this chart on the right, can go where there is no data for your own animal, you can have a moderate prediction when you incorporate information from relatives, giving an accuracy up around 80%, as there is lots of information that can be drawn upon even without your own data. It is a nice feature of the Wagyu data, that there is a good amount of data for all carcass traits and good pedigree connections. This means that animals that do not have data, can still have useful EBVs. One of the factors impacting the correlation between the eye muscle area and the EBVs is that there are a couple of measurement sites for eye muscle area depending on the processor, so the correlation between an animal's own record and the EBV is not quite as high, given the different site. The different site location is accounted for when we calculate the EBV, but not at the raw phenotype.

Yuandan Zhang

This chart shows the EBV accuracy for test animals in different cases. We tested based on whether the phenotype was included or not and then the genotype included or not. even without the phenotype information the genomic accuracy increased slightly. Not all animals had both genotype and phenotype so it is a relatively ideal case (on the left).

Rob Banks

A general point to make here, which relates more to evaluating animals that go on feed, so young animals without a phenotype. A genotype on a young animal the EBV you will get for marbling will have an accuracy of around 53% even without having a record itself which is a useful basis for identifying poor performers, or those that won't do as well. The third block of bars, breeders may think in the future, basically tells you the accuracy of the EBV if you genotype at birth. Clearly a newborn calf does not phenotype information, but as more carcass data increases in the breed, we can expect that the predictions on young animals will be more and more accurate.

Yuandan Zhang

The genomic vs non-genomic chart looks at the impact of the phenotype recording on accuracy. Left is with the phenotype with non-genetic factors, there is a slight increase, but for non-phenotype information, it is not as great.

We look at the confounding issue between the sire group and the contemporary group. For the eye muscle area, we have 9 sires, from the same contemporary group, those will affect the usefulness of the data. To demonstrate, the magnitude of the confounding group has more than 9000 records, but less carcass weight records, the confounding effect is less than 5%. Plotted on a scatter graph, in some cases the number of animals is around 10, but the effectiveness is near zero because the number of progeny in the contemporary group is low due to confounding in the contemporary group. For the other carcass traits, particularly marbling score, the confounding effect gives a small contemporary group. Looking at the magnitude for marbling score the number of records is extremely small, around 15%, makes the data not very useful.

Rob Banks

The general principle, as we go from raw phenotype data through the adjustment process (contemporary and age/weight), we end up with something less closely similar to the raw phenotype and more related to the ultimate

ABN 35 0037 00721



Suite 6, 146 Marsh Street,
Armidale NSW AUSTRALIA 2350

email office@wagyu.org.au
phone 02 8880 7700

EBV of the animal, which is what the process is designed to do – get rid of the effect of non-genetic effects.

The power of breeding prediction values for progeny is better as we improve adjustments for non-genetic factors, and improves again as we take in pedigree information through the relatives. It is clear in this analysis, and what we have done with other traits, you will always get more accuracy if you have phenotypes, and the genomic information is adding more power than pedigree along. So phenotype information is good, but genotype is very helpful.

The point about genomic breeding values, it is a better predictor for EBVs than pedigree for the traits we have looked at, but it does depend on the number of animals with genotype and phenotypes or the reference population, it also depends on the genetic heritability – how the genetics influence it, and also the genetic structure of the breed. The Wagyu pedigree structure for the breed is actually very helpful for the genomic breeding values. Effectively there is a lot of connection across the herds, due to the small number of ancestor animals. The coverage of genotypes is also very good.

The two things I hope you get out of today is firstly that the process of going from raw data to EBVs is a systematic one and essentially, we are removing non genetic effects in order to get a better picture of the genes and that is working nicely for these carcase traits. It doesn't work so obviously for marbling, and one reason is that once the animal gets to the age where marbling data is seen, there is less influence from factors such as age and other non-genetic factors. Good carcase data underpins good EBVs and relatively easy to do. Having said that, some of the data in data set is not contributing significantly. It is to do with the number of Sires we see in a group, and the variation within that group. The problem arises particularly with small groups where one sire is represented, or little variation in the trait of interest – marbling is a good example. So there is a subset within the Wagyu data that is not contributing very much at all to the EBVs of the animals or their relatives, and it is about 15% of the data. Some of those animals might be in a management group where they have been pulled out for slaughter because they have reached a particular endpoint, it makes sense for commercial applications, but limits the data for that particular group. It doesn't seriously impact overall on the data, but there is a subset that does not really contribute.

We will continue to look into ways of how to record data, assign animals to contemporary groups and recommendations are communicated more effectively. Please don't take the thought that there is a problem with the EBVs, they are working very very well. It is just a proportion of the data is not contributing. Defining management groups based on a small number of animals and slaughtering on the same day because they have reached some marbling for example, is not actually very helpful, it would be better to include the others who may not have reached that point where possible.

Matt McDonagh

Thanks Rob, for walking us through that. Harvesting: some people think it is where you get a dataset and only submit the high marble score data from that set, thinking that is of benefit. That is not what harvesting is. Harvesting is taking out the group of animals that are ready to go for slaughter. Effectively making the group small, single sire groups are a problem. The issue of submitting only high marble score animals from a larger dataset, only ends up hurting the person submitting the data, as it will come back with zero genetic variation, and we can't determine the differences between the animals, and the net affect is zero on the EBVs. We need all the variation in the dataset to get the genetic merit.

I would also like to highlight, that the AWA is quite unique in that we pay BREEDPLAN on behalf of our members for

ABN 35 0037 00721



Suite 6, 146 Marsh Street,
Armidale NSW AUSTRALIA 2350

email office@wagyu.org.au
phone 02 8880 7700

data submission, totally free to be a member of BREEDPLAN, which would otherwise be a cost to your business.

Dr Matt Wolcott will also be available at the end of today to answer any questions, but in the last webinar he gave an update on data audits and data quality for carcasses.

One of the points he raised last time, is that it can look like there is not a lot of data for carcass traits compared to the total, we actually get a high level of data contribution in relation to carcass traits. Getting carcass data is a difficult and sometimes expensive thing to do, and Wagyu seem to do it better than other breeds. As a portion of the records we need to drive a good BREEDPLAN we are doing a good job – and these are Matt's slide, not mine.

Members are doing a good job, and genomics presents an opportunity to use that high quality data because of the link to pedigree based on a small group of founder animals.

Matt also showed us that the quality of the carcass data is very high, so while some small contemporary groups are not used, or single sire groups, the quality score of 80, the cut off, you can see that carcass weight, eye muscle area and IMF and camera traits are high quality data. Good investment in people's time to capture that data as it has high economic value, and good impact within BREEDPLAN.

With Yuandan's slide on eye muscle area, I want to highlight that it is very difficult to correlate eye muscle area data with an EBV, and Rob did explain that eye muscle area can be measured at the 8th, 10/11th, or 12/13th rib sites, huge variation between processors on where that data is measured. The raw correlation for the EBV is difficult. But if you compare within a contemporary group, comparing within 5/6th rib site on the same day, then the genetic variation correlation compares well within that contemporary group. And why the eye muscle area is atypical correlation variations. With the carcass weight the correlation with EBVs is actually pretty good as most would grow the animals on average 900 - 990 days, so they have maximised their opportunity to show their genetic potential.

There is a relatively high starting point, but including the contemporary group tightens it up further. So really that is what the contemporary group is really doing, is taking into account a small effect. Matt will give us another update on genetic trends for economically important traits, and how influential breeders and bulls are impacting the genetics of the breed (all anonymous) – how the breeding direction for members has changed over time. And discussion on how to maximise the recording of data based on what we have learned today. To finish up, the carcass data in BREEDPLAN. We have carcass data on over 500 sires – some sires have more data than others and how it is distributed.

So you can see Michifuku on the left has the most carcasses submitted to BREEDPLAN, more than 500 for him alone, and a good chunk of those are used in BREEDPLAN (about 490), about 30 were discarded due to some issues with contemporary groups or statistical confounding in that group, for example if all the progeny in the contemporary group were sired by Michifuku – no genetic comparison. For the top 12 sires, the vast proportion, about 95% of the data submitted gets used by BREEDPLAN, so the blue bars on the graph is submitted, orange is data used. Sires 14-26, the numbers are still showing more than 100 carcass records for these animals.

We have the odd animal, all the data had a problem and consequently not used. This is very atypical and there can be some 'noise'. Take care with the data if it is commercially important to you over performance recording, to consider the structure of how your breeding. If you are breeding large groups from one sire, it doesn't matter if it is 100 animals, it won't be of much use. But you can see in this case, the 26th ranked sire still has more than 100 records. Jumping way down the tree to sires 120-139, you can see the animal has 20 carcass records submitted. And

ABN 35 0037 00721



Suite 6, 146 Marsh Street,
Armidale NSW AUSTRALIA 2350

email office@wagyu.org.au
phone 02 8880 7700

you can see that the majority of the records are still being used. But at 140th animal, you can still see that 20 records were submitted, even though the majority were not used in this case. this overview shows that we have a good spread of data, and some of these sires are really quite old, but it still has a really good flow on effect on pedigree, which is why the genomics work so well with the data.

The last bit of news I want to share is a bit exciting. We have been working on a new wagyu progeny test program in collaboration with MLA and producer members, gathering carcass data from industry recording, very valuable and high quality data. We have been working to develop a program that will run for about 10 years, with 20 sires tested per year, hoping to start within this financial year and hoping to achieve 1200 dams per year and over the ten year program gather 2380 slaughter progeny and 2520 females that have re-bred to test the efficacy of those females and their calves. It is a new initiative that the Board is working on at the moment. A great way to assist the data recording efforts of our members and add value, and to test some of the newer sires that are coming through. With Rob's help we can identify some of the new sires early on in life and target those to a program like this to expedite the accuracy and improvement and genetic testing of those animals.

Thank you everyone for listening in.

Questions:

The number of sires we have carcass records for?

RB: Reinforcing Matt's slide showing the sires out to 140 – simple answer must be you have 140 sires with at least 20 progeny and guessing that going to sires with 15 progeny, we could add another 100 sires. There is an enormous number of sires with a useful number of progeny carcasses. As Matt said, that is why the genomic breeding values is working well as there is wide coverage of the breed with good data. Quick calculation, probably is that 6,000 carcass records, half of which is coming from around 140 sires, who have a strong number of progeny with records themselves. It is a good structure for underpinning breeding values.

Q: marble score data – is there a way of improving that data?

RB: there are 2, 3 ways of recording marble score data and how they are related to each other, and they are all useful for generating genetics information.

MM: We know that until recently, there has been more graders comfortable with grading at 9+ and when BREEDPLAN called it 10, it gave it more range. But for Fullblood marbling records, we see an average of 7-8 marble score, but when there is a hard lid sitting on the measurement of 9, we lose a lot of the ability to see the genetics that give higher than a 9 or 10. When we have small groups marbling at 8 or 9, we don't have a lot of power to tease apart the genetics. So the MIJ camera is particularly useful in this case as it has a high correlation with eye muscle area, and the R&D by AGBU the genetic correlation between eye muscle area and IMF is almost 1. The camera does not stop at 9 – it can go out to IMF of 50% equivalent to marble score of 15 and 16. So the more camera data we get in more, the more ability we have to use that data. With the camera data we get both the IMF and marble score, both contribute to the EBV.

ABN 35 0037 00721