An evaluation of the use of video image analysis

to predict the classification and meat yield

of sheep carcases.

A report for the Devolved Bodies, E+V and the Welsh Country Food Group

April 2007

Executive Summary

Background

UK lamb carcases are classified subjectively according to the EUROP grid. Elsewhere, and particularly for beef, video image analysis (VIA) is used to provide a rapid means of automatically and objectively predicting the fat, conformation and saleable meat yield in carcases. VIA can limit any variability that may arise through the use of subjective human evaluation under the present EUROP classification.

The Study

A study using the VSS 2000 video image analysis equipment from E+V GmbH to predict conformation, fatness and saleable meat yield in lamb was carried out at Welsh Country Foods. A range of leg and shoulder presentations representing Industry practice were assessed and carcases across the EUROP classification grid were evaluated.

VIA was assessed against 3 criteria:

1. Professional judgement of acceptable accuracy using the current 5/7 point scale levels as follows:

Conformation	80% total agreement of classification awarded
	99% agreement within one class difference
	100% agreement within two classes difference
Fatness	70% total agreement of classification awarded
	90% agreement within one class difference
	99% agreement within two classes difference

2. VIA to achieve same level of agreement with the expert classifiers as the classifiers achieve with each other.
3. The level of agreement of VIA with the expert classifiers to match the level of agreement between the MLC in-plant classifiers and the expert classifiers.

Evaluations were also considered on the 15 point scale.

Results

Classification

- With regard to agreement between VIA classification and the median classification scores from the expert classifiers on the 5/7 point scales, VIA achieved the percentage targets for accuracy within 1 or 2 class(es) difference for conformation and fatness, but generally it did not meet the accuracy target for total agreement with the expert classifiers. However these accuracy targets were not achieved with respect to total agreement between the expert classifiers nor for total agreement between the MLC in-plant classifiers and the expert classifiers.
- For conformation, VIA did not achieve the same levels of total agreement with the expert classifiers as the classifiers achieved with each other, but VIA met or exceeded the targets for within 1 or 2 class(es) difference. For fatness VIA did not match the levels of agreement between the experts themselves at any of the target levels.
- VIA performed as well as the MLC in-plant classifiers for conformation, but less well with respect to fatness.

Meat Yield

- VIA can predict both meat yield and the weight of primal cuts and the results showed that VIA has greater precision in predicting meat yield than current EUROP class/sub class based systems.

Repeatability

- VIA was more consistent than the expert graders at classifying conformation, but less consistent than the experts on fat class

Conclusion

VIA offers lamb abattoirs a means of predicting meat yield and primal weights as well as providing objective assessment of carcase conformation and fatness. The present study provides data to allow industry to assess the potential that the equipment offers in these areas of operation.

Introduction

Sheep classification is not mandatory in the EU, however in the UK sheep carcases are assessed visually using the EUROP classification grid to establish their levels of fat and conformation. The subjective nature of the current classification can lead to individual anomalies and inconsistencies which are sometimes questioned by producers, hence there is a need to investigate the potential for objective systems which gives a more consistent and independent assessment of sheep carcases.

The potential for employing mechanical methods to classify carcases was first recognised in the 1980s. Denmark led the way with research initiatives to develop Video Image Analysis (VIA) but the technology has now been adopted more widely with a range of manufacturers offering systems for the objective assessment of carcases from beef, lamb, pork and poultry.

VIA provides a means of automatically and objectively assessing the fat, conformation and saleable meat yield in carcases. In general, the technology allows rapid, accurate and consistent assessment of carcases and limits any variability that may arise through the use of subjective human evaluation under the present EUROP classification. In addition to being automatic and objective, VIA operates at line speeds, and can provide accurate assessments of lean yield. This opens up the possibility of producers eventually being paid on the basis of yield.

In the abattoir, the VIA system is integrated into the slaughter line at a place usually near to the scale area. The suspended carcases are illuminated and digital video images of the carcase are captured and processed using specialised software to extract data relating to the carcase shape or conformation. The fat level is determined via interpretation of the colour or gray level across the carcase. The image information can also be used to make predictions on carcase yield.

Following validation work in 2000 and 2001 (Allen and Finnerty 2000, 2001) and official approval in 2004, VIA is now widely used in Eire to evaluate the conformation, fatness and meat yield of beef carcases. The trials in Eire showed that VIA can predict conformation and fat class based on the EUROP grid. However, the level of accuracy appeared to be higher for the estimation of conformation than for fat class, reflecting the greater challenges of determining fat class from the digital images. The ability of VIA to predict lean meat yield in beef was also assessed and data showed that saleable meat yield could be predicted with a high level of accuracy.

More recently work in the USA examined the use of VIA to predict carcase fabrication yield in lamb (Brady et al, 2003; Cunha et al 2004). The study validated the prediction equations and found that both the accuracy and precision of the prediction of bone in cut yields of lamb carcases was improved by the use of VIA compared with current conventional methods (Brady et al, 2003; Cunha et al 2004).

The prediction of lamb carcase grades from images of lamb chops has also been evaluated recently by a group from New Zealand (Chandraratne et al 2006). The study used geometric and texture feature extracted from the chops to predict lamb grade and achieved up to 79.4% overall classification accuracy which increased to 85% when hot carcase weight was added into the equations.

Clearly, VIA can offer considerable benefits to both the producer and processor. Although VIA is used by the meat industries in a number of countries, it is not currently in commercial use in the UK. Consequently, the present study was carried out to evaluate the use of one VIA system to classify and assess meat yield in sheep carcases.

Materials and methods

The project was based at Welsh Country Foods in Gaerwen, Wales and set out to evaluate the VSS2000 Video Image Analysis system developed by E+V technologies GmbH, Germany. Further technical and background information on the equipment is available from the EplusV website for VSS2000 (http://www.eplusv.de/start_E.htm). The project aimed to develop and validate equations which could predict the conformation, fat class and trimmed primal yield of sheep carcases.

Overall the project had 4 objectives.

Objectives

1) To ensure compatibility between the current subjective classification system and VIA.
2) To ensure that operational procedures e.g. suspension method, which may impact on industry uptake of VIA technology, are taken into consideration.
3) To develop an accurate generic prediction equation for conformation and fat class.
4) To demonstrate, based on a specific butchery method, the ability of VIA technology to predict meat yield.

Project Overview

Following the initial installation and setting up of the equipment in the plant, 500 carcases were classified and imaged to allow the development of the classification prediction equations. Then a further group of 800 carcases were classified and imaged to allow fine tuning of the classification and yield prediction equations (200 carcases) and the validation of the classification prediction equations (600 carcases). To evaluate the ability of VIA to predict yield 500 carcases were used.

Suspension and presentation methods

Unlike beef carcases which tend to be suspended as sides through the Achilles tendon prior to classification, sheep carcase suspension and carcase presentation varies within the UK. Some plants use conventional gambrels (legs apart), others use a single hook placing both legs together and some cross the legs (Figure 1). Normally legs are crossed for export, some plants therefore use two methods. There are also two methods of shoulder presentation. Some plants use elastic bands to tuck/hold in the forelegs, in others the fore legs are left un-banded. Consequently, the study design included the various carcase suspension and presentation methods (table 1) in the development of prediction equations.

Figure 1. The 3 main suspension methods used for lamb carcases. Note the shoulders are all banded in these images. The hind legs of the carcases are from the left ; crossed, together and gambrelled

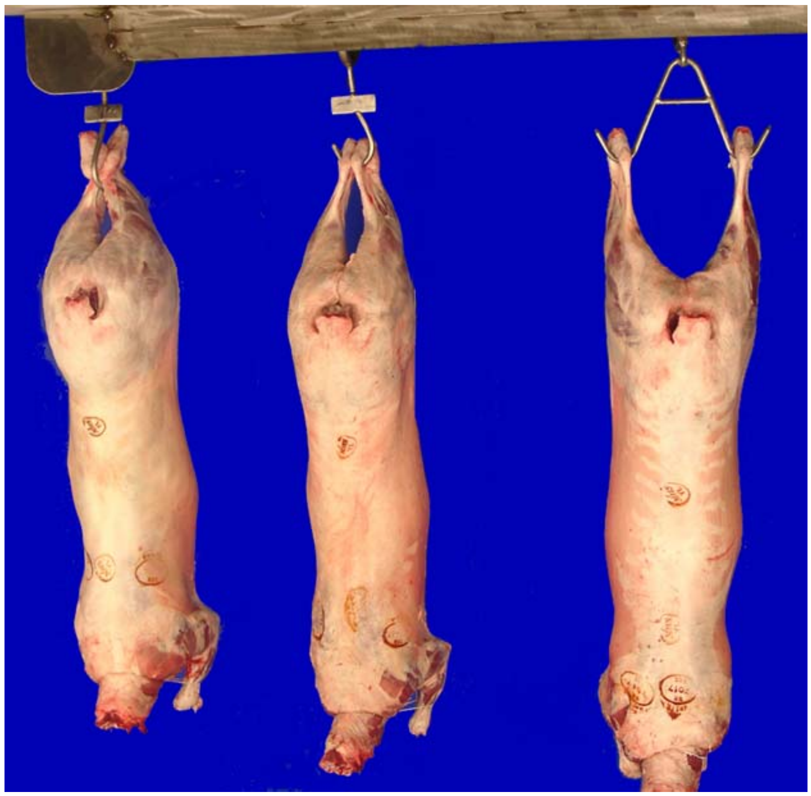


Table 1. The six carcase suspension/presentation combinations evaluated in the study

Shoulders	Legs
Banded	Gambrelled
Banded	Crossed
Banded	Together
Unbanded	Gambrelled
Unbanded	Crossed
Unbanded	Together

Dressing specification also varies across the industry. The present study developed prediction equations which were based on carcase weight exclusive of kidney knob and channel fat .

Development of the prediction equations

500 carcases were selected to develop the prediction equations. The carcases were assessed by a member of MLC's Authentication service on both the current 5 point conformation/7point fat class scale and the 15 point scale for both conformation and fatness. The 15 point scale was used in addition to the 5/7 point scale because the additional sub classes allow industry greater discrimination between carcases. The carcases were assessed by the MLC in-plant classifier on the 5/7 point scale.

The carcase selection was designed to provide as close as practically possible to equal numbers of carcases per classification cell as shown in table 2 below.

Table 2. Development of the prediction equations. The target distribution of 500 carcases within the EUROP grid designed to provide as near as possible to equal numbers of carcases per classification cell

		FAT CLASS							
		1	2	3L	3H	4L	4H	5	Total
Conformation	E	2	7	17	17	19	19	19	100
	U	5	10	15	15	15	17	23	100
	R	5	15	15	15	15	15	20	100
	O	20	15	15	15	15	15	5	100
	P	39	25	10	10	7	5	4	100
	TOTAL	71	72	72	72	71	71	71	500

Provision for specific requirements such as those for Northern Ireland, where fat class 3 is not split into low and high, can be made from the transformation of the prediction equations. Transformations may be made into any scale; transformations into any scale of less than 15 classes can be made whilst retaining accuracy.

The 500 carcases were suspended/resuspended to give the 6 combinations suspension/presentation methods previously outlined (table 1).

The following traits were also recorded for each individual carcase and the analysis included the variance of these values.

- Fat colour
- Bruising
- Poor dressing

Validation of the prediction equations for conformation and fat class

The study design required the classification of 800 carcases, which should have a frequency distribution the same as the distribution of carcases in the UK annual kill within the EUROP classification grid as shown in table 3.

Table 3. The target frequency distribution of 800 carcases within the EUROP grid to represent the distribution of the UK kill

		FAT CLASS							
		1	2	3L	3H	4L	4H	5	Total
Conformation	E	2	4	6	4	4	2	2	24
	U	2	16	72	35	11	6	2	144
	R	4	80	228	100	20	10	4	446
	O	4	52	86	17	7	2	2	170
	P	4	4	2	2	2	2	-	16
	TOTAL	16	156	394	158	44	22	10	800

A panel of 3 UK expert classifiers from MLC, RPA and SEERAD together with the MLC in-plant classifier were used to classify the carcases. Throughout the study a team of MLC in-plant classifiers classified the carcases, but only one classifier worked at any one time so the data produced were from individual classifiers not a team as for the experts.

Classification was carried out as follows and as shown in table 4:

- Each expert member was instructed to classify individual carcases independently without conferring with other members of the panel.
- The experts classified carcases on both the current MLC 5/7 and 15-class scales.
- The expert classifiers did not work at line speed but examined the carcases on a side line and only classified the carcases which were gambrelled and banded.
- The MLC in-plant classifiers worked at line speed (approx 800 carcases/hr), only assessed carcases that were gambrelled and unbanded and only assessed on the 5/7 point scale.
- The VIA equipment assessed all carcases in each of the 6 leg and shoulder combinations.

Table 4. The different carcase configurations classified by the various classifiers

Carcase configuration		Classifiers		
Shoulders	**Legs**	**Expert**	**MLC In-plant**	**VIA**
Banded	Gambrelled	✓		✓
Banded	Crossed			✓
Banded	Together			✓
Unbanded	Gambrelled		✓	✓
Unbanded	Crossed			✓
Unbanded	Together			✓

The two configurations assessed by classifiers (expert or MLC in-plant) and VIA are highlighted in grey.

Images, classification data and trait information were collected from all 800 carcases and were divided into two groups. The images and data for 200 carcases selected to represent the UK carcase EUROP grid distribution frequency were given to E+V to allow fine tuning of the prediction equations. The images and data from the remaining 600 carcases were held by an independent statistician to allow independent validation of the final prediction equations provided by E+V.

Study criteria

In order to assess the accuracy of VIA in predicting conformation and fat class, a set of criteria for the levels of agreement between manual graders and VIA were developed. Since there are no EU requirements for VIA accuracy for sheep carcase classification, the criteria used were based on a combination of current industry standards for MLC carcase classifiers and the standards for VIA for beef. Specifically, they were based on the Rural Payments Agency (RPA) criteria for classifier performance when undertaking checks under the EU regulations for beef. These regulations require a quarterly unannounced check of 40 carcases. The RPA allows up to 8 errors out of the 40 carcases of one class for both fatness and conformation on the undivided 5 class scales. This is equivalent to a requirement for 80% exact matches. This was relaxed to 70% for the fat class in setting the criteria for this study because the study was using the 7 class scale for fatness.

Hence, the criteria were as follows:

1. Professional judgement of acceptable accuracy using the current 5/7 point scale levels as follows:

Conformation	80% total agreement of classification awarded
	99% agreement within one class difference
	100% agreement within two classes difference
Fatness	70% total agreement of classification awarded
	90% agreement within one class difference
	99% agreement within two classes difference

2. VIA to achieve same level of agreement with the expert classifiers as the classifiers achieve with each other.

3. Level of agreement of VIA with the expert classifiers to match the level of agreement between the MLC in-plant classifiers and the expert classifiers.

Demonstration of the ability of VIA to predict meat yield

The ability of VIA to predict meat yield was assessed on a separate set of 500 carcases (gambrel/banded) selected as evenly as possible across the classification grid and butchered by a team of 4 butchers according to a standard butchery protocol. Prior to jointing, the accurate cold carcase weight of each carcase was recorded together with all primal (chump, leg, loin, shoulder, and breast) and subsequent butchery weights (trimmed primal cuts, lean trim, fat and bone trim and waste; see results). The fore-end/shoulder was removed by a cut between the $6^{th}/7^{th}$ ribs and the loin was separated from the leg and chump by a cut at the last lumbar vertebra. The primals were then part boned/trimmed according to the standard protocol and the fat on the loin was trimmed to a maximum of 6mm. The weights of trimmed primal, lean trim, fat trim and bone and waste were then recorded.

Data from a sample of 300 carcases distributed as evenly as possible across the classification grid were provided for E+V to develop the prediction equations. Data from the remaining 200 carcases were retained and used by the independent statistician to validate accuracy.

In addition to the main objectives of the trial, a further piece of analysis was undertaken as part of a PhD project to investigate the potential of VIA technology to predict meat yield in terms of saleable meat yield (SMY), saleable primal meat yield (SPMY) and the carcase components leg, chump and loin as compared with MLC scoring. Total SPMY was expressed as the sum of weights of all sub-primal cuts as a proportion of cold carcase weight (CCW), and SMY was the sum of weights of all sub-primal cuts plus the residual lean tissue of the trimming process as a proportion of CCW.

Demonstration of Repeatability

A separate set of 105 carcases were each classified twice by a single expert, using 5/7 and 15 point scales for conformation and fatness. They were similarly classified by VIA, three times for 80 of the carcases and five times for the remaining 25 carcases.

Results

Set up and integration of VIA in a UK abattoir

The equipment was set up on a by pass section of the line in the premises of Welsh Country Foods as shown in figure 2.

Figure 2. The E+V VSS2000 VIA system for sheep carcases

The successful installation of the equipment and its compatibility with the current subjective system met objective 1 of the study

Development of the prediction equation for fat class and conformation

The study set out to have an equal distribution of lambs within the EUROP grid. The target numbers are shown in black (on the left), the numbers actually achieved in each cell are shown in red (on the right).

Table 5. The target and achieved distribution of lamb carcases within the EUROP grid

Conformation		FAT CLASS							
		1	2	3L	3H	4L	4H	5	Total
	E	2 8	7 25	17 40	17 10	19 7	19 4	19 7	100 101
	U	5 1	10 25	15 52	15 18	15 20	17 25	23 19	100 160
	R	5 6	15 25	15 40	15 31	15 25	15 41	20 47	100 215
	O	20 24	15 47	15 37	15 22	15 41	15 13	5 1	100 158
	P	39 72	25 13	10 0	10 0	7 0	5 0	4 0	100 85
	TOTAL	71 111	72 135	72 169	72 81	71 66	71 83	71 74	500 719

E+V developed prediction equations based on linear statistics. Due to the change in shape associated with the various carcase presentation methods (figure 3), it was necessary to develop a range of equations.

Separate equations were developed for the following

- All 6 carcase presentation configurations
- 15 point scale
- 5/7 point scale
- Conformation and fat class
- Equations using 8 and 10 predictors

Figure 3. The effect of suspension on carcase shape: gambrelled (left) and crossed (right), the same carcase is suspended in each manner

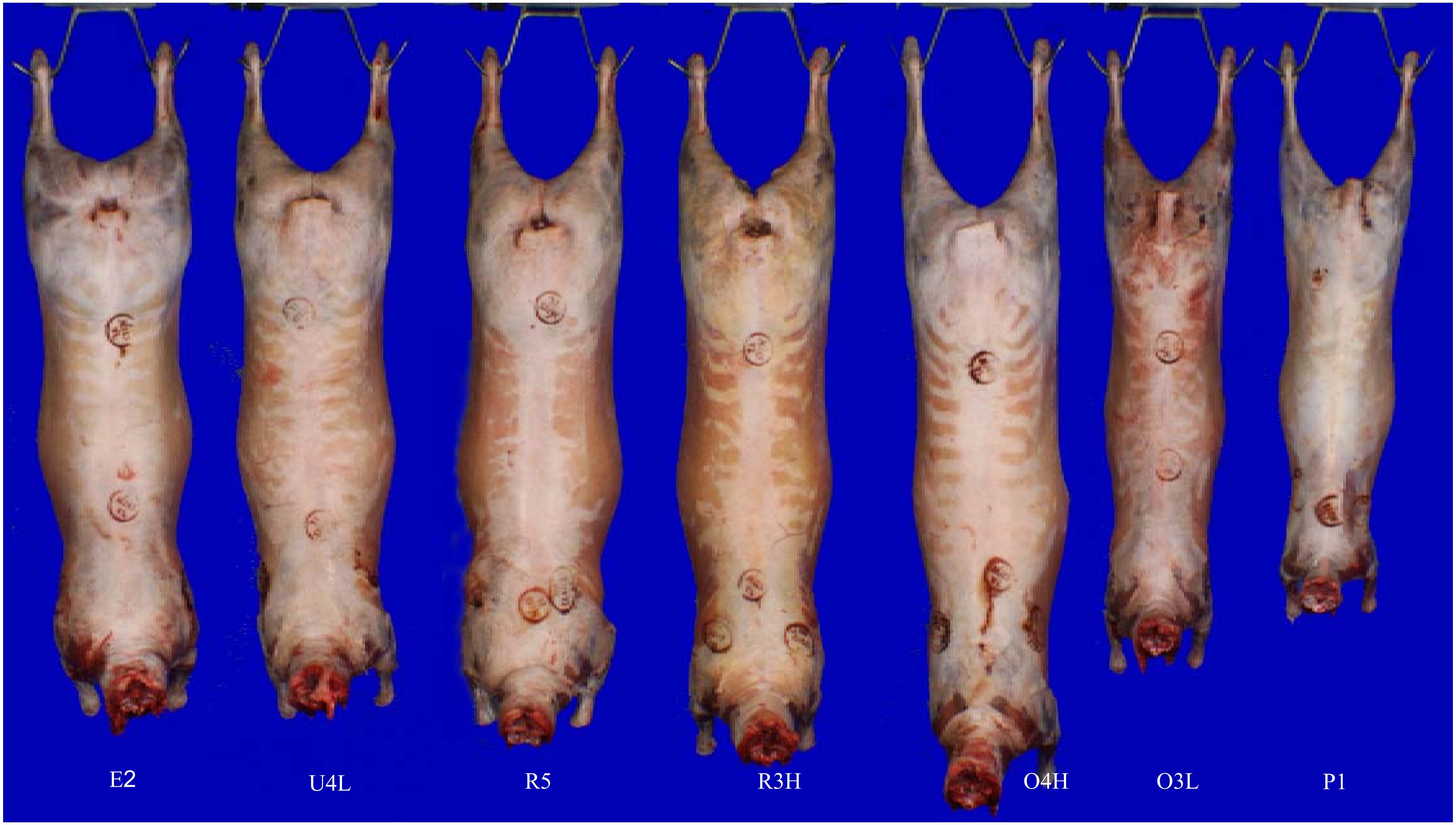
The consideration of operational procedures such as suspension method and shoulder presentation met objective 2 of the study.

The production of the prediction equations for conformation and fatness met objective 3 of the study

Prediction of conformation and fat class

800 lamb carcases for the calibration and validation of the VIA were selected from December to February across the weight range 8-28kg and, where practically possible, balanced by conformation/fat class to reflect as closely as possible the UK annual slaughterings. However, lean/poor conformation carcases were inevitably lighter and fatter/better conformation carcases heavier. Typical examples of lamb carcases in this study are shown in figure 4.

Figure 4. Typical examples of the types of carcases used in the study



The distribution of the population of 800 carcases that was achieved was similar to the target numbers, although numbers for the grades at the extremes of the ranges tended not to meet the required targets (Table 6). This was particularly the case for P conformation carcases, where all but P1 classification were not present in the selection and P1 carcases were over-represented (Table 6). It should be noted that the data on the distribution of P carcases by fat class were not available as national data do not separate P carcases by fat class.

Table 6. The target (black) and achieved distribution (red in brackets) of 800 lamb carcases used in the final calibration and validation of the prediction equations.

Conformation	Fat class 1	2	3L	3H	4L	4H	5	Total
E	2 **(0)**	4 **(3)**	6 **(10)**	4 **(6)**	4 **(3)**	2 **(0)**	2 **(0)**	24 **(22)**
U	2 **(0)**	16 **(23)**	72 **(58)**	35 **(38)**	11 **(14)**	6 **(7)**	2 **(5)**	144 **(145)**
R	4 **(4)**	80 **(71)**	228 **(230)**	100 **(115)**	20 **(34)**	10 **(10)**	4 **(3)**	446 **(467)**
O	4 **(3)**	52 **(51)**	86 **(81)**	17 **(16)**	7 **(4)**	2 **(1)**	2 **(0)**	170 **(156)**
P	4 **(10)**	4 **(0)**	2 **(0)**	2 **(0)**	2 **(0)**	2 **(0)**	- **(0)**	16 **(10)**
Total	16 **(17)**	156 **(148)**	394 **(379)**	158 **(175)**	44 **(55)**	22 **(18)**	10 **(8)**	800 **(800)**

The 800 carcases were divided into two groups one of 200 carcases and the other of 600 carcases which were used for final calibration and validation respectively. The distribution of the carcases in both groups is shown in tables 7a and 7b below.

Table 7. Breakdown of carcase distributions in the calibration and validation groups

Table 7a . Group of 200 carcases used for final calibration

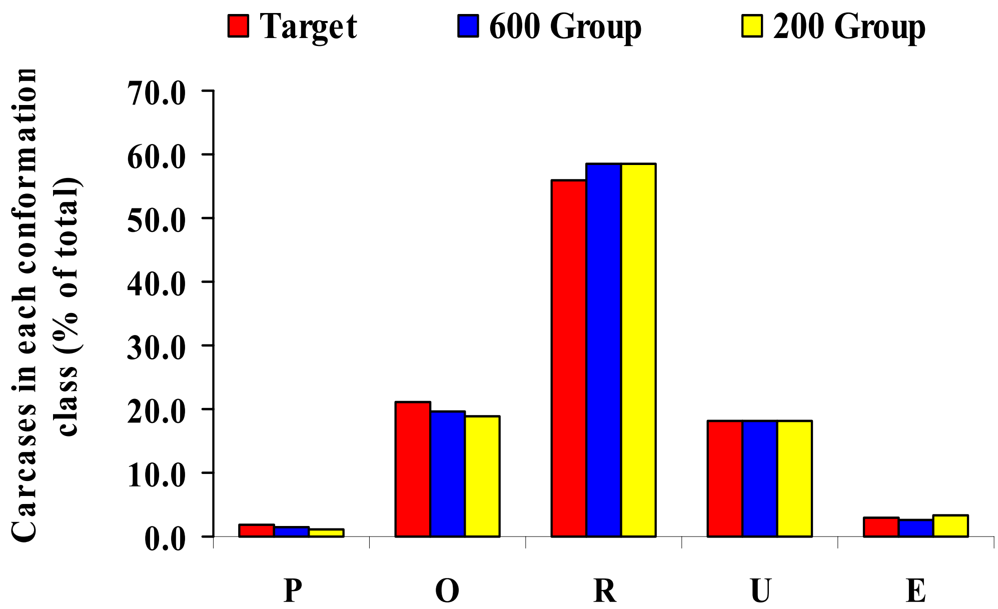
Conformation	Fat Class							
	1	**2**	**3L**	**3H**	**4L**	**4H**	**5**	**Total**
E	0	1	3	2	1	0	0	7
U	0	6	14	9	4	2	1	36
R	1	17	58	29	9	2	1	117
O	1	13	19	4	1	0	0	38
P	2	0	0	0	0	0	0	2
Total	4	37	94	44	15	4	2	200

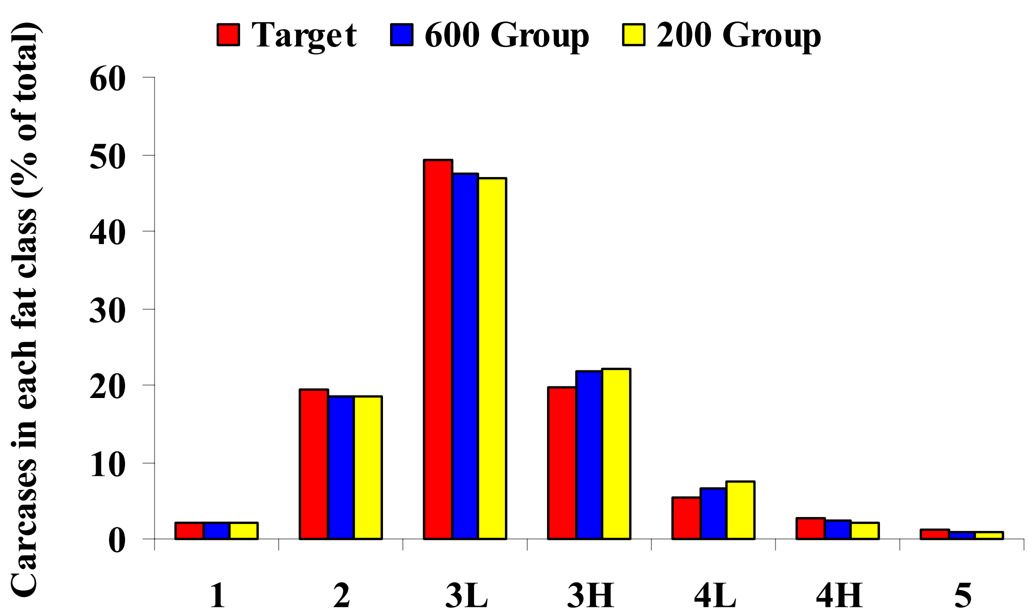
Table 7b. Group of 600 carcases used for validation

Conformation	Fat Class							
	1	**2**	**3L**	**3H**	**4L**	**4H**	**5**	**Total**
E	0	2	7	4	2	0	0	15
U	0	17	44	29	10	5	4	109
R	3	54	172	86	25	8	2	350
O	2	38	62	12	3	1	0	118
P	8	0	0	0	0	0	0	8
Total	13	111	285	131	40	14	6	600

The distribution of the carcases within the grid between the two groups was broadly similar, and largely met the target of the UK distribution frequency.

Figure 4. The frequency distribution of the carcases with respect to conformation class in the groups used for calibration and validation





Following the construction and fine tuning of the prediction equations by E+V, the data were analysed by an independent statistician, who assessed whether the VIA machine had met the 3 criteria set out at the start of the study.

The results of that analysis against the criteria of the study are shown below.

Criterion 1. Professional judgement of acceptable accuracy using the current 5/7 point scales as follows:

Conformation	80% total agreement of classification awarded
	99% agreement within one class difference
	100% agreement within two classes difference
Fatness	70% total agreement of classification awarded
	90% agreement within one class difference
	99% agreement within two classes difference

The statistician assessed the level of agreement between the VIA predictions and the median classification score of the 3 expert classifiers (identified as A, B and C). [Note the median is the middle value not the average (mean)]. The results for conformation and fatness on the 5/7 point scale are shown in Table 8. The data are presented as rounded down percentages for total agreement, agreement within one class and within two classes.

NOTE:

Data are presented for all six carcase configurations, although it should be noted that the expert classifiers only assessed banded and gambrelled carcases, and the MLC in-plant classifiers only assessed carcases that were unbanded and gambrelled as indicated by the highlighting in table 8. The remaining configurations were assessed only by VIA. For the purposes of the analysis the expert panel classification of banded and gambrelled carcases was regarded as the reference.

Table 8. Agreement between VIA classification and the median classification from the expert classifiers using the current 5/7 point scales for conformation and fatness

Shoulders	Legs	Conformation % agreements within class			Fatness % agreements within class		
		0	1	2	0	1	2
Banded	Gambrelled	72	100	100	50	93	98
Banded	Crossed	72	100	100	49	93	99
Banded	Together	68	99	100	48	94	99
Unbanded	Gambrelled	72	99	100	55	95	99
Unbanded	Crossed	74	99	100	51	95	99
Unbanded	Together	74	99	100	53	95	99
Target Accuracy		**80**	**99**	**100**	**70**	**90**	**99**

It is clear from table 8 that the target accuracy of 80% and 70% for total agreement for conformation and fatness respectively is not achieved. In contrast, the target accuracy for within 1 and 2 classes was achieved or exceeded in all cases for both conformation and fatness.

Using the 5/7 point scale

- VIA does not meet criterion 1 for total agreement for either conformation or fatness
- VIA does meet the criteria for both conformation and fatness within one and two classes.

Criterion 2. VIA to achieve same level of agreement with the expert classifiers as the classifiers achieve with each other.

Tables 9 and 10 show the level of agreement between VIA and the individual expert classifiers for conformation/fatness classifications on 5/7 point scales. As in Table 8, the numbers shown are rounded-down percentages in total agreement, within one class and within two classes, for the six carcase configurations.

NOTE: As in table 8 the carcase configurations assessed by both expert classifiers and VIA are highlighted. The remaining configurations were not assessed by expert classifiers but were assessed by VIA. The agreement figures for the configurations not assessed by the experts are based on the scores given for banded gambrelled carcases by the classifiers compared with the scores given by VIA using configuration specific prediction equations. The between classifier agreements are based on assessment of banded gambrelled carcases only.

Table 9. The level of agreement between VIA and expert classifiers for conformation on the 5/7 point scale

		Conformation % agreements within class scores								
Shoulders	**Legs**	**Classifier A**			**Classifier B**			**Classifier C**		
		0	1	2	0	1	2	0	1	2
Banded	Gambrelled	71	100	100	68	99	100	71	99	100
Banded	Crossed	70	100	100	71	99	100	71	100	100
Banded	Together	66	99	100	67	99	100	67	99	100
Unbanded	Gambrelled	71	99	100	69	99	100	70	99	100
Unbanded	Crossed	71	99	100	74	99	100	71	99	100
Unbanded	Together	73	99	100	72	99	100	72	99	100
Agreement	with B	80	99	100						
	with C	83	100	100	79	99	100			

Poorest level of agreement

The results show that for conformation, the poorest agreement between classifiers was 79, 99, 100% for total agreement, agreement within 1 class and within 2 classes respectively which was found between classifiers B and C. It should be noted that 79% is below the level of total agreement set for the study. Comparison with the VIA scores shows that the target of 79% was not met for total agreement, but the targets for within 1 and 2 classes were met or exceeded for conformation.

Table 10. The level of agreement between VIA and the expert classifiers for fatness on the 5/7 point scale

		Fatness % agreements within class scores								
Shoulders	**Legs**	**Classifier A**			**Classifier B**			**Classifier C**		
		0	1	2	0	1	2	0	1	2
Banded	Gambrelled	50	93	99	48	92	98	49	94	98
Banded	Crossed	50	93	99	46	91	99	47	93	99
Banded	Together	53	93	99	45	92	98	46	94	99
Unbanded	Gambrelled	55	95	99	51	95	99	55	95	99
Unbanded	Crossed	52	94	99	46	93	99	51	94	99
Unbanded	Together	53	95	99	51	94	99	51	95	99
Agreement	with B	64	99	100						
	with C	70	99	100	73	99	100			

Poorest level of agreement

For fatness the poorest agreement between classifiers was 64, 99, 100 % for total agreement, agreement within 1 class and within 2 classes respectively which was found between classifiers A and B. It should be noted that 64% is below the level of total agreement set for the study.

Comparison with the VIA scores shows that the target of 64 % was not met for total agreement, nor were the targets for within 1 and 2 classes for fatness.

Using the 5/7 point scale

- VIA did not meet criterion 2 for total agreement with the classifiers
- VIA did not meet the criteria for within one or two classes on the 5/7 point scale

The level of agreement with the expert panel was also assessed using the 15 point scale. Tables 13 and 14 show the agreements expressed as rounded-down percentages for total agreement, and within one to six sub classes of agreement, for the six carcase configurations. As in previous tables the configuration assessed by both experts and VIA is highlighted.

For conformation, the poorest agreement between classifiers across all six carcase configurations was considered by the statistician to be 40, 85, 97, 99, 100, 100, 100%, for total agreement and within one to six sub classes of agreement. This is a composite of A versus B which gave values of 40, 88, 98, 100, 100, 100, 100, % for total agreement and within one to five classes of agreement respectively and B versus C which gave 41, 85, 97, 99, 100, 100, 100% for total agreement and within one to six sub classes of agreement respectively.

For fatness, the poorest agreement was 40, 89, 99, 100, 100, 100, 100% for total agreement and within one to six classes of agreement respectively which was between classifiers A and B.

These results suggest that the poorest level of agreement of the experts for conformation and fatness on the 15 point scale is as follows

Table 11. Level of agreement between expert classifiers for conformation

Agreement within number of classes (%)						
0	1	2	3	4	5	6
40	85	97	99	100	100	100

Table 12. Level of agreement between expert classifiers for fatness

Agreement within number of classes (%)						
0	1	2	3	4	5	6
40	89	99	100	100	100	100

It should be noted that the target level of agreements for the study were set on the 5/7 point scale, and the relationship between the 5/7 point scale and the 15 point scale is not linear.

Table 13. Agreement scores for Conformation using the 15 point scale
B: Banded, U: Unbanded, G: Gambrelled, C: Crossed, T: Together

		Conformation % agreement within class scores																				
		Classifier A							Classifier B							Classifier C						
Shoulder	**Legs**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**0**	**1**	**2**	**3**	**4**	**5**	**6**
B	G	34	81	96	99	99	100	100	35	82	97	99	100	100	100	31	79	96	99	99	100	100
B	C	29	77	96	99	99	100	100	34	80	96	99	99	100	100	29	76	95	98	99	100	100
B	T	36	81	96	100	100	100	100	38	85	97	100	100	100	100	32	76	96	99	100	100	100
U	G	35	80	97	99	100	100	100	36	82	97	99	100	100	100	33	79	95	99	100	100	100
U	C	33	77	97	99	100	100	100	37	85	97	99	100	100	100	32	78	94	99	99	100	100
U	T	33	83	97	100	100	100	100	36	86	98	100	100	100	100	35	80	96	99	100	100	100
Agreement	with B	40	88	98	100	100	100	100														
	with C	45	88	98	100	100	100	100	41	85	97	99	100	100	100							

Poorest level of agreement

Table 14. Agreement scores for Fatness using the 15 point scale
B: Banded, U: Unbanded, G: Gambrelled, C: Crossed, T: Together

Shoulder	Leg	Fatness % Agreement within class scores																				
		Classifier A							**Classifier B**							**Classifier C**						
		0	**1**	**2**	**3**	**4**	**5**	**6**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**0**	**1**	**2**	**3**	**4**	**5**	**6**
B	G	33	79	93	98	99	99	100	29	76	93	98	99	100	100	31	76	94	98	99	99	100
B	C	31	73	91	97	99	99	100	27	69	90	97	99	99	100	30	73	92	98	99	99	99
B	T	33	75	93	98	99	100	100	31	75	92	98	99	100	100	31	74	94	99	99	99	100
U	G	32	76	93	98	99	99	99	29	72	92	97	99	99	100	30	77	93	97	99	99	99
U	C	31	74	93	98	99	99	99	29	73	91	98	99	99	100	32	75	94	99	99	99	99
U	T	33	78	94	99	99	99	100	31	75	95	99	99	100	100	32	80	95	99	99	99	100
Agreement with B		40	89	99	100	100	100	100														
Agreement with C		45	90	99	100	100	100	100	43	93	99	100	100	100	100							

Poorest level of agreement

In addition, to percentage agreement scores, the data can be presented as Mean Absolute Differences (MAD) of 15 point scale classifications for the six carcase configurations (tables 15 and 16). The MAD is the average absolute difference between two sets of classifications, so the smaller the value the better. A MAD of 0 means that two sets are in exact agreement.

Table 15. The Mean Absolute Difference for Conformation

		Conformation Mean Absolute Difference		
Shoulders	**Legs**	**Classifier A**	**Classifier B**	**Classifier C**
Banded	Gambrelled	0.93	0.89	0.97
Banded	Crossed	0.99	0.92	1.01
Banded	Together	0.89	0.84	0.98
Unbanded	Gambrelled	0.92	0.88	0.96
Unbanded	Crossed	0.95	0.83	0.98
Unbanded	Together	0.89	0.82	0.93
Agreement with B		0.73		
	with C	0.68	0.76	

Poorest level of agreement

For conformation, the poorest agreement between classifiers is a MAD of 0.76 seen between classifiers B and C. In all cases the MADs between VIA and individual classifiers exceed these values, suggesting that there was lower agreement between VIA and classifiers than there was between the classifiers themselves.

Table 16. The Mean Absolute Difference for Fatness

		Fatness Mean Absolute Differences		
Shoulders	**Legs**	**Classifier A**	**Classifier B**	**Classifier C**
Banded	Gambrelled	0.98	1.06	1.01
Banded	Crossed	1.10	1.18	1.09
Banded	Together	1.03	1.05	1.04
Unbanded	Gambrelled	1.04	1.13	1.05
Unbanded	Crossed	1.06	1.08	1.02
Unbanded	Together	0.98	1.01	0.96
Agreement with B		0.72		
with C		0.65	0.62	

Poorest level of agreement

For fatness the poorest agreement was a MAD of 0.72 between classifiers A and B. In all cases the MADs between VIA and individual classifiers exceed these values, suggesting that there was lower agreement between VIA and classifiers than there was between the classifiers themselves.

Using the 15 point scale

- No percentage levels of agreement were set for the 15 point scale and the relationship between the two scales is not linear.
- VIA achieved 99% and 100% agreement for conformation within 3 and 5 sub classes respectively.
- VIA achieved 90% and 99% agreement for fatness within 2 and 4 sub classes respectively.

Criterion 3. Level of agreement of VIA with the median value from the expert classifiers to match the level of agreement between the MLC in-plant classifiers and the expert classifiers.

The statistician compared the levels of agreement between the MLC in-plant classifiers, working a line speed of about 800 /h and assessing unbanded gambrelled carcases and the median of the three expert classifiers working off line and assessing banded gambrelled carcases on the 5/7 point scale. This was found to be 73, 99, 100% for total agreement, agreement within 1 class difference and within 2 classes difference respectively for conformation and 58, 94, 99% for total agreement, agreement within 1 class and within 2 classes respectively for fatness classification on 5/7 point scales.

Table 17. Level of agreement of VIA with the median value from the expert classifiers in relation to the level of agreement between the MLC in-plant classifiers and the expert classifiers

Shoulders	**Legs**	**Conformation** % agreements within class			**Fatness** % agreements within class			**Classifiers** Expert	In-plant	VIA
		0	1	2	0	1	2			
Banded	Gambrelled	72	100	100	50	93	98	✓		✓
Banded	Crossed	72	100	100	49	93	99			✓
Banded	Together	68	99	100	48	94	99			✓
Unbanded	Gambrelled	72	99	100	55	95	99		✓	✓
Unbanded	Crossed	74	99	100	51	95	99			✓
Unbanded	Together	74	99	100	53	95	99			✓
Target - Agreement with MLC in-plant classifier		**73**	**99**	**100**	**58**	**94**	**99**			

The configurations classified by the various classifiers are shown on the right.

Comparison of these values with those previously presented in Table 8 and reproduced above in table 17 shows that the agreement between VIA and the experts is generally at comparable levels for conformation. The presentation of banded carcases with legs together has slightly lower values for total agreement, but was only assessed by VIA, and was not assessed by the expert panel.

For fatness, the level for total agreement is not achieved (table 17), although levels are comparable for agreement within one or two classes difference.

- VIA largely met the levels of consistency against the MLC in-plant classifier with regard to conformation in terms of total agreement and agreement within 1 class and 2 classes difference.
- VIA did not achieve total agreement with the MLC in-plant classifier for fatness, although it did achieve consistency for within 1 class and 2 classes difference.

Meat Yield.

The ability of VIA to predict meat yield was assessed on 500 carcases with a weight range of 7 to 35 kg which had been presented in the banded/ gambrelled configuration. A standard butchery process was used for the trial so the data will be trial specific and plants purchasing a VIA system in the future will therefore need to develop their own plant specific meat yield equations which are designed for their own butchery specifications.

The 500 carcases were divided into 2 groups. One group of 200 was used for validation, remaining 300 carcases were used for calibration.

Table 18. Butchery sample selections

Calibration

Conformation	Fat Class							
	1	2	3L	3H	4L	4H	5	Total
E	2	12	14	4	1	2	1	36
U	4	9	11	5	9	13	13	64
R	7	7	9	13	26	20	19	101
O	14	10	8	10	10	12	3	67
P	23	8	0	0	0	0	0	31
Total	50	46	42	32	46	47	36	299

Validation

Conformation	Fat Class							
	1	2	3L	3H	4L	4H	5	Total
E	1	8	10	2	0	2	1	24
U	3	6	7	4	6	8	8	42
R	4	4	6	8	17	14	13	66
O	10	7	5	7	6	8	2	45
P	16	6	1	0	0	0	0	23
Total	34	31	29	21	29	32	24	200

The carcases in each group showed the frequency distribution across the grid for conformation and fat class as shown below.

Figure 5. The frequency distribution of the carcases in the calibration and validation groups with respect to conformation

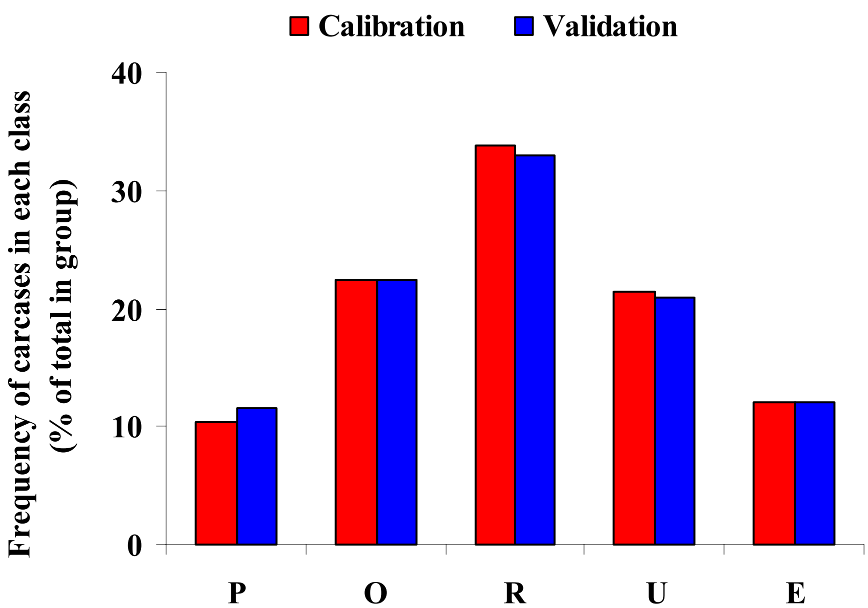
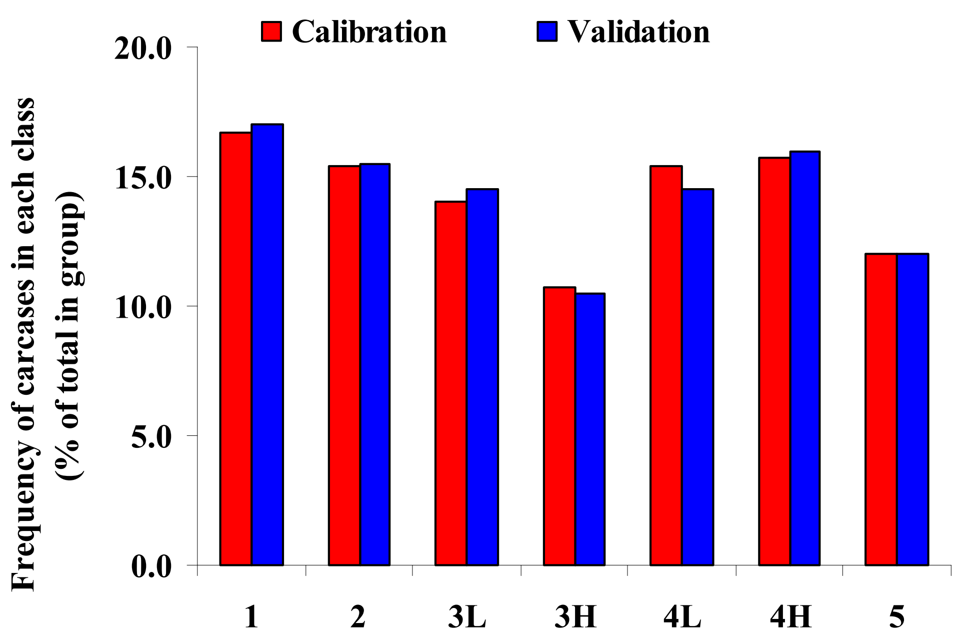


Figure 6. The frequency distribution of the carcases in the calibration and validation groups with respect to fat class



The distribution of carcases for both groups with respect to fat class was different to that seen for the carcases used in other parts of the trial, because of the deliberate aim to select equal numbers across the classification grid.

There was a small non-significant difference in the accurate cold carcase weight between the two groups, such that the calibration group was slightly lighter on average and showed a slightly greater range of values than seen for the validation group.

Table 19. Cold carcase weight in the calibration and validation groups for estimation of meat yield.

Cold carcase wt	Group	Wt (kg)
Mean	Validation	19.93
Stdev		3.81
median		19.98
max		29.42
min		9.62
Mean	Calibration	19.43
Stdev		4.02
median		19.68
max		35.36
min		7.00

The carcases were butchered according to the standard trial specification and weights of five primal cuts were recorded as shown in table 20.

Table 20. Opening weights and trimmed primal weight for carcases in calibration and validation groups.
*** Denotes that there is a statistically significant difference between the two groups $p<0.05$**

Leg and Chump			Trimmed Primals	
Weights in g	Group	Opening wt	Leg	Chump
Mean	**Validation**	**6921.64**	**4450.30**	**749.35**
Stdev		**1248.00**	**805.95**	**172.60**
median		6920.00	4437.00	743.00
max		10230.00	7050.00	1240.00
min		3360.00	2091.00	344.00
Mean	**Calibration**	**6758.90**	**4353.71**	**733.31**
Stdev		**1356.86**	**894.40**	**178.98**
median		6800.00	4340.00	744.00
max		12620.00	8536.00	1491.00
min		2960.00	1975.00	221.00

Fore Shoulder			Trimmed Primal
Weights in g	Group	Opening wt	Fore Shoulder
Mean	**Validation**	**7983.20**	**5136.27**
Stdev		**1454.38**	**1059.43**
median		8030.00	5116.00
max		12060.00	8555.00
min		3700.00	2361.00
Mean	**Calibration**	**7689.62***	**4964.28**
Stdev		**1563.96**	**1118.63**
median		7680.00	4926.00
max		13900.00	9575.00
min		3140.00	2026.00

Loin and Breast			Trimmed Primals	
Weights in g	Group	Opening wt	Loin	Breast
Mean	**Validation**	**5023.95**	**2897.03**	**1520.41**
Stdev		**1230.92**	**607.20**	**426.35**
median		5040.00	2924.50	1491.00
max		8060.00	4500.00	2730.00
min		2460.00	1470.00	736.00
Mean	**Calibration**	**4999.46**	**2882.09**	**1526.23**
Stdev		**1228.95**	**638.13**	**430.35**
median		5060.00	2918.00	1536.00
max		8800.00	5437.00	2843.00
min		1480.00	1002.00	339.00

Totals			Trimmed Primals
Weights in g	Group	Opening wt	Totals
Mean	**Validation**	**19928.79**	**14753.35**
Stdev		**3802.87**	**2907.61**
median		19980.00	14855.00
max		29400.00	23018.00
min		9640.00	7021.00
Mean	**Calibration**	**19447.97**	**14459.61**
Stdev		**4012.94**	**3082.76**
median		19680.00	14713.00
max		35320.00	27694.00
min		7580.00	5759.00

Examination of the summary data for the butchery study showed that there were some statistically significant differences between the calibration and validation groups. These differences were particularly evident for fore shoulder, and it may therefore be appropriate to regard the data for fore shoulder with some caution. It was also noted that although the carcases in the calibration group were slightly, but not significantly lighter than those in the validation group, the total saleable primal meat yield was significantly higher in the calibration group.

The trimmed primal weights were also predicted by VIA using the validation dataset. The predicted primal weights were compared with the recorded weights, by computing the root-mean-square error (RMSE) of prediction for each cut. Results are summarised in Table 21, expressed in grams. For comparison, the mean and standard deviation (SD) of recorded weights for the validation carcases are also given.

Table 21. The prediction of trimmed primal weights (grams) by VIA

Primal Cut	Mean	SD	RMSE
Leg	4456	810	168
Chump	749	172	53
Loin	2895	608	218
Breast	1517	420	169
Shoulder	5135	1061	260

(SD – a measure of variation; RMSE – root mean square error - a perfect score of 0.0 would mean there was no error in the prediction. An RMSE smaller than the SD indicates that there is a good prediction)

VIA predicted yield well, as the RMSEs are considerably smaller than the SDs. Most of the variability in primal cut weights was likely to be due to differing carcase weights.

In practice, it is likely that the saleable meat yield, saleable primal yield and the percentage of lean meat in the sub-primals may also be of commercial importance. Consequently. the data from 300 carcases (calibration set) were analysed as part of a PhD project to evaluate the ability of VIA to predict saleable meat yield ((SMY) (expressed as the sum of the weights of all the primals plus residual lean trim as a proportion of CCW)), saleable primal meat yield ((SPMY)(expressed as the sum of all the weight of the sub-primal cuts expressed as a proportion of CCW)) and the percentages of saleable meat yield in the sub-primal cuts leg, chump, and loin (expressed as a proportion of CCW) in comparison to predicting yield using calculations based on the use of the MLC classification from the MLC in-plant classifier and the accurate CCW.

The results of these analyses are shown in table 22 below.

Table 22. The prediction of SMY, SPMY and carcase components and its precision using VIA traits or MLC standard classification

Predicted Variables (%)	VIA		MLC classification	
	R^2	RMSE	R^2	RMSE
SMY	0.66	0.013	0.55	0.015
SPMY	0.66	0.012	0.54	0.014
Leg	0.81	0.008	0.62	0.011
Chump	0.59	0.002	0.48	0.002
Loin	0.31	0.008	0.30	0.010

(R^2 - a perfect score of 1.0 would mean that all the variations in that particular parameter could be explained by the prediction method used. RMSE – root mean square error -a perfect score of 0.0 would mean there was no error in the prediction).

Table 22 shows that both the prediction of the variables SMY, SPMY, leg, chump and loin was 20, 22, 31, 23 and 3% respectively higher using VIA than using the MLC standard classification. The precision, measured as reduction in RMSE of predicting SMY, SPMY, lean meat percentage of leg, chump, and loin was also higher using VIA than using MLC standard classification.

Since it is known that a variety of factors eg the gender of the animal may influence data from such studies, the data were corrected for such factors as shown in table 23.

Table 23. The prediction of SMY, SPMY and carcase components and its precision using VIA traits or MLC standard classification, and corrected for known influencing factors

Predicted Variables (%)	VIA		VIA corrected †		MLC classification		MLC classification corrected †	
	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE
SMY	0.66	0.013	0.71	0.012	0.55	0.015	0.65	0.013
SPMY	0.66	0.012	0.74	0.010	0.54	0.014	0.66	0.012
Leg	0.81	0.008	0.82	0.008	0.62	0.011	0.64	0.011
Chump	0.59	0.002	0.61	0.002	0.48	0.002	0.52	0.002
Loin	0.31	0.008	0.43	0.009	0.30	0.010	0.41	0.010

(R^2 - a perfect score of 1.0 would mean that all the variations in that particular parameter could be explained by the prediction method used. RMSE – root mean square error -a perfect score of 0.0 would mean there was no error in the prediction).

Table 23 shows that following adjustment for known influencing factors eg gender (†), the prediction of SMY, SPMY, lean meat of the leg, chump, and loin increased by 8, 12, 1, 3, and 39%, respectively, using VIA, whereas generally larger increases of 18, 22, 3, 8, and 37%, respectively, were observed using MLC classification. These data suggest the precision of predicting yield data was greater using VIA than using MLC classification. It was of interest to note that the adjustment for gender and slaughter date improved both methods of prediction.

The demonstration, based on a specific butchery method, of the ability of VIA to predict meat yield meets objective 4 of the study.

Assessment of Repeatability

Table 24 summarises the repeatability of the expert and VIA for each of the four classifications, expressed as rounded-down percentages in total agreement, and within one to four classes/ sub classes of agreement. The data show that VIA was slightly more consistent for conformation on both 5/7 and 15 point classification scales. However, with regard to measurement of fatness, the data suggest that the expert classifiers were considerably more consistent for fatness on both 5/7 and 15 point classification scales.

Table 24. A summary of the repeatability study

	Conformation						Fatness							
Scale	5		15				7			15				
	0	1	0	1	2	3	0	1	2	0	1	2	3	4
Expert	86	100	73	91	96	100	90	99	100	76	94	99	100	100
VIA	88	100	70	98	99	100	75	99	100	52	92	98	98	100

Table 25 shows the mean differences between the expert and VIA scores for the same carcases. The data show the differences for carcases 1-80 which were assessed 3 times each by VIA and for carcases 81-105 which were assessed 5 times each by VIA, as well as for all carcases together. Taking all carcases together shows, on average, a good agreement, except for fatness on a 15-point scale, where the expert averages one point higher than VIA.

Table 25. Mean Difference between expert and VIA scores

	Conformation		Fatness	
Scale	5	15	7	15
All carcases	0.11	0.12	0.12	1.05
Carcases 1-80	0.23	0.42	0.25	1.49
Carcases 81-105	-0.13	-0.46	-0.13	0.21

Examination of the separate groups of carcases suggests that the differences tended to be reduced in the groups assessed five times, particularly for fatness of the 15 point scale.

Discussion

The study set out to evaluate the use of video image analysis as a predictor of carcase classification and meat yield in sheep carcases.

The study had three objectives: to develop an accurate generic prediction equation for conformation and fat class for the E+V VIA system, to ensure compatibility between the current subjective classification system and VIA, and to demonstrate the ability of VIA technology to predict meat yield.

The study was successful in developing prediction equations for conformation and fat class. In addition to developing an equation for the most common configuration of carcase presentation, the E+V team also successfully developed prediction equations for five other configurations (see table 1) to meet the specific requirements of industry.

A range of tests was carried out to evaluate the comparability between subjective classification, on both 5/7 and 15 point scales, and VIA. The study set a range of criteria for acceptable accuracy for determining conformation and fatness on the 5/7 point scale. These comprised for conformation: 80% total agreement of classification awarded, 99% agreement within one class difference and 100% agreement within two classes difference, and for fatness: 70% total agreement of classification awarded, 90% agreement within one class difference and 99% agreement within two classes difference. The first agreement point was based on current training criteria for lamb classifiers, and the remaining points were based on VIA assessment studies in beef. The study set out to evaluate the performance of the equipment on the 5/7 point scale currently used for lamb, and set out 3 criteria against which the VIA was to be assessed. In addition the tests were repeated on the 15 point scale.

Whilst VIA did not achieve the criteria of 80% and 70% for total agreement for conformation and fatness, it is worth noting that these criteria were not achieved for total agreement between the expert classifiers nor for total agreement between the expert classifiers and the MLC in-plant classifiers. The poorest level of total agreement between the expert classifiers themselves was 79% for conformation, and 64% for fatness (tables 9 and 10), similarly, the poorest level of total agreement between expert and MLC in-plant classifiers was 73% and 58% for conformation and fatness respectively (table 17). Consequently, VIA could not be expected to meet the criteria set when the reference base of the expert panel itself did not achieve the target.

The comparison of VIA with the MLC in-plant classifier and expert assessor panel with respect to accuracy showed that VIA met all the criteria for agreement with respect to conformation, but only met the target criterion for fatness to within 1 and 2 classes difference. The results showed that VIA performed as well as the MLC in-plant classifier for conformation but less well with respect to fatness. The assessment of repeatability of VIA showed that the equipment was more consistent for conformation and less consistent for fat on both 5/7 point scales and 15 point scales. Close inspection of the data showed also that there was a tendency for consistency to be greater in the group of carcases that were assessed five times by VIA rather than three times.

In regard to the ability of VIA to predict meat yield and distribution, the results showed the importance of several parameters (e.g. carcase weight and gender) but that overall VIA has a better predictive precision of saleable meat yield (Tables 22 and 23) than current MLC classification.

When considering the implications of the findings a number of points may be pertinent:

First, for the purposes of this study, the expert panel assessments were considered to be the reference classification. However none of the carcase presentation methods was assessed by all three assessor methods viz: the expert panel, the MLC in-plant classifier and VIA; and four of the carcase presentations were assessed only by VIA (see Table 4).

Second, the MLC in-plant classifiers and expert classifiers were working under different conditions (the MLC in-plant classifiers were working at line speed - approximately 800 carcases/hour - the expert classifiers were working off line with no time constraints), and examining different carcase configurations. It is not clear whether these differences would affect the level of agreement between them.

Third, comparison of the results from the present study with those from previous trials of VIA in both beef and lamb is difficult due to methodological differences. For example, the recent studies of VIA in lamb (Brady et al, 2003; Cunha et al 2004) were based on the use of mean classifier scores rather than median scores. The mean score was used by E+V, to build the predictive equations, due to the higher precision with this approach particularly on border line grades and with only 3 experts (EU validation studies for beef require an expert panel of 5 classifiers). The statistical analysis was based on the use of the median value of the experts' scores.

Fourth, the present study used cold carcase weight as a variable in the model for estimating classification and yield parameters. Other studies tend not to use cold carcase weight either for classification or yield estimation. The major reason for this is that accurate cold carcase weight is not routinely available in commercial practice and so any equations dependant on this variable will either require a change in commercial practice or the equations will have to be re-derived based on hot carcase weight. This is important in the context of the present study and the implementation of VIA in the UK.

Finally, the data from the repeatability study suggest that assessing repeatability at the beginning of the trial rather than at the end might have been useful in developing more accurate equations, particularly for fat class.

Conclusion

VIA offers lamb abattoirs a means of predicting meat yield and primal weights as well as providing objective assessment of carcase conformation and fatness. The present study provides data to allow industry to assess the potential that the equipment offers in these areas of operation.

References.

Allen, P. & Finnery, N. 2001. Mechanical grading of beef carcases. National Food Centre Research Report, Teagasc 45, 1-26.

Allen, P. & Finnery, N. 2000. Objective beef classification. A report of a trial of three VIA classification systems. http://www.teagasc.ie/publications/beefgrading/beefgrading.pdf

Brady, A. S., K. E. Belk, S. B. LeValley, N. L. Dalsted, J. A. Scanga, J. D. Tatum, and G. C. Smith. 2003. An evaluation of the lamb vision system as a predictor of lamb carcase red meat yield percentage. J. Anim. Sci. 81: 1488-1498

Chandraratne, M.R., Kulasiri, D. Frampton,C. Samarasinghe,S. Bickerstaffe,R 2006 Prediction of lamb carcase grade using features extracted from lamb chop images. Journal of Food Engineering.. In press

Cunha B.C. N., Belk, K.E., Scanga, J.A, LeValley,S.B., Tatum, J.D., Smith, G.C. 2004. Development and validation of equations utilizing lamb vision system output to predict lamb carcase fabrication yields. J. Anim. Sci.. 82:2069–2076