final report

Project code: B.LSM.0020

Prepared by: Stuart Warner,
Department of Primary
Industries, Victoria
Emma Babiszewski &
Janelle Hocking Edwards
South Australian Research &
Development Institute
Kelly Pearce
Murdoch University

Date published: April 2012

PUBLISHED BY
Meat & Livestock Australia Limited
Locked Bag 991
NORTH SYDNEY NSW 2059

Development of a skin cleanliness scoring system

Meat & Livestock Australia acknowledges the matching funds provided by the Australian Government to support the research and development detailed in this publication.

Abstract

An excessively wet and dirty fleece contains high levels of bacteria. During processing if these bacteria contaminate the carcasses is has the potential to pose serious health risks to humans. Pre-transport curfewing of prime lambs and sheep on farm is commonly practised in an effort to reduce soiling of lamb and sheep pelts to improve carcass hygiene. This project developed a skin cleanliness scoring system in the major sheep production states of WA, Vic and SA that can be implemented on-farm and in lairage.

The three point skin cleanliness scoring system subjectively classifies lambs destined for slaughter as either Low (risk of contamination) including lambs with little to no soiling; Medium (risk of contamination), including lambs with moderate soiling or overlong wool; or High (risk of contamination), including lambs with heavy soiling. In applying the three point skin cleanliness scoring system assessors were able to confidently and consistently demonstrate that they were able to identify lambs representing the greatest risk of contamination. The assessors observed that there is no difference in skin cleanliness score of lambs sourced from saleyards or through direct consignment. Furthermore, initial skin cleanliness scores were correlated with skin cleanliness in lairage, regardless of timing and length of transport.

The robustness of this scoring system will assist further research into the many variables that influence skin cleanliness including breed, age, seasonal variation, weather, curfew times, truck cleanliness and individual sheep variation.

Executive Summary

Excessively wet and dirty fleece on sheep and lambs can cause microbiological contamination of carcasses post-slaughter[1] which may pose serious health risks to humans[2]. Pre-transport curfewing of prime lambs and sheep on farm is commonly practised in an effort to reduce soiling of lamb and sheep pelts to improve carcass hygiene. The likelihood of contamination can be predicted by visual appraisal of the level of soiling of a fleece, with dirtier skins resulting in a higher probability of contamination[3,4]. Consequently, visual scoring systems have been developed internationally which can be used on farm, at saleyards and in lairage at abattoirs to inform processors the likely risk of carcass contamination. Prior to this project no such system existed in Australia.

This project has developed a simple and commercially relevant skin cleanliness scoring system that is applicable to Australian production systems which can be utilised across states and processors. Assessors worked with collaborating processors including JBS Australia in Victoria, T&R Pastoral in South Australia and WAMMCO in Western Australia. Lambs were assessed pre and post transport at saleyards (pre transport to lairage), at property of origin (on farm) and at lairage at the processing works. Assessments of skin cleanliness were made in order to determine the following points:

1. Is there a high degree of correlation between the skin cleanliness score of a pen of lambs when scored at different points in the supply chain, e.g. on farm, at saleyards and at lairage.
2. Is it possible for assessors to consistently assess the same pen of lambs on consecutive occasions?
3. Is there significant variation between assessors when assessing the same pen of lambs?

Photos were also taken to be used for the development of a set of drawings identifying the key points of differentiation between skins representing a Low Risk, Moderate Risk and High Risk of carcass contamination.

The initial scoring system was developed at JBS Australia's Brooklyn plant, and further refined during the course of the project to reflect what the assessors agreed represented lambs of the three scores. The scoring system developed for this study was based on scoring systems currently used in the United Kingdom (Meat Hygiene Services) and in similar studies[5]. The scoring system differentiates between breech 'dag' contamination, wool length, wool wetness and contamination along key cutting lines which will adversely affect carcass hygiene. This assessment provides an overall skin cleanliness score that represents low, moderate and high risk of carcass contamination due to fold back along the cutting line.

The scoring system was then used to examine the changes in skin cleanliness during transportation from both saleyards and farms to lairage prior to slaughter. Lambs were assessed multiple times by the same assessor, and by multiple assessors to ensure

[1] Biss, M.E., Hathaway, S.C., (1996) *Journal of Applied Bacteriology*, Vol 81(6), 594-600
[2] Duffy L. L., Small A., Fegan N., (2010), *Australian Veterinary Journal*, Vol 88 (10), 399-404
[3] Byrne, B., Dunne, G., Lyng, J., Bolton, D.J., (2007), *Food Microbiology*, Vol 24 (3), 301-304
[4] Gill, C.O., (2004), *J. Food Protection*, Vol 67 (2), 413-419
[5] Hauge, S.J., Nafstad, O., Skjerve, E., Rotterud, O., Nesbakken, T., (2011), *Int. J. Food Microbiol*., Vol 150, 178-183

consistency and repeatability of the scoring system. Lambs were assessed on a 'per truck' or 'per pen' basis and a random sample of 50 animals (or 25 lambs if there were less than 100 lambs in the group) were assessed to give an accurate representation of the mob.

Twenty seven loads of lambs where assessed at saleyards and again in lairage prior to slaughter (N=10 Vic, N=10 SA; N=7 WA). The data was centrally analysed, and skin score was significantly correlated between initial and lairage score (r=0.93), between saleyard and lairage score (r=0.92) and between on-farm and lairage score (r=0.96). This shows that there was no discernable change in skin cleanliness scores within any of the loads measured. In addition, 17 loads of direct consignment lambs were scored on-farm and again in lairage prior to slaughter (N=7 Vic, N=6 SA; N=4 WA). The average skin cleanliness scores across all saleyards (1.5±0.03) was not different to the average skin cleanliness scores across all direct consignment farms (1.4±0.04) indicating that lambs sourced from the saleyards or delivered directly from farm has no impact on skin cleanliness score. Lambs assessed in lairage had a 0.2 of a score greater risk of carcass contamination (P=0.03), than lambs assessed prior to transport. This may be due to some lambs becoming stained on their bellies and/or brisket during transport.

To demonstrate that the scoring system is transferable across assessors with a high level of repeatability and to further refine the scoring system the project team undertook another group assessment day at Dublin Saleyards in SA. Analysis of this data showed that individual assessors are highly consistent within themselves and correlations of average pen scores across assessors were highly significant between all assessors, ranging from r=0.913 to r=0.789. However, there was variation between assessors. The view of the project team is that with further training and 'calibration' of assessors that this variation will be removed as the bulk of variation occurred in the subjective areas of degree of soiling. Of critical importance is that there was no variation across assessors when identifying those lambs with a high risk of contamination.

While the results show that there was no visual change in skin cleanliness during transportation, the production of a set of visual standards which are applicable to Australian production systems will allow for facilitated training to be undertaken with lamb buyers, lamb producers and lamb processors - and is a valuable tool for potential further research.

Despite extensive work on understanding the mechanisms and impacts of dehydration[6], minimising and alleviating the effects of dehydration[7] and possible welfare implications[8] during curfew prior to transport, there has been limited research into optimisation of curfew times. Currently, recommendations for curfew times vary greatly between States, ranging from 4 to 24 hours. As such, the scoring system developed in this study may be used as a tool in the optimisation of curfew times. Additionally, many variable factors may influence skin cleanliness including seasonal variation, weather, curfew times, truck cleanliness and individual sheep variation. With research into these variables, this scoring system has the potential to be part of a nationally recognised program for the management of lamb curfew periods prior to transport and the control of microbiological contamination of sheep meat.

[6] Hogan, J.P., Petherick, J.C., Phillips, J.C., (2007) Nutrition Research Reviews, Vol 20, 17-18
[7] Toohey, E.S., Hopkins, D.L., Nielsen, S.G., (2006) *Aust. J. Exp. Agric*., Vol 46 (6-7), 903-908
[8] Fisher, A. D., Colditz, I. G., Lee, C., Ferguson, D. M., (2009) J. *Vet. Behav. Clin. App & Res*., Vol 4 (4), 157-162

Contents

1 Background

An excessively wet and dirty fleece on sheep and lambs can cause microbiological contamination of carcasses post-slaughter[9]. *Salmonella*, *Campylobacter* and *Escherichia Coli* (E.Coli) among others, are commonly found in fleece and faecal matter, and carcass contamination with these pathogens can pose serious health risks to humans[10]. Previous studies have found a strong correlation between a visual appraisal of the level of soiling of a fleece and the likelihood of contamination, with dirtier skins resulting in a higher probability of contamination[11,12]. In addition, uncontrolled pelt removal and evisceration are a source of contamination throughout the slaughter process[13]. Accordingly, there is a demand to obtain carcasses with pelts that are visibly free of contaminants. Therefore, several countries, including Ireland[3,14], the United Kingdom (UK)[6], Finland[15] and Norway[6] have developed visual cleanliness scoring systems which can be used on farm, at saleyards and in lairage at abattoirs to exclude heavily soiled animals from slaughter and aid in microbiological control. One such system implemented in the UK uses a 0-3 scoring system where sheep and lambs scoring 2 or 3 are deemed excessively dirty and will not be slaughtered.

Pre-transport curfewing of prime lambs and sheep on farm is commonly practised in an effort to reduce soiling of lamb and sheep pelts to improve carcass hygiene and skin value. The length of time recommended for food and water withdrawal before transport varies greatly around Australia ranging from 4 to 24 hours. In the southern regions of eastern and central Australia (Victoria, South Australia), extended curfews are recommended by most processing companies. However extended curfews do create issues, including (MLA review LIVE.122A):

- Some European supermarket supply chains question the validity of extended curfews on animal welfare grounds.
- There is scope for considerable carcass weight loss (approximately 0.1%/hour after 10-12 hours of curfew; see curfew review 2008; LIVE.122A).
- Food safety experts recommend 24 hours curfew as a precaution against the expression of Salmonella (max time off feed of 48 hours).
- Compliance to Meat Standards Australia (MSA) can be compromised (max time from muster to slaughter is 48 hours).

An extensive review of feed and water curfews was completed for MLA in 2006 (LIVE.122A) and a number of recommendations were made based on how to best optimise food safety, environmental impacts, meat quality, carcass shrink, animal welfare and carcass and skin value. The major conclusions was that the curfew period should have water available (for animal welfare reasons) and the period of feed withdrawal should not exceed 24 hours and indeed could be reduced to 12 hours or less in many situations. Further research relating to animal welfare during transportation has supported these conclusions (AHW.055 - Animal welfare outcomes of livestock road transport practices).

[9] Biss, M.E., Hathaway, S.C., (1996), Journal of Applied Bacteriology, Vol 81(6), 594-600
[10] Duffy L. L., Small A., Fegan N., (2010), Australian Veterinary Journal, Vol 88 (10), 399-404
[11] Byrne, B., Dunne, G., Lyng, J., Bolton, D.J., (2007), Food Microbiology, Vol 24 (3), 301-304
[12] Gill, C.O., (2004), Vol 67 (2), 413-419
[13] Milios, K., Mataragas, M ., Pantouvakis, A., Drosinos, E. H., Zoiopoulos, P. E., (2011), International Journal of Food Microbiology, Vol 146 (2), 170-175
[14] Hauge, S.J., Nafstad, O., Skjerve, E., Rotterud, O., Nesbakken, T., (2011), International Journal of Food Microbiology, Vol 150, 178-183
[15] Ridell, J., Korkeala, H., (1993), Meat Science, Vol 35, 223-228

A robust skin cleanliness scoring system that is accepted nationally is necessary to underpin any work on curfews. The aim of this project is to establish a nationally accepted skin cleanliness scoring system that can be implemented on farm and at lairage. Once established, the skin scoring system would be used to optimise curfew times to produce higher value clean skins and while at the same time minimising curfew times to allow better eating quality outcomes.

2 Project objectives

The following objectives (in italics) have been taken directly from the original project schedule. These have been modified slightly during the course of the project. Where modification has occurred it has been highlighted below.

"*A skin cleanliness scoring system with associated photo standards will be developed that aims to be both simple but importantly commercially relevant. The scoring system needs to differentiate between breach 'dag' contamination, fresh faecal and urine contamination generally spread across the skin and also contamination along key cutting lines which will adversely affect carcass hygiene. This may mean an overall cleanliness score (i.e. 1-5) and a carcass hygiene score (1-5)".*

The project team decided to develop a single scoring system that encapsulated both skin cleanliness and skin hygiene. It was deemed impractical to differentiate between the two without mycological assessment of carcass contamination (to determine hygiene as opposed to cleanliness). The final skin cleanliness scoring system developed utilises three different carcass contamination risk scores: Low Risk, Moderate Risk and High Risk and for the remainder of the report is described as a skin cleanliness score.

3 Methodology

3.1 Development of a skin cleanliness scoring system

The methodology followed by the project team was outlined in the project objectives established by the Meat and Livestock Australia/Sheep CRC II Lamb Supply Chain group in conjunction with the project partners. This methodology is given below. During delivery of the project, the research team further modified the methodology to better match the needs of industry. As a result the overall outcome of the project was the development of a single three point skin cleanliness scoring system.

3.1.1 Five point skin hygiene scoring system

Assessors from Western Australia, Victoria and South Australia travelled to JBS Australia abattoirs in Brooklyn on 20 June 2011, Victoria where a skin hygiene scoring system was devised using visual appraisal of lambs in lairage pre-slaughter. The scoring system developed for this study was based on similar scoring systems used in the UK (Meat Hygiene Services) and in similar studies[16]. An overall cleanliness score (1-5) of lambs in lairage was developed and many photos were taken to document the scoring system.

3.1.2 Three point skin cleanliness scoring system

Skin cleanliness along key cutting lines that affected carcass hygiene was assessed on the kill chain. Photos were taken to demonstrate pictorially skin cleanliness along key cutting lines.

[16] Hauge, S.J., Nafstad, O., Skjerve, E., Rotterud, O., Nesbakken, T., (2011), International Journal of Food Microbiology, Vol 150, 178-183

An overall three point cleanliness score was developed and photos were taken to document the scoring system.
The project team met at Dublin Saleyards, SA to reach final agreement on the scores and assessments. Areas that were causing inconsistency across scorers were identified and these were clarified and a final three point skin cleanliness scoring system was produced.

3.1.3 Economic weight or penalty associated with skin scoring system
Researchers met with WAMMCO Ltd., JBS Swift Australia and T&R Pastoral to determine the economic weight or penalty associated with skin scoring system.

3.2 Repeatability of scoring system

In order to ensure statistically relevant and consistent data across the three States involved, the repeatability of scoring was assessed between and within assessors.

3.2.1 Repeatability of five point skin hygiene scoring system
The initial repeatability assessment was undertaken at JBS Australia Brooklyn processing plant and involved three assessors independently assessing three pens of lambs to ensure the consistency of evaluation. Lambs from Delegate (n=570), Tarcutta (n=420) and Dean (n=1260) were available for assessment. Lambs were assessed on a per pen basis with random samples of 10-20 animals per pen scored using the 1-5 skin cleanliness scoring system. A general modelling procedure for categorical data analysis was used to determine the repeatability of scores across assessors on a group basis (PROC CATMOD; SAS Institute Inc. Cary, NC, USA 2008).

3.2.2 Repeatability of three point skin cleanliness scoring system
Pens and loads of lambs and sheep at saleyards were scored by multiple assessors in South Australia and Western Australia. Each assessor scored the groups two to three times, leaving at least 10 minute intervals between repeats to reduce bias.

Eight assessors across Australia assessed multiple pens of lambs to determine repeatability within and between assessors (Table 1). At Muchea saleyards, three assessors scored using a percentage method of scoring the lambs. The assessors assigned a percentage of lambs to score 1, a percentage to score 2 and a percentage to score 3. At Muchea, Naracoorte and at JBS' feedlot at Yambinya assessors scored up to 50 lambs individually and calculated an average score for the pens.

Table 1 Number of pens of lambs scored by individual assessors, the number of repeated assessments of individual pens, location of assessment and the number of common pens between assessors.

Assessor	Location	Rep 1	Rep 2	Rep 3	Common pens
DP	Muchea	42	39		26
KP	Muchea	39	37		27
MB	Muchea	46	41		27
RJ	Muchea	39	39		27
MK	Muchea	37	2		
EB	Naracoorte	6	6	5	1
JE	Naracoorte	1	1	1	1
SW	Yambinya	6	6	6	

At Dublin Saleyards, four assessors scored 20 lambs individually and calculated the average score for the pens. Only 20 lambs were scored because the pen sizes were small (50-100 lambs/pen). If more lambs were scored there was the risk of repetitive scoring of individual lambs. Five pens were scored twice by each assessor. Chi-squared tests were used to test whether the pattern of scoring different between scorers, were the scorers independent and is the proportion per score different between scorers.

3.3 Variation of skin cleanliness scoring across source and timing of assessment

There were three hypotheses to be tested to determine the variation between on farm or saleyard scoring and lairage scoring and then the suitability of on farm or saleyard scoring:

1. Skin cleanliness score prior to transport (either on-farm or sale yard) is correlated with post-transport (lairage) skin cleanliness score;
2. There is no difference between initial (on-farm or sale yard) and lairage skin cleanliness scores; and
3. There is no difference between on-farm and sale-yard skin cleanliness score.

Skin cleanliness scores were recorded for a total of 27 loads of lambs from saleyard to lairage (10 in Victoria, 10 in South Australia and 7 in Western Australia), and 17 loads of directly consigned lambs from farm to lairage (7 in Victoria, 6 in South Australia and 4 in Western Australia). An additional 13 loads had a single assessment (4 loads of lambs assessed in lairage from direct consignments; 6 loads of lambs in the saleyards; 3 loads of lamb in lairage sourced from saleyards).

The skin cleanliness scores from the 44 loads of lambs that were measured prior to transport and in lairage were used to test Hypothesis 1 (Table 2.). Skin cleanliness scores from all 57 loads of lambs were used to test Hypotheses 2 and 3.

Table 2. Number of loads of lambs that had skin cleanliness scores assessed.

Location/Source	Direct	Saleyard	Total	*Hypothesis*
Initial & lairage (same lambs)	17	27	44	*1*
Initial	17	33	50	*2&3*
Lairage	21	30	51	
Total	38	63		

The effect of assessor (SW, EB, and KP), source (direct consignment or saleyard) and timing (pre-transport, lairage) on skin cleanliness was determined using generalised linear mixed models with maximum-likelihood methods (PROC GLIMMIX; SAS Institute Inc. Cary, NC, USA 2008). A multinomial model for ordinal data was used. Pen was included in the model as a random effect. Pearson correlation coefficients were calculated to describe the correlations between location of assessment.

4 Results

4.1 Development of scoring system

4.1.1 Five point skin hygiene scoring system

There was no significant difference between the three assessors (Table 3) when using a five point scoring system to score the skin hygiene across all lambs (P=0.1271), within lamb origin (P=0.1784) or within pens (P=0.2045), indicating that the skin hygiene scoring system was repeatable across the three assessors. This scoring system however was deemed too detailed and too complicated for application on a wider scale. This factor combined with the view that skin cleanliness can be adequately described with a three point system resulted in any further development of a five point skin hygiene system being abandoned.

Table 3. Weighted least square estimates (standard error) skin cleanliness scores of three assessors across all lambs, across origin and across pens.

Scorer	All	Origin	Pen
SW	2.18 (0.038)	1.92 (0.068)	2.15 (0.044)
EB	2.28 (0.052)	1.97 (0.052)	2.20 (0.054)
KP	2.10 (0.054)	1.98 (0.056)	2.20 0.058)

4.1.2 Three point skin cleanliness scoring system

Skin cleanliness along key cutting lines that affects carcass hygiene was assessed on the kill chain and it was concluded that the original skin cleanliness scoring system developed in lairage at JBS Brooklyn was too broad. Skin cleanliness is adequately described by a three score system, broadly divided into the following categories:

1. **Low Risk** of carcass contamination;
2. **Moderate Risk** of contamination; and
3. **High Risk** of contamination.

The scores range from 1 (least soiling) to 3 (most soiling) and are designed to represent low, moderate and high risk of carcass contamination due to fold back along the cutting line. A

three point scoring system was determined based on a subjective assessment of the degree of fleece contamination and the associated risk of carcass contamination. Due to other numeric scoring systems used to describe both the condition score and fat score of lambs as well as the skin categories of buyers the decision was made to use the more descriptive terms of Low Risk, Moderate Risk and High Risk when referring to the individual scores.

The project team met at Dublin Saleyards, SA to reach final agreement on the scores and assessments. The following areas were identified as causing the bulk of the inconsistent assessment across assessors.

- Wool length short enough to be classed as 'recently crutched'
- Bare breech
- Dag length sufficient to cause downgrading

Definitions of Dag Length and Wool Length of the breech were added to the original table to create the skin cleanliness scoring system and risk or carcass contamination table (Table 4).

Table 4: Skin cleanliness scoring system and risk of carcass contamination

	Score		
Risk of carcass contamination	**Low Risk**	**Moderate Risk**	**High Risk**
	Assessment Factors for consideration		
	Faecal/dirt		
Belly & Brisket	Nil	Medium	Heavy
Body	Nil	Medium	Heavy
Breech	Nil	Medium	Heavy
	Wool length		
Belly	<50mm	>50mm	>50mm
Body	<50mm	>50mm	>50mm
Crutch	<8mm	>8mm	>8mm
	Dags		
	Nil	Medium: Dags exist but are; • short, • exist at the breech only, • closely attached to the wool, • not swinging or hanging.	Heavy: Regardless of length dags will be swinging presenting a greater risk of contamination.
	Wetness		
Belly & Brisket	Dry	Damp to touch	Saturated
Body	Dry	Damp to touch	Saturated

A library of photos has been collected to illustrate the three risk contamination categories. Examples of these can be seen in Figures 1-3.

Figure 1. A score 1 (Low risk of contamination) animal with fleece that is dry and free from faecal and urine contamination and dags. Fleece length is less than or equal to 50mm, and will present a low risk of carcass contamination.

Figure 2. A score 2 (Moderate Risk of contamination) animal has fleece with moderate levels of faecal contamination on belly and brisket, and may by dry or slightly damp. Wool length is equal to or greater than 50mm in length and poses a moderate risk of carcass contamination.

Figure 3. A score 3 (High Risk of contamination) animal has fleece with heavy faecal and urine contamination on belly, brisket and legs which may extend to the back and heavy dags. The fleece will by damp or wet with wool length greater than 50mm and will pose a high risk of carcass contamination.

4.1.3 Economic weight or penalty associated with skin scoring system

All three processors interviewed indicated that it is unlikely that there will be any difference in skin value between skins classed using the skin cleanliness scoring system described in Table 4. This is due to the fact that, under normal circumstances, faecal and dirt contamination on the skins can be removed by washing and/or trimming. Therefore, though it is the principle factor in determining the cleanliness score of the skin, it is not a factor in determining the economic value of the skin. However, there is a possibility that some skins with a High Risk cleanliness score could be down graded, depending on the level of soiling

In the current skin cleanliness scoring system, a High Risk skin has long wool (>50mm) with significant soiling of the legs, belly, brisket and breech and may have moderate contamination on the back and sides (Table 4). For skin buyers, only skins with soiling of the entire length of the staple over the majority of the skin will be down graded to Skin Grade 3 or 4 (using the skin buyers grading system). This is likely to decrease the value of the skin by approximately $10 (approximately 50%) at current prices (2012). However not all lambs with a High Risk skin cleanliness score will have extensive soiling along the entire length of the staple or even over the entire body. Thus lambs with a High Risk skin cleanliness score would not consistently result in decreased in skin value.

Other characteristics such as breed (particularly “cleanskins”), fleece rot, seed (grass seeds and burr) and fleece density generally have a greater impact on skin value than the level of soiling. As such, some lambs with a Low Risk skin cleanliness score will be down graded due to the presence of grass, burrs etc. It is likely that lambs that are ‘over crutched’, perhaps to remove dags to increase skin cleanliness may have a lower skin value, though this depends on the market of the time.

4.2 Repeatability of three point skin cleanliness scoring system

In all assessments, there was no significant difference within individual assessors ($P=0.3674$), indicating good repeatability within the assessors. Furthermore, there were significant differences in skin cleanliness scores between pens of lambs, indicating that individual assessors were able to consistently score across a range of skin cleanliness scores. There was no effect of count or percent method on skin cleanliness score.

There was a significant difference between assessors in the skin cleanliness scores given to each pen of lamb ($P<0.0001$; Table 5). At the Muchea, KP scored more lambs with a moderate risk of contamination whereas DP, RJ, MK & MB scored more lambs with Low Risk of contamination (Score 1). All assessors at Muchea scored a similar portion of lambs with High Risk of contamination. There was no difference in average score and the proportion of lamb in each risk category in lambs scored at Naracoorte.

Table 5. Least square estimates (LSM; standard error) of the original skin cleanliness scores of eight assessors across all lambs, and the percentage of lambs that each assessor allocated to risk of contamination.

Assessor	Location	Method	LSM	Low Risk	Mod. Risk	High Risk
DP	Muchea	Percent	1.45 (0.061)	58%	32%	10%
KP	Muchea	Count	1.81 (0.061)	29%	61%	10%
MB	Muchea	Percent	1.54 (0.061)	49%	44%	7%
RJ	Muchea	Count	1.61 (0.061)	57%	30%	13%
MK	Muchea	Percent	1.44 (0.062)	44%	49%	7%
EB	Naracoorte	Count	1.65 (0.167)	37%	62%	1%
JE	Naracoorte	Count	1.56 (0.174)	40%	58%	2%
SW	Yambinya	Count	1.24 (0.183)	58%	32%	10%

Assessment of the repeatability of the reviewed skin cleanliness scoring system at Dublin confirmed the results of repeatability of the three point scoring system. There was no significant variation within assessor, indicating that individuals are consistent in the assessment of skin cleanliness. Correlations of average pen scores across assessors were highly significant between all assessors (Table 6) and ranged from r=0.913 between JE and SW to r=0.789 between JE and EB.

Table 6. Correlations between the four assessors of the reviewed skin cleanliness scores of across 11 pens of lambs. (Pearson Correlation Coefficients above the diagonal and level of probability below the diagonal)

Assessor	EB	JE	KP	SW
EB	1	0.789	0.864	0.747
JE	0.0039	1	0.802	0.913
KP	0.0006	0.003	1	0.702
SW	0.0083	<.0001	0.0161	1

There was a significant difference in average cleanliness score between assessors (Table 7; P<0.0001). The average difference in score across the four assessors when assessing the same lambs was less than 0.2 of one score.

There was a significant difference in the pattern of scoring between assessors, with SW and JE scoring lambs to have a lower risk of carcass contamination. This is due to EB and KP scoring fewer lambs as Low Risk and more lambs as Moderate Risk when compared to SW and JE. All assessors scored a similar number of lambs as having a High Risk of contamination. This is an important point as it identifies the ability of assessors to agree on those lambs which present the greatest risk of carcass contamination.

Table 7. Least square estimates (LSM; standard error) of the reviewed skin cleanliness scores of four assessors across all lambs, and the percentage of lambs that each assessor allocated to each level of risk of contamination.

Assessor	LSM	Low Risk	Moderate Risk	High Risk
SW	1.50 (0.101)	57%	35%	8%
JE	1.54 (0.101)	51%	43%	7%
EB	1.70 (0.101)	40%	48%	12%
KP	1.70 (0.101)	36%	57%	7%

4.3 Variation of skin cleanliness scoring across source and timing of assessment

There was a very strong positive correlation (P<0.001) between skin cleanliness scores at the source (either saleyard or on-farm direct) and skin cleanliness scores of the same animals in lairage prior to slaughter (Table 8), indicating there was little change in skin cleanliness during transport. This trend was consistent across assessors (and therefore states and processors), saleyards and farms. The transport times and times off feed varied dramatically between loads, with some direct consignment transport times as little as 30 minutes off feed prior to loading onto transport to lairage in contrast to some saleyard loads which had transport times in excess of 4 hours and were off feed for up to 48 hours prior to delivery to lairage.

Across all assessors, there were significant correlations between skin cleanliness scores when scored prior to transport and at lairage (Table 8). This was consistent regardless of whether initial scoring was undertaken at the saleyards or on farm prior to consignment to the processing plant. Furthermore, there was also a significant correlation within assessors for pre- and post transport skin cleanliness scores (Table 8).

Table 8. Pearson correlation coefficients between initial (saleyard or direct) skin cleanliness score and score at lairage.

	Initial-Lairage	Saleyard-Lairage	Direct-Lairage
All assessors	0.93 (P<0.001)	0.92 (P<0.001)	0.96 (P<0.001)
SW	0.87 (P<0.001)	0.90 (P<0.001)	0.91 (P<0.05)
EB	0.82 (P<0.001)	0.64 (P<0.05)	0.94 (P<0.01)
KP	0.94 (P<0.001)	0.75 (P=0.05)	0.96 (P<0.05)

The skin cleanliness score of lambs sourced from saleyards (1.52±0.034) was not significantly different to the skin score of direct consigned lambs (1.43±0.042; P=0.2474). This indicates that lambs sourced from the saleyards or delivered directly from farm has no impact on skin cleanliness score, at least in the season and conditions that these lambs were assessed.

Lambs assessed in lairage were 0.2 (P=0.03) of a score poorer, or expressed another way, presented 0.2 of a score greater risk of carcass contamination, than lambs assessed prior to

transport. This may be due to some lambs becoming stained on their bellies and/or brisket during transport.

There was no interaction between source and timing of assessment. This means that a similar change in skin cleanliness score occurs between initial measurement and measurement in lairage, regardless of whether lambs were initially measured on-farm or at the saleyards.

There was a significant difference between assessors ($P<0.0001$) and a significant interaction between assessors and timing of assessment ($P<0.0001$; Figure 4). It is not possible to determine the reason for the difference as the assessor is confounded with location. There were no common assessments made on these lambs, so differences in assessors could be due to location (Vic, SA, and WA).

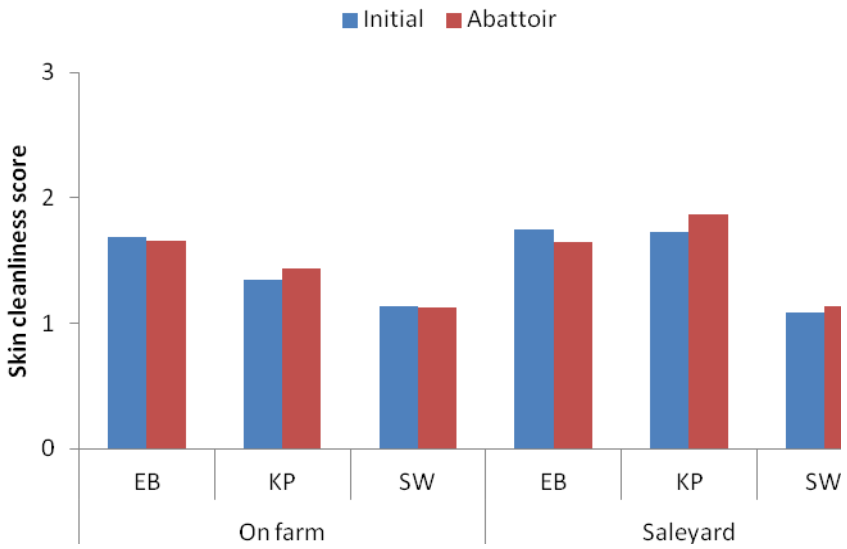

Figure 4. Skin cleanliness scores (LSM) at initial assessment (either saleyard or on-farm) and at lairage for each assessor (EB, KP, SW).

5 Discussion

The results from this project highlight the following key points:

- That when using a five point scoring system there is no significant difference across assessors, with the origin of the lamb or within pens. Therefore it is possible to get consistent repeatable scores using a visual appraisal scoring method. This is an important point as it indicates the opportunity to further develop the 3 point system to a level where across assessor consistency is achievable.
- The project team, through the development of the skin cleanliness scoring system have rejected the need to develop a separate skin hygiene score. This is due to the fact that the scores are virtually interchangeable and that carcass hygiene, rather than skin value, is the most important factor.

- There is no significant change in skin cleanliness during transport from farm or saleyards to the abattoir, regardless of timing and length of transport. As a result, if lambs are assessed on farm prior to transport or at lairage post transport, it is likely that the scores will be extremely similar, thus one scoring location is a good predictor of how the animals will score at another. However, anecdotal evidence would suggest that dirty lambs on farm have a greater likelihood of becoming more soiled throughout transport as a result of increased soiling of truck floor and therefore an increased likelihood of slipping during transport. The cumulative soiling of truck floors would also affect subsequent loads which may be clean on-farm. Further research into this area is required, and the use of this scoring system to classify skin cleanliness changes would further validate the reliability of the scoring system. Attempts were made to assess the cleanliness of trucks at lairage. As no scoring system was developed for assessing the cleanliness of trucks post unloading it is difficult to utilise this data. This is a further area for research to identify if the cleanliness of trucks post transport is related to the skin cleanliness score of lambs when delivered to lairage.

- There is no significant difference if lots of lambs are assessed utilising the approximation of the percentage of lambs falling within each score, or assessed utilising a system where the score of individual lambs within each pen is assessed and tallied into a cumulative percentage for each score. Therefore both methods are equally valid. However it was the unanimous view of the project team that a more thorough method of assessing the score of each pen was to utilise the individual score tally method. It was felt that this was a more robust method of assessment.

- Significant variation between assessors may occur, particularly when differentiating between lambs assessed as being of Low Risk of contamination and those of Moderate Risk.

- Assessors are confidently and consistently able to identify those lambs which they view as representing the greatest risk of contamination. Importantly, if feedback were to be given to suppliers of lambs utilising this scoring system, then different assessors will be consistent in identifying those lambs which present the greatest risk of contamination.

- The project was unable to identify why there was significant variation between assessors, particularly in the Low Risk and Moderate Risk of contamination scores. The fact that a number of the assessment criteria are objective, such as wool length, wetness and the determination of whether a lamb had been recently crutched or not, suggests that it was the degree of soiling of the skin that was the area of the most contention. It is important to highlight the point that the degree of assessment variation across all four assessors involved in the project across all lambs was only 0.2 of a score. While this may be a significant variation statistically, it was felt by the project team that it was not an untenable level of variation. It is also the view of the project team that through further development of visual standards and facilitated training, that variation in the areas resulting in the statistical variation between assessors could be reduced or removed.

- It is of note that the skin cleanliness score is not an indicator of skin value, and would not influence the value placed on skins by any of the processing works involved in the trial. This is not a rejection of the value of the scoring system. The system has been developed to enable further research to be conducted into curfew periods prior to transport and to determine if assessment at one point in the lamb delivery chain reflects the assessment at another point, meaning that the scoring system can be utilised for feedback purposes. The fact that scores at one point, either on farm, at the saleyards or in lairage reflect with a high degree of confidence the scores applied by the same assessor at another point in the chain, means that the scoring system can be utilised as a reliable measure of a lambs skin cleanliness at any point in the delivery chain. Therefore the system is robust enough to be utilised for further research into curfew lengths and critical control points for improving the skin cleanliness score of lambs.

While there has been extensive work on understanding the mechanisms and impacts of dehydration (Hogan et al., 2007), minimising and alleviating the effects of dehydration (Toohey et al., 2006) and possible welfare implications (Fisher et al., 2009) during curfew prior to transport, there seems to be limited research into optimisation of curfew times. Currently, recommendations for curfew times vary greatly between states, ranging from 4 to 24 hours. Although this parameter was not assessed in the present study, it is another area in which research is required and in which this scoring system would be a valuable research tool.

6 Conclusion

The project successfully developed a national skin cleanliness scoring system that was accepted by the three partner lamb processing organisations. It was shown that the assessment, when made utilising the guidelines here, is consistent within an assessor and at different times through the lamb chain. The project also proved that an assessment at one point in the chain is provides an accurate estimation of the skin cleanliness at another point in the chain.

It was not possible during this project to demonstrate that consistent assessment across assessors could be achieved to statistically acceptable level. However the degree of variation that existed between assessors was only 0.2 of one score and therefore it is the view of the project team that with further training and calibration this level of variation would be removed. As the greatest area of contention between assessors was describing the degree of contamination on a skin; the development of an objective measurement may help in this area. However any such objective measurement would need to be repeatable in the field. It is therefore suggested that facilitated training be utilised to remove any such variation, in place of further objective measurement development.

Due to the fact that the assessment of lambs at one point in the chain accurately reflects their cleanliness at another level, it is possible to use this assessment method as a direct feedback tool to producers. Importantly if the same assessor is conducting the assessment at both points in the chain there will be a high degree of predictability in how the lambs will score.

As there is no economic penalty that easily aligns with the skin cleanliness scores, the system is not useful for determining economic penalties with each score. This is not a failure of the system as it was not intended to be utilised as such.

The system is robust enough to be utilised for further research and development purposes, particularly for the research into reducing curfew times prior to transport.

7 Bibliography

Biss, M.E., Hathaway, S.C., (1996) Microbiological contamination of ovine carcasses associated with the presence of wool and faecal material, *Journal of Applied Bacteriology*, Vol 81(6), 594-600

Byrne, B., Dunne, G., Lyng, J., Bolton, D.J., (2007) The development of a 'clean sheep policy' in compliance with the new Hygiene Regulation (EC) 853/2004 (Hygiene 2), *Food Microbiology*, Vol 24 (3), 301-304

Duffy L. L., Small A., Fegan N., (2010) Concentration and prevalence of Escherichia coli O157 and Salmonella serotypes in sheep during slaughter at two Australian abattoirs, *Australian Veterinary Journal*, Vol 88 (10), 399-404

Fisher, A. D., Colditz, I. G., Lee, C., Ferguson, D. M., (2009) The influence of land transport on animal welfare in extensive farming systems, *Journal of Veterinary Behaviour: Clinical Applications and Research*., Vol 4 (4), 157-162

Gill, C.O., (2004) Visible contamination on animals and carcasses and the microbiological condition of meat, *Journal of Food Protection*, Vol 67 (2), 413-419

Hauge, S.J., Nafstad, O., Skjerve, E., Rotterud, O., Nesbakken, T., (2011) Effects of shearing and fleece cleanliness on microbiological contamination of lamb carcasses, *International Journal of Food Microbiology*, Vol 150, 178-183

Hogan, J.P., Petherick, J.C., Phillips, J.C., (2007) The physiological and metabolic impacts on sheep and cattle of feed and water deprivation before and during transport, *Nutrition Research Reviews*, Vol 20, 17-18

Milios, K., Mataragas, M ., Pantouvakis, A., Drosinos, E. H., Zoiopoulos, P. E., (2011) Evaluation of control over the microbiological contamination of carcasses in a lamb carcass dressing process operated with or without pasteurizing treatment, *International Journal of Food Microbiology*, Vol 146 (2), 170-175

Ridell, J., Korkeala, H., (1993) Special treatment during slaughtering in Finland of cattle carrying an excessive load of dung; meat hygienic aspects, *Meat Science*, Vol 35, 223-228

Toohey, E.S., Hopkins, D.L., Nielsen, S.G., (2006), Effectiveness of electrolytes at alleviating dehydration and improving carcass characteristics of adult sheep in lairage *Australian Journal of Experimental Agriculture*, Vol 46 (6-7), 903-908