TITLE: Probabilistic modelling of events at evisceration during slaughtering of pigs using expert opinion: Quantitative data in support of stochastic models of risk of contamination

AUTHORS: M. Crotta; E. Luisi; N. Dadios; J. Guitian

JOURNAL: Microbial Risk Analysis

PUBLISHER: Elsevier

PUBLICATION DATE: 26 October 2018 (online)

DOI: https://doi.org/10.1016/j.mran.2018.10.001

Probabilistic modelling of events at evisceration during slaughtering of pigs using expert opinion: Quantitative data in support of stochastic models of risk of contamination

Matteo Crotta[a], Elena Luisi[b], Nikolaos Dadios[a], Javier Guitian[a]

[a] Veterinary Epidemiology, Economics and Public Health Group. Department of Pathobiology and Population Science. The Royal Veterinary College. Hawkshead Lane. North Mymms. AL9 7TA, Hatfield, UK

[b] Pig Health Scheme assessor, HallMark Ltd

ABSTRACT

The evisceration stage is one of the most critical steps in the slaughtering process of pigs when considering the risk of carcass contamination. Unfortunately, it is also characterized by a number of fundamental quantitative data gaps preventing modellers from reproducing events in probabilistic terms. Recognising the practical difficulties that a systematic data collection would imply, in this study we modelled the answers of structured questionnaires submitted to eleven veterinarians (official veterinarians/meat hygiene inspectors) working in pig abattoirs to provide ready-to-use probability distributions in support of future quantitative risk assessments. The questions were aimed at modelling the occurrence of ruptured gut (P_{GUT}) and gallbladders (P_{GALL}) during evisceration procedures, the amount of faecal (F_L) and bile (B_L) contamination dropping on the carcass, the probability of internal cavities (P_{IF-B}) and external surface (P_{EF-B}) being contaminated and the conditional probability of partial condemnation of the carcasses (as unfit for human consumption) as a function of the level of contamination (P_C^{Sa}). The answers were weighted according to the level of confidence each expert had in their own estimation. Out of 10,000 simulated values, P_{GUT} and P_{GALL} were higher in small (Mean=0.048 and 0.035) compared to high (Mean=0.021 and 0.016) or middle (Mean=0.025 and 0.019) throughput abattoirs. The cumulative distributions describing F_L and B_L produced 50th and 90th percentile values of 24.5g and 19.9g (50th percentile) and 88.7g and 68.8g (90th percentile), indicating the level of contamination is generally low. The distributions describing both P_{IF} and P_{EF} and those describing P_{IB} and P_{EB} show comparable shapes suggesting there are no significant differences in the likelihoods of those events when considering the faecal and bile contamination respectively. Finally, the results obtained for P_C^{Sa} suggested that common linear or nonlinear relationships are not adequate to describe the probability of a carcass being partially condemned as a function of the dose. Highly contaminated carcasses are not unlikely to be detained for manual removal of visible contamination rather than partially condemned, indicating that factors other than the amount of contamination are driving this relationship. With this study, we made use of the experience of eleven Meat Hygiene Inspectors/Official Veterinarians to provide quantitative information on the key events occurring during evisceration. As presented, the probability distributions can be directly used to inform and integrate probabilistic models aimed at estimating to the risk of human exposure to foodborne pathogens through consumption of pork products.

1. INTRODUCTION

Pork meat represents 9.0 % of the total agricultural output of the European Union (EU) and is the major type of meat produced in the 28 EU Member States [1]. At global level, the per capita consumption of pig meat in both high and low-income countries steadily ranks in first position since the 1980s ahead beef and poultry [2]. According to predictions from the Food and Agriculture Organization of the United Nations (FAO), consumption of poultry meat will only overtake pig meat in 2030 [2]. Pork meat and pork-derived food products are a well-known source of foodborne illness in humans, particularly due to gastrointestinal pathogens such as *Salmonella spp.*, *Yersinia enterocolitica, [3, 4]* and more recently, Hepatitis E Virus [5, 6].

Quantitative Microbial Risk Assessment (QMRA) models are increasingly recognized as a powerful approach in support of transparent and science-based decision-making in relation to food safety [7, 8]. In recent years, QMRAs have been widely used by international agencies such as FAO and the European Food Safety Authority (EFSA) to aid the identification of critical intervention points in farm-to-fork pathways of human exposure to different zoonotic pathogens [4, 9, 10]. In the last two decades, numerous probabilistic models aimed at quantify the risk of human exposure to foodborne hazards and/or inform strategies for surveillance and control have been implemented and published [11-14]. As any mathematical model, the degree of credibility of a QMRA is inevitably dependent on the quality and quantity of the data and the assumptions made [15, 16]. This is particularly relevant for very comprehensive farm-to-fork models composed by several sub-modules such as “farm”, “slaughter”, “production”, and “consumer” and in which data related to one or more key inputs are often unavailable to researchers.

QMRAs modelling the fate of pathogens harboured by pigs along the pork chain face the challenge of representing the occurrence of events related to carcass’ contamination and cross-contamination during evisceration at the abattoir [17]. These events can be very influential in the final risk estimates and empirical data to support their detailed representation is almost invariably absent. Several studies [18-20] have identified evisceration as a critical stage during the slaughtering process with respect to microbial contamination of the carcasses (this is mainly due to the possible rupture of the intestine and internal organs). However, because of the practical difficulties in obtaining quantitative estimations about key events such as the occurrence of gut ruptures and the amount of faecal contamination of the carcass following a rupture, the models published so far used distributions to describe the author's best guess values or the opinion obtained from the quality manager of the slaughterhouse [17, 21]. How often guts and gallbladder are accidentally ruptured during belly-opening/evisceration, where and to what extent the resulting leakage of faeces/bile is more likely to occur and what is the likely fate of the contaminated carcass are key quantitative parameters for the development of pork-related QMRAs if the slaughtering stage is to be considered [17].

Recognizing a substantial lack of data for those key inputs and their importance when modelling the fate of foodborne pig pathogens, the aim of this study was to provide quantitative data and ready-to-use probability distributions describing the events occurring at evisceration to help the parameterization of future QMRAs related to gastrointestinal pathogens in pigs. Data were obtained by modelling the answers to structured questionnaires submitted to veterinarians working in different pig abattoirs in the UK as Official veterinarians (OVs) or meat hygiene inspectors (VMHIs). Both OVs and VMHIs are trained to ensure all hygiene standards for the meat production industry are met to protect consumers; this includes post-mortem inspection for detection of abnormalities (including contaminations) of all the carcasses and offal of slaughtered animals. For this reason, and the fact that they work/have worked in different abattoirs, these professionals were considered as the ideal subjects to interview for the scope of this study. Through the combination of the working experience of different official veterinarians in pig abattoirs, this study does not have the ambition to provide accurate point estimate values but rather to capture the plausible range of these values and their associated likelihoods/probabilities

2. MATERIAL AND METHODS.

Data used in this study were obtained from a structured questionnaire submitted to 11 OVs/VMHIs working in pig abattoirs in UK. The questions were devised to specifically explore the events and related consequences occurring during the belly opening/evisceration stage, namely:

(i) the occurrence of ruptured guts/gallbladders during belly opening/removal of gastro-intestinal tract at evisceration,
(ii) the amount of faeces/bile leaking outside in the event of a rupture,
(iii) the location of the contamination,
(iv) the dose-conditional fate of the faecal/bile-contaminated carcasses.

In order to ensure that: (i) interviewees were asked to estimate quantities they had real knowledge of, (ii) the questions were not ambiguous and (iii) easily answerable; all the questions were written in consultation with two VMHIs. The parameters of interest are, by their own nature, characterized by high variability and uncertainty, therefore, the intent of the questions was not to get accurate estimations of unknown fixed parameters, but instead, at reducing the overall lack of knowledge surrounding these values. For this reason, interviewers were invited to express their quantitative estimations using qualitative terms or as ranges such as "*from xx to xx*" or "*from xx to xx but most likely xx*". Moreover, considering that some values might be difficult to estimate even for experienced veterinarians, interviewees were given the opportunity to provide the level of confidence from 1 to 5 they had in their own estimations (i.e. not confident, low confidence, fairly confident, confident, very confident). These confidence scores were used to weight their answers and assign more emphasis to the more confident interviewees.

2.1. Recruitment to the study

Potential recruits were contacted by email between April and June 2017 and asked for consent to participate in this study. Considering that the OVs and VMHIs working in abattoirs are not necessarily and/or regularly assigned to slaughterhouses processing pigs, a convenient pre-selection of the interviewees was considered appropriate to ensure the questionnaire was sent to suitable persons. To this end, the veterinarians deployed on the Pig Health Scheme on behalf of Agriculture and Horticulture Development Board (AHDB) working on rotation in different slaughterhouses in UK and veterinarians with considerable experience in pig's abattoirs (personal knowledge of co-authors) were selected to be included and contacted. To ensure recruitment of interviewees with an appropriate level of experience, as preliminary information they were asked to indicate the approximate number of carcasses they had seen in their professional experience in pig slaughterhouses; all the veterinarians ultimately enrolled in this survey answered "from hundreds of thousands to one million" or "more than one million".

2.2. Expected occurrence of ruptured guts and gallbladder

The first question of the questionnaire was aimed at estimating how often the gut and the gallbladder are accidentally ruptured/damaged during the evisceration procedure. In order to capture differences that may originate from the capacity of the abattoir, interviewees were asked to consider four different scenarios:

1) High throughput slaughterhouse (HTS) processing from 3,000 to 4,000 pigs/day

2) High throughput slaughterhouse with robotic belly opener (HTS-r) processing from 3,000 to 4,000 pigs/day
3) Middle throughput slaughterhouse (MTS) processing from 400 to 600 pigs/day
4) Low throughput slaughterhouse (LTS) processing from 10 to 50 pigs/day.

The second scenario was included to consider that high throughput slaughterhouses might make use of automated robotic belly openers and explore possible differences due to the presence of this technology.

For each scenario, respondents were asked to estimate the number of accidental ruptures of the gut (s_{GUTi}) and gallbladder (s_{GALLi}) it would be reasonable to expect in a typical working week, considering a number of processed animals equal to 20,000 for HTS and HTS-r, 2,500 for MTS and 100 for LTS. In each scenario, the answers were used to parameterize a pert distribution describing the perceived variability in the number of ruptures (s_{GUTi}) and (s_{GALLi}):

$$s_{GUTi} and\ s_{GALLi} \sim Pert(Min_i;\ Max_i;\ Mlik_i) \qquad \text{Eq. 1}$$

Where Min_i, Max_i and $Mlik_i$ are the Minimum, Maximum and Most likely number of ruptured gut and gallbladder expected by each *i*th expert in each scenario respectively.

To account for differences in the answers and the level of confidence amongst interviewees' responses, the pert distributions were combined to obtain a distribution representing the uncertainty in the perceived variability of the expected number of ruptured guts (s_{GUT}) and gallbladders (s_{GALL}). To this end, individual estimations were included into the discrete distribution: $Discrete(\{x\};\{p\})$ where $\{x\}$ is the vector of the expert opinions (i.e. pert distributions) and $\{p\}$ that of the weights (w_i) given to each opinion according to the expert's confidence score. This way, values from each pert distribution are sampled with a frequency proportional to the confidence score of the pert distribution they belong to:

$$s_{GUT}\ and\ s_{GALL} \sim Discrete(Pert_1, Pert_2 \ldots Pert_{11}; w_1, w_2 \ldots w_{11}) \qquad \text{Eq. 2}$$

Where $Pert_i$ represents the estimation of the i^{th} interviewee, and w_i is the confidence score (from 1 to 5) of the i^{th} interviewee in his/her own estimation.

Finally, the probability of rupturing the gut (P_{GUT}) or the gallbladder (P_{GALL}) in each scenario was described by the beta distribution:

$$P_{GUT}\ and\ P_{GALL} \sim Beta\ (s+1; N-s+1) \qquad \text{Eq.3}$$

Where s is the estimated number of ruptured guts or gallbladder ($s_{GUT}\ or\ s_{GALL}$) and N the total number of animals that were considered in each scenario (i.e. 20,000. 2,500 and 100 for HTS/HTS-r, MTS and LTS respectively).

To fit the final probability distributions describing P_{GUT} and P_{GALL}, 10,000 values of s_{GUT} and s_{GALL} for each scenario were simulated (Equation 2) and used to obtain a dataset of 10,000 values of P_{GUT} and P_{GALL} (Equation 3). In the case of bimodal or multimodal-shaped frequency distributions of the generated datasets of P_{GUT} and P_{GALL} obtained for each scenario (i.e. HTS, HTS-r, MTS and LTS), k-mean cluster analysis ("kmeans" function in R) was used to infer the optimal number of clusters (i.e. distributions) to consider and assign each simulated value to one of the identified clusters. To parameterize the distributions identified by the cluster analysis, the Maximum Likelihood Estimation (MLE) for a beta distribution was used. Assuming that a given set of data can be described by a certain distribution (e.g., beta), the MLE estimates the distribution's parameter(s) so that the joint probability of the observed data under the resulting distribution is maximized, formally:

$$logL(X|\varphi) = \sum \log(f(x_i, \varphi)) \qquad \text{Eq.4}$$

where φ represents the parameter(s) of the distribution of the likelihood function (α and β of the beta distribution) and $logL(X|\varphi) = \sum \log(f(x_i, \varphi))$ is the likelihood of observing the *x* observations given φ. Values ranged from 0 to 1 and the beta distribution was chosen as it describes the uncertainty around the probability of occurrence in a binomial process.

Finally, for each scenario, the number of observations belonging to each cluster was used to inform a multinomial process and allow the model to sample the values from each beta distribution a number of times that is proportional to the occurrence of the each respective cluster. Recognising that P_{GUT} and P_{GALL} could in reality be described by unique underlying probability distributions, the presence of clusters in this case is representative of the different perceptions, experience and degree of confidence the interviewees had for this phenomenon.

2.3. Amount of faeces/bile leaking onto carcass

The second question aimed at estimating the overall amount of faeces and bile leaking outside the intestine and gallbladder following accidental rupture.

This question was presented to interviewees using an interactive Web app developed using the R Package "Shiny" [22] available at: https://mcrvc.shinyapps.io/expmra/. Experts were asked to move the two slidebars in order to reach the desired numbers (XX) in the statements below, so that the statements were as close as possible to their belief:

"I expect the overall amount of faeces remaining on carcass following accidental rupture of the gut to be less than XX grams about 50% of the time";

"9 ruptures in 10 will result in a contamination smaller than XX grams and 3 in 10 smaller than XX grams";

Using the slidebars, experts were unconsciously shaping the "most likely/mode" and the "shape" parameters of a modified pert distribution as defined by the R package "mc2d" [23] with 300 and 125g for faeces and bile respectively selected as assumed maximum values.

Interviewees were then asked to report in the questionnaire the values of the slidebars together with their level of confidence. Those values were used to reproduce their pert distributions and final distributions representing the variability in the amount of faeces (F_L) and bile (F_B) were obtained using the same logic as described in section 2.2 (Equation 2):

$$L_F \text{ and } L_B \sim Discrete(Pert_1, Pert_2 \dots Pert_{11}; w_1, w_2 \dots w_{11}) \qquad \text{Eq.5}$$

Where $Pert_i$ represents the estimation of the i^{th} interviewee, and w_i is the confidence score (from 1 to 5) of the i^{th} interviewee in his/her own estimation. Again, the cumulative distributions describing F_L and B_L were obtained running 10,000 iterations.

2.4. probability of internal and external surface of the carcass being contaminated

For the quantitative modelling of faecal and bile contamination, it is important to understand where contamination of the carcass is more likely to occur following the rupture of the gut or the gallbladder. Recognizing that the strong stochastic component of this event prevents reasonable predictions of the accurate spatial distribution of the contamination; the purpose of the third question was limited to understanding how often the internal cavities and the external surface of the carcass are expected to be contaminated in case of accidental rupture of the gut and gallbladder.

For this question, interviewees were asked to estimate the likelihood that faecal and bile contamination drops on the internal and external surface of the carcass following accidental rupture of the gut or the gallbladder.

Respondents could choose from a Likert scale including five qualitative answers ranging from "Extremely unlikely/Never" to "Almost certain". In order to standardize the answers, a numerical reference was provided for each qualitative term.

For modelling purpose, each qualitative term was also translated into a probability distribution using the uniform distribution: $Uniform\,(Min; Max)$ where Min and Max are the minimum and maximum values of each considered range (Table 1). Again, respondents were asked to provide the level of confidence they had in their own estimations (w_i).

The probabilities of internal surface being contaminated by faeces and bile ($P_{IF}\ and\ P_{IB}$) were computed simulating 10.000 values extracted from the eleven uniform distributions. At each iteration, the probability of P_{IF} and P_{IB} being sampled from each expert's distribution was proportional to the corresponding level of confidence:

$$P_{IF} \sim Discrete\ (P_{IF_1}, P_{IF_2} \dots P_{IF11};\ w_1, w_2 \dots\ w_{11}) \qquad \text{Eq.6}$$

Where: P_{IFi} are the individual probabilities sampled in each iteration from the uniform distributions corresponding to the qualitative experts' believe of faecal contamination dropping on the internal surface of the carcass. The same procedure was applied for P_{IB} and for the probabilities of the external surface being contaminated by faeces and bile ($P_{EF}\ and\ \ P_{EB}$)

In order to provide practical probability distributions to be used in QMRAs, the simulated values for P_{IF}, P_{IB}, P_{EF} and P_{EB} where fitted to continuous probability distributions using the same MLE fitting method for a beta distribution described in section 2.2 (Equation 4).

Table 1 Numerical meaning of the qualitative frequency terms used in the questionnaire and the assumed distributions for modelling purposes

Qualitative estimation	Probability	Distribution
Almost certain	≥ 95%	$Uniform\ (0.95; 1)$
Very probable	80% - 95%	$Uniform\ (0.80; 0.95)$
Probable	50%-80%	$Uniform\ (0.50; 0.80)$
Fairly good possibility	30%-50%	$Uniform\ (0.30; 0.50)$
Some possibility	10%-30%	$Uniform\ (0.10; 0.30)$
Unlikely	5%-10%	$Uniform\ (0.05; 0.10)$
Very unlikely	>0%-5%	$Uniform\ (0; 0.05)$
Extremely unlikely/Never	0%	0

2.5. Partial condemnation of the carcass

When the contamination resulting from a rupture event is extended and/or involves internal cavities to a degree that corrective actions (i.e. trimming) cannot -or are not convenient- to be undertaken, the contaminated portion of the carcass may be removed as a whole (partial condemnation of the carcass) rather than trimmed (i.e. manual removal of the visible contamination). From a modeling perspective, accounting for a dose-dependent probability of partial condemnation of the carcass is critical because during simulation, this relationship has a direct impact on the number of iterations generating a carcass with visible contamination that is either detained for trimming or partially condemned.

The main practical difference is that because of the detection limit of the human eye, it is reasonable to assume the presence of not-visible contamination on the area surrounding the trimmed portions. In the case where the whole portion is removed, the presence/amount of undetected contamination in the area is not relevant anymore .

Therefore, the objective of the fourth question was to estimate what is the probability of a portion of a carcass being partially condemned when contaminated by different levels of visible faeces/bile.

When trying to describe phenomena via expert opinion, it is critical to consider the context in which the professional experience that is necessary to provide the answer lies. From preliminary consultation with veterinarians working on the slaughter line, it was clear that the perception of "degree of contamination" on the carcass, for this particular question, could not be presented as an "amount". In fact, what the veterinarians notice in their daily activity is the extent of the contamination, not the amount.

Following this consideration and in order to present the questions in a way that is suitable given the experience of the experts with this issue, the measure unit for this question was the surface area rather than amount. Therefore, this conditional relationship was estimated asking the interviewees how likely it is for a carcass to be partially condemned if the overall level of contamination dropping on the internal and external surface contaminates up to about 100 cm^2, 100-200 cm^2, 200-500 cm^2, 500-800 cm^2 or > 800 cm^2. Following preliminary consultations, it was considered reasonable to assume that for contaminations extended less than 50cm^2, carcasses are never partially condemned.

Similar to the first question, respondents could choose from the qualitative answers reported in Table 1**.** For this question, Interviewees were also instructed to consider the exact location of the visible contamination as unknown; therefore, their final estimation is implicitly representative of the likelihoods of partial condemnation *given* the specified extent of visible contamination and *given* the portion of the carcass that is more likely to be contaminated according to each expert's opinion.

For each combination: external/internal surface and extent of contamination, the probability of partial condemnation the carcass (P_C) was computed as follows:

First, for each ith expert, each extent of contaminated area was assigned to the corresponding uniform distribution associated to the qualitative term describing the conditional probability of partial condemnation (Table 1); this way, for each contaminated surface area (*Sa*), there is a paired $P_{C_i}^{Sa}$ by each expert. As each individual $P_{C_i}^{Sa}$ is associated with the expert's level of confidence (w_i), the overall area-dependent P_C^{Sa} is obtained from the discrete distribution:

$$P_C^{Sa} \sim Discrete\,(P_{C_1}^{Sa}, P_{C_2}^{Sa} \dots P_{C_{11}}^{Sa}; w_1, w_2 \dots w_{11}) \qquad \text{Eq.6}$$

Where: $P_{C_i}^{Sa}$ are the individual area-conditional probabilities or partial condemnation and w_i the associated expert's confidence score.

This procedure was used to estimate both the probabilities of partial condemnation conditional to the extent of the contaminated areas on the external $(P_{C_E}^{Sa})$ and internal surface $(P_{C_I}^{Sa})$ of the carcass.

Having used the surface rather than the amount as measure unit, these estimates cannot be directly converted to model the *dose*-conditional probability of partial condemnation.

However, in order to evaluate the potential meaning of those estimates in relation to the level of contamination, it was considered that 1g of faeces or bile can reasonably contaminate from 1 to 10 cm^2 (Author's best guess) depending on the density and the degree of squashing/spreading that have occurred.

As both of these parameters are not predictable, three different scenarios with 1g of faeces and bile approximated to 1 and 5 and 10cm^2 of contaminated surface were assessed.

In all the scenarios, $P_{C_E}^{Sa}$ and $P_{C_I}^{Sa}$ are reported for the fixed levels of contamination corresponding to the 5th, 25th 50th 75th and 95th percentiles of the cumulative distributions describing F_L and B_L. Probabilities were obtained performing 10.000 iterations and reporting the proportion of partially condemned carcasses for each combination dose-scenario.

2.6. Software

The software @Risk (version 7.0.1 for Excel, Palisade Corporation, Newfield, NY) was used to perform simulations and obtain the model outputs; R 3.4.1 was used to implement cluster analysis, produce charts and develop the web App.

3. RESULTS

3.1. Expected occurrence of ruptured guts and gallbladder

Results of the simulated and fitted occurrence of ruptured guts and gallbladders in each of the scenarios are reported in figure 1 and 2 respectively. The parameters of the fitted beta distributions together with the absolute frequency of the simulated values within each cluster that can be used to implement and reproduce the multinomial process are reported in table 2.

Table 2 Fitted beta ditributions and absolute frequency of the number of simulated values falling within each identified cluster. Values are reported for each considered scenario: High Throughput Slaughterhouse (HTS), High Throughput Slaughterhouse with robotic belly opener (HTS-r) Middle Throughput Slaughterhouse (MTS), Low Throughput Slaughterhouse (LTS) for the the events involving the gut and the gallbladder.

Scenario	N# clusters	Fitted distribution	Frequency	Median (Cluster)	5th perc (Cluster)	95th perc (Cluster)	Median (Scenario)	5th perc (Scenario)	95th perc (scenario)
HTS	Cluster 1	$Beta\ (72.25; 1169.77)$	1111	0.057	0.047	0.069	0.018	0.0014	0.058
	Cluster 2	$Beta\ (1.98; 565.92)$	3616	0.002	0.0006	0.008			
	Cluster 3	$Beta\ (27.49; 1382.32)$	5273	0.019	0.013	0.026			
HTS-r	Cluster 1	$Beta\ (112.29; 5064.31)$	3083	0.021	0.018	0.025	0.0081	0.0009	0.023
	Cluster 2	$Beta\ (3.13; 1460.78)$	3530	0.002	0.0006	0.004			
	Cluster 3	$Beta\ (18.99; 2194.41)$	3387	0.008	0.005	0.012			
MTS	Cluster 1	$Beta\ (111.21; 1617.53)$	1865	0.064	0.054	0.074	0.013	0.002	0.067
	Cluster 2	$Beta\ (16.08: 847.96)$	2827	0.018	0.011	0.026			
	Cluster 3	$Beta\ (60.93: 1293.80)$	1401	0.044	0.036	0.054			
	Cluster 4	$Beta\ (3.67: 735.46)$	3907	0.004	0.001	0.010			
LTS	//	$Beta\ (0.89; 17.54)$	10000	//	//		0.033	0.002	0.147
GALLBLADDER									
HTS	Cluster 1	$Beta\ (22.69; 992.32)$	1369	0.022	0.015	0.030	0.0075	0.0017	0.079
	Cluster 2	$Beta\ (2.53; 377.13)$	7560	0.005	0.001	0.014			
	Cluster 3	$Beta\ (372.77; 4329.47)$	1071	0.079	0.072	0.085			
HTS-r	Cluster 1	$Beta\ (56.43; 5006.16)$	2520	0.011	0.008	0.013	0.0047	0.0005	0.012
	Cluster 2	$Beta\ (18.65; 3523.84)$	3823	0.005	0.003	0.007			
	Cluster 3	$Beta\ (2.10; 1432.76)$	3657	0.001	0.0002	0.003			
MTS	Cluster 1	$Beta\ (25.31; 1102.10)$	2706	0.022	0.010	0.030	0.017	0.0018	0.045
	Cluster 2	$Beta\ (62.30; 1469.59)$	2625	0.040	0.032	0.049			
	Cluster 3	$Beta\ (2.79; 532.68)$	4669	0.004	0.001	0.011			
LTS	//	$Beta\ (0.96; 26.21)$	10000	//	//	//	0.025	0.001	0.105

Results indicate that P_{GUT} ranged from 0 to a maximum of 0.08 in both HTS and MTS, from 0 to 0.03 in HTS-r and from 0 to 0.30 in LTS respectively. Similarly, P_{GALL} ranged from 0 to 0.09, in HTS, from 0 to 0.017 in HTS-r, from 0 to 0.06 in MTS and from 0 to 0.20 LTS respectively.

The cluster analysis on the simulated results identified three clusters for HTS and HTS-r, four clusters for MTS when considering the expected occurrence of ruptured guts and three clusters in HTS, HTS-s and MTS when considering the expected occurrence of ruptured gallbladders. When considering the expected occurrence of ruptured guts and gallbladder in LTS, the shape of the distributions where unimodal; no further actions were taken.

An illustrative example as to implement the multinomial process by means of nested beta distributions is available in the electronic supplementary material

3.2. **Amount of faeces/bile leaking onto carcass**

The cumulative distributions describing the variability in F_L and B_L are outlined in figure 3 and 4 respectively.

The distribution for F_L indicates that following accidental rupture of the gut, the variability in the amount of faeces leaking on the carcass can be described by a cumulative distribution showing values of 2.9, 11.5, 24.5, 49 and 131.4g at 5th, 25th, 50th, 75th and 95th percentile respectively. Similarly, the cumulative distribution for B_L describing the variability in the amount of bile leaking on the carcass following accidental rupture of the gallbladder showed 0.7, 5.4, 19.9, 41.7 and 88.1g at the same percentiles. The complete datasets to reproduce the cumulative distributions are available in the supplementary material.

3.3. probability of internal surface of the carcass being contaminated

The fitted beta distributions describing the probability of faecal and bile contamination dropping on internal (P_{FI} and P_{BI}) and external (P_{FE} and P_{BE}) surface of the carcass following accidental rupture of the gut and gallbladder are:

$$P_{FI} \sim Beta(1.48; 2.56) \quad \text{Eq.7}$$

$$P_{\boldsymbol{BI}} \sim Beta(0.72; 2.26) \quad \text{Eq.8}$$

$$P_{\boldsymbol{FE}} \sim Beta(2.93; 5.30) \quad \text{Eq.9}$$

$$P_{\boldsymbol{BE}} \sim Beta(1.63; 4.17) \quad \text{Eq.10}$$

The median value for P_{FI} and P_{FE} was 0.34 for both the distributions but the values at the 5th, 25th and 95th percentile (0.05, 0.19 and 0.75 for P_{FI} and 0.11, 0.23 and 0.63 for P_{FE}) indicate a slightly more right-skewed distribution for P_{FE}. Both the distributions describing P_{BI} and P_{BE} were strongly right-skewed with a median value of 0.19 and 0.25 for P_{BI} and P_{BE} respectively and values at the 5th, 25th and 95th percentile of: 0.007, 0.06 and 0.68 for P_{BI} and 0.05, 0.14 and 0.60 for P_{BE}.

3.4. Partial condemnation of the carcass

The probabilities of partial condemnation of the carcass because of the level of faecal and bile contamination dropping on the external and internal surfaces are reported in Table 3.

Table 3 dose-dependent conditional probabilities of partial condemnation. The probabilities of a carcass being partially condemned as a result of faecal or bile contamination on the internal and external surface are reported for the three scenarios (SC1, SC2, SC3) in which 1 gram of faeces and bile was approximated to 1, 5 and 10cm^2 respectively. As it is assumed that carcasses are never partially condemned when less than 50cm^2 are contaminated, the probability in those cases is always zero. Some values have the same probability because they fall within the same range for which the probabilities were calculated.

Faecal contamination				Bile contamination			
	dose (g)	pe	pi		dose (g)	pe	pi
SC1 1g=1cm^2	2.6	0.00	0.00	SC1 1g=1cm^2	0.7	0.00	0.00
	11.4	0.00	0.00		5.5	0.00	0.00
	25.6	0.00	0.00		24.1	0.00	0.00
	60.4	0.04	0.07		44.5	0.00	0.00
	≥232.9	0.47	0.65		≥87.8	0.04	0.06
SC2 1g=5cm^2	2.6	0.00	0.00	SC2 1g=5cm^2	0.7	0.00	0.00
	11.4	0.04	0.06		5.5	0.00	0.00
	25.6	0.16	0.25		24.1	0.16	0.25
	60.4	0.47	0.65		44.5	0.47	0.65
	≥232.9	0.65	0.79		≥87.8	0.47	0.65
SC3 1g=10cm^2	2.6	0.00	0.00	SC3 1g=10cm^2	0.7	0.00	0.00
	11.4	0.16	0.25		5.5	0.04	0.14
	25.6	0.47	0.65		24.1	0.47	0.65
	60.4	0.65	0.79		44.5	0.47	0.65
	≥232.9	0.65	0.79		≥87.8	0.65	0.79

Highly contaminated carcasses are more likely to be partially condemned; however, the actual probability is never above 0.79, not even in the worst case scenario in which 1g contaminates 10cm^2.

Results in Table 3 also highlight how critical the extent of surface that can be contaminated by 1g is. Although there is not a convincing way to obtain accurate predictions of the extent 1 gram can contaminate, identification of a reasonable practical limit (i.e. 10cm^2) allows approximation of the likely range describing the probability of partial condemnation for specific values. For example, according to results in Table 3, considering the extent of surface that is possibly contaminated by 1g (i.e. from 1 to 10cm^2), for 25g of faeces dropping on the external surface, that carcass has a probability ranging from 0 to 0.65 of being partially condemned.

An illustrative example on how to obtain $P_{C_E}^{Sa}$ and $P_{C_I}^{Sa}$ as a function of the amount/extent of contamination is available in the electronic supplementary material.

4. DISCUSSION

In this study, we use the opinion of experts to characterize, as probability distributions, the likelihood of a series of key events occurring during the evisceration procedures in pig slaughterhouses.

When evaluating the occurrence of ruptured guts and gallbladders in abattoirs of different capacity, it became evident that multimodal distributions would be appropriate to reflect the combined effects of the different perceptions and degree of certainty amongst the interviewees. After weighting the variability distributions (Eq.1) by the scores (Eq.2), both the distributions of S_{GUT} and S_{GALL} were multimodal in HTS, HTS(r) and MTS (results not shown). Distributions describing the perceived "weighted" variability were themselves uncertain, and this was reflected by the multimodality of the simulated uncertainty distributions P_{GUT} and P_{GALL} (Eq.3). The LTS resulted in the settings where ruptures of gut and gallbladder are in general more likely to occur, on the other hand, in HTS and MTS those probabilities were of the same magnitude; always less than 0.09 but with marked uncertainty in the modes due to the underlying multinomial process. As previously discussed (section 2.2.), the true –unknown- variability in S_{GUT} and S_{GALL} and the associated uncertainty distributions P_{GUT} and P_{GALL}, are likely to be unimodal in each abattoir (i.e. every abattoir has its own variability and uncertainty distributions capturing the characteristics of that setting). In this context, the results in figure 1 and 2 can be interpreted as the ranges within the true uncertainty distributions of P_{GUT} and P_{GALL} are more likely to fall in each setting (i.e. HTS, HTS(r), MTS and LTS). As the uncertainty distributions were simulated, the presence of clusters is representative of the different perceptions, professional experiences and confidence the interviewees had around the variability in the number of events (i.e. ruptures) in a given setting.

Noticeably, when considering both the rupture of the gut and the gallbladder, the clusters become increasingly less evident when moving from the high to the small throughput settings. From this evidence, it can be concluded that the interviewees tended to expect the same variability when asked to consider small throughput abattoirs.

Although the boundaries of the multimodal distributions are all narrow (the wider ranges from 0 to 0.1) from a modelling perspective it might be useful to evaluate the impact of the uncertainty distributions describing P_{GUT} and P_{GALL} on the final outcome (e.g. number of ruptures in a given population of slaughtered animals). One way of doing this would be by separating the variability from the uncertainty by means of second order modelling [24]. This can be done for example by comparing *n* cumulative distributions of *n* $Binomial{\sim}(N; p_n)$ where N is the total population (fixed) and p_n are either the n^{th} Gamma distributions within each clusters (to separate the uncertainty in the variability), n^{th} values corresponding to random percentiles of the Gamma distributions within each cluster (to separate the overall uncertainty) or combinations of both.

The cumulative distributions describing the levels of faecal and bile contamination leaking on a carcass following accidental rupture of the gut or gallbladder ranged from a few grams to a maximum of 297g and 123g for the faeces and bile respectively. The shape of the distributions are the result of each individual estimation combined with the paired level of confidence.

The level of residual bile in gallbladders of slaughtered animals is difficult to handle considering the liver physiologically produces this fluid continuously [25]. On the other hand, that of faeces can be modulated by preventing animals from eating before arriving at the abattoir. The benefits of withdrawing feed from pigs prior to slaughter are well-known [26, 27], in fact, animals with full stomachs at slaughter are more often subjected to intestinal rupture or lacerations. However, it should be remarked that the cumulative distributions for F_L and B_L were not intended to describe the level of residual faeces and bile residing in the guts and gallbladders of all the animals entering the slaughterhouse but the amounts extruded following a rupture event only. The amount of faeces contaminating the carcasses was estimated to be greater than ~60 and ~100g only above the 81th and 91th percentile suggesting that the overall amount of visible faecal contamination on the carcass is likely to be low most of the times.

The evisceration procedure is a highly standardized sequence of accurate cuts and movements performed by trained personnel [28]; the extent to which the ruptures are actually induced/facilitated by the presence of residual faeces is unknown and not directly considered in this study.

The distributions describing the uncertainty in the occurrence of faecal contamination dropping on the internal and external surface of the carcass had comparable median values and similar shapes; the two distributions describing the uncertainty in the occurrence of bile contamination were also similar between them. Marked differences in the likelihoods of internal and external contamination were not perceived when considering faecal and bile contamination. This is probably a reflection of the strong random nature of this event.

The dose dependent probability of partial condemnation generated results indicating that, as expected, the chance for a carcass of being partially condemned are strongly influenced by the level/

extent of the contamination and for the same amount, carcasses are less likely to be partially condemned if the contamination is on the external surface.

However, the probability of partial condemnation was relatively low in scenarios with a considerable extent of contamination (i.e. according to SC3 in table 3, for an extent of contamination of ~250cm^2 on the external surface there is a probability of 0.47 for the carcass of being partially condemned). This evidence suggest that the relationships between the level of contamination and the probability of partial condemnation cannot be simplified to a linear/non-linear combination of these two parameters only and that there are other factors that explain the probability of partial condemnation.

Following consultation with OVs/MHIs, it seems that unless the damage that provoked the high level of contamination is particularly severe or the management gives specific or exceptional instructions; in general, it is still economically convenient to keep the carcass as a whole by trimming the visibly contaminated areas. In this regard, pig carcasses are easily manageable/workable and trained personnel could quickly remove extended areas even in high-speed processing lines. The fact that trimmed carcasses are kept in the food chain is particularly relevant when assuming that in the presence of visible contamination (removed with trimming), the simultaneous presence of not-detectable contamination cannot be excluded [17]. The amount of contamination that is visible and the relationship between visible and not-visible contamination is critical in terms of public health implications as it explains the level of –undetected- contamination on apparently clean carcasses. However, the parameters that are likely to govern the relationship between visible and non-visible contamination (e.g. geometrical spread of contamination on the carcass, detection limit of human eye) could not be addressed by means of expert opinion in this study.

The objective of this work was to provide practical information in support of QMRAs modelling the fate of pathogens in pig slaughterhouses. The study focused on the events occurring during the belly opening/removal of gastro intestinal tract at evisceration, which is considered one of the most critical steps with respect to carcass contamination [19, 21, 29] but has several quantitative data gaps. During processing of pigs carcasses, other events could lead to faecal contamination, particularly the leakage from the rectum/anus during dehairing or polishing where a small amount of faeces can extrude from the interior of the carcass (if the anus is not plugged) following the vigorous action of the machines [17]. However, this aspect was not taken into account in the present study; from empirical observation, the amount of faeces extruded from the anus at those stages appeared to be characterised by very low variability. For the purpose of modelling we believe the value provided by Swart et al.[17] by means of expert opinion (i.e. 10g) is already a good approximation.

The distributions we present in this study reflect the experience (weighted by the confidence) of eleven veterinarians working in pig slaughterhouses, therefore, we believe they represent a reliable approximation for parameters that would be impractical/challenging to be measured or accessed. In fact, acquisition of an appropriate number of empirical data to parameterise the distributions/relationships considered in this manuscript would require access to sensitive data from multiple slaughterhouses and would represent an excessive hurdle of the working routine due to the presence of external people in specific areas (typically characterised by very limited space). Should the much needed experimental data be available in the future, these could be used to validate, compare or integrate our findings.The outputs made available from this study can be directly included in QMRA models in which the events occurring during the belly opening/removal of gastro-intestinal tract stage need to be included. In fact, for carcasses processed in each setting, each carcass would have a probability of having the gut and/or the gallbladder ruptured, a consequent probability of a given amount of faecal/bile contamination dropping on the internal/external surface of the carcass and a dose-conditional probability of partial condemnation. Probabilistic modelling of those events inevitably requires assumed independency amongst sequential carcasses for this stage (i.e. results of the previous does not have an impact on that of the following).

Main assumption and limitations

A limitation of this study was the subjective selection of the number of levels in the Likert scales and uniform distributions chosen to translate the qualitative estimations of the experts' answers (Table 1). Recognizing that this selection has an influence on the shape of the output distributions, it should be considered that expert/informed opinion should be elicited in such a way that interviewees can answer easily and without ambiguity. From a modelling perspective, it would have been ideal to include additional levels in the qualitative scales; but considering the nature of the questions and the parameters under investigation, this would have created difficulties in experts' perception with the risk of less comparable answers. Therefore, the choice of the levels in the Likert scales was a compromise between the need to put experts in a situation in which they could easily answer the questions and the desired accuracy of the outputs. This was also the main reason for having approximated "Extremely unlikely" to 0 with the category "Extremely unlikely/Never". Inclusion of an additional lower level (e.g. "Extremely unlikely"= 0% - 1%), could have created difficulties in discriminating between "extremely unlikely" and "very unlikely", and, in our previous experience in face-to-face interviews, it would have encouraged experts to consider eventualities that are so remote that in practice would never happen. We believe this choice helped interviewees to discriminate more genuinely rare events from those that might happen purely by chance only under extreme circumstances.

This work is based on opinions gathered from 11 experts. While in general, the more experts are interviewed the better, in our case, experts were pre-selected not only on the basis of their experience but also personality. Considering the type of questions and the length of the questionnaire, we aimed interviewing experts who we knew would have dedicated appropriate time and efforts to our study. From a technical perspective, Meyer and Booker (2001) noted that "*less than five experts reduces the chances of providing adequate diversity or information to make inferences*" [30] and from his experience, Aspinal (2010) concluded that "*between 8 to 15 experts is a viable number*" [31]. For these reasons, we believe the number of experts interviewed here is reasonable for the scope of the study.

As a general limitation that should be recognized when dealing with expert elicitations is that although interviewees were selected with the intent of gathering estimations from individuals being able to provide knowledgeable opinions; the outputs presented in this study are entirely dependent upon their expertise which was weighted only considering their self-assigned level of confidence. Considering that OVs and VMHIs have similar training, background and experience on hygiene standards and the aspects under investigation, opinions from OVs and VMHIs were treated the same. In fact, it should be noted that typically OVs served as VMHIs before being OVs and the main difference between the two lies in the additional responsibilities of the former in terms of animal and carcass' inspection.

Another limitation is due to the challenges of exploring, by means of expert opinion, dependency amongst certain events. As previously mentioned, the probability of ruptures could in reality be related to the amount of residual intestinal content [26], which might in turn be strictly related to the actual amount of visible faecal contamination leaking on the carcass following accidental rupture. Therefore, with the exception of the probabilities of partial condemnation, where the conditional dose-probability dependencies were modelled, all the other events discussed in this study were modelled as independent from each other.

Finally, it should be noted that the microbiological contamination of the carcasses can be due to hygiene errors as well as environmental and cross-contaminations [19, 32]. However, as this work was mainly intended to provide estimations about "how often" some key events are likely to happen using as main study unit the visible faecal/bile contamination (the only contamination the interviewees could actually see) actual changes in microbial loads could not be estimated by means of expert opinion.

Ethical approval

Ethical approval for this survey has been given by the Social Science Research Ethical Review Board at the Royal Veterinary College. Reference number: URN SR2017-1339.

Acknowledgments

We are extremely grateful to Diego Sprekelsen from HallMark Ltd for his invaluable help in facilitating contact with veterinarians and to the Official Veterinarians and Meat Hygiene Inspectors who kindly answered the questionnaires. The study was part of project funded by the Agriculture and Horticulture Development Board [ref. 5140004020]

REFERENCES

1. EUROSTAT. *Meat production statistics*. 2014 February 2018]; Available from: http://ec.europa.eu/eurostat/statistics-explained/index.php/Meat_production_statistics.
2. FAO. *World agriculture: towards 2015/2030*. 2003 February 2018]; Available from: http://www.fao.org/docrep/005/y4252e/y4252e05b.htm.
3. EFSA, *Scientific Opinion on the public health hazards to be covered by inspection of meat (swine).* EFSA J, 2011. **9**(10): p. 2351.
4. EFSAa, *Scientific Opinion on a Quantitative Microbiological Risk Assessment of Salmonella in slaughter and breeder pigs. .* EFSA J, 2010(4): p. 1547-1637.
5. EFSA, *Public health risks associated with hepatitis E virus (HEV) as a food-borne pathogen.* EFSA J, 2017. **15**(7): p. 4886-4975.
6. Crotta, M., et al., *Quantitative risk assessment of hepatitis E virus: modelling the occurrence of viraemic pigs and the presence of the virus in organs of food safety interest.* Microbial Risk Anal, 2017. **9**: p. 64-71.
7. WHO, *Food safety risk analysis. A guide for national food safety authorities*. FAO food and nutrition paper. Vol. 87. 2006, Rome.
8. WHO/FAO, *Codex Alimentarius Commission: Procedural Manual*. 19 ed. Joint FAO/WHO Food Standards Programme. 2010, Rome.
9. EFSAb, *Scientific Opinion on Quantification of the risk posed by broiler meat to human campylobacteriosis in the EU.* EFSA J, 2010. **8**(1): p. 1437.
10. FAO/WHO, *Risk assessment of Listeria monocytogenes in ready-to-eat foods*, in *Microbiological Risk Assessment*. 2004.
11. Costa, E., et al., *A Stochastic Model to Assess the Effect of Meat Inspection Practices on the Contamination of the Pig Carcasses.* Risk Anal, 2017. **37**(10): p. 1849-1864.
12. Bollerslev, A.M., et al., *A risk-based approach for evaluation of hygiene performance at pig slaughter.* Food Control, 2017. **75**: p. 116-125.
13. Bollerslev, A.M., et al., *A risk modelling approach for setting microbiological limits using enterococci as indicator for growth potential of Salmonella in pork.* Int J Food Microbiol, 2017. **240**: p. 102-107.
14. Crotta, M., M. Georgiev, and J. Guitian, *Quantitative risk assessment of Campylobacter in broiler chickens–Assessing interventions to reduce the level of contamination at the end of the rearing period.* Food Control, 2017. **75**: p. 29-39.
15. Crotta, M., et al., *Consumers' behavior in quantitative microbial risk assessment for pathogens in raw milk: Incorporation of the likelihood of consumption as a function of storage time and temperature.* J Dairy Sci, 2016. **99**(2): p. 1029-1038.
16. Crotta, M., et al., *Knowledge gaps in host-parasite interaction preclude accurate assessment of meat-borne exposure to Toxoplasma gondii.* International journal of food microbiology, 2017. **261**: p. 95-101.

17. Swart, A., et al., *Modeling of Salmonella Contamination in the Pig Slaughterhouse.* Risk Anal, 2016. **36**(3): p. 498-515.
18. Delhalle, L., et al., *Risk factors for Salmonella and hygiene indicators in the 10 largest Belgian pig slaughterhouses.* J Food Prot, 2008. **71**(7): p. 1320-9.
19. Pearce, R.A., et al., *Studies to determine the critical control points in pork slaughter hazard analysis and critical control point systems.* Int J Food Microbiol, 2004. **90**(3): p. 331-339.
20. Berends, B.R., et al., *Identification and quantification of risk factors regarding Salmonella spp. on pork carcasses.* Int J Food Microbiol, 1997. **36**(2-3): p. 199-206.
21. Alban, L. and K.D. Stark, *Where should the effort be put to reduce the Salmonella prevalence in the slaughtered swine carcass effectively?* Prev Vet Med, 2005. **68**(1): p. 63-79.
22. RStudio, I. *Shiny: Web Application Framework for R*. 2013 May 2018]; Available from: http://shiny.rstudio.com.
23. Pouillot, R. and M.L. Delignette-Muller, *Evaluating variability and uncertainty separately in microbial quantitative risk assessment using two R packages.* Int J Food Microbiol, 2010. **142**(3): p. 330-40.
24. Vose, D., *Risk analysis: a quantitative guide*. 2008, England: John Wiley & Sons.
25. Baggot, J.D., *The physiological basis of veterinary clinical pharmacology*. 2008, United Kingdom, : Blackwell Science.
26. Miller, M.F., et al., *Microbiology of pork carcasses from pigs with differing origins and feed withdrawal times.* J Food Prot, 1997. **60**(3): p. 242-245.
27. Bidner, B., F. McKeith, and E. Berg. *Feed Withdrawal Prior to Slaughter: Effects on Pork Quality and Safety*. Facts. National Pork Board. 1998; Available from: http://old.pork.org/filelibrary/factsheets/pigfactsheets/newfactsheets/12-03-03g.pdf, Accessed: July 2017.
28. FSA. *Slaughter licensing* 2017 July 2018]; Available from: https://www.food.gov.uk/business-guidance/slaughter-licensing.
29. Wheatley, P., E.S. Giotis, and A.I. McKevitt, *Effects of slaughtering operations on carcass contamination in an Irish pork production plant.* Irish veterinary journal, 2014. **67**(1): p. 1.
30. Meyer MA and Booker JM, *Eliciting and analyzing expert judgment: a practical guide.* 2001, Alexandria, Virginia, USA: American Statistical Association and the Society for Industrial and Applied Mathematics.
31. Aspinall, W., *A route to more tractable expert advice.* Nature, 2010. **463**(7279): p. 294.
32. Botteldoorn, N., et al., *Salmonella on pig carcasses: positive pigs and cross contamination in the slaughterhouse.* J Appl Microbiol, 2003. **95**(5): p. 891-903.

CAPTIONS

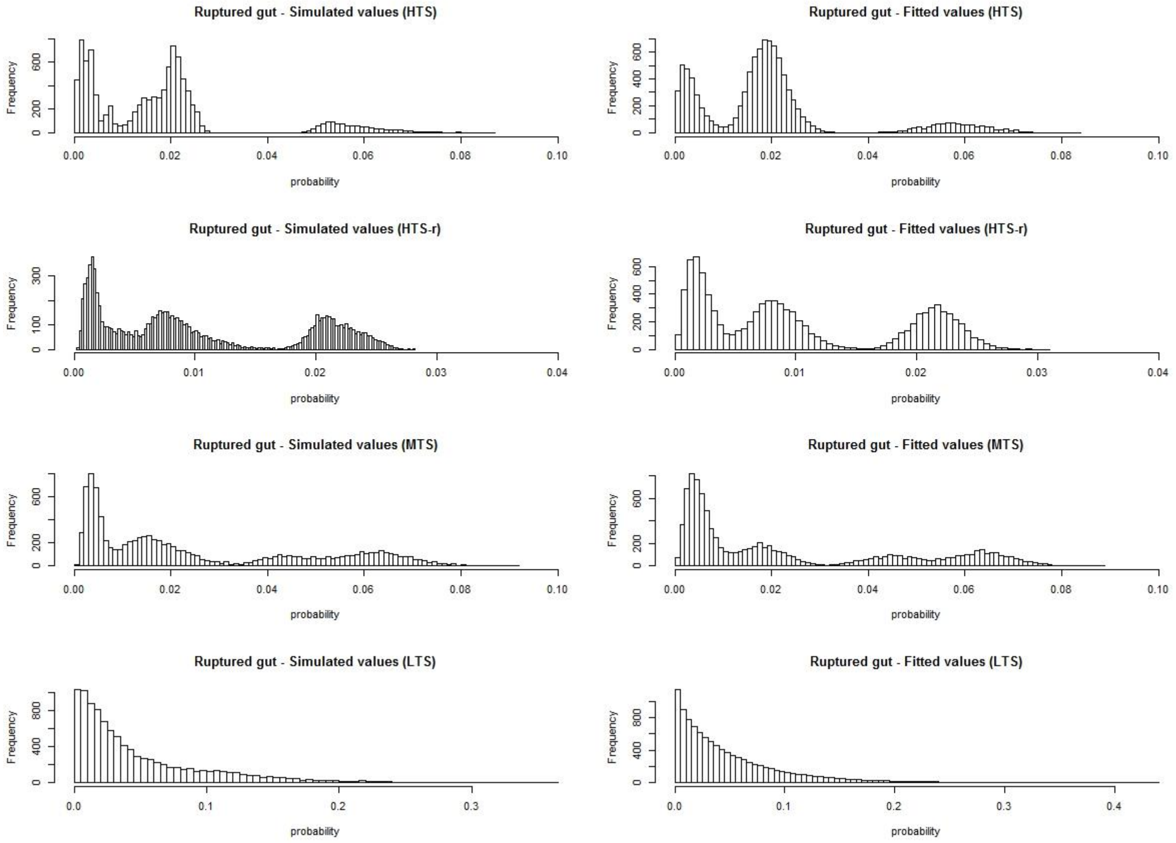

Figure 1. Paired comparisons between simulated and fitted values for the expected occurrence of ruptured guts during evisceration in High Throughput Slaughterhouse (HTS), High Throughput Slaughterhouse with robotic belly opener (HTS-r), Middle Throughput Slaughterhouse (MTS) and Low Throughput Slaughterhouse (LTS).

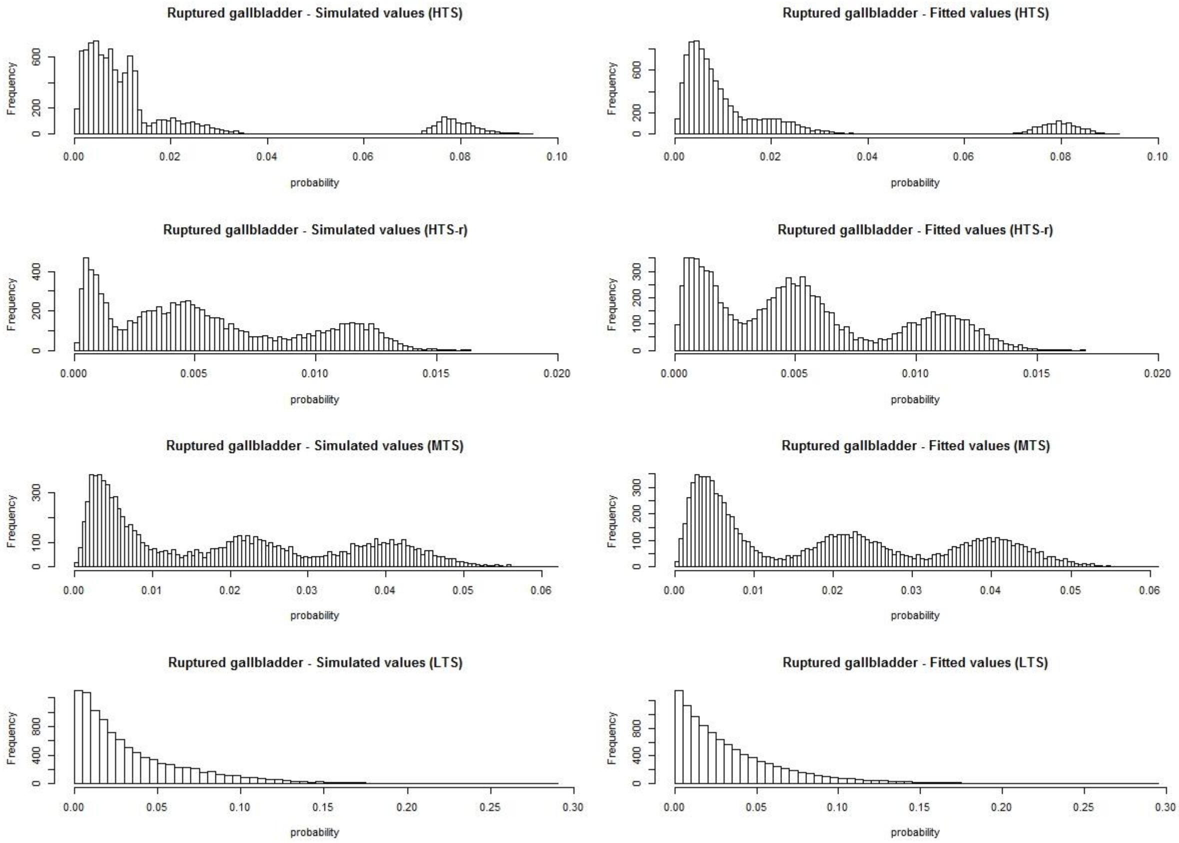

Figure 2. Paired comparisons between simulated and fitted values for the expected occurrence of ruptured gallbladder during evisceration in High Throughput Slaughterhouse (HTS), High Throughput Slaughterhouse with robotic belly opener (HTS-r), Middle Throughput Slaughterhouse (MTS) and Low Throughput Slaughterhouse (LTS).

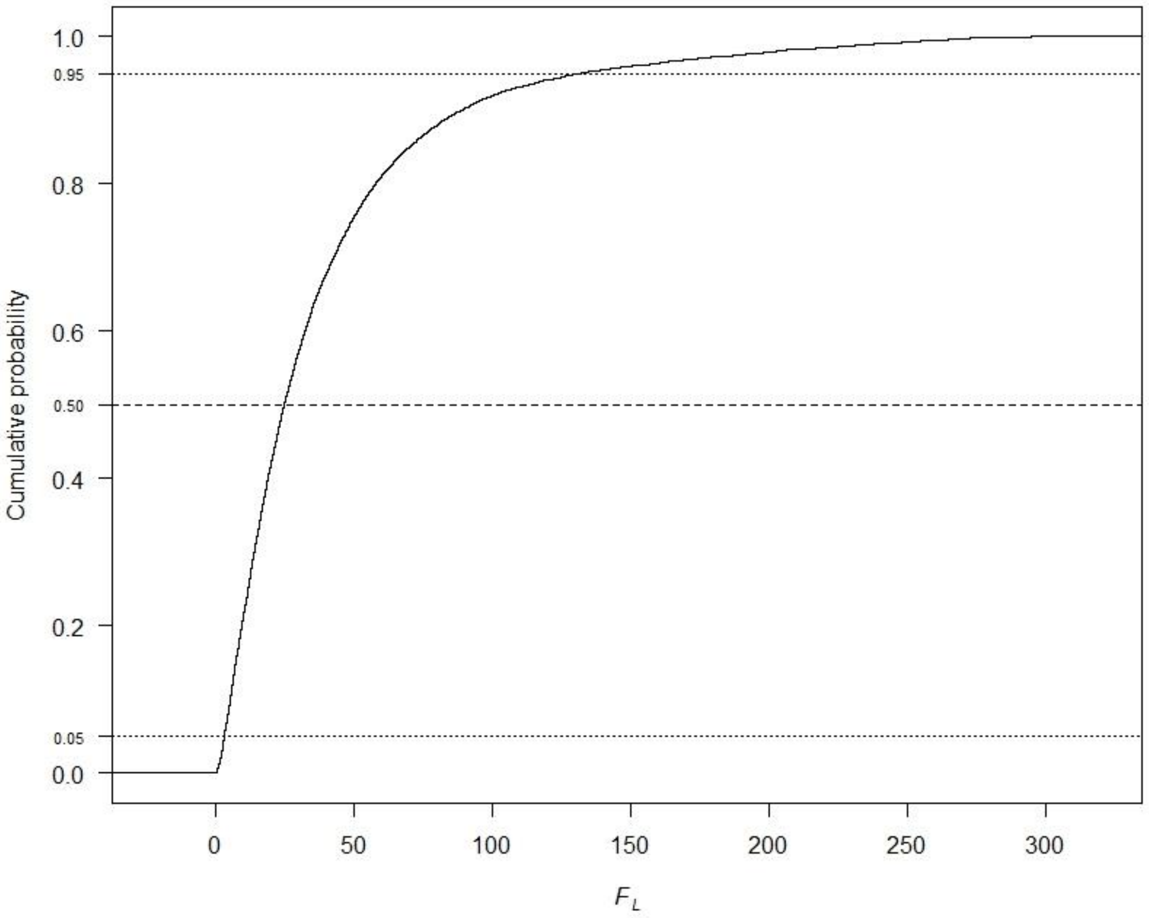

Figure 3. Cumulative distribution describing the variability in the amount of faeces (F_L) *leaking on a carcass following accidental rupture of the gut.*

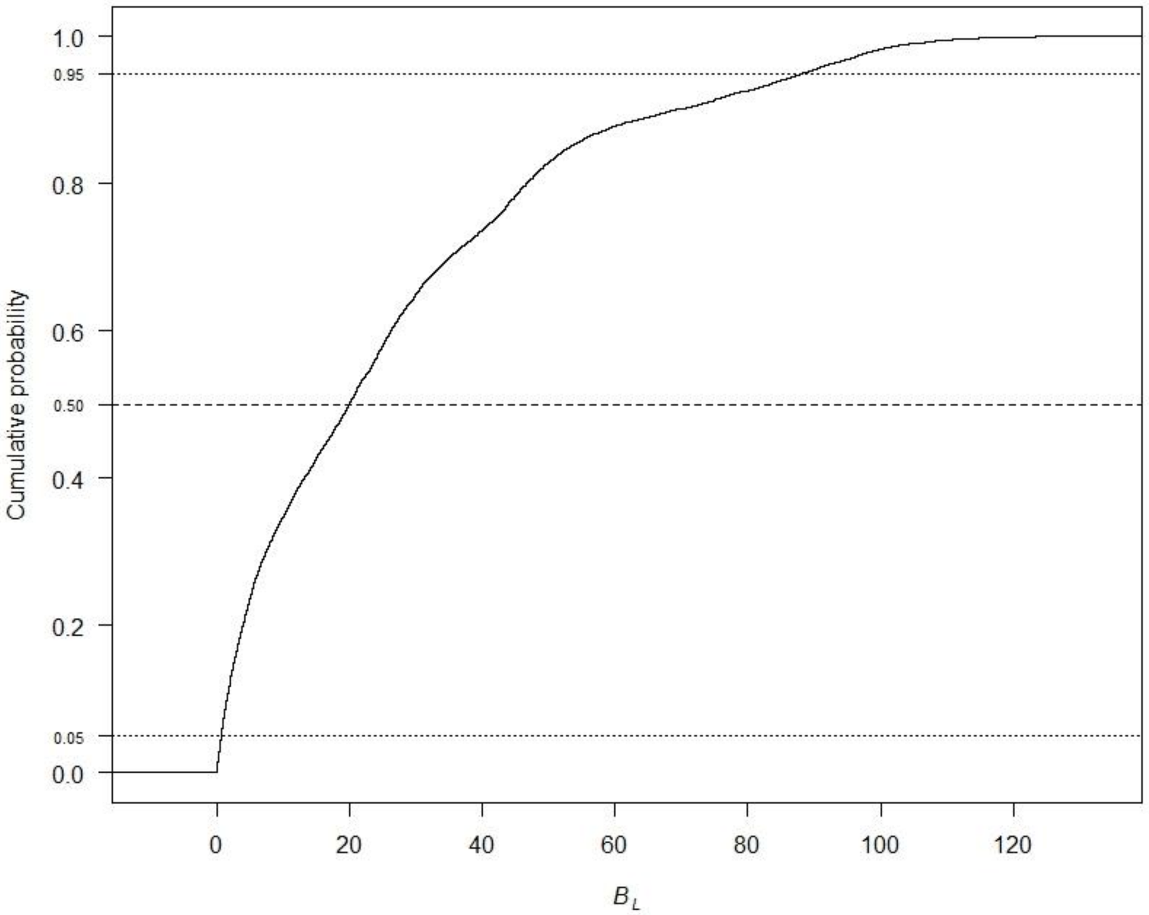

Figure 4. Cumulative distribution describing the variability in the amount of bile (B_L) *leaking on a carcass following accidental rupture of the gallbladder.*