Software for evaluating sampling strategies and error rates in the identification of mixed-meat products

I.VETHARANIAM and G. H. SHACKELL[1]

AgResearch Limited, Private Bag 3123, Hamilton, New Zealand

ABSTRACT

DNA verification of the origin of meat within traceability systems involves the comparison of DNA samples taken at a reference stage with samples taken further down the supply chain. While such verification is well established when whole cuts of meat are involved, this is not the case for mixed-meat products made from an unknown number of contributing animals. Spreadsheet based software has been developed to analyse one DNA verification scheme for determining whether a returned product did indeed originate from a particular manufacturing batch. The scheme involves the comparison of *k* individual DNA samples (comprising an unknown number of individuals) from the product, with a set of *s* DNA samples (comprising an unknown number of individuals) taken randomly before manufacture of the batch in question. The software gives the error rate for correctly matching the product to the batch, for different sampling rate combinations of the product and at the factory.

Key words: DNA verification; mixed meat; traceability; sampling.

INTRODUCTION

The safety, quality and integrity of meat are of increasing importance at all levels of the food chain as consumers demand knowledge of the origin, the pre-market treatment and the production methods of their food (Arana *et al*., 2002). This has been fuelled in part by a number of crises in the meat industry such as bovine spongiform encephalopathy (BSE; Smith *et. al.*, 2004), foot-and-mouth disease (FMD), labelling scandals, and *E. coli* outbreaks (Arana *et al*., 2002; Clemens, 2003; Cagney *et al*., 2004), as well as an increasing public awareness of animal welfare issues (Hobbs *et al.*, 2002) and negativity towards genetically modified food (Miraglia *et al*., 2004). As a result of these concerns, meat traceability schemes that document the origin and movement of animals and their disassembled carcasses along all or part of the food supply chain have been established. Traceability schemes have been adopted for benefits such as brand protection and/or enhancement in niche markets or during a food crisis (Arana *et al*., 2002, Cavani & Petracci, 2004), for quality control, as an aid to identifying and recalling contaminated products from the market (Caporale *et al*, 2001), or, in some regions such as Japan (Clemens, 2003, Ozawa, *et al.,* 2005) and the European Union (Cavani and Petracci, 2004), because of government legislation.

One weakness of most paper-based traceability schemes is that they require an implicit trust that the documentation is both correct and free from manipulation and in the absence of an independent (empirical) verification system to confirm (or refute) claims against producers, manufacturers, wholesalers or retailers (Arana *et al*., 2002) they cannot offer unequivocal guarantees. DNA technologies offer the potential to effectively identify meat from individual animals by providing the ability to augment and verify a traceability paper trail (Meghen *et al*., 1998; Cunningham and Meghen, 2001; Shackell *et al*., 2001; Tate, 2001), or to detect fraud (Arana *et al*., 2002; Vázquez *et al*., 2004). These technologies involve taking DNA from a meat sample at any point along the supply chain and comparing it with reference DNA samples taken prior to, or immediately after slaughter, to provide a conclusive means of identification. Such schemes are not so readily applicable to mixed meat products, which use ground meat from an unknown number of unidentified animals. The traceability of blended meat products is problematic since the origin of the source material is difficult to establish (McKean, 2001), and because of this ground beef and processed products are generally excluded from traceability legislation (e.g., *Japan*, Clemens, 2003; *European Union*, Regulation (EC) No 1760/2000). However, the traceability of ground meat products is arguably more important than that of whole meat cuts from a food safety aspect. Meat grinders can act as distributors of microbial contamination (Flores & Tamplin, 2002). Furthermore, ground meat is more susceptible to contamination by *E. coli* (Schroeder *et al*., 2003) and *Salmonella* (Stock & Stolle, 2001) than whole meal cuts and may pose a greater health risk than whole cuts. A technology that would allow a reliable verification of the origin of ground meat would extend the benefits of traceability to improving the quality and safety of mixed meat products.

In order to provide DNA-based traceability for a batch of ground product, it is necessary to be able to match individuals identified from within the product to individuals known to have contributed to the batch of origin of that product. As the DNA of each animal is unique, individuals can be identified as long as enough markers are used to eliminate chance matches (Arana *et al*., 2002). We have analysed some of the sampling issues surrounding the DNA identification of mixed ground meat products in the context of verifying (or refuting) that a product has been returned to the correct

[1] AgResearch Invermay, Private Bag 50034, Mosgiel, New Zealand

manufacturer. Spreadsheet-based software for mixed meat products ('Sampling Calculator') has been designed to be part of a DNA verification scheme. It is intended for use at the time of manufacture of mixed-meat products when DNA samples are being taken for reference in the event that some of the product is returned. The Sampling Calculator is underpinned by a quantitative analysis of cutting room data (Vetharaniam, Shackell & Upsdell, unpublished), and in this paper we describe the use of the software.

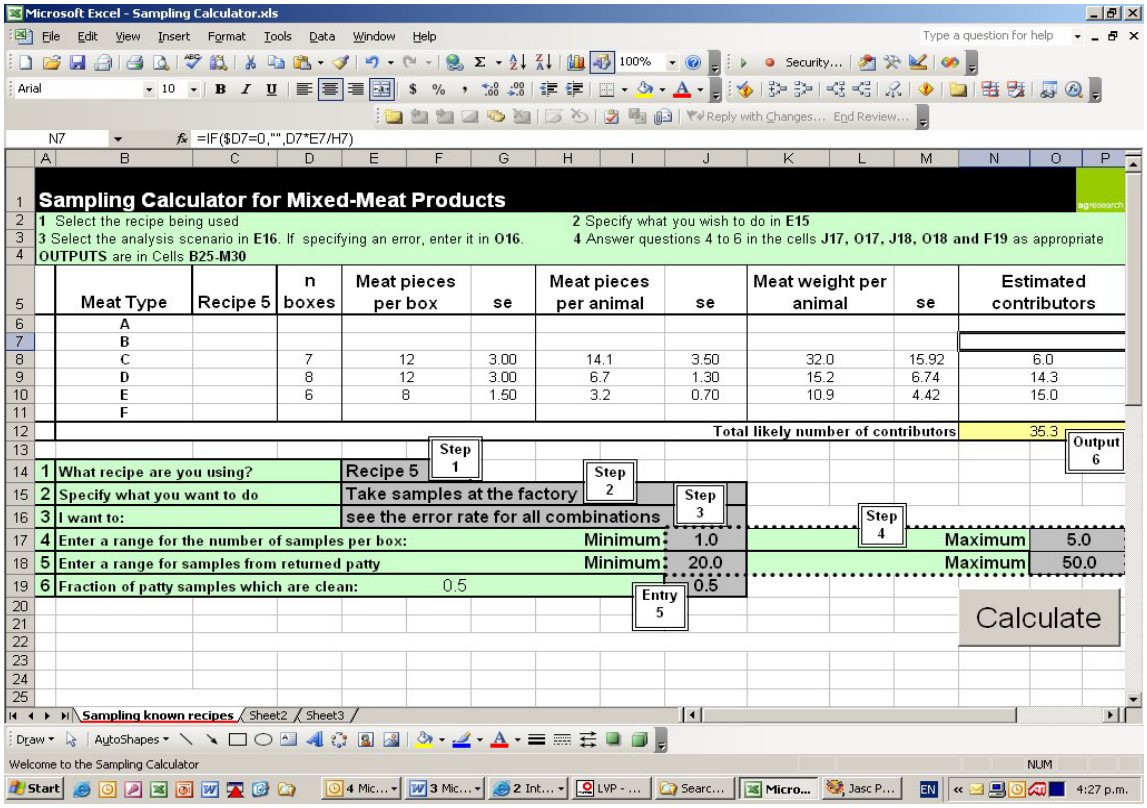
Description of the Sampling Calculator

It is assumed that the meat used in the manufacture of mixed-meat products is supplied in boxes containing particular types of trimmed meat (say Type A, Type B, … Type F). Each product made will have its own specific recipe which will specify the meat types to use, and the number of boxes of each. A specified number of DNA samples (say s) are taken from each box of meat to give a total of $Q = s \times M$ reference samples, where M is the total number of boxes of meat used. These Q samples are stored in case the product is returned and verification of its origin is required. In this event k samples will be taken from the returned product and sent with the Q samples for a comparative DNA analysis.

When a sufficiently strong match of genotypes is made between the reference samples and the products samples, the returned product is deemed to be from the batch in question. Otherwise, the manufacturer can disown responsibility for the returned product. As the number of contributors to each batch of mixed product increases, and if boxes of trim meat are distributed by the processor to more than one manufacturer, the required strength of this match increases. There is negligible probability that genotype matching will wrongly ascribe responsibility of the product to the manufacturer. However, if too few genotypes are sampled at the factory/and or from a returned sample, it may not be possible to make a robust genotype match between a returned product sample and its manufacturing batch. This Type I error can be reduced by increasing sampling frequency, but this would entail a cost at either end. Increasing the sampling rate from the boxes would require extra labour on the factory floor, and potentially could slow the manufacturing process. Sampling from a well-mixed blended meat product is time consuming since the individual samples must be cleanly separated so that they represent only one animal, to prevent genotypes from different individuals contaminating each other.

The Sampling Calculator allows a degree of optimisation of the sampling process by calculating the likely error rate for different sampling strategies, or finding strategies for a given error rate. An operator can choose an appropriate strategy, taking into account the cost of implementation. In addition, the Sampling Calculator can estimate the likely error of a negative result once DNA analyses have been performed. The Sampling Calculator uses Microsoft Excel (Fig 1).

FIGURE 1: Screen shot of Sampling Calculator



To use the calculator, an operator goes through the following steps:

1. Select the meat product being used from a drop-down list of pre-specified recipes (Step 1 in Figure 1 where Recipe 5 has been selected). A table is displayed showing how many boxes of each meat type are used for that recipe along with pre-specified characteristics of the packaging of that meat type, which are important to the sampling results.
2. Choose a sampling scenario from a drop-down list (Step 2 in Figure 1):
   a. take reference samples at the factory,
   b. take samples from a returned product sample.
3. Choose an analysis scenario from a drop-down list at (Step 3 in Figure 1). The operator can either specify a target error rate and have the Calculator find required sampling rates, or see error rates for a range of sampling combinations (at the factory and product ends).
4. Answer simple questions appropriate to the choices in 2 and 3 (Step 4 in Figure 1), click the calculate button and view the results.

In Figure 1, the operator is taking samples prior to manufacture of the mixed product and is estimating error rates for a range of sampling combinations, choosing from 1 to 5 samples per box at the factory and considering from 20 to 50 samples from the product in the even that it is returned. The variable 0.5 in Entry 5 of Figure 1 assumes that on average only 50% of samples taken from the product may be 'clean' enough to yield a single genotype. The size of this variable would need to be validated for each testing laboratory. Results for the analysis in Figure 1 are shown in Table 1 and suggest that if only one DNA sample is taken per box of meat prior to manufacture, the error in not matching a returned product would be 3.6%, 0.2% and 0.004% if respectively 20, 33 or 50 samples are taken from the returned product.

If, say, 2 samples were taken per box, these error rates would be respectively 0.2%, 0.001% and 4×10^{-6} % for the same number of product samples. An appropriate sampling strategy would take into account the costs of sampling at both ends of the production chain together with the likely return rate of products. The above values are for Recipe 5 (respectively 7, 8 and 6 boxes of meat type C, D, and E in the table in Figure 1) which has an estimated 35 individual animals contributing to the batch (Output 1 in Figure 1). If an alternative recipe (say 8 boxes of type C and 14 of type D, averaging 42 individuals per batch) was used, then for one (two) factory samples per box the corresponding error rates for 20, 33, and 50 product samples are 31% (6.5%), 5.5% (0.3%) and 0.6% (0.006%) respectively, showing that the type of meat used has a huge bearing on the error rate, and thus sampling strategies must be tailored to each recipe.

A very small error rate resulting from having appropriate sampling strategies in place would arguably underpin the position of a manufacturer when refuting a claim for compensation, thus providing an alternative use for the Sampling Calculator. If the operator decided, for the above recipe, to specify a target error rate, a box would appear for entry of this value. The results of a required error rate of 0.01% are shown in Table 2.

If one, two or three samples per box are taken at the factory, then, respectively 45, 26 and 21 samples are needed from the returned product. The calculator will not search outside the bounds specified by the minimum and maximum values, although smaller values may suffice. If the target error rate cannot be reached in the specified sampling range, the Calculator notifies the operator of this.

If the operator is sampling returned product at the post-manufacture stage, the operation of the Sampling Calculator is similar to the scenario discussed above, with the operator having similar choices. However, since the factory sampling has already been performed, rather than specifying a range of values to be considered, the actual number of samples per box is used as an input. The operator can then see error rates for different numbers of product samples, or find the required number of samples to meet a target error rate.

TABLE 1: Example of output from the Sampling Calculator showing the error rate in matching a returned product with its batch for different sampling frequencies at the factory (1 to 5) and from the product (20 to 50).

	Samples taken from product						
Factory samples per box of meat	20	26	30	36	40	46	50
1	0.036	0.0081	0.0031	0.0008	0.0003	8E-05	4E-05
2	0.0015	0.0001	3E-05	4E-06	9E-07	1E-07	4E-08
3	0.0002	1E-05	2E-06	1E-07	3E-08	3E-09	7E-10
4	5E-05	2E-06	3E-07	2E-08	4E-09	3E-10	7E-11
5	3E-05	9E-07	1E-07	8E-09	1E-09	1E-10	2E-11

TABLE 2: Output from the Sampling Calculator to show the number of product samples needed to meet a target error rate of 0.0001% for a range of factory sampling frequencies.

Target error rate:	0.0001	Samples required from returned product
Factory samples per box	1	45
	2	26
	3	21
	4	20
	5	20

DISCUSSION

Results from the calculator suggest that the efficiency of verification of mixed-meat products decreases rapidly as the number of individuals contributing to the batch increases, and that sampling rates must be increased to maintain acceptable error rates. The average yield of different grades or types of meat per individual may vary considerably, and therefore, products made predominantly from high-yield type meats would be more reliably matched to their batch of manufacture than products made predominantly from low-yield meat types. In the latter case, the number of individuals contributing to a batch will be greater, but a smaller fraction of them is sampled. Thus, it is necessary to take into account the meat types used in a recipe when developing sampling strategies. The calculator provides a fast and convenient way to determine appropriate sampling strategies for DNA identification of mixed meat products, involving the matching of individual genotypes from different levels of the food supply chain. The data suggest that DNA verification of batching using individual contributors may be more efficient in situations where batches have relatively small numbers of contributors. Where manufacture is on a very large scale, the sampling levels required will have a large effect on the cost-benefit of using this technology.

For a successful traceability system to be established using the Sampling Calculator, a number of practical issues must be addressed. Apart from a major contamination scare, returns are usually likely to constitute a very small proportion of the original batch. Whereas a batch may contain several thousand ground beef patties, the most likely traceability scenario would be that only a single product sample is returned. Therefore, it is crucial that the mixture is homogeneous to ensure that every piece of contributing meat that a reference sample was taken from has an equal opportunity of being present in every unit of product. Furthermore, the presentation of the meat when reference samples are being taken is important to ensure that the reference samples are in fact representative of the individual contributors to the batch. In terms of laboratory analyses, such a system would also rely on the ability to separate samples out from a mixture, where each sample is irrefutably from an individual. Any contamination with DNA from another individual renders a genotype useless for matching to a reference sample.

ACKNOWLEDGEMENT

We thank M.D. Rollo and A.B. Pleasants for helpful advice with the preparation of this manuscript.

REFERENCES

Arana, A.; Soret, B.; Lasa, I.; Alfonso, L. 2002: Meat traceability using DNA markers: application to the beef industry. *Meat science 61:* 67-373

Cagney, C.; Crowley, H.; Duffy, G.; Sheridan, J.J.; O'Brien, S.; Carney, E.; Anderson, W.; McDowell, D.A.; Blair I.S.; Bishop, R.H. 2004: Prevalence and numbers of *Escherichia coli* O157:H7 in minced beef and beef burgers from butchers shops and supermarkets in the Republic of Ireland *Food microbiology 21:* 2003-212

Caporale, V.; Giovannini, A.; Di Francesco, C.; Calistri, P. 2001: Importance of the traceability of animals and animal products in epidemiology. *Revue Scientifique et Technique de l'Office International des Epizooties 20*: 372-378

Cavani, C.; Petracci, M. 2004: Rabbit meat processing and traceability. *In* Proceedings of the 8th World Rabbit Congress. pp 1318-1336. Available online by Center for Agricultural and Rural Development Iowa State University at http://www.dcam.upv.es/8wrc, accessed on 30/01/05

Clemens, R. 2003: Meat traceability in Japan. *Iowa Agriculture Review* 9(4). Available online at http://www.card.iastate.edu, accessed on 30/01/05

Cunningham, E.P.; Meghen, C.M. 2001: Biological identification systems: genetic markers. *OIE scientific and technical review 20:* 491-499

Flores, R. A.; Tamplin, M. L. 2002: Distribution patterns of *Escherichia coli* O157:H7 in ground beef produced by a laboratory-scale grinder. *Journal of food protection 65:* 1894-1902

Hobbs, A.L.; Hobbs, J.E.; Isaac, G.E.; Kerr, W.A. 2002: Ethics, domestic food policy and trade law: assessing the EU animal welfare proposal to the WTO. *Food policy 27:* 437-454

McKean, J.D. 2001: The importance of traceability for public health and consumer protection. *Revue Scientifique et Technique de l'Office International des Epizooties 20:* 363-371

Meghen, C.N.; Scott, C.S.; Bradley, D.G.; MacHugh, D.E.; Loftus, R.T.; Cunningham, E.P. 1998: DNA based traceability techniques for the beef industry. *Animal genetics, 29 (Supplement 1):* 48-49

Miraglia, M.; Berdal, K.G.; Brera, C.; Corbisier, P.; Holst-Jensen, A.; Kok, E.J.; Marvin, H.J.P.; Schimmel, H.; Rentsch, J.; van Rie, J.P.P.F.; Zagon, J. 2004: Detection and traceability of genetically modified organisms in the food production chain. *Food and chemical toxicology 42:*1157-1180

Ozawa, T., Lopez-Villalobos, N. and Blair, H.T. 2005: An update on beef traceability regulations in Japan. *Proceedings of the New Zealand Society of Animal Production 65:* 80-84

Schroeder, C.M.; White, D.G.; Ge, B.; Zhang, Y.; McDermott, P.F.; Ayers, S.; Zhao, S.; Meng, J. 2003: Isolation of antimicrobial-resistant *Escherichia coli* from retail meats purchased in Greater Washington, DC, USA. *International journal of food microbiology 85:* 197-202

Shackell G.H.; Tate M.L.; Anderson, R.M. 2001: Installing a DNA-based traceability system in the meat industry. *Proceedings of the Australasian Association for the Advancement of Breeding and Genetics 14*: 533-536

Smith, P.G.; Cousens, S.N.; Huillard d'Aignaux, J.N.; Ward, H.J.T.; Will, R.G. 2004: The epidemiology of Variant Creutzfeldt-Jacob Disease. *Current topics in microbiology and immunology, 284*: 161-191

Stock, K.; Stolle, A. 2001: Incidence of *Salmonella* in minced meat produced in a European Union-approved cutting plant. *Journal of food protection 64(9)*:1435-1438

Tate, M.L. 2001: Traceability of meat products – application of DNA technology. *Proceedings of the New Zealand Grasslands Association 63*: 255-257

Vázquez, J.F.; Pérez, T.; Ureña , F.; Gudín, E.; Ablornoz, J.; Domínguez, A. 2004: Practical application of DNA fingerprinting to trace beef. *Journal of food protection 67:* 972-979