Biotracing microbial contaminants - *Salmonella* on pork

Joost H. Smid

Joost H. Smid, 2012
Biotracing microbial contaminants - *Salmonella* on pork
Dissertation Utrecht University, Faculty of Veterinary Medicine
ISBN: 978-94-6191-256-5

Printing: Ipskamp Drukkers, Enschede

The reserach presented was conducted at the national Institute for Public Health and the Environment (RIVM), Bilthoven, The Netherlands

Biotracing microbial contaminants - *Salmonella* on pork

Biotraceren van microbiële besmetting
– *Salmonella* op varkensvlees

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht
op gezag van de rector magnificus, prof.dr. G.J. van der Zwaan,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op
dinsdag 5 juni 2012 des ochtends te 10.30 uur

door

Joost Hubert Smid

geboren op 8 december 1979
te Hoogeveen

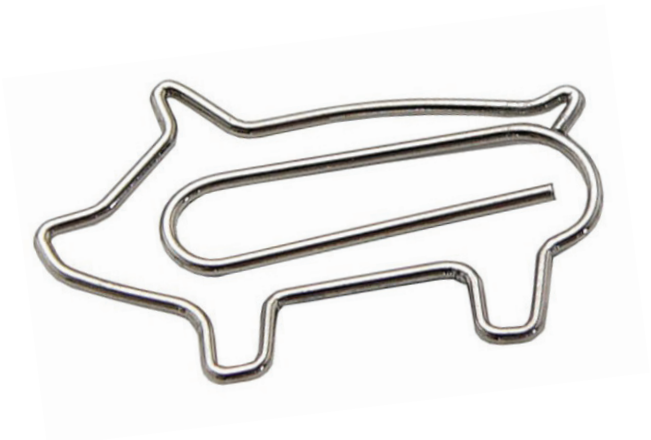
Promotoren:	Prof. dr. ir. A.H. Havelaar Prof. dr. F. van Knapen
Co-promotor:	Dr. A. Pielaat

This work presented in this thesis was supported by the European-Union-funded Integrated Project BIOTRACER (contract #036272) under the 6th RTD Framework.

Contents

Chapter 1

General introduction



Section 2 of the General Introduction is a shortened version of:
An introduction to biotracing in food chain systems
G.C. Barker [a], N. Gomez-Tomé [a], J.H. Smid [b,c]

[a] Institute of Food Research, Colney, Norwich, NR4 7UA, UK
[b] RIVM, A. van Leeuwenhoeklaan 9, 3720 BA Bilthoven, The Netherlands
[c] IRAS, Yalelaan 2, 3584 CM Utrecht, The Netherlands

Trends in Food Science and Technology 20: 220-226 (5)

1. SALMONELLA

1.1 *Salmonella*

Salmonella is a genus of Gram-negative, rod-shaped, non spore forming, in general motile enterobacteria with a diameter of 0.7 to 1.5 µm. *Salmonella* is a member of the *Enterobacteriaceae* family and was first reported by (and named after) Daniel Elmer Salmon(1850-1914), an American veterinary pathologist. There are more than 2400 known *Salmonella* serovars, grouped in two species (*bongori* and *enterica*) and 6 subspecies (*houtenae, arizonae, diarizonae, enterica, salamae* and *indica*). Serovars can again be subdivided in a large number of phage types, which indicate subsets of one serovar that are susceptible to the same bacteriophages (Fig. 1, Wareing & Fernandes, 2007).

Salmonella species are capable of colonizing a wide variety of warm and cold blooded animals, including humans. Well-known reservoirs for *Salmonella* spp. are farm animals such as chicken, cows and pigs, pets, (avian) wildlife, several reptile species and insects. Furthermore, *Salmonella*e are constantly released from infected animals into the environment, often through infected feces (Baudart et al., 2000). In feces, water and soil they can survive during extended periods of time, e.g. up to 60 days (Parker & Mee, 1982; Semenov et al., 2009). Also, *Salmonella* may be transferred from the (products of) farm animals to the industrial plant environment during processing of these (products), where they can also survive and multiply and be involved in complex contamination cycles (Botteldoorn et al., 2004).

1.2 *Salmonella* and the human health risks

Salmonella spp. cause a relatively high burden of disease in the population (7.7 DALY per 100,000 Dutch inhabitants, Havelaar, et al., submitted). The host-adapted serovars *S.* typhi and *S.* paratyphi can cause serious illness in humans, such as typhoid fever. Non-typhoidal *Salmonella* infections can be transferred between humans and animals. The name of an infection with *Salmonella* bacteria is salmonellosis. Human salmonellosis is most often contracted through the consumption of contaminated poultry, eggs, pork and beef. The consumption of and exposure to undercooked meat and raw eggs are reported as risk factors (Doorduyn et al., 2006). Also, many other foods such as contaminated chocolate, cheese or vegetables contaminated via manure are found to be associated to infections. Playing in a sandbox was reported as a risk factor for children. More than 80% of all salmonellosis cases are individual cases rather than outbreaks, although outbreaks typically receive most attention in the media.

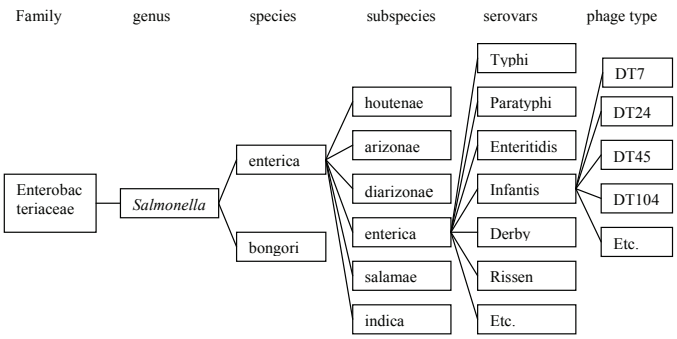


Fig. 1: *Schematic representation of Salmonella classification and diversity (Hermans, 2007)*

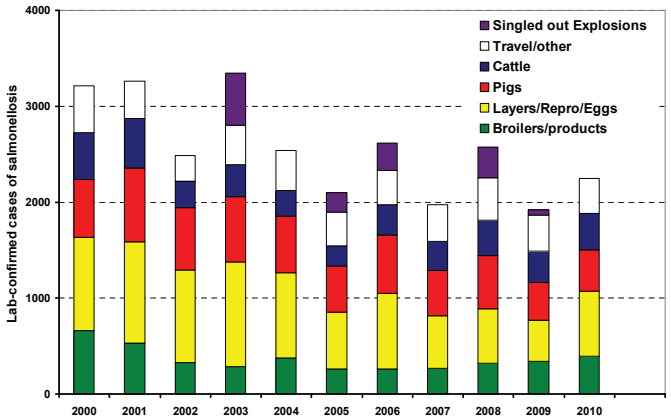


Fig. 2: *Overview of the total number of lab-confirmed salmonellosis cases and the estimated total number of salmonellosis in the general population in the Netherlands per year in the period 2000-2010 (RIVM, 2010). Estimates of the number of cases that are caused by the different sources of Salmonella are also indicated.*

The incubation period for human salmonellosis is typically 6-48 hours, depending on the condition of the host and the *Salmonella* serovar. Salmonellosis can cause an acute lower gastrointestinal tract disease in humans. The acute symptoms of this disease are nausea, vomiting, diarrhea, abdominal cramps, fever and headaches. Usually antibiotic treatment is not required and the acute symptoms disappear after 1 to 2 days. However, depending on the condition of the patient and the virulence of the *Salmonella* serovar, the disease can persist longer. In some cases, invasive forms of *Salmonella* infections may develop, in which *Salmonella* is no longer restricted to the gastrointestinal tract and may reach the bloodstream, from which other body organs may be affected with severe consequences. Sequellae may be expressed by symptoms of reactive arthritis and irritable bowel syndrome (IBS), 3 to 4 weeks after the acute symptoms.

In the past 10 years, the number of lab-confirmed cases of salmonellosis in The Netherlands has declined from around 3200 in 2000 to around 2250 in 2010 (Fig. 2). This agrees with a reduction of an estimated number of 51000 cases in 2000 to 36000 cases in 2010 in the general population. Most of this decline is credit to a successful policy of reducing *Salmonella* prevalence in laying hens and, consequently, in eggs. The estimated number of cases caused by the remaining sources has much less decreased. Therefore, pork has become the second-most important source of salmonellosis for humans in 2010, having caused around 430 lab-confirmed cases of salmonellosis (or around 7000 cases in the general population, Anonymous, 2011). It was estimated that 19% of all human cases of salmonellosis in the Netherlands in 2010 were associated with the consumption of pork. In Denmark and Germany, also relatively high attributable proportions of cases to pork (15-20%) were found (Anonymous, 2008; Borch et al., 1996; Steinbach & Kroell, 1999).

1.3 Salmonella in pork

Salmonella can colonize the intestinal tract of pigs if these eat contaminated feed, or be in contact with a contaminated on-farm environment (Blackman et al., 1992; Shapcott, 1984). Most often, pigs

are healthy carriers of *Salmonella* and, usually, *Salmonella* infections produce no severe disease in pigs. It is estimated that between 5-30% of the pigs will still excrete *Salmonella* at the end of the finishing period (Berends et al., 1996). In the Netherlands, pig suppliers take part in the so-called IKB chain quality system (Integrale Ketenbeheersing, certified as EN45011), in which blood samples are taken to categorize finishing pigs in risk categories for *Salmonella*. Farms in the high risk categories must implement measures to reduce *Salmonella*. Yet, this categorization is not further used by most large slaughterhouses in the Netherlands (among which the slaughterhouse that was studied for this thesis). Moreover, it has been shown in several studies that massive cross-contamination may take place between pigs during transport from the farm to the slaughterhouse and during lairage (Boes et al., 2001; Rostagno, et al., 2003; Warriss, 2003). Therefore, contaminated pigs may introduce *Salmonella* from the farm into the slaughterhouse.

Slaughterhouses have a long history of controlling microbiological hazards and *Salmonella* is one of the most studied and notorious zoonotic pathogens in meat. Good slaughtering hygiene has, therefore, always been important in the pork industry. The steps of the slaughter chain are similar for the majority of slaughterhouses in North-West Europe, which are typically large and modern, with an infrastructure to avoid cross-contamination e.g. hanging carcasses, decontamination facilities and biosecurity (EFSA, 2010). The following description of these steps applies to the particular slaughterhouse that was studied for this thesis, but it applies equally to most large slaughterhouses.

Several steps in the slaughter line (especially in the first part of the process) contain elements by which *Salmonella* is inactivated, such as hot water treatments, mechanical hair removal and flaming (Chapter 3). According to EU regulation, cutting knives are inserted into water of 82°C after cutting every 10 carcasses (Eustace & Midgley, 2007). The floors and slaughtering machines are sanitized after every slaughtering day using warm water, cleaning and disinfection. Also the cutting knives of the robots are replaced by clean knives after every day. Yet, it has been shown that *Salmonella* is in some instances not completely eradicated by these steps. At the end of slaughtering, carcasses may have become contaminated through contact with feces from the intestinal tracts, the lymphoid tissues or from the skins of contaminated pigs (Quirke, et al., 2001; Sorensen et al., 2004). Also the slaughterhouse environment may become contaminated by *Salmonella* from these sources and cross-contaminate other carcasses passing through this environment (Botteldoorn et al.,; Pearce et al., 2004; Swanenburg, 2000; Warriner et al., 2002). Several studies have found that *Salmonella* may also survive in certain niches in the slaughterhouse, become part of the house flora of that slaughterhouse and persist over an extended period of time (Giovannacci et al., 2001; Swanenburg, 2000; Vieira-Pinto et al., 2006). Therefore, house flora can also be considered as a potential source of *Salmonella* for carcasses after slaughter.

To verify how many carcasses are contaminated with *Salmonella* after slaughtering, it is standard practice to take microbiological samples from one in 1000 carcasses the next day in the cooling area. These samples are subsequently analyzed for total viable counts (TVC), enterobacteriaceae and for *Salmonella* prevalence. Any violation of the standards, such as described in EU regulation (EC/2073/2005), should lead to appropriate measures to improve the microbiological performance. In general, findings of a high *Salmonella* prevalence or elevated enterobacteriaceae and/or TVC are indications of an anomaly in hygienic procedures of the slaughterhouse. In case of an anomaly, the quality managers of the

slaughterhouse should do a root cause analyses and steer corrective and preventive actions. Typically, problems are caused by insufficient cleaning or improper hygienic procedures of the slaughter house worker, including fecal contamination from the rectal contents of the pig and insufficient dehairing. In automated slaughterhouses increased levels of *Salmonella* can also indicate resident *Salmonella* in the automatic equipment of the slaughter line. A good way to solve the problems is, therefore, by disinfecting the robots. Subsequently, samples from carcasses after slaughter are again analyzed to verify the effects of the instituted measures. Extra analyses such as comparing *Salmonella* serotypes from the carcasses and potential sources can also be done to reduce the quality manager's uncertainty about the source of contamination. Prioritizing measures is based on the nature of the contamination problem, the contamination history of the specific slaughterhouse and on the experience and commitment of the management in that slaughterhouse. In practice, finding the source of contamination can be a time-consuming and inefficient process so that the contamination problem can in exceptional cases persist over several weeks.

The importance of pork for human salmonellosis makes it clear that more insight is needed where *Salmonella* is introduced in the pork chain, how its dynamics are in the chain, and how *Salmonella* can be traced back to its point of introduction in the chain. From a perspective of public health, it is important to develop a more systematic approach to the prevention and tracing of *Salmonella* in pig slaughterhouses. To this purpose, the use of *biotracing* models will be discussed in this thesis. Biotracing is a novel discipline in food safety science. Its concept will be defined in the next section and placed in the context of other disciplines in microbial food safety.

2. BIOTRACEABILITY

2.1 Introduction to biotraceability

In relation to modern food supply chains, tracing is a well established and an important concept. A traceability system relies on labeling of individual food items and on record keeping at the point of origin and often at intermediate steps in the chain. The label is often a bar code or another kind of (macroscopic) marker that uniquely identifies a space time point for the source of the food materials and is irreversibly attached to the food item. The irreversibility of the association between the food and the label ensures that the information content through the food chain is conserved. In principal, tracing is perfect and it is not compromised by the complexity of the chain. Food traceability plays a large part in modern food quality assurance systems and has been the subject of considerable recent academic, technological and regulatory developments e.g. European Commission (2007), Golan et al. (2004), Food Standards Agency (2002).

The word 'biotraceability' will be used to indicate an area of research that is distinct from, but related to, traceability and is less well established. Biotracing concerns the spatial and temporal distribution of sources of harmful agents (particularly bacterial pathogens, such as *Salmonella*, but also pathogenic toxins) in the food chain to indicate the origin of a potential food safety problem. Pathogens that contaminate food follow the manufacturing chain and the dispersal patterns of the food vehicle but, unlike the actual food units, pathogens are not easily monitored, cannot be labeled uniquely and record keeping, usually based on sampling, is often sparse, imperfect and not immediate. Moreover,

the combination of the chain and population dynamics is typically not completely understood. This ensures that the biotracing process is very different to the tracing process. It involves background understanding of the dynamics of the pathogen and statistical inference about their points of introduction (sources) rather than direct deterministic monitoring and, for this task, it integrates food chain expertise, food safety information, rapid detection methods and mathematical models.

Clearly biotracing is related to many established elements of food safety management and food security including hazard analysis and quality control. It also includes a statistical perspective relating to operating experience within a food chain and to interpretation of multiple streams of uncertain evidence. Although the concept of causal inference, in relation to biological hazards, is reasonably well established, the process of biotracing in food chain systems remains poorly defined. It is clear that biotracing is distinct from inferential schemes like 'diagnosis' and 'tracing', and might be seen as including elements of both processes at once (possibly with additional influence from other auditing and fault finding processes). A simple working definition of biotracability will be used, which might apply to many food chain hazard scenarios.

"Biotraceability is the ability to use down stream information to point to materials, processes or actions within a particular food chain that can be identified as the source of undesirable agents".

Biotracing offers many benefits. Most simply a biotrace could be used to target immediate food safety interventions and hence could help prevent the recurrence of specific hazardous events following an observed incident. However, more strategically, biotracing could operate alongside established control systems to give pre-emptive indications of potential weaknesses in a food chain and hence could be a driving force for a steady improvement in food safety. Additionally, each of these mechanisms could be used to limit downtime or to reduce the number of recalls associated with safety incidents and therefore could act to drive improved operational performance.

2.2 Elements of a biotraceability system

The main elements of a biotracing system are an information set, based on chain measurements, and a belief system, based on the chain structure and the expected dynamics of the pathogen in the chain. The quantitative picture for the dynamics of the pathogen embraces the concept of causality, e.g. Pearl (2000). A biotraceability system supports the identification or prioritization of potential initiation points or processes if hazards or hazard markers are detected. Operation of biotracing in a food chain might use an inline monitor or trigger to provide an 'action' signal that, in turn, triggers an investigation based on established statistical (uncertain) information. The investigation would point to the points of introduction of the pathogens (sources), or to potential unobserved failures, that have priority for follow on actions aimed at improved food safety. Connectivity between the trigger, the investigation and the consequent activity is a central feature of biotracing. The posterior nature of the information that constitutes a biotrace and the probabilistic structure of the domain suggest the use of Bayesian methods, which include an explicit representation of uncertainties, as the basis for quantitative understanding of biotraceability.

A particular conditional probability, the probability for source strengths given information about the

food chain and given evidence from a set of monitoring observations, may be used to symbolize biotraceability. In general the set of sources includes all the possible sites and times at which pathogens may enter the system, i.e. contamination events, but may also include processes and actions, localized in space and time, that lead to rapid multiplication or unintentional survival of the population of pathogens, i.e. failure events. A complete identification of prospective sources is usually impossible; because the number of possible sources is often unknown.

Observational evidence supporting biotracing corresponds to any events which may be attributed to the presence of pathogens or to an active source. A case of illness or an identified exposure event may be the most obvious evidence but, within the context of food safety, out of specification population sizes during quality control monitoring, or other in line tests, are more appropriate. In abstract terms the evidence set is used to make a diagnosis amongst hypothesized sources. In addition to being conditional on evidence the beliefs about sources are conditional on the specific food chain structure. The structure of a particular food chain interacts with the dynamics of pathogens and so is reflected in the inference. Generally it was assumed that the food chain structures are static (do not depend explicitly on time) but their elements may be time ordered and may be stochastic (i.e. lead to dispersion of individual food units); for many food chains these are acceptable approximations.

The probabilistic structure of a biotracing system also helps to identify elements that impede the process of biotracing. Most clearly the biotracing problem is very complex; as well as the large number of sources and the large number of events there are large numbers of variables that are associated with the chain structure. In order to establish inference it is essential to quantify the relevant variables in the food chain as well as their relationships or dependencies. Additionally information relating to the populations of pathogens in food is necessarily uncertain, because it is obtained indirectly, and often it is incomplete (so expert opinions are included). Clearly it is essential to fully understand the information collection processes in order to make quantitative inference relating to biotraceability.

2.3 Relation to established methods in food safety

Published scientific literature related to biotracing can be found in many different disciplines including outbreak investigations, risk assessment, microbial forensics and source attribution.

2.3.1 Outbreak investigations

Outbreak investigations (e.g. Reingold, 1998) are possibly the most familiar structures used for the association of a (food) source with an observed food safety failure. In the majority of outbreak investigations the outcome is the identification of a particular food that could be associated with the group of confirmed cases of illness.

An outbreak investigation can be considered either as a direct, semi-quantitative, evaluation of the biotracing probability or, more correctly, as a test of a hypothesis which asserts that a particular source is an active source and is the origin for the evidence. In an outbreak investigation the food safety event is a group of cases that are potentially connected. The relevant chain of events is usually very simple and relates only to the recent consumption for each of the individual cases. Crucial to an outbreak investigation is an assumption that the group of cases can be associated with a single source. The

evidence usually includes an observation of 'matching' for a particular biomarker, such as a pathogenic subtype, at the source and cases. When the biomarker is very discriminating, this means that the matching information, for a group of cases and a single source, is very strong and is sufficient to make a unique association between the cases and a single source. Clearly the assignment of a source probability requires assumptions about the probability for a match and the probability of unidentified active sources.

Although an outbreak investigation often ends with the strong assignment of a source it does not fit with many of the criteria associated with biotracing. The investigation usually takes place after the event and therefore is not a real time process. In most cases the investigation does not extend beyond the association of a food type with a group of cases and therefore does not provide information about preceding elements of the food chain.

2.3.2 Risk assessments

The majority of Quantitative Microbial Risk Assessments (QMRA, e.g. EFSA (2010)) provide estimates for the joint probability of the variables within the hazard domain i.e. the probability that the sources are active and that intermediate processes are successful and that the evidence is observed. From this probability it is possible to establish, by appropriate numerical methods, all of the marginal probabilities for exposure, for event sizes etc. that are usually the end points of risk assessment.

The expression for the joint probability, which encapsulates the risk assessment process, indicates a route to biotraceability information. This route will be described in more detail in Chapter 1 and 2 of this thesis. In principle Bayes' theorem (See Chapter 1) can be used to convert risk assessment into biotracing. The Bayesian inversion, in the complex domain that represents a food chain, is non trivial but is practical with some modern statistical techniques, e.g. Kjaerulff & Madsen (2008), and it is an integral element of Bayesian belief network modelling for complex systems. Increasingly Bayesian belief network models are applied in QMRA (Barker, 2004).

2.3.3 Microbial forensics

Microbial forensics is a relatively new activity concerning evidence collection and source identification for biocrime; e.g. American Academy of Microbiology (2003). A forensic approach evaluates the probability of one particular evidence set for two prospective sets of active sources to support a decision regarding the most likely source (often based on the value for the ratio of the two probabilities – likelihood ratio). However it is important to realize that although the likelihood evaluation can aid discrimination (a hypothesis test) it does not, on its own, provide a posterior probability assignment for prospective active sources (biotrace).

2.3.4 Source attribution

Source attribution (e.g. Hald et al., 2004; Pires, 2008; Wilson et al., 2008) consists of the quantification of the contributions of different (food) animal reservoirs to human illness. To this purpose, differences in the relative frequency of occurrence of pathogenic subtypes in individual sources can be used. Attribution takes the form of statistics to establish time averaged association (not direct causation) between cases and sources. Attribution can be considered as an averaged form of biotracing.

2.3.5 Summarizing the biotracing concept

Biotracing will rely on optimized record keeping aimed largely at the continuous updating of prior beliefs concerning source profiles. Many existing quality control operations provide large volumes of information about sources that is underused and can potentially support this development. Within a food chain system some behaviour that is distinct from normal performance should alert a biotracing system. Information from a biotrace should facilitate remedial actions that can be targeted more accurately or interventions that are not completely indiscriminate. The identification of appropriate monitors and signals is chain specific and can be guided by the top level biotracing questions. Many countries maintain extensive disease outbreak identification systems for both public health and bioterror scenarios (e.g. Smith et al. (2006)). Although these systems are not constrained by food chains and do not include an actual biotrace they indicate the nature of anomaly detection for time series data that should be an essential element, a trigger (Chapter 3), in an operational biotracing system.

It is thought that operational biotracing could lead to improved efficiency, improved confidence in controls, improved decision making and, ultimately, improved food safety. Biotracing suggests that systematic surveillance and analyses can soon become superior to precedent and subjective judgments in decisions concerning the sources of food borne hazards.

3. AIM AND OUTLINE OF THIS THESIS

This thesis was developed as a part of the BIOTRACER-IP (www.biotracer.org), in which international participants collaborated on developing microbiological test methods and mathematical computer models to improve the tracing of accidental and deliberate contamination of food and animal feed. A number of food chains and pathogens in these chains were defined in the project, for which proofs-of-concepts for biotracing-systems were developed. *Salmonella* in the pork slaughter chain was one of these chains; other chains that were considered were *Salmonella* in feed (Binter et al., 2010), *Staphylococcus aureus* in milk (Barker & Goméz-Tomé, 2011), *Listeria monocytogenes* in cheese and *Campylobacter* in chicken. The research presented in this thesis describes the development of a biotracing system for *Salmonella* in the pork slaughter chain. The main goal of the research presented in this thesis is to answer the research question: Can a model be developed to systematically trace back sources of *Salmonella* by which pig carcasses have become contaminated after slaughter? An answer to this question will guide the discussion whether the development of such a system is -strategically- a good alternative to existing methods to trace anomalies in hygienic procedures of the slaughterhouse. Studies related to modeling methods, *Salmonella* in pork and the dynamics of *Salmonella* during slaughtering will be described. These studies were designed to meet the following objectives:

To establish a precise working definition of the biotraceability concept.

To develop an appropriate modeling framework for answering biotracing questions.

To develop a generic model for biotracing *Salmonella* in a pig slaughterhouse.

To use *Salmonella* data in the slaughterhouse to generate a biotracing model for that specific slaughterhouse.

As a first step in the BIOTRACER-IP, the concept of biotracing was developed and placed in the

context of existing methods for the improvement of microbial food safety. A summary of the developed theoretical framework was given in the previous section. The remainder of the thesis describes the development of quantitative methods for biotracing *Salmonella* in the pork slaughter chain, using different modeling approaches and qualitative and quantitative *Salmonella* data from a Dutch pork slaughter plant.

In Chapter 2, we discuss different aspect of Bayesian belief networks and the more traditionally used Monte Carlo simulation models in microbial risk assessment. Based on this discussion, we describe a practical framework in which a Monte Carlo model, where sequential events in the chain are simulated, is converted into a Bayesian belief network, which is a natural framework for answering biotracing questions. In Chapter 3, the proposed framework is illustrated with a biotracing model of *Salmonella* in the pork slaughter chain, based on a recently published Monte Carlo simulation model (EFSA, 2010). In Chapter 4 we discuss the setup of an experiment in which microbiological samples of carcasses before and after slaughter and of slaughter robots were collected in a Dutch pig slaughter plant. These samples were analyzed for *Salmonella* prevalence, concentrations and serotypes. We present the data and discuss the transmission paths of *Salmonella* in a pig slaughter plant. The *Salmonella* concentration data is subsequently used as inputs for the biotracing model. We discuss biotracing results in Chapter 5, and show how the data is used to gradually improve certain parameter estimates of the model. Inspection of the estimated parameter values showed that the rates of cross-contamination remained uncertain after inputting *Salmonella* concentration data. Furthermore, the results of a sensitivity analysis showed that the exact values of these rates are important for the conclusion of the biotracing model. Therefore, a laboratory experiment is described in Chapter 6 in which these rates are measured. In Chapter 7, we used the serotyping data and a number of statistical approaches to compare to the results of the biotracing model described in Chapter 5 and to obtain additional evidence for the relative importance of the different sources of *Salmonella* that play a role in the contamination of carcasses after slaughtering. In the final Chapter 8, the results will be discussed jointly and future challenges for biotracing *Salmonella* in the pork slaughter line will be provided.

REFERENCES

American Academy of Microbiology (2003). Microbial forensics: A Scientific Assessment.

http://www.asm.org/ASM/files/CCPAGECONTENT/docfilename/0000018026/FOREN%20REPORT_BW.pdf

Anonymous. (2008). Scientific Opinion of the Panel on Biological Hazards on a request from EFSA on Overview of methods for source attribution for human illness from food borne microbiological hazards. The EFSA journal, 764, 1-43.

Anonymous. (2011). Staat van Zoonosen 2010. RIVM report 330291007/2011. Bilthoven: RIVM.

Barker, G. C. (2004). Application of Bayesian Belief Network models to food safety science. Bayesian statistics and quality modelling in the Agro-Food production chain (Vol. 3, pp. 117-130). Wageningen: Kluwer Academic Publishers.

Barker, G. C., & Goméz-Tomé, N. (2011). A risk-assessment model for enterotoxigenic Staphylococcus aureus in pasteurized milk: a potential route to source-level inference. Risk Analysis

Baudart, J., Lemarchand, K., Brisabois, A., & Lebaron, P. (2000). Diversity of Salmonella strains isolated from the aquatic environment as determined by serotyping and amplification of the ribosomal DNA spacer regions. Appl Environ Microbiol, 66(4), 1544-1552.

Berends, B. R., Urlings, H. A. P., Snijders, J. M. A., & van Knapen, F. (1996). Identification and quantification of risk factors in animal management and transport regarding Salmonella spp. in pigs. Int J Food Microbiol, 30, 37-53.

Binter, C., Straver, J. M., Haggblom, P., Bruggeman, G., Lindqvist, P. A., Zentek, J., & Andersson, M. G. (2010). Transmission and control of Salmonella in the pig feed chain: A conceptual model. Int J Food Microbiol, 145, S7-S17.

Blackman, J., Bowman, T., Chambers, J., Kisilenko, J., Parr, P., St-Laurent, A. M., & Thompson, J. (1992). Controlling Salmonella in Livestock and Poultry Feeds. Plant Products Division. Ontario, Canada: Centre for Food and Animal Research of Agriculture.

Boes, J., Dahl, J., Nielsen, B., & Krog, H. H. (2001). Effect of separate transport, lairage, and slaughter on occurrence of Salmonella typhimurium on slaughter carcasses. Berl Munch Tierarztl Wochenschr, 114(9-10), 363-365.

Borch, E., Nesbakken, T., & Christensen, H. (1996). Hazard identification in swine slaughter with respect to foodborne bacteria. Int J Food Microbiol, 30(1-2), 9-25.

Botteldoorn, N., Herman, L., Rijpens, N., & Heyndrickx, M. (2004). Phenotypic and molecular typing of Salmonella strains reveals different contamination sources in two commercial pig slaughterhouses. Appl Environ Microbiol, 70(9), 5305-5314.

Botteldoorn, N., Heyndrickx, M., Rijpens, N., Grijspeerdt, K., & Herman, L. (2003). Salmonella on pig carcasses: positive pigs and cross contamination in the slaughterhouse. J Appl Microbiol, 95(5), 891-903.

Doorduyn, Y., Van Den Brandhof, W. E., Van Duynhoven, Y. T., Wannet, W. J., & Van Pelt, W. (2006). Risk factors for Salmonella Enteritidis and Typhimurium (DT104 and non-DT104) infections in The Netherlands: predominant roles for raw eggs in Enteritidis and sandboxes in Typhimurium infections. Epidemiol Infect, 134(3), 617-626.

EFSA. (2010). EFSA Quantitive Microbial Risk Assessment on Salmonella in Slaughter and Breeder pigs: Final Report.

European Commission (2007). Food traceability.

http://ec.europa.eu/food/food/foodlaw/traceability/factsheet_trace_2007_en.pdf

Food Standards Agency (2002). Traceability in the Food Chain.

Eustace, I., & Midgley, J. (2007). An alternative process for cleaning knives used on meat slaughter floors Int J Food Microbiol, 113(1), 23-27.

Giovannacci, I., Queguiner, S., Ragimbeau, C., Salvat, G., Vendeuvre, J. L., Carlier, V., & Ermel, G. (2001). Tracing of Salmonella spp. in two pork slaughter and cutting plants using serotyping and macrorestriction genotyping. J Appl Microbiol, 90(1), 131-147.

Golan, E., Krissoff, B., Kuchler, F., Calvin, L., Nelson, K., & Price, G. (2004). Traceability in the U.S. Food Supply: Economic Theory and Industry Studies. USDA Agricultural Economic Report 830.

Hald, T., Vose, D., Wegener, H. C., & Koupeev, T. (2004). A Bayesian approach to quantify the contribution of animal-food sources to human salmonellosis. Risk Anal, 24(1), 255-269.

Havelaar, A. H., Haagsma, J. A., Mangen, M. J., Kemmeren, J. M., Verhoef, L. P., Vijgen, S. M., Wilson, M., Friesema, I. H., Kortbeek, L. M., Van Duynhoven, I. T., & Van Pelt, W. (submitted). Disease burden of foodborne pathogens in The Netherlands, 2009. Int J Food Microbiol.

Hermans, A. P. M. H. (2007). Stress response and virulence in Salmonella Typhimurium: a genomics approach. Wageningen University.

Kjaerulff, U. B., & Madsen, A. L. (2008). Bayesian Networks and Influence Diagrams. New York: Springer.

Parker, W. F., & Mee, B. J. (1982). Survival of Salmonella adelaide and fecal coliforms in coarse sands of the swan coastal plain, Western Australia. Appl Env Microbiol, 43, 981-986.

Pearce, R. A., Bolton, D. J., Sheridan, J. J., McDowell, D. A., Blair, I. S., & Harrington, D. (2004). Studies to determine the critical control points in pork slaughter hazard analysis and critical control point systems. Int J Food Microbiol, 90(3), 331-339.

Pearl, J. (2000). Causality: Models Reasoning and Inference. Cambridge: Cambridge University Press.

Pires, S. M., Evers, E., van Pelt, W., Avers, T., Scallan, E., Angulo, F.J., Havelaar, A.H. & Hald, T. (2008). Attributing the human disease burden of foodborne infections to specific sources. Report Med-Vet-Net WP28.

Quirke, A. M., Leonard, N., Kelly, G., Egan, J., Lynch, P. B., Rowe, T., & Quinn, P. J. (2001). Prevalence of Salmonella serotypes on pig carcasses from high- and low-risk herds slaughtered in three abattoirs. Berl Munch Tierarztl Wochenschr, 114(9-10), 360-362.

Reingold, A. L. (1998). Outbreak Investigations – A Perspective. Infectious Diseases, 4, 21-27.

RIVM. (2010). Staat van Zoönosen. RIVM Rapport (330291007/2011 ed.). Bilthoven: RIVM.

Rostagno, M. H., Hurd, H. S., McKean, J. D., Ziemer, C. J., Gailey, J. K., & Leite, R. C. (2003). Preslaughter holding environment in pork plants is highly contaminated with Salmonella enterica. Appl Environ Microbiol, 69(8), 4489-4494.

Semenov, A. V., van Overbeek, L., & van Bruggen, A. H. (2009). Percolation and survival of Escherichia coli O157:H7 and Salmonella enterica serovar Typhimurium in soil amended with contaminated dairy manure or slurry. Appl Environ Microbiol, 75(10), 3206-3215.

Shapcott, R. (1984). Practical aspects of Salmonella control: progress report on a programme in a large broiler integration. International symposium on Salmonella. New Orleans, USA: American Association of Avian Pathology.

Smith, G. E., Cooper, D. L., Loveridge, P., Chinemana, F., Gerard, E., & Verlander, N. (2006). A national syndromic surveillance system for England and Wales using calls to a telephone helpline. Euro. Surveill., 11, 667.

Sorensen, L. L., Alban, L., Nielsen, B., & Dahl, J. (2004). The correlation between Salmonella serology and isolation of Salmonella in Danish pigs at slaughter. Vet Microbiol, 101(2), 131-141.

Steinbach, G., & Kroell, U. (1999). Salmonella infections in swine herds - epidemiology and importance for human diseases. Berl. Tierärztl Wochenschr, 106, 282-288.

Swanenburg, M. (2000). Salmonella in the pork production chain: sources of Salmonella on pork. PhD thesis, University Utrecht, Utrecht, the Netherlands.

Vieira-Pinto, M., Tenreiro, R., & Martins, C. (2006). Unveiling contamination sources and dissemination routes of Salmonella sp. in pigs at a Portuguese slaughterhouse through macrorestriction profiling by pulsed-field gel electrophoresis. Int J Food Microbiol, 110(1), 77-84.

Wareing, P., & Fernandes, R. (2007). Micro-Facts (1 ed.). Cambridge: RSC Publishing.

Warriner, K., Aldsworth, T. G., Kaur, S., & Dodd, C. E. (2002). Cross-contamination of carcasses and equipment during pork processing. J Appl Microbiol, 93(1), 169-177.

Warriss, P. D. (2003). Optimal lairage times and conditions for slaughter pigs: a review. Vet Rec, 153(6), 170-176.

Wilson, D. J., Gabriel, E., Leatherbarrow, A. J., Cheesbrough, J., Gee, S., Bolton, E., Fox, A., Fearnhead, P., Hart, C. A., & Diggle, P. J. (2008). Tracing the source of campylobacteriosis. PLoS Genet, 4(9), e1000203.

Chapter 2

Strengths and weaknesses of Monte Carlo simulation models and Bayesian belief networks in microbial risk assessment

J.H. Smid [a,d], D. Verloo [b], G.C. Barker [c], A.H. Havelaar [a,d]

[a] RIVM, A. van Leeuwenhoeklaan 9, 3720 BA Bilthoven, The Netherlands
[b] European Food Safety Authority, Largo N. Palli 5/A, I43100 Parma, Italy
[c] Institute of Food Research, Colney, Norwich, NR4 7UA, UK
[d] IRAS, Yalelaan 2, 3584 CM Utrecht, The Netherlands

International Journal of Food Microbiology 2010; 139: 57–63

International Journal of Food Microbiology 139 (2010) S57–S63

Contents lists available at ScienceDirect

International Journal of Food Microbiology

journal homepage: www.elsevier.com/locate/ijfoodmicro

Strengths and weaknesses of Monte Carlo simulation models and Bayesian belief networks in microbial risk assessment

J.H. Smid [a,d,*], D. Verloo [b], G.C. Barker [c], A.H. Havelaar [a,d]

[a] RIVM, A. van Leeuwenhoeklaan 9, 3720 BA Bilthoven, The Netherlands
[b] European Food Safety Authority, Largo N. Palli 5/A, I43100 Parma, Italy
[c] Institute of Food Research, Colney, Norwich, NR4 7UA, UK
[d] IRAS, Yalelaan 2, 3584 CM Utrecht, The Netherlands

ARTICLE INFO

Keywords:
Microbial risk assessment
Bayesian belief network
Monte Carlo analysis
Food safety

ABSTRACT

We discuss different aspects of farm-to-fork risk assessment from a modelling perspective. Stochastic simulation models as they are presented today represent a mathematical representation of nature. In food safety risk assessment, a common modelling approach consists of a logic chain beginning at the source of the hazard and ending with the unwanted consequences of interest. This 'farm-to-fork' approach usually begins with the hazard on the farm, sometimes with different compartments presenting different parts of the production chain, and ends with the 'dose' received by the consumer or in case a dose response model is available the number of cases of illness. These models are typically implemented as Monte Carlo simulations, which are unidirectional in nature, and the link between statistics and simulation model is not interactive. A possible solution could be the use of Bayesian belief networks (BBNs) and this paper tries to discuss in an intuitive way the possibilities of using BBNs as an alternative for Monte Carlo modelling. An inventory is made of the strengths and weaknesses of both approaches, and an example is given showing an additional use of BBNs in biotracing problems.

1. Introduction

The backbone of every microbial risk model is the risk pathway and a common modelling approach consists of a logic chain beginning at the source of the hazard and ending with the unwanted consequences of interest. In food safety risk assessment, this approach is often referred to as the 'farm-to-fork' approach, and here the logic chain (and risk pathway) usually begins with the hazard on the farm, and ends with the 'dose' received by the consumer or the number of human illness cases through a dose response model. These models aim to answer questions like "what is the Log reduction of bacteria at the end of the chain when imposing an extra heating step in the chain?", or "given all steps in the food chain, what is the number of people getting ill?", or "given a failure of a decontamination step, what is the augmented risk of getting ill?". Examples of these approaches are (Alban and Stark, 2005; Havelaar et al., 2007; Hill et al., 2003; Ivanek et al., 2004; Sauli et al., 2005; Titus, 2007). An overview of the FAO/WHO-funded risk assessments can be found on (AGNS-Risk Assessments, 2009). A framework for all risk assessments, including QMRA, is well established (National Research Council, 2009).

Most current food chain models are based on the quantitative microbial risk assessment (QMRA) methodology. In general, QMRA methodology is based on the question of how microbial dynamics in a food chain evolve throughout the chain. Often many parameters are involved, describing process features such as time and temperature of decontamination steps.

The QMRA methodology typically splits up a food chain in a certain number of modules, using the Modular Process Risk Modelling (MPRM) paradigm. MPRM is derived from process risk modelling and was first documented in (Nauta, 2002). MPRM identifies the different stages of a process in a structured way, and it incorporates uncertainty and variability in the model. Variability represents a true inhomogeneity in a population, and is inherent to the population. On the other hand, uncertainty reflects our lack of knowledge of the exact value of a parameter (Vose, 2000).

The MPRM technique subdivides the food pathway into different stages called "modules", in which events called "processes" are explicitly modelled. These modules are sub-models in themselves and represent a well defined part of the real-world problem. Six processes exist which describe the dynamics of growth, inactivation, partitioning, cross-contamination, mixing and removal. Here the first two processes refer to bacterial dynamics, and the other four to food-handling. For a QMRA model aiming to describe the dynamics of bacteria in a food chain, the variable of most interest describes the number of bacteria in a unit food item, i.e. the *concentration* of the

* Corresponding author. RIVM, A. van Leeuwenhoeklaan 9, 3720 BA Bilthoven, The Netherlands. Tel.: +31 302744284.
E-mail address: joost.smid@rivm.nl (J.H. Smid).

0168-1605/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijfoodmicro.2009.12.015

bacteria. This variable is dependent on the variable describing the number of bacteria at a previous stage (i.e. a previous module), and various process parameters, like the temperature, time, acidity, etc. In each module one or more of the basic microbiological processes and processing stages of inactivation, growth, partitioning, mixing, removal and cross-contamination are modelled. These processes are defined by functional relations having the process parameters and the concentrations of the bacteria at a previous module as arguments of the functions. The final model is obtained by linking the modules, passing information from one module to the next.

We can implement these QMRA models, based on a MPRM technique, using two (and possibly more) distinct modelling techniques: unidirectional simulation models, most often based on a Monte Carlo (MC) sampling algorithm, and Bayesian belief network (BBN) models. The main difference is that the simulation techniques concentrate on marginal distributions as opposed to elements in the partition in a joint probability. It is important to notice that, although this paper focuses on BBNs such as defined in (Parsons et al., 2005; Uusitalo, 2006), most statements about BBNs made in this paper equally apply to hierarchical simulation models based on a Markov Chain Monte Carlo (MCMC) algorithm (Gelman et al., 2004), because also in MCMC models the Bayesian paradigm is the underlying one, and invariably they can be used to address the same problems.

The core of this paper is split up in three parts. First, we discuss MC simulation models. We start with a description of these models, and then give their strengths and weaknesses. Then we discuss BBNs following the same structure. An overview of the strengths and weaknesses of both approaches is given in Tables 1 and 2. We illustrate through an example how the BBNs have a possible additional use in farm-to-fork risk assessment models and in the last section we present a general discussion.

2. Monte Carlo simulation models

2.1. Description

Mathematically, in MC simulation models an output of interest y can be expressed as $y = f(x,\theta)$, where x denotes a vector of input values, θ a vector of parameter values and $f(\cdot)$ a set of nested functions working element wise on entries of x and θ. The parameters and input values are the independent variables that are involved in the model, and their values are usually prescribed by marginal distributions reflecting the uncertainty and/or the variability of the parameter. They are estimated from data by statistical analyses of epidemiological or experimental data and/or from expert opinions. Two main frameworks for data analysis can be considered. In the frequentist framework, whether parametric or non-parametric, parameters are estimated based solely on the available data. In contrast, by Bayesian parameter estimation prior knowledge can be incorporated in the statistical analysis and used together with actual data. This prior knowledge could be based on historical data or could also be expert opinion. Although Bayesian parameter estimation is a natural element in BBN methodology, it may be used in all farm-to-fork models.

Table 1
Strengths and weaknesses of Monte Carlo simulation models.

Unidirectional simulation models	
Strengths	Weaknesses
Can model complex systems (including coupled degrees of freedom)	No interactive link between data and parameters
All kinds of probability distributions can be modelled	Approximate technique (because simulation-based)
Intuitive and relatively easy to implement	Unidirectional
Time to results reasonably short	Validation often ad hoc (by means of resampling)
Widely accepted and used	Does not allow "backwards reasoning" (see biotracing example)

Table 2
Strengths and weaknesses of Bayesian belief networks.

Bayesian belief networks	
Strengths	Weaknesses
Interaction data and parameters	Size network is restricted (restriction on storage of joint distribution)
Natural framework for inclusion of expert opinions	Simulation or discretization is needed
Natural framework for updating parameters continuously	Construction of complex conditional probabilities may be computationally intensive
Model validation relatively simple	
Graphical user interface makes it straightforward to use for non-experts	
Allows "backwards reasoning" (see biotracing example)	

The dynamics of the dependent variables is given by $f(\cdot)$ and written in terms of the known inputs and estimated parameters. The margins of the dependent variables are (most often) simulated via a MC modelling framework (Vose, 2000). The basic idea behind an MC procedure is that the simulation is iteratively refined during a number of iterations, giving feasible representations of a real-world situation. At each iteration, numbers are drawn from the probability distributions representing variability. The number of iterations is preset, or dependent on the convergence behaviour of the simulation. The results of each iteration are stored. In a 2D simulation also uncertainty is simulated. At the end, statistical methods may be employed in order to extract summary data from the results. For example, means, standard deviations or percentiles may be insightful to the risk assessor or risk manager. Equally correlations between input values and output variables of interest (such as can be quantified using automated sensitivity analysis routines in stochastic simulation software) can provide useful insight. Note that any model inputs that are considered not to be variable or uncertain, are assumed to be included in the definition of $f(\cdot)$.

MC models quantify, or predict, the fate of microbes in products as a function of time, which (as far as we know it in practice) only ever moves forward. Substantially this means a simulation of time ordered events; for this reason we may refer to this type of models as "forward" models. However not a unique property of MC models, they allow for a calculation of risk reductions due to interventions in the chain, which can be simulated by changes of the parameter values.

2.2. Strengths

MC methods are useful for modelling phenomena with significant uncertainty and variability in inputs, which we assume is the case in most farm-to-fork models. These methods are especially useful in studying systems with a large number of coupled degrees of freedom, i.e. correlation and other inter-dependencies are modelled in a very natural manner. This means that a very complex domain can be modelled by using MC simulations. Because of the great flexibility of a simulation approach, it is for example possible to model individual food products in the production chain. The major strength of the simulation approach is that these models are relatively easy to implement using commercial software. The distributions of the model's dependent variables do not have to be approximated in any way, because they are obtained by simulations. These models are widely accepted: currently several commercially available software packages exist to automate the tasks involved in the simulation, and most of them yield satisfying results reasonably rapidly (Robert and Casella, 2004; Vose, 2000).

2.3. Weaknesses

Basically in MC farm-to-fork modelling a probabilistic/stochastic model is set up describing the process from the beginning to the end of those steps that the model considers. To assess the uncertainty and variability of the parameters used in the simulation model statistics are used on available data. Both processes (the simulation model and the statistics) are only related in the sense that what comes out of the statistical data analysis is used as an input in the simulation model. Technically this amounts to a limited incorporation of complex dependency. The relationship is unidirectional and data available for estimating a parameter in point A of the model will not have an influence on the parameter in a subsequent point B of the model which was estimated from other data, hence there is no interaction between the different parts of the model. The following simple example will be used for illustration. At the farm level module of a simulation model the data available on the diagnostic accuracy of a test shows that in a small test validation trial the diagnostic test had a sensitivity of 100% (the test was positive for all 10 diseased animals tested), while the specificity was 80% (8 of the 10 non diseased animals tested negative, the other 2 appeared to be false positive). These values are used as the input for sensitivity and specificity in the simulation model. Suppose now that, for the estimation of the observed prevalence parameter needed in the stochastic model, survey data is available where a big number of animals (>1000) were tested leading to an observed prevalence of 1%. To simulate in the model the number of diseased animals out of *N*, one could assess the true prevalence as a function of the observed prevalence and the sensitivity and specificity of the diagnostic tests. However, due to the fact that based on the specificity an average of 20% false positives is expected while the survey data does not confirm this, the straightforward calculation of the true prevalence is not possible (or would lead to negative true prevalence estimates). In an MC simulation model it is highly likely that the tests specificity will be ignored and the 1% observed prevalence will be used as the binomial parameter to simulate the number of diseased animals out of *N*.

In farm-to-fork modelling the idea of a network where one event is connected with another event is current practice. However an effect observed (data) at one stage of a network can have its origin (parameter) in a prior stage. Practically this means that data observed at a certain stage could be used to estimate parameters which are occurring 'before'. Coming back to the example above a possible explanation could be that the estimated test specificity was too low or the sensitivity was too high. As the above simulation model is only fed by the statistics no interaction is possible to adjust the sensitivity and specificity estimated based on the observed outcome of the survey.

In practice data is used in MC models as a means of validating the outputs. If necessary, the model parameters and/or the models will be modified to match observations. So although one might argue that the absence of interaction between the data and the statistics of the parameters results in a possible loss of information, the information in the data may still be used to update the parameter values. However, this is not straightforward and requires a careful analysis of the model equations and parameter sensitivity analyses. It is often an ad hoc approach, requiring many resamplings: first to find those parameters that need to be re-estimated, and then to find those parameter values that match with the observed data. Once the uncertain values of the parameters, which are the inputs of the simulation model, are re-estimated, often a re-sampling of the entire model has to be performed, which may be time consuming. Moreover, validation in these models is often arbitrary, because the acquired result may be explained by several combinations of parameter values. Note that in a deterministic model tools like Solver or Goal Seek could be used to determine what initial (unknown) conditions lead to the later (observed) state of the system being modeled, under the assumed or known causal relationships in the model.

MC simulation models are constructed to answer questions in only one direction, namely the *forward* direction of time or chain processing order. They aim to infer statements about future processes or variables (i.e. $y = f(x,\theta)$) that are effects of the quantities that can occur earlier (i.e. draws from distributions x and θ). MC modelling approaches cannot easily yield answers targeted at unobserved variables at the beginning of the chain based on downstream observations (data), since not all relations between the variables can be inverted. Only marginal distributions of the variables are obtained after sampling, and not the joint distribution of the model, which compromises making probabilistic inferences. Further on in this paper we discuss an example of a situation where backward reasoning in the chain is explicitly needed.

Note that, although MC is often criticised as being an approximate technique, in theory at least, any required level of precision can be achieved by simply increasing the number of iterations in a simulation (Vose, 2000).

3. Bayesian networks

3.1. Description

A Bayesian belief network (BBN, or a belief network) is a compact model representation for reasoning under uncertainty. It is a probabilistic graphical model that represents a set of variables and their probabilistic dependencies. A diagrammatic network structure is used which partitions the joint probability $P(x,\theta,y)$ onto a directed acyclic graph. In this graphical network representation, variables are represented by nodes, and the relations that relate the variables to each other are represented by directed arcs. Quantitatively, these relations are expressed in conditional probability tables (CPTs). The whole set of variables and relations between these (i.e. the joint distribution) is represented as a directed acyclic graph. This is a succinct representation of the dependency between relevant variables and, as such, a powerful complexity reduction of the multivariate statistical problem. Note that the graphical structure is not a unique property of BBNs. Analytica (www.lumina.com) is an MC simulation software that uses a graphical interface for model construction and representation.

The methods to construct the joint probability are based, to a large extent, on beliefs concerning causality (Pearl, 2000). For example models assume that heat processes and chemical conditions cause changes in the size of the pathogen population and that consumption of large concentrations increase the probability of illness in the consumer. The models also express the quality of the information content by including parameter uncertainty and population variability. This inclusion aids the quantification of sensitivities and confidence in the outcomes of the assessment.

A BBN uses Bayes' rule to compute the posterior distribution of variables given evidence, which is termed as probabilistic inference. BBNs have some attractive features that have made them popular in a wide range of applications (e.g. finance, banking, medicine, robotics, civil engineering, geology, geography, genetics, forensic science and ecology). However apart from some publications, e.g. (Barker et al., 2005, 2002; Delignette-Muller et al., 2006; Ranta and Maijala, 2002), until now the use of Bayesian approach in QMRA has been fairly limited to make inference on uncertainty distributions of separate input parameters. In most of the cases this is done in a straightforward manner by fitting, for example, beta distributions to model the uncertainty of a proportion.

One can choose to model the whole domain in a BBN after making some assumptions, often involving simplifications of the domain model, i.e. the connected set of variables and the relationships between them (Barker, 2004). The marginal distributions that are formulated to construct the MPRM model, can be attributed to the independent parent nodes in the BBN, and distributions for dependent

nodes can be implemented by conditioning on the independent ones via the MPRM causal relations. This way, the MPRM model defines the BBN structure and quantifies the relations between the variables.

In a BBN a parameter could be updated though different data sources occurring before or after in the chain. Several softwares exist for computing the marginal distributions before and after making inferential queries, but they all do this via the joint distribution. Storage of a joint distribution is computationally complex and is done via secondary computational structures, called 'junction trees' (Lauritzen and Spiegelhalter, 1988). Several software packages for decision support systems have been developed for this task and can be found on the website of Murphy (2005), or in Uusitalo (2006). BBNs represent one branch of Bayesian modelling, the other major approach being hierarchical simulation-based modelling in which the full joint distribution is sampled using computationally intensive Markov Chain Monte Carlo (MCMC) algorithms (Gilks et al., 1996), e.g. WinBUGS (Lunn et al., 2000). Note that the only major distinction between network approaches and MCMC approaches concerns the way in which components of the joint probability are estimated and stored. MCMC models are also based on the Bayesian paradigm, and may also be presented in terms of a graphical structure (e.g. the Doodle menu in WinBUGS). Unless indicated otherwise, the strengths and weaknesses described below equally apply to MCMC-based models.

3.2. Strengths

In a BBN the joint distribution of all possible values of the parameters, the observable and the non-observable variables are included. Via Bayes' rule this joint distribution is instantaneously updated when new data for the observable variables becomes available. This implies that in a BBN interaction between the statistics of the parameters and the simulation model is possible, unlike in MC farm-to-fork models (Barker et al., 2002). We come back to our simple example. In a BBN the specificity and sensitivity network parameters occurring (in the model) before the parameter obtained from the survey data would both be updated by the survey data while the estimated true prevalence now will be based on those updated values of sensitivity and specificity and the survey data (sensitivity would decrease while specificity would increase and the posterior of the true prevalence would be on average higher than its prior, however in this case this could have been a non-informative prior). It is obvious that this prevalence estimate could also be influenced by prevalence or other data concerning later parts of production chain which also could be taken into account in the BBN. By a consistent application of Bayes' rule in a directed graph, multiple pieces of evidence (observations on variables elsewhere in the network) can be used to update the statistics of the parameters of the model. Hence, a BBN is a natural framework for combining results of a risk assessment with results from epidemiological studies.

The opportunity to use probability distributions as subjective representations of belief, is appealing, especially if data is scarce. If proper elicitation techniques are applied (Cooke and Goossens, 2000; O'Hagan, 2007), expert opinions can be included in the model. The synthesizing of multiple sources of information may reduce uncertainties, but of course one should be cautious with including subjective probabilities, and strive to do it in a formal and consistent way. It is important to recognize however that the subjective probabilistic framework does not give the analyst complete freedom. The subjective probabilities are constrained to behave according to the laws of probability and these can be fairly strong constraints (Siu and Kelly, 1998).

BBNs are in fact complicated statistical models. For 'relatively simple' BBNs techniques are available to validate the model (how does the model describe the data) and to do model comparison (which is the simplest model describing the data the best). In a BBN rapid algorithms for sensitivity analyses (SA) can be implemented, which need the joint probability for calculations (Coupé and Gaag, 2002). One can think of evidence SA, investigating how sensitive the results of a belief update (propagation of evidence) is to variations in the set of evidence (observations, likelihood, etc.), or of parameter SA, which is the analysis of how sensitive the results of a belief update (propagation of evidence) is to variations of the value of a parameter of the model. In many BBN softwares these algorithms are directly implemented as automated routines.

BBNs often describe a very complicated domain, namely that of the joint distribution over all variables, but the way in which the variables are represented (in a network-form including direct causal relations) is rather intuitive. This makes it easy for non-experts to understand and build this kind of knowledge representation. The role as a communication device is a particular strength of BBNs — non-modellers can see "explicit reasoning" as opposed to numerical assertion. In MCMC-based software this advantage may be minor because it is less common to use a network representation in these models.

Like in MC models, also the causal structure of BBNs provides a strong link to predictive microbiology, a field in which conditional dependencies are often experimentally assessed (e.g. inactivation rate varying with temperature of decontamination). This is a big advantage in QMRA. Due to the interactive nature BBNs are also flexible to integrate risk assessment with economics (risks with benefits) to move from risk assessment to decision analysis in food safety. This would include the different factors of the decision making process into the model. Practically this means inclusion in the model of an appropriate utility function as well as the risk management actions that could be taken, modelled by an explicit decision node. Once the network is constructed, determining the conditional distribution of relevant parameters and outcomes given a certain decision is straightforward (Jensen, 2001). Inclusions of these aspects into the model would increase the transparency of the decision making process and optimize the model as a decision making tool for risk managers. However, this approach may contrast with the current interpretation on the separation between risk assessment and risk management. For an example of an automated decision making process using BBNs see the website of Fernandes and Barker (2009).

BBNs have a strong statistical foundation and, as such, are easily linked with progressive statistical techniques such as machine learning. Although many of the original network implementations of causal networks were based on prescribed structures, modern constructions rely more on automated learning from data; both structural learning and parameter learning from case data bases. An example of a learned and validated network, for hazards associated with bovine mastitis is given by Steeneveld et al. (2009). BBNs have the potential to be very dynamic models, also suitable for problems where reasoning backward in the chain is required, such as in the example further on in this paper.

3.3. Weaknesses

In farm-to-fork models often variables have a continuous domain. When using BBN tools based on the Lauritzen–Spiegelhalter (LS) algorithm, these variables have to be discretized, because this algorithm cannot (yet) handle continuous probability distributions different from Gaussian or mixtures of Gaussians. After discretizing a possibly less realistic model is obtained, for a practical example see Parsons et al. (2005). This is especially important for problems where the distribution tails are critical, as it is difficult to accurately model these with discrete distributions. Also, it is hard to discretize in any other way than in an ad hoc manner, which may yield inaccurate results, given the combined effect of approximations for large numbers of variables (Uusitalo, 2006). It is difficult to check this overall effect in practice. Furthermore, discovering appropriate discretization is time consuming and no optimal discretization methods for BBNs have been found (Uusitalo, 2006). In theory, any

required level of precision can be achieved by simply increasing the number of intervals for discretization, but in practice computational problems might arise when we use large CPTs (i.e. as an effect of discretization) or dense junction trees (i.e. as a result of large complexities). Note that this is only problematic during the construction phase of the model; once it is built, updating the joint probability is (almost) instantaneous.

When using MCMC methods one does not need to divide continuous variables into discrete ranges because these methods use numerical integration to approximate probability distributions. However the numerical approximations that are obtained may be dependent on the convergence properties of the sampling algorithm. A method for determining the number of iterations required for convergence can be found in Brooks (1998). Also here any level of precision can be achieved by increasing the number of iterations in a simulation.

Cross-contamination is a microbiological process occurring in a food chain that is often computer-intensive to simulate. In a certain stage in a food chain an individual item out of a population might be contaminated and contaminate other individuals from that population, either directly or indirectly. E.g. a carcass in a slaughterhouse might transfer contamination to a piece of slaughter equipment, which in its turn contaminates a subsequent carcass at the same stage in the slaughter procedure. This process can be modelled as $y_k = f(x_k, \theta, y_{k-1})$, where x_k is the number of bacteria on carcass k before this specific stage in the chain, y_k is the number of bacteria on carcass k after this stage and θ is some process parameter. In a MC framework these dependencies between subsequent carcasses can be modelled quite easily, e.g. see (Nauta et al., 2005). In a BBN this cannot be done so easily because if a model includes feedback then in principle it cannot be represented in an acyclic structure. To put the recursive relations in a BBN would require a dynamic BBN (Soudant et al., 1997), which explicitly captures the dynamics over time. It might grow very large because k could be very large (in the example, in one day, up to 5000 pigs could be slaughtered).

The cross-contamination example indicates that with the software available today, it is not realistic that a complex farm-to-fork chain could be modelled as one gigantic BBN, using the same causal relations as would be used when modelling the chain in a unidirectional MC based approach. The way forwards chosen by most BBN models is to split up the model in a set of models parametrized by time (time slices), in which every time slice represents the state of the variables at a certain moment. The inputs of a time slice are the states of the variables at a previous moment and the outputs those states at the next moment. This gives a new representation in each time slice — clearly the level of coarse graining time that is appropriate determines practical models. It is interesting to note that Object Oriented net software has been built for precisely this reason (most BBNs used for DNA sequencing employ this time sliced structure). Note that in MCMC-based models recursive relations can be modeled more easily by implementing *for*-loops in the computer code.

4. Example: a biotracing network of growth in a food chain

In this example we discuss a *biotracing* model, where we define *biotraceability* as the ability to use down stream information to point to materials, processes or actions within a particular food chain that can be identified as the source of "undesirable agents" (Barker et al., 2009). In these non-typical applications of a farm-to-fork model the purpose is to recover the origins of food-borne hazards. A model is then needed that can answer questions in the reverse direction of the chain processing order. It is preferable to use a model that captures the joint probability distribution over all variables, because the food chain may include many uncertainties and variabilities. BBN models could fulfill this need. We show this possible additional use of BBNs in farm-to-fork risk assessment, and the relationship between domain modelling and biotraceability. The model uses the BBN software HUGIN (HUGIN Expert A.S., Aalborg, Denmark).

We have five sequential events (processes or time steps) in a food chain ($t = 0,\ldots,4$); every time step is characterized by a population-distribution of a pathogen ("LogN0" to "LogN4", the logarithms of the pathogen populations sizes at each step). At every time step, we have a potential source of contamination (Source 0 to 4), each with prior probability of contamination of 0.5%. In this example the pathogen concentration that every source transfers is not variable and always equal to -1 Log. Population sizes only depend on the previous population size, the uncertain growth factor and the activity of the current source. In the diagram (see Fig. 1), dependency-relations are expressed as directed arcs. For example, the number of pathogens at $t = 1$ ("LogN1") depends on the numbers that were present at $t = 0$ ("LogN0"), on a potential (0.5%) contributing source of contamination (Source 1) and on how fast the pathogen can grow at that step (growth-factor). In this form the model could be used to compute the probability (a belief) for "LogN4", however the whole joint probability is included in the network. We use the model to find the source of contamination with the greatest posterior probability, given an observed concentration of the pathogen at the final stage of the food chain, i.e. given an observed value of variable "LogN4".

The expression for the joint probability, which encapsulates the risk assessment process, indicates a route to biotraceability information. The conditional probability included in the risk assessment is the reverse conditional for a biotrace. In principle Bayes' theorem can be used to connect these two conditional probabilities and hence can be used to convert risk assessment into biotracing. That is the reverse conditional probability, in combination with evidence from observed events and prior information about sources and chain variables, can be used to compute posterior beliefs regarding the sources.

When we enter the observation that the contamination at the last time step, "LogN4", is between 1.25 and 1.5 Log, the information propagation computed by the BBN software updates the beliefs on the initial contamination by Bayesian inversion. In Fig. 1 it can be seen that the contamination probably occurred at stage 2 in the chain; the posterior probability that contamination came from Source 2 is 94.96%. This model includes an option for the precision of knowledge of the growth factor (variable: "Precision"). In the example above we have assumed that this knowledge is very certain, the state of the Precision variable can be set to high. On the other hand when we change the "Precision" variable to low, so that the amount of population growth attributed to the individual steps along the chain is less well defined, then the ability of the model to lead to a definite inference (a dominant source) is reduced: in this case Bayesian inversion includes significant probabilities for at least three of the five possible sources, see Fig. 2.

In this example the population size is a very informative marker of the contamination source, provided that the growth factor is sufficiently well known. Once the transmission of information includes uncertainties, the inference concerning the source is compromised. This example is sufficient to conclude that for QMRA to be converted into biotracing a reduction of the uncertainties is crucial; incorporation of additional information streams, such as biomarker information, is one scheme for compensating for information uncertainty by adding additional information to the model. Although in this example we assumed that there was no variability, note that in general (irreducible) variability of variables compromises inference concerning the source, hence a deterministic model is often not sufficient for biotracing purposes.

5. Discussion

Prior to constructing any model, its purpose should be specified. Although in some cases the model may provide answers to questions outside of its primary scope, in general this is not the case. We have

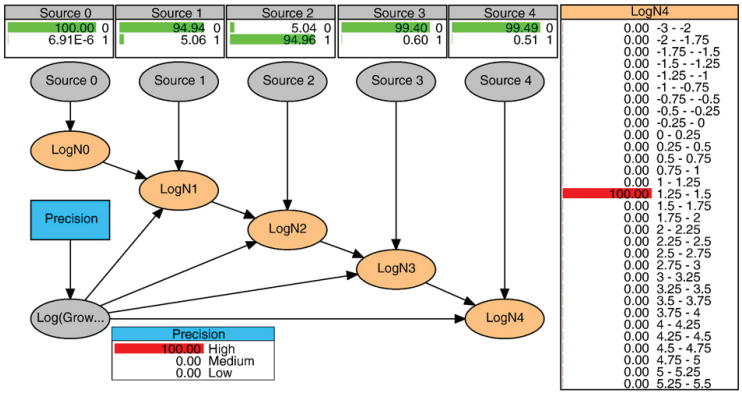


Fig. 1. Bayesian inversion, given an adequate knowledge of population growth, enables a precise trace of contamination.

described two approaches for farm-to-fork modelling: unidirectional MC simulations and BBNs. We have given strengths and weaknesses of both methods, indicating the purposes for which the models can be used. Note that we have not given boundary conditions of these approaches by defining strengths and weaknesses: strengths imply possibilities but weaknesses do not imply impossibilities, only impracticalities. For example, the values of the parameters may be validated based on data in the chain using a MC model of a food chain, but this is more straightforward using a BBN. One may argue that the difference between the two modelling methods is one of method of implementation. However we have shown that MC models concentrate on marginal distributions whereas BBNs concentrate on elements in the partition in a joint probability, from which marginal distributions can be derived. This means that if the same dynamics of pathogens in a food chain is implemented in both approaches, then the approach using BBNs is really more informative.

However, when translating a MC model into a BBN, often approximations have to be made to prevent complex conditional probabilities which may result in computational problems. This is especially true if the modeller has to deal with (many) continuous variables (complex domain), a lot of coupled degrees of freedom (correlations) and/or recursive relations (time dependency), all three often being the case in farm-to-fork models which are complex by nature. For an example of a risk assessment study where an MC model is translated into a BBN where some of these assumptions are highlighted, see Pouillot et al. (2004). In that study approximations were made in an ad hoc manner; an open question is how to make these in a more structured way.

A possible approach to overcome computational problems is to use different modules, each or some modules being a BBN which are connected to each other by 'simple' MC simulation. Data available within a specific BBN compartment would influence the relevant parameters within the compartment but only influence the outcome of the other compartments at later times. It should be noted that a pure beginning to end simulation could also be conducted over a BBN (for example to illustrate the effects of different scenarios). In O'Hagan, (2004) a so-called "emulator" is described, which is essentially a statistical model of the model. Typically it is much simpler than the original model. As a result, this approach is particularly advantageous for understanding uncertainty in large and complex models of the type commonly encountered in QMRA. Possibly this approach could also generate a framework for structurally approximating a farm-to-fork model, for implementation into a BBN.

We recommend a tiered, fit-for-purpose approach. First, the modeller should assess the purpose of the model, and then depending on this purpose, the modelling approach should be chosen: in some cases this is a simulation model, in other cases a BBN. A similar conclusion was drawn in Zwietering (2009), where it was advocated that the amount of complexity of a model should be seen as a function of needs and prior knowledge.

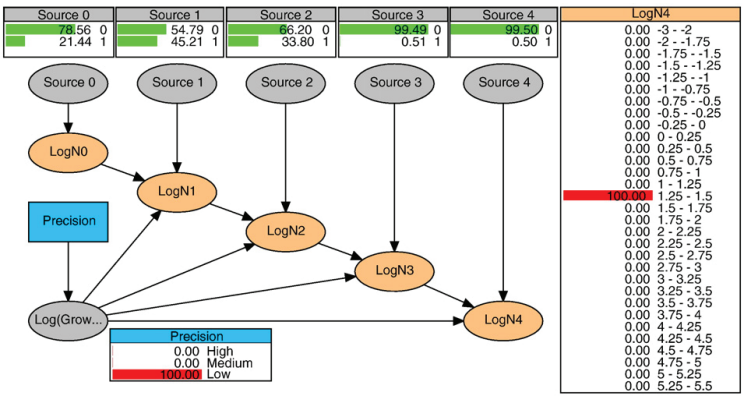


Fig. 2. If knowledge about population growth is uncertain, the trace is compromised.

However, we have shown that in general the use of BBNs in farm-to-fork risk assessment modelling has some clear advantages above using MC models. An important one is the possibility of incorporating independent data from validation protocols or epidemiological studies to update uncertainty and variability distributions of the parameters in an interactive manner, which is less straightforward in MC models, and is often done arbitrarily. In Latimer et al. (2008) an MC model of *Salmonela* spp. in egg products was described. The term "anchoring" was introduced to indicate the process of adjusting the results from the risk assessment to fit epidemiologic surveillance data, by scaling all outputs of the MC model (for any assumed scenarios) with a fixed ratio. Because it is not likely that the (combination of) parameter estimates have a linear effect on the output, such an approach may be regarded as arbitrary and erroneous. The advantages of BBNs for validation with data were also shown in Pouillot et al. (2003) and Delignette-Muller et al. (2006), where the Bayesian approach was suggested for the estimation of uncertainty and variability in bacterial growth.

By means of an example we have shown an additional use of BBNs in biotracing problems, but more applications could be thought of, such as the definition of more realistic Food Safety Objectives (FSO) and Performances Objectives (PO). Downstream data in a food chain must reflect these Objectives, and by parameter sensitivity analysis the influence of parameters leading to the FSOs and POs can be studied. An example illustrating this can be found in Cox and Popken (2004) where it was shown that health risks from antimicrobial resistance through use of antibiotics in food animals are highly sensitive to the human prescription rate of similar antibiotics. BBNs constitute a natural framework for doing these sensitivity analyses because in most BBN software automated routines are implemented to do this.

Acknowledgement

The work was supported by the European Union funded Integrated Project BIOTRACER (contract 036272) under the 6th RTD Framework.

References

AGNS-Risk Assessments, 2009. Links to JEMRA pathogen–commodity risk assessment pages. http://www.fao.org/ag/agn/agns/jemra_riskassessment_en.asp.

Alban, L., Stark, K.D.C., 2005. Where should the effort be put to reduce the *Salmonella* prevalence in the slaughtered swine carcass effectively? Preventive Veterinary Medicine 68, 63–79.

Barker, G.C., 2004. Application of Bayesian belief network models to food safety science. Bayesian Statistics and Quality Modelling in the Agro-Food Production Chain. Kluwer Academic Publishers, Wageningen, pp. 117–130.

Barker, G.C., Talbot, N.L.C., Peck, M.W., 2002. Risk assessment for *Clostridium botulinum*: a network approach. International Biodeterioration and Biodegradation 50, 45–47.

Barker, G.C., Malakar, P.K., Del Torre, M., Stecchini, M.L., Peck, M.W., 2005. Probabilistic representation of the exposure of consumers to *Clostridium botulinum* neurotoxin in a minimally processed potato product. International Journal of Food Microbiology 100, 345–357.

Barker, G.C., Gomez, N., Smid, J., 2009. An introduction to biotracing in food chain systems. Trends in Food Science and Technology 20, 220–226.

Brooks, S., 1998. Markov Chain Monte Carlo method and its application. Statistician 47, 69–100.

Cooke, R.M., Goossens, L.H.J., 2000. Procedures guide for structured expert judgement. Radiation Protection Dosimetry 90, 303–309.

Coupé, V.M.H., Gaag, L.C.v.d., 2002. Properties of sensitivity analysis of Bayesian belief networks. Annals of Mathematics and Artificial Intelligence 36, 323–356.

Cox Jr., L.A., Popken, D.A., 2004. Bayesian Monte Carlo uncertainty analysis of human health risks from animal antimicrobial use in a dynamic model of emerging resistance. Risk Analysis 24, 1153–1164.

Delignette-Muller, M.L., Cornu, M., Pouillot, R., Denis, J.B., 2006. Use of Bayesian modelling in risk assessment: application to growth of *Listeria monocytogenes* and food flora in cold-smoked salmon. International Journal of Food Microbiology 106, 195–208.

Fernandes, M. and Barker, G.C., 2009. Institute of Food Research, Norwich. http://bbn.ifrn.bbsrc.ac.uk:8080/gary/bbnjp1.html.

Gelman, A., Carlin, J., Stern, H., Rubin, D., 2004. Bayesian Data Analysis. Chapman and Hall/CRC.

Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1996. Markov Chain Monte Carlo in practice: interdisciplinary statistics. Chapman&Hall/CRC, Boca Raton, Florida.

Havelaar, A.H., Mangen, M.J., de Koeijer, A.A., Bogaardt, M.J., Evers, E.G., Jacobs-Reitsma, W.F., van Pelt, W., Wagenaar, J.A., de Wit, G.A., van der Zee, H., Nauta, M.J., 2007. Effectiveness and efficiency of controlling *Campylobacter* on broiler chicken meat. Risk Analysis 27, 831–844.

Hill, A., England, T., Snary, E., Cook, A., Kelly, L., Evans, S., Wooldridge, M., 2003. A 'Farm-to-Consumption' Risk Assessment for *Salmonella Typhimurium* in Pigs, Final Report. Department of Risk Research, Veterinary Laboratories Agency, Weybridge.

Ivanek, R., Snary, E.L., Cook, A.J.C., Grohn, Y.T., 2004. A mathematical model for the transmission of *Salmonella typhimurium* within a grower–finisher pig herd in Great Britain. Journal of Food Protection 67, 2403–2409.

Jensen, F.V., 2001. Bayesian Networks and Decision Graphs. Springer-Verlag, New York.

Latimer, H.K., Marks, H.M., Coleman, M.E., Schlosser, W.D., Golden, N.J., Ebel, E.D., Kause, J., Schroeder, C.M., 2008. Evaluating the effectiveness of pasteurization for reducing human illnesses from *Salmonella* spp. in egg products: results of a quantitative risk assessment. Foodborne Pathogenic Diseases 5, 59–68.

Lauritzen, S.L., Spiegelhalter, D.J., 1988. Local computations with probabilities on graphical structures and their application to expert systems. Journal of the Royal Statistical Society 50, 157–224.

Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D., 2000. WinBUGS — a Bayesian modelling framework: concepts, structure, and extensibility. Statistics and Computing 10, 325–337.

Murphy, K., 2005. Department of Computer Science, University of British Colombia. http://people.cs.ubc.ca/ murphyk/Bayes/bnsoft.html.

National Research Council, 2009. Science and Decisions: Advancing Risk Assessment. The National Academies Press, Washington.

Nauta, M.J., 2002. Modelling bacterial growth in quantitative microbiological risk assessment: is it possible? International Journal of Food Microbiology 73, 297–304.

Nauta, M., van der Fels-Klerx, I., Havelaar, A., 2005. A poultry-processing model for quantitative microbiological risk assessment. Risk Analysis 25, 85–98.

O'Hagan, A., 2004. Bayesian analysis of computer code outputs: a tutorial. Reliability Engineering and System Safety 91, 1290–1300.

O'Hagan, A., 2007. Research in elicitation. In: Upadhyay, S.K., Singh, U., Dey, D.K. (Eds.), Bayesian Statistics and its Applications. Anamaya Publishers, New Delhi, pp. 375–382.

Parsons, D.J., Orton, T.G., D'Souza, J., Moore, A., Jones, R., Dodd, C.E., 2005. A comparison of three modelling approaches for quantitative risk assessment using the case study of *Salmonella* spp. in poultry meat. International Journal of Food Microbiology 98, 35–51.

Pearl, J., 2000. Causality: Models Reasoning and Inference. Cambridge University Press, Cambridge.

Pouillot, R., Albert, I., Cornu, M., Denis, J.B., 2003. Estimation of uncertainty and variability in bacterial growth using Bayesian inference. Application to *Listeria monocytogenes*. International Journal of Food Microbiology 81, 87–104.

Pouillot, R., Beaudeau, P., Denis, J.B., Derouin, F., 2004. A quantitative risk assessment of waterborne cryptosporidiosis in France using second-order Monte Carlo simulation. Risk Analysis 24, 1–17.

Ranta, J., Maijala, R., 2002. A probabilistic transmission model of *Salmonella* in the primary broiler production chain. Risk Analysis 22, 47–58.

Robert, C.P., Casella, G., 2004. Monte Carlo Statistical Methods. Springer-Verlag, New York.

Sauli, I., Danuser, J., Geeraerd, A.H., Van Impe, J.F., Rufenacht, J., Bissig Choisat, B., Wenk, C., Stark, K.D.C., 2005. Estimating the probability and level of contamination with *Salmonella* of feed for finishing pigs produced in Switzerland — the impact of the production pathway. International Journal of Food Microbiology 100, 289–310.

Siu, O.N., Kelly, D.L., 1998. Bayesian parameter estimation in probabilistic risk assessment. Reliability Engineering and System Safety 62, 89–116.

Soudant, D., Beliaeff, B., Thomas, G., 1997. Dynamic linear Bayesian models in phytoplankton ecology. Ecological Modelling 99, 161–169.

Steeneveld, W., van der Gaag, L.C., Barkema, H.W., Hogeveen, H., 2009. Providing probability distributions for the causal pathogen of clinical mastitis using naive Bayesian networks. Journal of Dairy Science 92, 2598–2609.

Titus, S., 2007. A novel model developed for Quantitative Microbial Risk Assessment in the pork food chain (Thesis). Massey University, Palmerston North, New Zealand

Uusitalo, L., 2006. Advantages and challenges of Bayesian networks in environmental modelling. Ecological Modelling 203, 312–318.

Vose, D., 2000. Risk Analysis, A Quantitative Guide. John Wiley and Sons, Chichester.

Zwietering, M.H., 2009. Quantitative risk assessment: is more complex always better? Simple is not stupid and complex is not always more correct. International Journal of Food Microbiology 134, 57–62.

Chapter 3

A practical framework for the construction of a biotracing model: application to *Salmonella* in the pork slaughter chain

J. H. Smid [a,b], A. N. Swart [a], A. H. Havelaar [a,b] and A. Pielaat [a]

[a] RIVM, A. van Leeuwenhoeklaan 9, 3720 BA Bilthoven, The Netherlands
[b] IRAS, Yalelaan 2, 3584 CM Utrecht, The Netherlands

Risk Analysis 2011; 31(9): 1434-1450

Risk Analysis, Vol. 31, No. 9, 2011 DOI: 10.1111/j.1539-6924.2011.01591.x

A Practical Framework for the Construction of a Biotracing Model: Application to *Salmonella* in the Pork Slaughter Chain

J. H. Smid,[1,2,*] A. N. Swart,[1] A. H. Havelaar,[1,2] and A. Pielaat[1]

A novel purpose of the use of mathematical models in quantitative microbial risk assessment (QMRA) is to identify the sources of microbial contamination in a food chain (i.e., *biotracing*). In this article we propose a framework for the construction of a biotracing model, eventually to be used in industrial food production chains where discrete numbers of products are processed that may be contaminated by a multitude of sources. The framework consists of steps in which a Monte Carlo model, simulating sequential events in the chain following a modular process risk modeling (MPRM) approach, is converted to a Bayesian belief network (BBN). The resulting model provides a probabilistic quantification of concentrations of a pathogen throughout a production chain. A BBN allows for updating the parameters of the model based on observational data, and global parameter sensitivity analysis is readily performed in a BBN. Moreover, a BBN enables "backward reasoning" when downstream data are available and is therefore a natural framework for answering biotracing questions. The proposed framework is illustrated with a biotracing model of *Salmonella* in the pork slaughter chain, based on a recently published Monte Carlo simulation model. This model, implemented as a BBN, describes the dynamics of *Salmonella* in a Dutch slaughterhouse and enables finding the source of contamination of specific carcasses at the end of the chain.

KEY WORDS: Bayesian network; biotracing; pork; risk assessment; *Salmonella*

1. INTRODUCTION

In quantitative microbial risk assessment (QMRA), mathematical models are used to quantitatively describe the transmission of pathogens in production chains.[1–4] A modular process risk modeling (MPRM) approach[5] may then be used, by which the risk pathway is subdivided in a number of stages of processing (modules). The number of bacteria on or in a product unit after every module (the *output* of that module) is calculated as a function of the model inputs and of the number of bacteria on the product unit before the module. The final model is obtained by linking the modules, passing information from one module to the next. Because of variability in the bacterial population and its interaction with the matrix and because of the uncertainty about the true values of several quantities in the chain, in most cases a probabilistic structure is required. Monte Carlo (MC) sampling is typically used for iteratively refining probability distributions for the outputs of the model by generating samples from the input distributions of the model based on sequences of random numbers between 0 and

[1]Laboratory for Zoonoses and Environmental Microbiology, National Institute for Public Health and the Environment, Bilthoven, The Netherlands.
[2]Division of Veterinary Public Health, Institute for Risk Assessment Sciences, Utrecht University, Utrecht, The Netherlands.
[*]Address correspondence to Joost Smid, Laboratory for Zoonoses and Environmental Microbiology, National Institute for Public Health and the Environment, PO Box 1, 3720 BA, Bilthoven, The Netherlands; tel: +31 30 2744284; joost.smid@rivm.nl.

1 and sequentially propagating these through the model.

Typical purposes of QMRA models are predicting the number of human cases caused by bacteria that are introduced into the chain, or estimating the effectiveness of interventions in the production process. For example, during the EFSA QMRA on *Salmonella* in slaughter and breeder pigs,[6] an MC simulation model of the pork production chain has been constructed to predict the number of cases per year attributed to consumption of pork, and to investigate the effectiveness of interventions at the farm, transport, or slaughter level.

A more novel purpose of QMRA is to identify the sources of bacterial contamination. In this article we describe the construction of a *biotracing* (BT) model, where BT is defined as "the ability to use down stream information to point to materials, processes or actions within a particular food chain that can be identified as the source of undesirable agents."[7] Such a model needs to answer questions in the reverse direction of the chain processing order. In the context of our application, a relevant BT question is whether the source of an observed case of contamination of pork was *Salmonella* introduced via the farm or *Salmonella* introduced via the slaughterhouse equipment, given a downstream observation of the bacterial load on the carcass. A high load may, for example, suggest post-heat-treatment contamination or a failure of a decontamination step in the process, whereas a lower load suggests pre-heat-treatment contamination.

The BT inference is based on an information set and a belief system[8] that includes variability and the uncertainties about the scenario that might have led to the observed case of contamination. The information set includes (statistical) information about different variables that have been observed in the past. These not only may be process input variables, such as time and temperature of decontamination measures, but also microbial numbers at specific stages in the chain. The belief system includes knowledge of the chain, bacterial dynamics, and measurement characteristics of analytical methods, such as detection limits.

The purpose of this article is to convey a practical framework for the construction of a BT model of an industrial process that can be characterized by the following:

(1) a discrete number of product units are consecutively processed in the chain, which are possibly contaminated during processing;
(2) changes in the bacterial populations on these product units are described by known (but perhaps uncertain and variable) dynamics;
(3) multiple potential sources of contamination can be identified in the chain, of which some or all may act on the same product unit.

For simplicity we replace "product unit" by "product" in the remainder of the text, although strictly speaking a product unit may only be a part of a complete product. The third characteristic above implies that the bacterial population on the product may contain bacteria originating from different sources. This means that the concept of a unique source is not always existent. In this article we define a source as the *dominant source* if that source is responsible for the majority of bacteria on the product at the end of the processing line. The BT model proposed in this article is constructed to identify the dominant source for each individual product. The proposed framework is suitable for many food production chains, such as slaughtering, vegetable production, chocolate manufacturing, etc.

Partially, we follow the methodology described by Albert *et al.*,[9] who propose a two-step approach for including downstream data in risk assessment: the first step being devoted to construct a predictive tool without data (the core model), and the second step to improve this model by using downstream data, using a Bayesian approach. We apply these ideas to the specific purpose of BT. We illustrate the proposed framework by showing the exact steps we took to convert an MC simulation model of *Salmonella* in the pork slaughter chain, based on the EFSA QMRA on *Salmonella* in pork,[6] to a BT model.

2. THEORETICAL FRAMEWORK

The proposed methodology comprises seven successive steps, which are described in the following sections. Steps 1 and 2 describe the construction of an MC simulation model of a production chain. Steps 3–6 describe the conversion of this model into a BT model. It encapsulates the joint probability distribution over all variables, that is, the probability of multiple events occurring simultaneously, which can conveniently be implemented in a Bayesian belief network (BBN). In step 7, the actual use of the model for BT is demonstrated. The steps are summarized in Table I.

In this article we use the term *variables* for quantities in the model that may have varying values. *Input variables*, or simply inputs (e.g., times,

Table I. Description of the Steps to Construct a BT Model, the Actions for Achieving These Steps, and the Results Obtained After Every Step; SA = Sensitivity Analysis

Step	Description	Actions	Results
1. Domain model	Identify the processing steps in the chain and collect information about possible sources. Identify the BT question, collect information about the relevant bacterial dynamics in the chain, and identify data (gaps).	Literature research, field work, expert consultation	Set of information
2. MPRM model	Establish MPRM causal relations using the information from 1 and quantify the inputs and parameter distributions of the model.	Model construction	MPRM equations and input distributions
3. SA of MPRM modules	Run each module of the model to assess which input variables are redundant, that is, have limited impact on the outputs of each module. Remove those.	Model simulation	Reduced MPRM model
4. BBN	Implement relations into BBN and simulate recursive relations. Identify one or more triggers, and implement one or more report variables.	Model construction	BBN, generic tracing model
5. SA BBN	Do global sensitivity analysis to analyze which input variables have prior distributions that are influential on the distribution of the report variable.	Model testing	Insight in sensitivities inputs BBN
6. Learn parameters	Assign weights to uncertain parameters, expressing the level of knowledge, and use (downstream) data to update their probability distributions to be plant-specific.	Model construction	BBN with reduced uncertainty, plant-specific tracing model
7. Start biotracing	Insert new observations, and assign to each observation its dominant source.	Model using	Tracing estimates individual products

temperatures), are variables that have values that are not conditioned on the values of other variables. However, they can be correlated with each other, in which case a correlation structure needs to be enforced. *Parameters* (e.g., means, variances) are used to define the probability distribution of an input variable, representing variability. Parameter values can be fixed or be defined by a probability distribution. In the models described in this article, a distribution for a parameter reflects uncertainty about its true value. An input variable may in practice be equal to a parameter; then its distribution only represents uncertainty.

Data are defined as direct measurements of variables. Other types of information, such as expert knowledge or data summaries obtained from the literature, are not treated as data. Direct measurements of input variables are termed *upstream data*; direct measurements of the other, conditioned variables are termed *downstream data*.

2.1. Domain Model

Construction of a BT model commonly starts with identifying the stages in the production chain and the microbial processes occurring at these stages. An inventory is made of all likely sources through which the bacteria can be introduced into the chain. Dependent hereof, a precise BT question is formulated. The relevant BT task for the type of models described in this article is to find the dominant source of contamination for an individual product, for which downstream data, typically collected at the end of the chain (and possibly also upstream data) are available. The begin- and endpoints of the chain model are defined and a subset of the stages in the chain that is relevant for answering the BT question is identified. Information such as growth, inactivation, and cross-contamination rates is collected. Relevant are prevalence and number of bacteria at the different stages of production, but also characteristics of analytical methods, such as sensitivity, specificity, and detection limits, specifying the uncertainty of the system. All this information can be obtained using various sources such as literature, websites of equipment manufacturers, or expert opinion. The information set can eventually be organized into a *conceptual model*, which includes causality, and may be represented by a flow diagram. For example, Binter *et al.*[10] propose a conceptual model

describing the transmission of *Salmonella* in the feed chain.

2.2. MPRM Model

Next, an MPRM model of the chain is constructed as a series of mathematical equations. These equations are based on the MPRM basic processes,[5] describing the changes of the bacterial numbers on a product from one stage in the chain to the next. The equations, but not the complete model, are implemented in appropriate MC software. Note that the implementation of the complete model in MC software would be more complex and requires a two-dimensional framework of variability and uncertainty, separating these two phenomena.[11]

The MPRM basic processes defining the dynamics of bacteria are growth, inactivation, mixing, partitioning, removal, and cross-contamination. We define $N_{j,k}$ as the number of bacteria on the kth product after the jth stage of production (the *output* of that module). Let f_j be a positive function for growth and negative for inactivation, often depending on time and temperature. Then *growth* and *inactivation* can mathematically be described by:

$$\log(N_{j,k}) = \log(N_{j-1,k}) + f_j. \tag{1}$$

Removal that is proportional to the initial bacterial load on the product (e.g., washing) may also be described by Equation (1), in which case f_j is negative.

Mixing can be formulated as summing over the bacterial contributions $N_{j-1,k,1}, \ldots, N_{j-1,k,n}$ of the separate n raw materials that are the components of one product:

$$N_{j,k} = \sum_{i=1}^{n} N_{j-1,k,i}. \tag{2}$$

Partitioning is distributing $N_{j-1,k}$ bacteria on one product over n portions $N_{j,k,1}, \ldots, N_{j,k,n}$. If p_i is the relative size of portion i and if a homogeneous distribution of bacteria over the portions is assumed, then a multinomial distribution can be used to describe the variability between the numbers on the portions:

$$(N_{j,k,1}, \ldots, N_{j,k,n}) \sim \text{MultiNom}(N_{j-1,k}, p_1, \ldots, p_n). \tag{3}$$

If the interest is only to estimate the level of contamination of *one* specific portion and if equal-sized portions are assumed, then a Poisson distribution can be used:

$$N_{j,k,i} \sim \text{Pois}(N_{j-1,k}/n).$$

Cross-contamination can be direct or indirect transmission of cells from one product to one or many others, lowering the average bacterial load per product, but increasing the prevalence of contaminated products. Let $W_{j,k}$ be the number of bacteria in the production environment (e.g., a machine, hands of personnel) after processing product k during stage j, β_j the fraction of bacteria that is transferred from the product to the environment, and γ_j the fraction of bacteria that is transferred from the environment to the product during stage j. Nauta[5] suggested to implement cross-contamination by a system of recursive equations, linking numbers of bacteria on a product to numbers of bacteria on a subsequently processed product:

$$\begin{aligned} N_{j,k} &= (1-\beta_j)N_{j-1,k} + \gamma_j W_{j,k-1}, \\ W_{j,k} &= \beta_j N_{j-1,k} + (1-\gamma_j)W_{j,k-1}. \end{aligned} \tag{4}$$

2.3. Sensitivity Analysis (SA) of MPRM Modules

In this step, individual modules of the MPRM model are run using the implementation of the MC algorithms from Section 2.2 to investigate the sensitivity of the output of a module to changes in the value of the input variables associated with that module. An approach that may be used is nominal range SA (NRSA),[11] which is a local SA used to evaluate the effect of varying only one of the input variables of a module across its entire range of plausible values on the outputs of that module while holding all other input variables at their nominal or base-case values.[11] The results of NRSA are most valid when applied to a linear model. Conditional NRSA can be done for nonlinear input-output relations, in which NRSA is applied to different combinations of input values.[12]

If an output of a module is non- or negligibly sensitive to changes in the value of an input variable, then the probability distribution assigned to this variable can be substituted by a point value, for example, its mean. This is important in step 4, where the MC simulation model is converted to a BBN. In BBNs, complex conditional probabilities may result in computational problems,[13] and one way to reduce this complexity is by replacing probability distributions of input variables by their mean values. By an NRSA, a selection is made of input variables that may be replaced without significantly altering the shapes of the marginal output distributions of each module.

Much complexity in the model may stem from the equations describing cross-contamination (Equation (4)). However, in some cases where cross-contamination is expected, it may be of no practical

importance, in which case the recursions should be removed from the model. This can be investigated in an NRSA by analyzing the values of the transfer input variables β_j and γ_j in Equation (4). Assuming that the environment is clean before each processing day ($W_{1,0} = 0$), then Equation (4) can be written into one equation as the summed total of all contributions of previous products via the processing environment:

$$N_{j,k} = (1 - \beta_j)N_{j-1,k} + \sum_{i=0}^{k-2} \gamma_j \beta_j (1 - \gamma)^i N_{j-1,k-i-1}. \quad (5)$$

Here, the first term stands for the number of bacteria that have remained on product k after stage j. The other terms stand for the contributions of previously processed products transferred to product k via the environment. In the case that, for all $i : 0, 1, \ldots, k-2$, typical values of $\gamma_j \beta_j (1 - \gamma_j)^i$ are small, then Equation (5) can be approximated by:

$$N_{j,k} = (1 - \beta_j)N_{j-1,k}. \quad (6)$$

Practically, this means that in some cases cross-contamination may be approximated by removal.

2.4. BBN

In classical MC approaches, samples from the probability distributions of the input variables are used to generate samples from variables that have a functional relation with these inputs and consequently construct their marginal probability distributions. As a next step, samples from these conditioned variables are used to generate samples of variables that have, in their turn, a functional relation with the first set of conditioned variables. This procedure is repeated from the beginning to the end of those steps that are considered in the model. Because of computational reasons, samples drawn from the probability distributions of variables in one stage of the chain are usually not linked to samples generated in all other stages, implying that not all possible combinations of values that might occur (the joint probability) are included in an MC simulation model. This makes these models not appropriate for making inferences in the reverse direction of chain processing order.[13] Inclusion of the whole joint probability of all variables acting in the chain would allow inversion of the probabilistic causal relations between variables, using Bayes' rule (see the Appendix). A BBN provides such a model implementation.

BBNs are a special class of probabilistic graphical models, in which the variables are partitioned onto a directed acyclic graph. Underlying the graphical representation, each of the variables is distributed according to a probability distribution, representing uncertainty and/or variability of the variable. BBNs provide a compact representation of the joint probability distribution of the process, in which the correlations between all variables are captured.[14] The variables are represented by nodes, and the directed arcs represent conditional causal dependence assumptions. Applying Bayes' rule in these models can intuitively be seen as inverting the direction of the arcs in the BBNs.

Note that the domain of continuous variables has to be divided into discrete ranges (in case software is based on the Lauritzen-Spiegelhalter algorithm[15]), or their distributions have to be numerically approximated (in case of MCMC methods[16]). Still, in principle, any level of precision can be achieved by increasing the number of intervals for discretization or number of iterations in a simulation.

2.4.1. Recursions

To model recursive relations that cannot be simplified according to step 3, a dynamic BBN[17] would be required, explicitly capturing the dynamics over time or processing order. In these networks the variables are indexed by time or product number and replicated in the BBN.[18] Such dynamic BBNs can become very large if the number of products is large. It may not be realistic to implement a BBN of this size because of computational complexities. Therefore, we may choose to replace the explicit equation defining the dynamics based on individual products by a stochastic relation defining the dynamics based on a random product, that is, we calculate with random variables instead of a set of deterministic variables. The conditional probabilities linking a random product before a cross-contamination step to a random product after this step can be estimated by drawing numbers from the input distributions of the product before the step, feeding these numbers into the recursive equations, calculating the output variables, and then estimating the conditional probabilities linking input to output variables via the expectation-maximization (EM) algorithm,[19] which is an iterative method for finding the maximum likelihood estimator for the joint probability of the outputs (see Fig. 1).

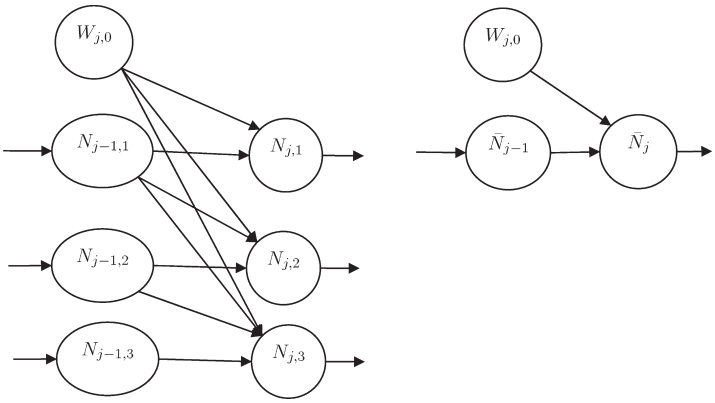


Fig. 1. Feedback relations are common in food chain modeling. The figure on the left shows a part of a dynamic BBN, in which the conditional dependency $N_{j,k}|N_{j,k-1}, \ldots, N_{j,1}, W_{j,0}$ (with $N_{j,k}$ the number of bacteria on the kth product after the jth stage) is given by the explicit recursive relation of Equation (4). The parameters β_j and γ_j are not shown. This scheme may be approximated by the conditional dependency $\bar{N}_{j,1}|\bar{N}_{j-1}, W_{j,0}$, which can be obtained by simulating many inputs $W_{j,0}$ and $N_{j-1,k}$ and outputs $N_{j,k}$ for $k:1, 2, 3$, and finding a stochastic dependency between these (using the EM algorithm). $\bar{N}_j$ is the fitted distribution over all simulated values $N_{j,1}$, $N_{j,2}$, and $N_{j,3}$ and may be described as the number of bacteria on a random product.

2.4.2. Trigger and Report Variable

Variables of which data are obtained that serve to trigger the process of BT are termed the *trigger*s. For the purpose of this article, these are variables indicating the total number of bacteria on a product at the end of the chain. The *report variable* is a discrete variable having as states the names of the different sources. It returns the probability of each possible source of being the dominant one, by comparing the (marginal) probability distributions of the contributions from each source to the product at the end of the line, and evaluating the probability that any source makes the largest contribution to the (observable) total, given by the trigger. In practice, bacteria that are transferred from different sources to one product mix during the chain process. However, we can keep track of the origin of the ultimately observed bacteria on the product only by separate bookkeeping of the bacterial dynamics from the source through the processing chain.

In Fig. 2, the report variable is represented as *Report*. After observing the downstream variable *Trigger*, the posterior distribution of *Report* indicates which source is responsible for the majority of the bacteria on the end product. Mathematically, the purpose of the model is to identify source s for which $P(Report = s|Trigger = x)$ is largest, where x is an observed number of bacteria.

2.4.3. Location of the Contaminant

In most MPRM models the physical location of the bacteria on the product is not considered, and the smallest unit that is modeled is often a product. However, this location may provide information for BT,

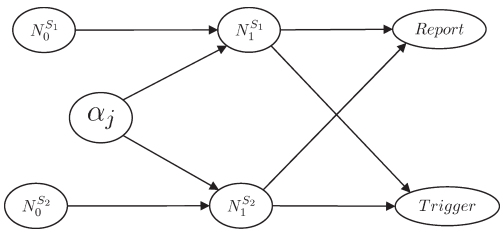


Fig. 2. A simple example of the BT model of a production chain, consisting of one processing step and two potential sources. The upper row of nodes represents the dynamics of bacteria originating from source 1 (S_1), the lower row of nodes the dynamics of bacteria originating from source 2 (S_2); $N_j^{S_i}$ is the number of bacteria on the product after stage j that have originated from S_i; α_j is a parameter describing the bacterial dynamics in stage j (here $j = 1$), for example, inactivation. Variable *Trigger* represents the total number of bacteria on the product, and is calculated by $Trigger = N_2^{S1} + N_2^{S2}$. Variable *Report* shows for each source the probability that it is the dominant one. The discrete conditional probability distribution of this variable is given as follows: $Report \sim dist[\mathrm{S_1}, \mathrm{S_2}, \text{none}; 0 \neq \mathrm{N_1^{S1}} \geq \mathrm{N_1^{S2}}, 0 \neq \mathrm{N_1^{S2}} \geq \mathrm{N_1^{S1}}, \mathrm{N_1^{S1}} = \mathrm{N_1^{S2}} = 0]$. Here the first list of arguments (before the semicolon) are the discrete states of *Report*; the probabilities of those states are the proportion of simulations for which the inequalities in the last list hold.

for example, parts of the product that are in contact with equipment during processing are more likely to contract bacteria originating from the equipment. Keeping track of these locations again needs a strategy of separated bookkeeping for a selected number of locations, similar to what was described earlier.

2.5. SA of the BBN

The ultimate purpose of the model is to make BT conclusions by observing the posterior distribution of

the report variable. Therefore, we want to determine how sensitive this distribution is to the prior probabilities of the input values. This is different from the analyses in step 3, first, because a global SA is done instead of a series of local SA, and second, because the sensitivity of the *distribution* of a variable to changes in an input *distribution* is tested instead of testing the sensitivities of the *value* of a variable to changes in an input *value*.

Coupé *et al.*[20] noted that the probability of any variable in a BBN having a given value is a linear function of the probability that any single input variable in the model has a given value, if all other conditional probabilities are kept fixed. This one-to-one correspondence enables a global SA. If additional evidence in the BBN is also accounted for, the *sensitivity function*[20] may be defined as a quotient of two linear functions, mapping points in the probability space of an input value to the probability space of the value of the variable that is analyzed (here: the report variable). The slope of this function is termed the *sensitivity value*[21] (SV). SVs can be calculated for all possible values of the input variable and all possible values of the report variable, but the (overall) sensitivity of the report variable to an input variable may be summarized by their minimum, maximum, and average SV. If an SV is close to zero, then the report variable does not react strongly to changes in the prior probabilities of the input variable. Otherwise, input variables revealing a large sensitivity to the report variable (large absolute value of the SV) may heavily influence the BT statement made by the model and therefore their priors should be chosen carefully.

The approach described above is a one-way SA, in which the network's conditional probabilities are investigated one at a time. Also multidimensional SA can be performed to determine the joint effect of varying the probability assessments for multiple input values simultaneously on the distribution of the report variable. In the case of two inputs, this is known as a two-way SA, and the sensitivity function describes a surface.[22]

In some BBN software, such as HUGIN (HUGIN Expert, Aalborg, Denmark), automated routines are implemented to calculate the SVs. If the prior probability assignment of an input variable reveals a large influence on the report variable, then the modelers should assess how confident they are about this prior probability. If not so confident, then additional study should be focused on obtaining better estimates for these inputs.

2.6. Learn Parameters from Downstream Data

Through the Bayesian paradigm, the prior distributions of model parameters are replaced by posteriors after inserting a set of downstream data. In a procedure called *sequential adaptation*,[19] these posteriors can subsequently be used as updated priors as new sets of data become available. As more data are sequentially fed to the model, the uncertainty about the true parameter values of the model is gradually reduced: we say that the parameters are learned. To learn a parameter in a BBN, weights, known as *experience counts*,[19] should be assigned to them. Such weights are indicative of the sample size of the study in which the values of the parameters were estimated or—more subjectively—on the degree of certainty of the expert estimating the parameter value. In one step of sequential adaptation, the posterior distributions of parameters are used to update their prior distributions in the light of the new data and the initial experience counts. See the Appendix for an example of how this works.

Note that the link between downstream data and parameters described here is absent in MC simulation models.[9,13] In practice, data are used in MC simulation models as a means of validating the outputs, but this is not straightforward and requires a careful analysis of the model results and of the data, including their uncertainties.

2.7. Start BT

Steps 1–6 describe the construction of the model. This final step highlights the actual biotrace: if a new data point for the trigger is available, then the posterior of the report variable indicates which source is most likely to have caused the observed case of contamination. After this analysis, the new data point can be added to the database of observations. This means that new observations can be used for BT statements, but also for parameter learning (i.e., by repeating step 6). This way we have constructed a BT model that gradually makes better BT statements the longer it is in use.

3. *SALMONELLA* IN THE PORK SLAUGHTER CHAIN

We will now illustrate the steps of the proposed framework with a practical example of a BT model of *Salmonella* in a Dutch pig slaughterhouse. The model is based on the EFSA QMRA on *Salmonella*

in slaughter and breeder pigs,[6] in which an MC simulation model was proposed that predicts the number of cases per year attributed to the consumption of pork for different EU regions. This model describes the microbial behavior of *Salmonella* by modeling the dynamics of *Salmonella* through the chain, from farm up until consumption. In the first stages of the chain only *Salmonella* prevalences are considered; changes in numbers of *Salmonella* (concentrations) are modeled during the slaughter and consumption phase. The model deals with variability of the number of bacteria over carcasses and variability of process inputs over EU regions, but does not consider uncertainty. In the following we describe the steps taken to modify the slaughter phase of this MC simulation model to a BT model representing one specific slaughterhouse in the Netherlands.

3.1. Domain Model

At the start of the slaughterhouse procedure, live pigs are unloaded from transport and kept between 4–6 hours in the lairage environment, where massive cross-contamination is likely to occur.[23] Subsequently, the pigs are stunned and their major arteries and veins cranial to the heart are cut manually with a knife that is dipped in water of 82 °C after each cut. After this, a conveyor belt leads the carcasses to a scalding bath of approximately 60 °C in which they are immersed for eight minutes. This loosens the hairs and reduces bacterial population sizes.[24,25] It also allows for subsequent removal of hairs during dehairing, a mechanic process in which revolving metal pedals scrape the hair off the carcass. If the dehairing equipment becomes contaminated with *Salmonella*, it may cross-contaminate carcasses that follow.[26] Morgan *et al.*[27] suggested that pressure of the machine onto the carcass might cause an amount of feces to emanate from the pig, leading to contamination of the dehairing machine. After dehairing, the carcasses are hung on an overhead rail by their hind limbs. They are transported on this rail to undergo a flaming treatment, which aims to remove remaining hairs. A side effect is the inactivation of bacteria that are present on the carcass surface.[25,26] Polishing, by scraping the carcass, aims to remove remaining dirt on the carcass. In this process, cross-contamination via the machine could occur and also fecal extrusion due to the pressure of the machine.[26] In our Dutch slaughterhouse, the processes of flaming and polishing are subsequently repeated. Upon entering the "clean area," the belly is opened by an automated cutting apparatus. There is a small risk of puncturing the colon, thereby recontaminating the carcass or machine. During the subsequent processes of bung dropping and pluck removal the anus and intestines are removed from the carcass. During splitting the carcass is split into two. After cutting each carcass, the knife is dipped into water of 82 °C for one second.[28] Although this reduces bacterial populations, this time-temperature combination is probably insufficient to inactivate all bacteria, leading to cross-contamination via the knife. After each slaughtering day the cutting robot is disinfected. Yet, expert opinion suggests that the inside of the robot is hard to clean, so bacteria may persist and/or multiply in small nooks. Several studies have indicated that the slaughtering environment should indeed be seen as an independent reservoir.[29–32] Next, the carcass halves are inspected for visual signs of fecal contamination, which is removed. Finally, the carcasses are hung overnight in a cooling room, until the carcasses are transported to a cutting plant.

Many different sources of *Salmonella* can be identified in the chain. When entering the slaughterhouse, pigs may carry *Salmonella* directly on the skin or in feces. Also house flora, defined as *Salmonella* on the slaughtering robots that are never completely removed by sanitation, may spread to carcasses. We let the BT question be:

Given these sources and given concentration data of *Salmonella* on the exterior and the cut section of a specific carcass after meat inspection, can we infer the dominant source?

Based on this question, the start and end phases (killing and meat inspection) were chosen and the relevant modules in between (see the conceptual model in Fig. 3). Inputs for the model are concentrations of *Salmonella* on the skin of the pig that is introduced into the slaughterhouse and concentrations of *Salmonella* in the gastrointestinal tract of the pig that may be transferred to the carcass during dehairing, polishing, evisceration, and bung removal. In Ref. 6 these inputs are outputs of the farm module, which is not discussed here, but which summary statistics were directly borrowed as inputs for this model. No data were available to quantify the spread of house flora to carcasses but it was expected that this spread was limited; therefore, modest values were assumed (see Section 3.2). Other inputs of the model are (distributions of) process inputs. We used the expert opinion of an expert in the slaughterhouse that was modeled to estimate the expert's

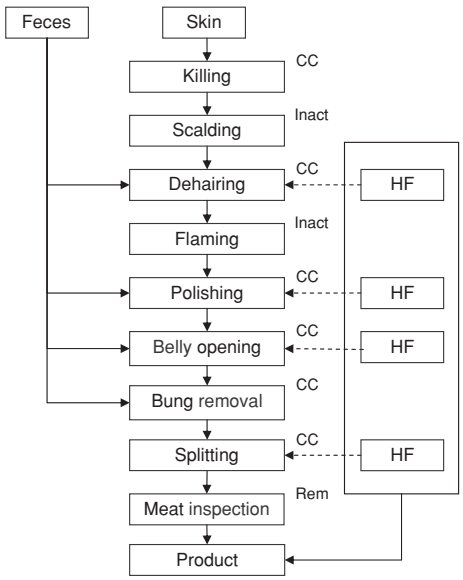


Fig. 3. An overview of the stages included in the domain. The narrow boxes indicate the sources: bacteria on the skin of the incoming pigs, bacteria in their feces, or bacteria from house flora (HF) in the slaughterhouse. The wide boxes indicate the stages of production. The microbial processes occurring in these stages are cross-contamination (CC), inactivation (Inact), and removal (Rem). The inputs associated with the bacterial transfer of house flora are uncertain and may differ per machine. Therefore, we assume in the model that house flora may act as one unknown source that explains observations on the product that cannot be explained by the other sources. The parameters of the model are estimated from downstream data.

values, and if this expert opinion was not available we used the values established in Ref. 6 for the Dutch region. These values are shown in Table III.

3.2. MPRM Model

See Table III for a detailed description of the implementation of most stages in the chain. Modifications of Ref. 6 as described later were made to adjust the model for one specific Dutch slaughterhouse. They were suggested by an expert in that slaughterhouse.

Puncturing of the gut during belly opening is a rare event, yet when it does happen, fecal soiling occurs in large quantities. We assumed that this is always detected by visual inspection and that such contaminated carcasses are directly removed from the line. Therefore, this event was not included in the model. However, bung removal, which was not implemented in Ref. 6, was implemented in this model. Fecal contamination occurs more regularly here, but in low quantities of around 1 gm, which is harder to detect. A very high probability of detection of fecal contamination during meat inspection was assumed in Ref. 6 so that in the model practically all feces was assumed to be removed from the skin. In our model this is unrealistic because of the low quantities of feces that may be extruded during bung removal. In absence of data, we assumed a linear relationship between the amount of feces on the carcass and the probability of detecting it during meat inspection.

The assumption that the hot water treatment after cutting operations will inactivate all bacteria on the knife was made in Ref. 6. In this model this assumption was not made,[33,34] and a combination of limited inactivation of bacteria on the knives and, with that, cross-contamination via the knives was implemented for the belly opening, bung removal, and carcass splitting operations, using Equation (4).

Four sources of house flora (the dehairing, polishing, belly opening, and carcass splitting machines) were assumed in Ref. 6. To be able to better discriminate between house flora as a source "in general" and the other sources, we lumped these and treated it as a single source (which, in the context of BT, is a category of contamination causing cases that cannot be explained by the other sources). As in Ref. 6, we assumed that the numbers of *Salmonella* transferred from house flora to passing carcasses were low.

3.3. SA of MPRM Modules

Input-output relations of individual modules of the model were simulated via MC sampling, and the sensitivities of the outputs to values of the input variables were assessed by NRSA. The probability distributions of some inputs were found redundant, such as those of the D- and z-values of *Salmonella* on pork skin, which are used to calculate inactivation rates. These were replaced by their mean values. For other redundancies in the estimates of the input variables originally taken from Ref. 6, see Table III. Also, cross-contamination during scalding was found not to contribute significantly to carcass contamination; therefore, the recursive equation describing cross-contamination was replaced by an equation describing removal, that is, transfer of *Salmonella* from the skin to the water.

3.4. BBN

We used the MPRM equations established in step 2 to construct the BBN. The conditional probabilities relating two or more variables are directly based on these equations if they are explicit and nonrecursive. Otherwise, they were generated by simulating them using MC sampling. The physical locations where *Salmonella* is found are valuable pieces of information for BT: *Salmonella* on the carcass's cut section are likely to stem from the cutting machines, whereas *Salmonella* on the carcass's skin are more likely to stem from the live pig. Therefore, a strategy is employed of separate bookkeeping of the dynamics of bacteria on the skin and on the cut section: these are separate variables in the BBN.

Two independent triggers were identified in the model, representing the total number of bacteria on the cut section (interior: Tot_i) and on the skin surface (exterior: Tot_e) after meat inspection. Two report variables were introduced to indicate the dominant source of contamination on the cut section (Rep_i) and the skin surface (Rep_e). The variables and equations that are used to construct the BBN are summarized in Table III. The BBN is shown in Fig. 4. Our implementation of this BBN employs the HUGIN software (HUGIN Expert, version 7.4, Aalborg, Denmark).

3.5. Sensitivity Analysis BBN

To gain insight into the sensitivity of the report variable to the prior probabilities of the uncertain input variables, a one-way SA was performed (see Table II). The report variable Rep_i proved to be most sensitive to the input variables indicating the probability of the pig carrying positive feces at the onset of slaughter and the bacterial concentration in it, the expected number of *Salmonella* transferred to the cut section of a carcass as house flora, the temperature of the decontaminating knife during bung dropping, and the probability of cutting the intestines during bung dropping. Rep_e was found to be most sensitive to the input variable quantifying the *Salmonella* concentration on the skin of a positive pig, the transfer and inactivation rates during dehairing, flaming and polishing, and the expected number of *Salmonella* transferred to the skin of a carcass as house flora. We only did a one-way SA here. Varying the prior probabilities for values of some different inputs simultaneously may have a joint effect on the distribution of the report variable beyond the effects of their separate variation, due to their synergistic influence.[22] For now, this was beyond the scope of this study.

3.6. Learn Parameters

Low experience counts of 10^{-6} (see the Appendix) were assigned to the parameters (Table III) because their initial assessments were very uncertain. Downstream data on the triggers (Tot_i and Tot_e) were used to refine the priors of the parameters. Note that in principle, also additional, upstream data can be used, such as observations of the bacterial populations on pigs entering the slaughter process, or registered values for time and temperature of processes in the chain. As more data are used, the experience counts of these parameters increase, that is, the certainty about their true value increases. The result of this step, where we have included *Salmonella* concentration data obtained in one specific slaughterhouse in the Netherlands, will be described in Smid *et al.* (in preparation).

3.7. Start BT

After reducing uncertainty of the model predictions by including downstream data in the model in Section 3.6, the model is ready for doing an actual biotrace, of which an example is shown in Fig. 4. Note that, due to the methodological character of this article, step 6 was not implemented here, so there is still considerable uncertainty in the model predictions, especially about the contribution of house flora (see Section 3.6). The model only reflects our prior information and, therefore, the numerical results of this example should be considered as illustrative only.

4. DISCUSSION

We have described a practical framework to construct a BT model, and we have illustrated this framework by an example of *Salmonella* in the pork slaughter chain. The EFSA QMRA on *Salmonella*[6] was used as the basis for this model because it is the most complete MPRM model of *Salmonella* in pork published thus far. Our baseline results (i.e., the computed *Salmonella* concentrations at each stage of slaughtering) are in good agreement with those produced by the MPRM model[6] (results not shown). This good agreement was expected because the modifications in the implementation of the dynamics, proposed in step 2, were only minor and the simplifications, proposed in step 3, were tested to have

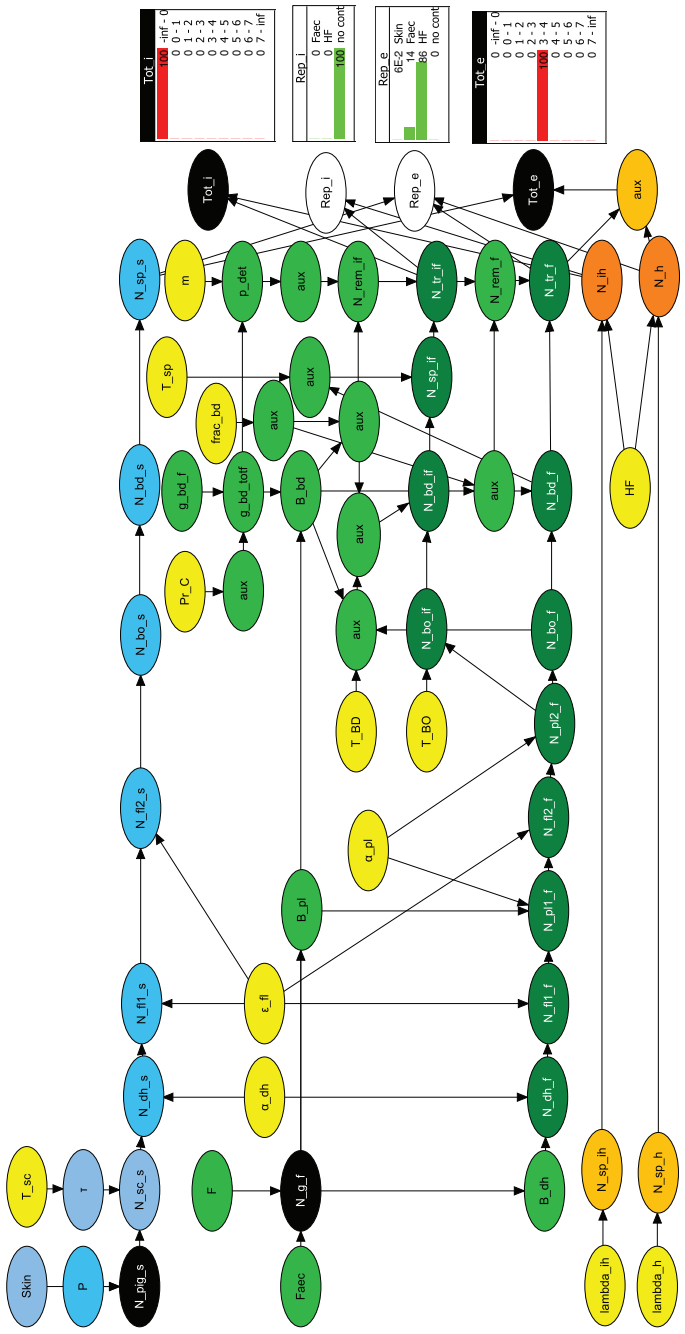


Fig. 4. BBN of the *Salmonella* in pork slaughter chain. The different colors (only visible in the online version of the article) represent the dynamics of *Salmonella* coming from different sources: the cluster of blue nodes represents the dynamics of *Salmonella* originating from the skin of the incoming pig, the cluster of green nodes represents the dynamics of *Salmonella* originating from the colon of the incoming pig, the cluster of orange nodes represents the transfer of *Salmonella* originating from house flora. The yellow nodes represent the parameters in the model. The black nodes are the triggers (Tot_i resp. Tot_e), representing the number of *Salmonella* on the cut section resp. skin of the carcass after meat inspection. The white nodes are the report variables (Rep_i resp. Rep_e). Note the bifurcations in the network after belly opening and polishing, indicating a distinction between *Salmonella* on the cut section and on the skin of the carcass. In this BT example, we have measured zero bacteria on the cut section ($Tot_i = 0$) and between 10^3 and 10^4 bacteria on the skin ($Tot_e \in [3, 4]$) of a carcass after meat inspection. Then, with 86% certainty, house flora was the dominant source and, with 14% certainty, fecal contamination was the dominant source of the observed case of external contamination; see the posterior distribution of Rep_e.

Table II. The Values in the Table Indicate How Sensitive the Probabilities for the Values of the Report Variables Rep_i and Rep_e Are to Initial Assessments (Prior Probabilities) of the Input Variables

Report Variable	Rep_i			Rep_e		
Input Variable	MIN SV	MAX SV	AVG SV	MIN SV	MAX SV	AVG SV
Skin	0	0	0	−1.3E−3	1.3E−3	1.3E−3
P	0	0	0	−2.8E−4	3.0E−2	7.1E−3
T_{sc}	0	0	0	−2.0E−4	3.3E−4	1.9E−4
F_{aec}	−1.1E−2	1.1E−2	1.1E−2	−2.6E−3	2.6E−3	2.6E−3
F	−2.3E−3	4.2E−3	2.0E−3	−6.8E−4	5.8E−4	3.4E−4
α_{dh}	−2.8E−5	1.1E−5	1.1E−5	−7.3E−3	1.3E−2	6.8E−3
ϵ_{fl}	−8.7E−5	1.5E−4	1.0E−4	−6.9E−3	7.6E−3	5.1E−3
α_{pl}	−4.9E−5	7.3E−5	3.2E−5	−6.3E−3	1.2E−2	5.0E−3
T_{bo}	−5.9E−4	7.6E−4	5.6E−4	−0.0E−0	0.0E−0	0.0E−0
P_{rc}	−1.9E−3	3.1E−2	7.2E−3	−4.8E−6	7.7E−5	1.8E−5
g^f_{bd}	−7.3E−5	5.2E−5	3.9E−5	−7.0E−6	4.5E−6	3.6E−6
T_{bd}	−2.1E−3	2.2E−3	1.5E−3	0	0	0
$frac_{bd}$	−1.4E−3	4.2E−4	7.8E−4	−5.3E−6	1.8E−5	9.9E−6
T_{sp}	−5.9E−4	7.6E−4	5.6E−4	0	0	0
m	−4.2E−4	8.3E−4	4.0E−4	−8.8E−6	4.4E−6	4.2E−6
λ_h	0	0	0	−8.7E−3	1.6E−2	5.1E−3
λ_{ih}	−4.9E−3	8.2E−3	2.7E−3	0	0	0

Notes: These sensitivities are summarized by a minimum, maximum, and average sensitivity value, indicating the minimum, maximum, and average slopes of the derivative of the sensitivity functions, taken over all states of the report variable and over all states of the input variable. The larger the absolute value of the SV, the more careful the prior distribution for an input should be assessed because it may heavily influence the BT decision made by the model.

negligible effects on the outcomes of each module. However, the BT model is implemented in a BBN, which has a graphical user interface and is, therefore, more straightforward to use for nonexperts than models based on MC sampling. Moreover, the BT model is more flexible as it includes the interactions between the inputs of the different modules.[13] Downstream data can be used to reduce uncertainty of the parameters in the model, thereby enabling BT.

The proposed framework may be used for the construction of a BT model in many product chains of which the structure is related to the pork chain (discrete products, multiple sources acting on the same product). However, note that by this framework we do not claim to have exhaustively prescribed all steps one needs to make to build a BT model for all types of product chain and all types of tracing question. For example, in a recently developed BT model of *Staphylococcus aureus* in the milk chain (Barker *et al.*, in preparation), the BT question was to trace the source of enterotoxins in a milk product at the point of sale. Source assignments were defined as discriminating between preformed toxin in the milk tank, pasteurization failure, or postprocess contamination. In this chain, for one case of contamination only one source is assumed to have been its cause. Also, milk from different suppliers is mixed in bulk tanks, so discrete products cannot be identified in all stages of the chain. This leads to several differences with the *Salmonella* in pork model, for example, the description of the dynamics of the toxin does not need to be labeled by source, and cross-contamination is a continuous process here and cannot be described by Equation (4). Nevertheless, also in this chain many steps of the proposed framework remain valid.

Tracing the source of contamination based on downstream observations of the bacterial numbers on products is possible if the microbial dynamics and the inputs of the dynamics are implemented in a causal framework. These inputs are, in general, only partially known, so uncertainties have to be accounted for. However, uncertainty hinders the BT inferences, and consequently it has to be reduced. This could be done by incorporating expert opinion in the network, such as was done by Donald *et al.*,[35] where beta priors estimated by experts were used to assess the prior probabilities of binary outcomes. To obtain less uncertain estimates for the inputs, also measurements could be made in the chain, or experiments could be done, such as predictive microbiological experiments. To assess the expected number of

Table III. Description of the Variables Used in the BT Model

Stages	Variable	Description	Expression for Conditional Probabilities	Mod.
Killing	*Skin*	Probability incoming pig has *Salmonella* on skin	*dist*[0, 1; 0.8, 0.2]	*SS3*
	P	Log #S on positive incoming pig	*trunc*(0, *Normal*[2.8, 0.511])	*SS3*
	N^s_{pig}	#S on incoming pig	$if(Skin = 0, 0, 10^P)$	
Scalding	β_{sc}	Fraction of bacteria moving from pig to water	0.02	*N3*
	t_{sc}	Time of scalding, in minutes	8.66	*E3, N3*
	T_{sc}	Temperature of scalding, in $^\circ$C	*Unif*[58, 60.5]	*E3, N3*
	τ	Fraction of *Salmonella* inactivation on the pig	$0.92\ 10^{(T_{sc}-60)/5.4}$	
	N^s_{sc}	#S on pig after scalding	$(1-\beta_{sc})N^s_{pig}e^{-\tau t_{sc}}$	
Dehairing	*Faec*	Probability incoming pig has *Salmonella* in feces	*dist*[0, 1; 0.845, 0.155]	*SS3*
	F	Log #S/g in feces positive pig	*trunc*(0, *Normal*[2.36, 4.39])	*SS3*
	α_{dh}	Rate of *Salmonella* transfer from machine to pig	$10^{Unif[-1.5,0]}$	
	β_{dh}	Rate of *Salmonella* transfer from pig to machine	0.18	*N3*
	t_{dh}	Time of dehairing, in minutes	*Unif*[0.48, 2.13]	*N3*
	g^f_{dh}	Amount of feces extruded, in grams	10	*N3*
	N^f_g	#S in feces per gram	$if(Skin = 0, 0, 10^F)$	
	B_{dh}	#S in spilt feces	$g_{dh} * N^f_g$	
	$N^{s(,k)}_{dh}$	#S on (kth) pig after dehairing, source = *skin*	$(e^{\alpha_{dh}t_{dh}}-1)\frac{\beta_{dh}}{1-\beta_{sc}}\sum_{j=1}^{k}N^{s,k-j+1}_{sc}e^{-j\alpha_{dh}t_{dh}}$ $+(1-\frac{\beta_{dh}}{1-\beta_{sc}})N^{s,k}_{sc}$	*R5*
	$N^{f(,k)}_{dh}$	#S on (kth) pig after dehairing, source = *feces*	$(e^{\alpha_{dh}t_{dh}}-1)\frac{\beta_{dh}}{1-\beta_{sc}}\sum_{j=1}^{k}B^{k-j+1}_{dh}e^{-j\alpha_{dh}t_{dh}}$	*R5*
Flaming1	t_{fl1}	Time of flaming, in minutes	22/60	*E3, N3*
	ϵ_{ft}	Rate of inactivation during flaming	*Unif*[5, 11]	*E3*
	N^s_{fl1}	#S on pig after 1st flaming, source = *skin*	$N^s_{dh}e^{-\epsilon_{ft}t_{fl1}}$	
	N^f_{fl1}	#S on pig after 1st flaming, source = *feces*	$N^f_{dh}e^{-\epsilon_{ft}t_{fl1}}$	
Polishing1	α_{pl}	Rate of *Salmonella* transfer from machine to pig	$10^{Unif[-0.5,0]}$	
	t_{pl}	Time of polishing, in minutes	*Unif*[0.47, 1.58]	*N3*
	g^f_{pl}	Amount of feces extruded	1.0	*N3*
	B_{pl}	#S in spilt feces	$g_{pl} * N^f_g$	
	N^s_{pl1}	#S on pig after 1st polishing, source = *skin*	N^s_{fl1}	
	$N^{f(,k)}_{pl1}$	#S on pig after 1st polishing, source = *feces*	$(e^{\alpha_{pl}t_{pl}}-1)\sum_{j=1}^{k}(B^{k-j+1}_{pl}+N^{k-j+2}_{fl})e^{-j\alpha_{pl}t_{pl}}$	*R5*
Flaming 2	t_{fl2}	Time of flaming, in minutes	25/60	*E3, N3*
	N^s_{fl2}	#S on pig after 2nd flaming, source = *skin*	$N^s_{pl1}e^{-\epsilon_{ft}t_{fl2}}$	
	N^f_{fl2}	#S on pig after 2nd flaming, source = *feces*	$N^f_{pl1}e^{-\epsilon_{ft}t_{fl2}}$	
Polishing2	N^s_{pl2}	#S on pig after 2nd polishing, source = *skin*	N^s_{fl2}	
	$N^{f(,k)}_{pl2}$	#S on pig after 2nd polishing, source = *feces*	$(e^{\alpha_{pl}t_{pl}}-1)\sum_{j=1}^{k}N^{k-j+2}_{fl2}e^{-j\alpha_{pl}t_{pl}}$	*R5*

(*Continued*)

Table III. (*Continued*)

Stages	Variable	Description	Expression for Conditional Probabilities	Mod.
Belly opening	T_{bo}	Temperature reached during decontamination, in °C	*Unif*[50, 70]	*D3, E3*
	t_{bo}	Time spent sterilizing	1.0	*D3, N3*
	α	Transfer rate from knife to pig	*Pert*[0.42, 0.94, 1]	*D3, N3*
	β	Transfer rate from pig to knife	0.21	*S4, N3*
	γ	Inactivation rate on the knife	$\frac{\ln(10)}{3} 10 \frac{-(60-T_{bo})}{8}$	*D3, N3*
	d_{bo}	Fraction *Salmonella* moving from pig to knife	$2.4 * 10^{-4}$	*S4, N3*
	N^f_{bo}	#*S* on skin after belly opening, source = *feces*	$(1-d_{bo})N^f_{pl2}$	*D3*
	N^{if}_{bo}	#*S* on cut section after belly opening, source = *feces*	$\alpha \sum_{j=0}^{k-2}(1-\alpha)^j (d_{bo} N^{f,k-j-1}_{pl2}) e^{-(j+1)\gamma t_{bd}} + (1-d_{bo})N^f_{pl2}$	*D3,R5*
	N^s_{bo}	#*S* on skin pig after belly opening, source = *skin*	$(1-d_{bo})N^s_{pl2}$	*D3*
Bung removal	T_{bd}	Temperature reached during bung removal, in °C	*Unif*[50, 70]	*D3, E3*
	Pr_c	Probability of cutting colon	*Beta*[4, 98]	*D3, E3*
	$frac_{bd}$	Fraction of feces transferred to exterior, if colon is cut	*Beta*[81, 21]	*D3, E3*
	g^f_{bd}	Amount of feces released if cut	*Unif*[0, 1]	*D3, E3*
	g^{totf}_{bd}	Amount of feces released	$if(Pr_c = 0, 0, g^f_{bd})$	*D3*
	d_{bd}	Fraction *Salmonella* moving from pig to knife	$3.3{*}10^{-5}$	*D3, S4*
	δ	Proportion of leaking fecal material moving to knife	0.02	*D3, N3*
	B_{bd}	Number in *Salmonella* in spilt feces	$g^{totf}_{bd} * N^f_g$	*D3*
	N^s_{bd}	#*S* on exterior after bung dropping, source = *skin*	$(1-d_{bd})N^s_{bo}$	*S4*
	$N^{if(,k)}_{bd}$	#*S* on cut section after bung dropping, source = *feces*	$\alpha \sum_{j=0}^{k-2} 1-\alpha^j (d_{bd} N^{f,k-j-1}_{bo} + \delta \mathrm{B}^{k-j-1}_{bd}) e^{-(j+1)\gamma t_{bd}} + Bernoulli[1-frac_{bd}](1-\delta)B_{bd}$	*D3, R5*
	N^f_{bd}	#*S* on skin after bung dropping, source = *feces*	$Bernoulli[frac_{bd}](1-\delta_{bd})B_{bd} + (1-d_{bd})N^f_{bo}$	*D3*
Carcass splitting	d_{sp}	Fraction *Salmonella* moving from pig to knife	$2.4{\times}10^{-4}$	*S4, N3*
	t_{sp}	Time spent sterilizing	1.0	*D3, N3*
	T_{sp}	Temperature reached during decontamination	*Unif*[50, 70]	*D3, N3*
	$N^{f(,k)}_{sp}$	#*S* on cut section after splitting, source = *feces*	$(1-d_{sp})N^f_{bd}$	*D3*
	$N^{if(,k)}_{sp}$	#*S* on cut section after splitting, source = *feces*	$\alpha \sum_{j=1}^{k=1}(1-\alpha)^{j-1}(d_{bo} N^{f,k-j+1}_{bd}) e^{-j\gamma t_{sp}} + N^{if}_{bd}$	*D3, R5*
	N^s_{sp}	#*S* on skin after splitting, source = *skin*	$(1-d_{sp})N^s_{bd}$	*D3, S4*
Meat inspection	$pdet$	Probability of detection of faecal contamination	$mim(1, mg^f_{bd})$	*D3*
	m	Multiplication factor amount of feces	$unif[0, 5]$	*D3, N3*
	N^{if}_{rem}	#*S* removed from interior if feces detected, source = *feces*	$Bernoulli[p_{\mathrm{det}}] * Bernoulli[1-frac_{bd}] * g^{totf}_{bd} * N^f_g$	*D3*
	N^f_{rem}	#*S* removed from skin if feces detected, source = *feces*	$Bernoulli[p_{\mathrm{det}}] * Bernoulli[frac_{bd}] * g^{totf}_{bd} * N^f_g$	*D3*
	N^{if}_{tr}	#*S* on cut section pig after trimming, source = *feces*	$N^{if}_{sp} - N^{if}_{rem}$	
	N^f_{tr}	#*S* on skin pig after trimming, source = *feces*	$N^f_{bd} - N^f_{rem}$	

(*Continued*)

Table III. (*Continued*)

Stages	Variable	Description	Expression for Conditional Probabilities	Mod.
House flora	HF	Probability that particles are released from house flora	1	
	N^h	#S from biofilm to skin	$LogN[\lambda_h, 1]$	*D3*
	N^{ih}	#S from biofilm to cut section	$LogN[\lambda_{ih}, 1]$	*D3*
	λ_h	Expected log #S from biofilm to skin	$unif[0, 2]$	*D3*
	λ_{ih}	Expected log #S from biofilm to cut section	$unif[0, 2]$	*D3*
Totals	Tr_i	Total #S on cut section after trimming	$N_{tr}^{if} + N^{ih}$	*Tr*
	Tr_e	Total #S on skin after trimming	$N_{sp}^{s} + N_{tr}^{f} + N^{h}$	*Tr*
	Rep_i	Report variable, returns probability assignment for source of contamination of cut section	$dist[Faec, BF, none;$ $0 \neq N_{tr}^{if} \geq N^h, 0 \neq N^{ih} \geq N_{tr}^{if},$ $N_{tr}^{if} = N_{ih} = 0]$	*Rep*
	Rep_e	Report variable, returns probability assignment for source of contamination of skin	$dist[Faec, Skin, BF, none;$ $N_{tr}^{f} \neq 0 \& N_{tr}^{f} \geq N_{tr}^{s} \& N_{tr}^{f} \geq N^h,$ $N_{tr}^{s} \neq 0 \& N_{tr}^{s} \geq N_{tr}^{f} \& N_{tr}^{s} \geq N^h,$ $N_h \neq 0 \& N^h \geq N_{tr}^{s} \& N^h \geq N_{tr}^{f},$ $N^h = 0 \& N_{tr}^{s} = 0 \& N_{tr}^{f} = 0]$	*Rep*

Notes: #S means "number of *Salmonella.*" The codes in the last column refer to steps that were taken to modify the QMRA model proposed in Ref. 6 to a BT model, the numbers in these codes refer to the steps in the framework where these modifications were proposed. *SS3:* Summary statistics of quantities borrowed from Ref. 6. *I3*: Inputs used in Ref. 6. *E3:* Parameter values different from Ref. 6; differences suggested by expert opinion. *D3*: Implementation of dynamics different from Ref. 6; different dynamics are described in words in this article. *N3:* variables not included as separate nodes in the BBN. *S4*: Sensitivity analysis of the forward model has indicated that the probability distributions of a parameter can be described by its mean value, or that parts of recursive equations could be removed. *R5*: the conditional probabilities of this variable are obtained by MC simulations that were made of recursive equations to generate conditional dependencies for the BBN. Empty entries in the last column indicate that the conditional probabilities were directly borrowed from Ref. 6. *Tr*: Trigger variable. *Rep*: Report variable.

bacteria transferred from house flora to each carcass in our model, measurements could be made directly before and after the cutting step where house flora is expected. However, in practice, it is often more convenient to obtain data at the end of the chain, for example, at meat inspection, where more time for sampling is available. Using a BBN, these downstream data are readily propagated backwards in the chain, refining the parameter distributions of the model.

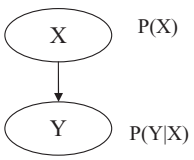
We believe that it is of vital importance to progressively include more (downstream) data to reduce uncertainties in the model and to obtain more reliable tracing estimates. By using data from one particular processing plant, the original model with generic parameter estimates is progressively adapted to reflect that specific plant. We will do this and provide detailed tracing results in a next paper (Smid *et al.*, in preparation).

ACKNOWLEDGMENTS

This work was supported by the European-Union-funded Integrated Project BIOTRACER (contract #036272) under the 6th RTD Framework.

APPENDIX

A1. Bayes' Rule



The figure above shows a very simple BBN. The variables X and Y are represented by nodes. The directed arc between the variables represent causal dependence relations, pointing from cause to consequence. The node X has only arcs pointing out, which means that it is conditionally independent. Describing the range of values that the parameter can take and the probabilities of taking these values, we assign to it the probability distribution $P(X)$, termed the *prior* distribution to indicate that it does not take into account any information on Y. Variable Y is causally dependent on X, and the underlying dependency is given by the probablistic relationship $P(Y|X)$. The pipe sign indicates that the variable in front of this sign is conditioned on the variable behind the pipe

sign. Also the probability distribution $P(Y)$ can be assigned to node Y, which describes the range of values that Y can take and the probabilities of taking these values, accounting for uncertainty/variability of X; $P(Y)$, called the marginal distribution of Y, is constructed by *marginalizing* over all values of the parameters, which can be mathematically written as $P(Y) = \sum_X P(Y|X)P(X)$. The *joint probability* that is described by this BBN is $P(X,\mathrm{Y})$, where $P(X, Y) = P(Y|X)P(X,Y)$. In unidirectional models the quantities of interest are $P(X)$ and $P(\mathrm{Y})$ (termed the *margins*), whereas in BBNs the quantity of interest is $P(X,\mathrm{Y})$. Once the value for Y is specified, by Bayes' rule the direction of the arc can be inverted, which is mathematically expressed as $P(X|Y) = \frac{P(Y|X)P(X)}{\sum_X P(X,Y)}$. This probability is termed the *posterior* distribution for X because it depends on the specified value of Y.

A2. Sequential Adaptation

If the state of only some of the variables in a BBN is known, then the marginal probability distributions for the remaining variables are calculated. Updating the remaining variables given evidence on other variables is called *propagation*. Each time this is done, we have described a *case*. Such cases can be used for reducing uncertainty in the model, thereby improving it.

The probability distributions of parameters in the model are often assessed by expert assessment. For example, in the case of a parameter A with two states a and b, a possible expert opinion is that the probability of a is 0.35. To model second-order uncertainty about this assessment (e.g., an expert opinion of the type "the probability of a is somewhere between 0.3 and 0.4, but most probably around 0.35"), a procedure of using experience counts was suggested.[36] These counts reflect the (hypothetical) number of observations on which this belief is based: a smaller number of observations leads to a larger variance around the expected probabilities of a and b. For A, a beta distribution with mean 0.35 and standard deviation 0.05 can be used to reflect this second-order uncertainty. The shape parameters of this beta distribution (α and β) may be interpreted as the number of times that each state of A has been observed and the sum $\alpha + \beta$ is the total sum of observations of A, called the *experience count*.[21] More generally, for a parameter with N states, we assume that the second-order probabilities for each state of the parameter follow a so-called Dirichlet distribution. The sum of the n shape parameters of this distribution($\sum_{i=1,\ldots,n} \alpha_i$) is in this case the experience count. The experience count and the probabilities of the parameter completely specify its distribution, including second-order uncertainty, and both are updated through sequential adaptation.[19]

Technically, adaptation consists of several steps, in which approximations are made by the BBN software to avoid combinatorial explosion. We will not go into mathematical detail and refer to Ref. 38 for an exact description of the calculations. For each parameter that must be adapted, the following steps are subsequently repeated for each case:

(1) start with the prior joint probability distribution;
(2) propagate a set of data points (i.e., a case) through the network, through which a posterior joint probability distribution is obtained;
(3) calculate the updated prior probability distribution of the parameters as a particular function of their initial experience counts and their posterior marginal probability distributions (the formula for doing this is given in Ref. 38);
(4) retract the case from the network and calculate the updated prior joint probability distribution.

In practice, if the initial degree of certainty about the true values of parameters is low, low initial experience counts (e.g., 10^{-6}) can be assigned to their priors. Note that not all input distributions in the model reflect uncertainty, and therefore they should not all be learned with data. Input distributions reflecting variability must be conserved during sequential adaptation, which can be achieved by not assigning initial experience counts to these. The sequential adaptation algorithm is included in the HUGIN software.

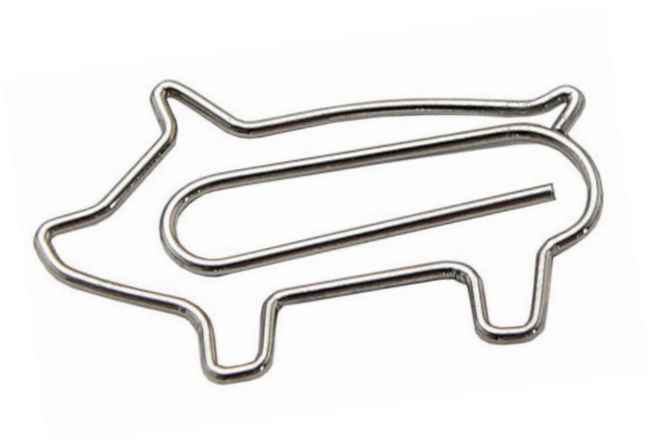
REFERENCES

1. Nauta MJ. A modular process risk model structure for quantitative microbiological risk assessment and its application in an exposure assessment of Bacillus cereus in a REPFED. Bilthoven, the Netherlands: RIVM Report, 2001.
2. Lindqvist R, Sylven S, Vagsholm I. Quantitative microbial risk assessment exemplified by *Staphylococcus aureus* in unripened cheese made from raw milk. International Journal of Food Microbiology, 2002; 78(1–2):155–70.
3. Rosenquist H, Nielsen NL, Sommer HM. Quantitative risk assessment of human campylobacteriosis associated with thermophilic Campylobacter species in chickens.

International Journal of Food Microbiology, 2003; 83(1): 87–103.
4. Cassin MH, Lammerding AM, Todd EC. Quantitative risk assessment for *Escherichia coli* O157:H7 in ground beef hamburgers. International Journal of Food Microbiology, 1998; 41(1):21–44.
5. Nauta M. The Modular Process Risk Model (MPRM): A structured approach to food chain exposure assessment. Pp. 99–136 in Microbial Risk Analysis of Foods. Washington, DC: ASM Press, 2008.
6. EFSA Quantitive Microbial Risk Assessment on *Salmonella* in Slaughter and Breeder Pigs. Final Report, CFP/EFSA/BIOHAZ/2007/01, March 9, 2009.
7. Barker GC, Gomez N, Smid J. An introduction to biotracing in food chain systems. Trends in Food Science and Technology, 2009; 20(5):220–226.
8. Pearl J. Probabilistic Reasoning in Intelligent Systems. Networks of Plausible Inference. San Francisco, CA: Morgan Kaufmann Editors, 1988.
9. Albert I, Grenier E, Denis JB. Quantitative risk assessment from farm to fork and beyond: A global Bayesian approach concerning food-borne diseases. Risk Analysis, 2008; 28(2):557–571.
10. Binter C, Straver JM, Haggblom P. Transmission and control of Salmonella in the pig feed chain: A conceptual model. International Journal of Food Microbiology, Epub 2010.
11. Cullen AC, Frey HC. Probabilistic Techniques in Exposure Assessment. New York: P Press, 1999.
12. Frey HC, Mokhtari A, Tanwir D. Evaluation of Selected Sensitivity Analysis Methods Based upon Applications to Two Food Safety Process Risk Models. Raleigh, NC: Report Prepared for Office of Risk Assessment and Cost-Benefit Analysis, U.S, 2003.
13. Smid JH, Verloo D, Barker GC. Strengths and weaknesses of Monte Carlo simulation models and Bayesian belief networks in microbial risk assessment. International Journal of Food Microbiology, 2010; 139:S57–S63.
14. Jensen FV. Introduction to Bayesian Networks. New York: Springer-Verlag, 1996.
15. Lauritzen SL, Spiegelhalter DJ. Local computations with probabilities on graphical structures and their application to expert systems. Journal of the Royal Statistical Society, 1988; 50(2):157–224.
16. Gilks WR, Richardson S, Spiegelhalter DJ. Markov Chain Monte Carlo in Practice: Interdisciplinary Statistics. Boca Raton, FL: Chapman&Hall/CRC, 1996.
17. Soudant D, Beliaeff B, Thomas G. Dynamic linear Bayesian models in phytoplankton ecology. Ecological Modelling, 1997; 99(2–3): 161–169.
18. Needham CJ, Bradford JR, Bulpitt AJ. A primer on learning in Bayesian networks for computational biology. PLOS Computational Biology, 2007; 3(8):1409–1416.
19. Spiegelhalter D, Lauritzen SL. Sequential updating of conditional probabilities on directed graphical structures. Networks, 1990; 20: 579–605.
20. Coupé VMH, Van Der Gaag LC. Practicable sensitivity analysis of Bayesian belief nets. Pp. 81–86 in Joint Session of the 6th Prague Syposium of Asymptotic Statistics and the 13th Prague Conference on Information Theory, Statistical Decision Functions and Random Processes. Prague; 1998.
21. Kjaerulff UB, Madsen AL. Bayesian Networks and Influence Diagrams. New York: Springer, 2008.
22. Coupé VMH, Gaag LC, Habbema JDF. Sensitivity analysis: An aid for belief-network quantification. Knowledge Engineering Review, 2000; 15(3):215–232.
23. Warriss PD. Optimal lairage times and conditions for slaughter pigs: A review. Veterinary Record, 2003; 153(6):170–176.
24. Borch E, Nesbakken T, Christensen H. Hazard identification in swine slaughter with respect to foodborne bacteria. International Journal of Food Microbiology, 1996; 30(1–2):9–25.
25. Berends BR, Van Knapen F, Snijders JMA. Identification and quantification of risk factors regarding Salmonella spp. on pork carcasses. International Journal of Food Microbiology, 1997; 36(2–3):199–206.
26. Pearce RA, Bolton DJ, Sheridan JJ. Studies to determine the critical control points in pork slaughter hazard analysis and critical control point systems. International Journal of Food Microbiology, 2004; 90(3):331–339.
27. Morgan IR, Krautil FL, Cravel JA. Bacterial-populations on dressed pig carcasses. Epidemiology and Infection, 1989; 1:15–24.
28. Eustace I, Midgley J. An alternative process for cleaning knives used on meat slaughter floors. International Journal of Food Microbiology, 2007; 113(1):23–27.
29. Swanenburg M, Berends BR, Urlings HAP, Snijders JMA, Knapen FV. Epidemiological investigations into the sources of Salmonella contamination of pork. Berliner und Muenchener Tieraerztliche Wochenschrift, 2001; 114:356–359.
30. Hanssen EJM, Swanenburg M, Berends BR, Urlings HAP, Snijders JMA, Knapen F, Van Maassen CBM. The Dutch Salmonella monitoring programme for pigs and some recommendations for control plans in the future. Pp. 169–172 in SafePork. Verona, 2007.
31. Jessen B. Biofilm and disinfection in meat processing plants. In International Symposium on Disinfection & Hygiene. Wageningen, 2001.
32. Joseph B, Otta SK, Karunasagar I. Biofilm formation by salmonella spp. on food contact surfaces and their sensitivity to sanitizers. International Journal of Food Microbiology, 2001; 64(3):367–372.
33. Peel B, Simmons GC. Factors in the spread of Salmonellas in meatworks with special reference to contamination of knives. Australian Veterinary Journal, 1978; 54:106–110.
34. Warriner K, Aldsworth TG, Kaur S. Cross-contamination of carcasses and equipment during pork processing. Journal of Applied Microbiology, 2002; 93(1):169–177.
35. Donald M, Cook A, Mengersen K. Bayesian network for risk of diarrhea associated with the use of recycled water. Risk Analysis, 2009; 29(12):1672–1685.
36. Olesen KG, Lauritzen SL, Jensen FV. aHUGIN: A system creating adaptive probabilistic networks. Pp. 223–229 in Eighth Conference on Uncertainty in Artificial Intelligence, Stanford, CA 1992.

Chapter 4

A quantitative approach towards a better understanding of the dynamics of *Salmonella* spp. in a pork slaughter-line



A.H.A.M. van Hoek [a,b], R. de Jonge [b], W.M. van Overbeek [a,b], E. Bouw [a,b], A. Pielaat [b], J.H. Smid [b], B. Malorny [c], E. Junker [c], C. Löfström [d], K. Pedersen [d], H.J.M. Aarts [a,b], L. Heres [e]

[a] RIKILT, Wageningen UR, Akkermaalsbos 2, 6708 WB Wageningen, The Netherlands

[b] RIVM, A. van Leeuwenhoeklaan 9, 3720 BA Bilthoven, The Netherlands

[c] BFR, National Salmonella Reference Laboratory, 12277 Berlin, Germany

[d] National Food Institute, Technical University of Denmark, Mørkhøj Bygade

International Journal of Food Microbiology 2012; 153: 45–52

International Journal of Food Microbiology 153 (2012) 45–52

Contents lists available at SciVerse ScienceDirect

International Journal of Food Microbiology

journal homepage: www.elsevier.com/locate/ijfoodmicro

A quantitative approach towards a better understanding of the dynamics of *Salmonella* spp. in a pork slaughter-line

Angela H.A.M. van Hoek [a,b,*], Rob de Jonge [b], Wendy M. van Overbeek [a,b], El Bouw [a,b], Annemarie Pielaat [b], Joost H. Smid [b], Burkhard Malorny [c], Ernst Junker [c], Charlotta Löfström [d], Karl Pedersen [d], Henk J.M. Aarts [a,b], Lourens Heres [e]

[a] RIKILT, Institute of Food Safety, Wageningen UR, Akkermaalsbos 2, 6708 WB Wageningen, The Netherlands
[b] National Institute for Public Health and the Environment (RIVM), Centre for Infectious Disease Control (CIb), Antonie van Leeuwenhoeklaan 9, 3721 MA Bilthoven, The Netherlands
[c] Federal Institute for Risk Assessment (BFR), National Salmonella Reference Laboratory, 12277 Berlin, Germany
[d] National Food Institute, Technical University of Denmark, Mørkhøj Bygade 19, 2860 Søborg, Denmark
[e] VION Fresh Meat West, Boseind 10, 5281RM Boxtel, The Netherlands

ARTICLE INFO

Article history:
Received 8 June 2011
Received in revised form 10 October 2011
Accepted 24 October 2011
Available online 30 October 2011

Keywords:
Salmonella
Pigs
PCR
Quantitative
Resident flora
Slaughterhouse

ABSTRACT

Pork contributes significantly to the public health disease burden caused by *Salmonella* infections. During the slaughter process pig carcasses can become contaminated with *Salmonella*. Contamination at the slaughter-line is initiated by pigs carrying *Salmonella* on their skin or in their faeces. Another contamination route could be resident flora present on the slaughter equipment. To unravel the contribution of these two potential sources of *Salmonella* a quantitative study was conducted. Process equipment (belly openers and carcass splitters), faeces and carcasses (skin and cutting surfaces) along the slaughter-line were sampled at 11 sampling days spanning a period of 4 months.

Most samples taken directly after killing were positive for *Salmonella*. On 96.6% of the skin samples *Salmonella* was identified, whereas a lower number of animals tested positive in their rectum (62.5%). The prevalence of *Salmonella* clearly declined on the carcasses at the re-work station, either on the cut section or on the skin of the carcass or both (35.9%). Throughout the sampling period of the slaughter-line the total number of *Salmonella* per animal was almost 2log lower at the re-work station in comparison to directly after slaughter.

Seven different serovars were identified during the study with *S.* Derby (41%) and *S.* Typhimurium (29%) as the most prominent types. A recurring *S.* Rissen contamination of one of the carcass splitters indicated the presence of an endemic 'house flora' in the slaughterhouse studied. On many instances several serotypes per individual sample were found.

The enumeration of *Salmonella* and the genotyping data gave unique insight in the dynamics of transmission of this pathogen in a slaughter-line. The data of the presented study support the hypothesis that resident flora on slaughter equipment was a relevant source for contamination of pork.

1. Introduction

Salmonellosis is an important cause of food-borne human gastro-enteritis in most European countries (EFSA, European Food Safety Authority, 2010; Valkenburgh et al., 2007). Farm animals and foods of animal origin form an important source of human *Salmonella* infections. In various European countries a significant number of human cases of salmonellosis (up to 25%) are described to be related to the consumption of pork and pork products (EFSA, European Food Safety Authority, 2006; van Pelt et al., 2000; Valdezate et al., 2005).

Carrier pigs are a predominant source of *Salmonella* contamination of pig carcasses during the slaughtering process (Alban and Stärk, 2005; Baptista et al., 2010; Berends et al., 1997; Borch et al., 1996). Pigs may already have *Salmonella* on their skin before entering a slaughterhouse and, despite stringent hygiene procedures during carcass processing, cross contamination to both *Salmonella* positive and negative carcasses can occur. The slaughter-line itself can become contaminated by faeces of carrier pigs. In addition, the presence of endemic 'house flora' of *Salmonella* has been described for several slaughterhouses (Baptista et al., 2010; Hald et al., 2003; Visscher et al., 2011; Warriner et al., 2002).

European data on the prevalence of *Salmonella* contaminated carcasses and on serotypes of *Salmonella* on the carcasses is available in

* Corresponding author at: National Institute for Public Health and the Environment, Centre for Infectious Disease Control, Laboratory for Zoonoses and Environmental Microbiology, P.O. Box 1, 3720 BA Bilthoven, The Netherlands. Tel.: +31 30 2747058; fax: +31 30 2744434.
E-mail address: angela.van.hoek@rivm.nl (A.H.A.M. van Hoek).

0168-1605/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.ijfoodmicro.2011.10.013

various papers. For example, Hald et al. (2003) documented that the prevalence of *Salmonella* contaminated carcasses varied between 0 and 8.5% among 1623 carcasses examined from five different countries. An EFSA study (26 countries; 5736 carcass samples) reported a prevalence of *Salmonella* positive carcasses of 0–20% (EFSA, European Food Safety Authority, 2008). The most frequently isolated serotype in both studies was *S.* Typhimurium.

The aim of this study was to investigate the dynamics of *Salmonella* in a pig slaughtering process and to assess the origin of carcass contamination. Hereto, the prevalence of *Salmonella* contaminated carcasses was determined. In addition, the concentration of this pathogen was measured at different sites on the pork meat and slaughtering equipment throughout the slaughtering-line by sampling individual carcasses at exsanguination up to the re-work station. *Salmonella* isolates were serotyped and genotyped.

2. Materials and methods

2.1. Slaughterhouse characteristics

The Dutch slaughterhouse investigated in this study was partly automated with robots for pre-cutting, belly opening, rectum drilling, splitting, leaf lard removal, neck cutting and marking. The capacity of the slaughterhouse is 650 pigs per hour, and 5000–6000 animals per day. The waiting time for the pigs at the slaughterhouse was as short as possible (not more than 2 h). Before entering the slaughter-line pigs were electrically stunned, sticked on a table, scalded in a tank, dehaired, flamed, wet polished, flamed and wet polished for a second time.

The belly opener cuts open the belly of a carcass and then cleaves the breastbone into two symmetrical parts. The carcass splitter cuts a carcass into two equal halves with a double knife, without cutting the head.

2.2. Sampling strategy

Carcass and equipment samples were collected on 11 days over a period of 4 months. Different herds were sampled on one sampling day, with a preference of two animals per herd, to account for herd variability. In total, 118 pigs and their carcasses were sampled at two steps of the slaughter process (see Fig. 1 for exact sampling sites). Directly after exsanguination, skin and rectal samples were taken for the detection, enumeration and typing of *Salmonella*. Immediately after exsanguination 4 cork borer samples were obtained from the shoulder of the animal. A sterile hand held cork borer was used to make four incisions on the shoulder. With a sterile scalpel and forceps slices of 5 cm^2 with a thickness of approximately 5 mm, were cut from the carcass. The four tissue samples, representing a total of 20 cm^2, were collected in one sterile plastic bag, constituting one sample. In addition, a rectal sample was taken from the same animal with a sterile swab (Transwab, Medical Wire and equipment Co. Ltd., Corsham, Wilts., England), which was immediately placed in 6 ml Buffered Peptone Water (BPW; bioTRADING Benelux B.V., Mijdrecht, The Netherlands).

The carcasses sampled at exsanguination were tracked in the slaughter-line and sampled again after meat inspection at the re-work station. From the cutting site, ham, back before pelvis, sternum and shoulder muscle were sampled with the cork borer. From the lard side, samples were taken with the cork borer from the back, the jowl, the ham and the belly. These interior and exterior samples were collected separately in two sterile plastic bags. In this way a paired set of 2×2 different samples were obtained from each animal; two at exsanguination (shoulder (EE), faeces (FS)) and two after final meat inspection at the re-work station (exterior (RE), interior (RI)).

In the slaughter-line the sets of parallel operating belly openers (BO) as well as the carcass splitters (CS) were sampled prior to the start and at the end of the day, immediately after finishing with the slaughtering process. Blades and other easy to reach contact surfaces from the belly openers and the splitting robots were swabbed on both sites using the Meat/Turkey carcass sampling kit (Nasco, Fort Atkinson, WI). In addition, sterile flexistem brushes were used for sampling of parts of the equipment which were less accessible with the carcass sampling kit.

All samples were cooled on site and transported to the laboratory to be analysed on the same day of collection.

2.3. Detection of Salmonella

Cork borer samples were weighed after arrival in the laboratory and an equal volume of BPW was added. To rectal swabs, equipment

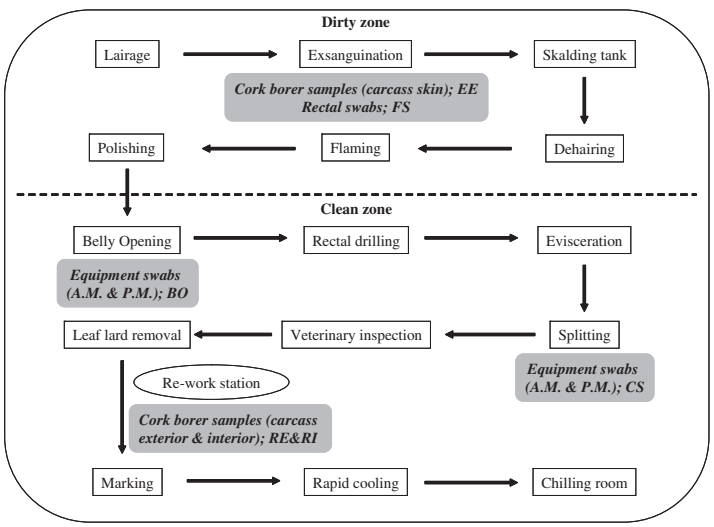


Fig. 1. Locations of the various cork borer samples, rectal and equipment swabs taken during the slaughtering process in the pig slaughterhouse investigated.

swabs and flexistem brushes 6, 20, and 40 ml of BPW, respectively, was added. Cork borer and equipment swab samples were homogenised for 1 min with a Stomacher 400 (Seward, Worthing, United Kingdom). Rectal swabs and equipment samples taken with a flexistem brush were vortexed for 30 s. A 5 ml aliquot was removed from each sample and stored at 4 °C for enumeration later (see next section). After addition of 90 ml BPW to the cork bore samples, rectal and equipment swabs, all samples were incubated without shaking at 37 °C for 18 to 20 h.

DNA was isolated from a 1 ml aliquot of the enriched culture, using a Chelex-100 suspension (50–100 mesh; Bio-Rad Laboratories B.V., Veenendaal, The Netherlands) according to the manufacturer's instructions. From the final DNA solution, a 5 µl aliquot was directly used as template in the PCR assay described below.

The *Salmonella* real-time assay described by Malorny et al. (2004), except for the internal amplification control, was used to determine the presence of DNA of this pathogen in the various samples. The 50 µl PCR mixture contained 0.4 µM of the primers ttr-4 and ttr-6, 0.25 µM ttr-5 probe (5′-FAM, 3′-BHQ1), 1×Universal Mastermix (Diagenode sa, Liège, Belgium) and a 5 µl aliquot of the sample DNA. Conditions for the real-time PCR were 95 °C for 1 min followed by 45 cycles of 95 °C for 15 s and 65 °C for 30 s. PCR tests were performed on a iQ™5 Cycler (Bio-Rad Laboratories B.V., Veenendaal, The Netherlands) and data was analysed using the Bio-Rad iQ5 software (Version 2.0).

Samples that were found positive by PCR were considered to be true positives for the assessment of the *Salmonella* prevalence (the cut off value was set at threshold cycle Ct 40 as result of an internal house validation process).

2.4. Enumeration of Salmonella

The most probable number (MPN; de Man, 1983) method was used to estimate *Salmonella* numbers in the samples identified as positive by PCR. Three subsequent 10-fold serial dilutions were prepared from the stored 5 ml of the original samples. In triplicate 1 ml of each dilution was added to 9 ml of BPW and enriched for 18 ± 2 h at 37 °C. Three separate and equally spaced drops of incubated BPW (total 100 µl) were pipetted onto the surface of a Modified Semi-solid Rappaport Vassiliadis (MSRV) medium base plate (Merck B.V., Schiphol-Rijk, The Netherlands) supplemented with Novobiocin (20 mg l^{-1}) (Oxoid B.V., Badhoevedorp, The Netherlands) in a triangular configuration. MSRV plates were incubated at 41.5 °C and examined after 24 and 48 h for suspect *Salmonella* growth. A sterile loop (1 µl) was dipped into the edge of any opaque growth and streaked onto SM® ID2 agar plates (BioMérieux SA, Marcy l'Etoile, France) which were incubated at 37 °C for 24 h for the confirmation of *Salmonella*.

To compute the MPN per ml of BPW, it was assumed that all *Salmonella* were detached from the cork borer sample of the carcass surface and brought into the BPW during stomaching. The MPN per ml was converted to MPN per cm^2. Hereto, it was assumed that the bacteria were homogeneously spread over the carcass skins. *Salmonella* numbers per gram of faeces from the rectal swab data were also assessed. The amount of faeces on the swab was not determined during the sampling experiment. Therefore, a small study was performed afterwards in which 50 swabs were weighted before and after insertion into pigs' rectums. The mean amount of faeces that was found on a swab was used to estimate the number of *Salmonella* per gram faeces, using the MPN per swab.

2.5. Statistical data analysis

A beta distribution was used to describe uncertainty about the prevalence estimates of *Salmonella* on site or at day level (Vose, 2000). For further analysis of the quantitative *Salmonella* data, the hypothesis that the variation in the Log of all MPN data at one sampling site for Salmonella-positive carcasses can be expressed by a Normal distribution was verified by visually checking its fit to Normality in a quantile–quantile plot. If, by this test, no deviations from normality could be seen, then the *per day* variation in the MPN data from one sampling site was expressed by a Log-Normal $(\hat{\mu}, \hat{\sigma})$ distribution. The parameters of this distribution were estimated using maximum likelihood estimation, yielding the estimators $\hat{\mu}$ (mean) and $\hat{\sigma}$ (standard error). Samples that were positive by PCR, but in which no *Salmonella* was detected in the dilution series for the MPN assessment, were taken into account and regarded as censored positives. For the censored numbers, the cumulative Log-Normal $(\hat{\mu},\hat{\sigma})$ distribution function was used to represent the probability of being an observation below detection limit (Gelman et al., 2004). Such concentration distributions could, however, not be assessed for all days. If most, or all, samples were negative in the MPN dilution series on one day, then $\hat{\mu}$ and $\hat{\sigma}$ could not be estimated. For such data sets only the upper limit of the expected concentration $\hat{\sigma}$, as provided by the minimal MPN, is given.

2.6. Sero- and genotyping of Salmonella

Depending on the *Salmonella* concentrations, one to a maximum of five (representative) isolates from each sample was randomly selected. All isolates were stored at -70 °C until use.

The multiplex PCR described by Lim et al. (2003) was used to discriminate between *S.* Typhimurium and non-Typhimurium serotypes in the numerous isolates from the slaughterhouse. The non-Typhimurium isolates were subsequently serotyped by slide and tube agglutination following the Kauffmann–White scheme (Grimont and Weill, 2007).

Multiple-locus variable-number of tandem-repeat analysis (MLVA) was performed on the (monophasic) *S.* Typhimurium isolates as described previously (Torpdahl et al., 2007) to determine whether the isolates were epidemiologically related. Only one (monophasic) *S.* Typhimurium isolate per sample was analysed by MLVA. The MLVA repeats were calculated and named according to the method described by Lindstedt et al. (2004).

Pulsed-field gel electrophoresis (PFGE) was carried out on *S.* Derby and *S.* Rissen isolates with the XbaI restriction enzyme according to the Pulse-Net protocol (Ribot et al., 2006). Gels were analysed using BioNumerics 6.5 software. A dendrogram was produced using the Dice coefficient and the unweighted pair-group method (UPGMA) with a 1.5% tolerance limit and 1.5% optimisation.

3. Results

3.1. Salmonella screening and enumeration

Salmonella was identified on the skin surfaces of 96.6% of all carcasses sampled at exsanguination (Table 1). The estimated mean concentration $(\hat{\mu})$ of *Salmonella* per day in the samples at this site varied between 0.04 and 1.75 log MPN cm^{-2} (Table 2). Of the rectal swabs taken directly after exsanguination 62.5% was identified positive, whereas the average number of *Salmonella* was 1.88 ± 1.42 log MPN $g.^{-1}$. At the re-work station, 16.2% and 29.9% of the exterior and interior samples, respectively, were tested positive for *Salmonella* (Table 1). In addition, the pathogen counts were lower in comparison to samples taken at exsanguination, with maximum estimated numbers of *Salmonella* of 0.11 and -0.13 log MPN cm^{-2} on the carcass surface (exterior) and cut section (interior), respectively (Table 2). Of all the samples taken in this study, 44.5% (265/596) were identified as *Salmonella* positive. The prevalence of *Salmonella* on the different carcass sampling sites varied between sampling days (Table 1). For the carcass samples collected at the re-work station, an increase in *Salmonella* prevalence was observed from around the second half of the sampling period (08-06-2009 till 16-06-2009), especially for samples collected from the interior part of the carcass. The prevalence

Table 1
Number of *Salmonella* positive samples per sampling date and per sampling site determined by real-time PCR.

Date	Robots — before slaughter[a]				Exsanguination		Re-work station		Robots — after slaughter[a]			
	BO1	BO2	CS1	CS2	Carcass skin	Faeces	Exterior	Interior	BO1	BO2	CS1	CS2
14-04-2009	nd	nd	nd	nd	6/6	nd	2/6	1/6	0/1	0/1	0/1	0/1
20-04-2009	0/1	0/1	0/1	0/1	12/12	10/12	0/12	0/12	0/1	0/1	0/1	0/1
11-05-2009	0/1	0/1	0/1	0/1	8/8	5/8	0/8	0/8	0/2	0/2	0/2	0/2
25-05-2009	0/1	0/1	0/1	0/1	11/12	5/12	2/11	0/11	0/2	0/2	0/2	2/2
02-06-2009	0/1	0/1	0/1	0/1	12/12	6/12	0/12	3/12	0/2	0/2	0/2	2/2
08-06-2009	0/1	0/1	0/1	0/1	12/12	11/12	5/12	8/12	0/2	0/2	0/3	3/3
09-06-2009	0/1	0/1	0/1	0/1	8/8	6/8	2/8	3/8	0/2	1/2	0/2	2/2
15-06-2009	0/1	0/1	0/1	0/1	12/12	8/12	2/12	8/12	0/2	0/2	0/4	4/4
16-06-2009	nd	nd	nd	nd	12/12	8/12	2/12	7/12	1/2	1/2	0/3	3/3
13-07-2009	0/2	0/2	0/2	2/2	12/12	8/12	3/12	2/12	0/2	0/2	0/2	2/2
14-07-2009	0/2	0/2	0/2	2/2	9/12	3/12	1/12	3/12	0/2	0/2	0/2	2/2
Total	0/11	0/11	0/11	4/11	114/118	70/112	19/117	35/117	1/20	2/20	0/24	20/24

Note: nd = not determined.
[a] BO = Belly opener; CS = Carcass splitter.

Table 2
Estimated parameters (sample mean, $\hat{\mu}$, and standard error, $\hat{\sigma}$) of the Log-Normal probability distribution representing the concentration of positive samples.

Date	Exsanguination				Re-work station			
	Carcass skin (log MPN/cm^2)		Faeces (log MPN/g)		Exterior (log MPN/cm^2)		Interior (log MPN/cm^2)	
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\mu}$	$\hat{\sigma}$
14-04-2009	1.75	0.59	nd		<-0.51		<-0.35	
20-04-2009	0.47	0.49	2.71	0.98				
11-05-2009	0.42	0.71	1.91	1.19				
25-05-2009	0.26	0.75	2.31	1.36	<-0.79			
02-06-2009	0.46	0.84	2.7	0.68			<-0.61	
08-06-2009	0.04	0.91	2.11	1.02	0.11	0.53	-0.13	1.12
09-06-2009	0.52	0.71	2.35	0.78	<-0.52		-0.47	1.04
15-06-2009	0.60	0.61	-3.32	4.96	-0.42	0.82	-0.31	0.72
16-06-2009	0.92	1.33	2.75	1.43	<-0.80		-0.32	0.43
13-07-2009	0.59	1.00	2.61	1.22	-0.98	0.47	-0.37	0.99
14-07-2009	0.34	0.34	2.65	0.6	<-0.83		<-0.59	
Average	0.58	0.75	1.88	1.42	-0.43	0.61	-0.32	0.86
Per month:								
April–May	0.75	0.80	2.43	1.08	<-0.51		<-0.35	
May–June	0.31	0.83	2.33	1.00	-0.38	0.69	-0.49	1.09
June–July	0.63	0.95	2.25	1.54	-1.30	1.03	-0.39	0.64

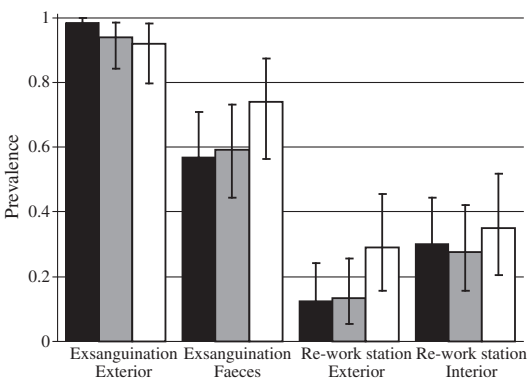


Fig. 2. *Salmonella* prevalence data at the different carcass sampling sites determined by real-time PCR. The black bars represent the samples taken at approximately 11 AM; the grey ones indicate the samples taken at approximately 1 PM; the white bars show the samples taken at approximately 3 PM.

declined again towards the end of the experiment. Within one day no clear increase of *Salmonella* positive samples could be demonstrated, i.e. the prevalence of this pathogen in samples taken in the morning were not different from those obtained in the afternoon (Fig. 2).

Before slaughter, no *Salmonella* could be demonstrated on either belly openers, whereas at the end of slaughter 3 out of 40 samples (7.5%) were tested positive. On one sampling day, *Salmonella* was identified on both belly openers (Table 1).

Samples taken from the carcass splitters were more frequently found to harbour *Salmonella*. More specifically, carcass splitter number 2 (CS2) was repeatedly contaminated with this pathogenic microorganism. In total, during eight out of the eleven sampling days *Salmonella* was identified on this robot after the end of slaughter. Moreover, on two consecutive days *Salmonella* was already found on CS2 at the beginning of the slaughtering process (Table 1). In all cases, equipment swabs and flexistem brushes had equal test results.

3.2. Salmonella serotypes

In total, 620 *Salmonella* isolates were obtained from all samples taken during this study. Because *S.* Typhimurium was expected to be the most prevalent serovar in pigs (Hald et al., 2003; EFSA, European Food Safety Authority, 2008), the multiplex PCR described by Lim et al. (2003) was used to discriminate *S.* Typhimurium isolates from other serovars. The PCR results revealed that 67.5% of all salmonellae isolated at the slaughterhouse were non-Typhimurium

Table 3
Salmonella serovars per sampling day determined by multiplex PCR and serotyping.

Date	Serovar[a]							
	BDY	BEG	DRB	INS	mSTM	RSN	STM	Unknown
14-04-2009	0	0	4	6	14	0	3	0
20-04-2009	0	0	55	1	2	0	1	0
11-05-2009	0	0	20	0	6	0	8	0
25-05-2009	0	0	4	0	0	0	16	0
02-06-2009	0	0	35	0	1	5	11	0
08-06-2009	5	16	33	0	0	13	26	1
09-06-2009	0	32	12	0	0	4	14	0
15-06-2009	1	6	41	0	0	20	12	0
16-06-2009	0	0	27	0	0	6	42	1
13-07-2009	3	30	5	0	0	13	24	0
14-07-2009	0	0	19	0	0	2	22	0
% of total	1.4%	13.5%	41.0%	1.1%	3.7%	10.1%	28.8%	0.3%

Note: The multiplex PCR has been described by Lim et al. (2003). The serotyping was performed by slide and tube agglutination following the Kauffmann–White scheme.
[a] BDY: *S.* Bredeney; BEG: *S.* Brandenburg; DRB: *S.* Derby; INS: *S.* Infantis; mSTM: monophasic *S.* Typhimurium; RSN: *S.* Rissen; STM: *S.* Typhimurium.

Table 4
Salmonella serovars per sampling site determined by multiplex PCR and serotyping.

Sample place[a]		Serotype[b]							
		BDY	BEG	DRB	INS	mSTM	RSN	STM	Unknown
Robots — before slaughter	BO1	0	0	0	0	0	0	0	0
	BO2	0	0	0	0	0	0	0	0
	SP1	0	0	0	0	0	0	0	0
	SP2	0	0	2	0	0	6	0	0
Exsanguination	Skin	7	63	118	6	17	0	97	1
	Faeces	2	17	47	1	2	0	60	0
Re-work station	Carcass exterior	0	3	5	0	3	10	8	0
	Carcass interior	0	1	37	0	1	12	8	1
Robots — after slaughter	BO1	0	0	1	0	0	0	0	0
	BO2	0	0	0	0	0	0	6	0
	CS1	0	0	0	0	0	0	0	0
	CS2	0	0	45	0	0	35	0	0

Note: The multiplex PCR has been described by Lim et al. (2003). The serotyping was performed by slide and tube agglutination following the Kauffmann–White scheme.
[a] BO = Belly opener, CS = Carcass splitter.
[b] BDY: *S.* Bredeney; BEG: *S.* Brandenburg; DRB: *S.* Derby; INS: *S.* Infantis; mSTM: monophasic *S.* Typhimurium; RSN: *S.* Rissen; STM: *S.* Typhimurium.

isolates. Because of this very large set, it was decided to serotype the main part (64%). When not all isolates from one sample were typed, the result of the subset of typed isolates was assumed to reflect the serotypes of the non-typed ones.

Overall, seven different serotypes were identified, i.e. *S.* 4,5,12:i:– (from here on called monophasic *S.* Typhimurium), *S.* Bredeney, *S.* Brandenburg, *S.* Derby, *S.* Infantis, *S.* Rissen and *S.* Typhimurium (Table 3). Six serotypes were characterised from the animals entering the slaughterhouse, whereas only five different serovars were identified on the carcasses after slaughtering, and only three serotypes were isolated from the slaughterhouse equipment sampled. The most prominent serovars identified at the carcass at exsanguination and their rectal swabs were *S.* Derby (38%), *S.* Typhimurium (36%) and *S.* Brandenburg (18%) (Table 4). The serotypes frequently isolated from the carcasses at the end of the slaughter-line were *S.* Derby (47%) and *S.* Rissen (25%), whereas *S.* Typhimurium was only found

Table 5
MLVA types distribution among the various samples.

Serovar	Allele string	Date	Exsanguination		Re-work station		BO2[a]
			Carcass skin	Faeces	Exterior	Interior	
Monophasic *S.* Typhimurium	02-03-19-14-02	02/06/2009		1			
	02-06-04-00-02	14/04/2009	4		3	1	
		20/04/2009	1				
	02-07-06-00-02	11/05/2009	2	1			
S. Typhimurium	02-02-05-00-02	09/06/2009	3		1		1
		13/07/2009			1		
		14/07/2009	7			1	
	02-03-19-01-02	02/06/2009	1				
		13/07/2009	1				
	02-03-19-14-02	25/05/2009	1				
		02/06/2009	1	4			
		08/06/2009	5	5			
		09/06/2009		1			
		15/06/2009	4	2			
		13/07/2009	5	1	1		
	02-05-05-00-02	11/05/2009	5	2			
	02-05-06-00-03	25/05/2009		2			
		08/06/2009		1			
	02-05-20-00-02	16/06/2009	9	3			
	02-06-04-00-02	14/04/2009	1				
	02-07-09-08-03	16/06/2009	2	2	1	1	1
	02-07-10-08-03	16/06/2009		1			
	02-07-11-06-03	02/06/2009	1				
		08/06/2009	1				
	02-08-09-05-03	20/04/2009		1			
	02-11-06-00-03	14/07/2009	1				
	02-17-05-00-02	09/06/2009	1				
	03-02-04-13-02	08/06/2009		2			
	03-03-20-05-02	14/07/2009				1	
	03-04-04-22-02	14/04/2009	2				
		02/06/2009				1	
		16/06/2009				2	
		14/07/2009		1			
	03-08-13-19-02	14/07/2009			1		
	04-01-17-14-02	25/05/2009	8	3	2		
	06-03-00-00-01	13/07/2009		1			

Note.
[a] BO = Belly opener.

in 18% of the cases. The predominant *Salmonella* serotype isolated at the slaughterhouse varied by day of the study.

Although *S.* Typhimurium was prominently present on the carcasses at exsanguination and to a lesser extent at the re-work station, this serovar was not isolated from the carcass splitters. In contrast on the belly openers *S.* Typhimurium was found in two out of three occasions. Carcass splitter 2 (CS2) was frequently contaminated with serovars Derby (56%) and Rissen (44%).

In 15% of all *Salmonella* positive incidences multiple serovars were isolated from individual samples. This was especially true for carcasses at exsanguination (data not shown).

3.3. Salmonella genotypes

At least one *S.* Typhimurium or monophasic *S.* Typhimurium isolate from each individual swab or carcass sample (80 animals, 119 isolates in total) positive for these serovars was typed by multiple-locus variable-number of tandem-repeat analysis (MLVA). Nineteen and three different MLVA types could be distinguished among the *S.* Typhimurium and monophasic *S.* Typhimurium isolates analysed, respectively (Table 5).

In 18 cases the same MLVA type was detected in both the rectal swab and exterior sample at exsanguinations, whereas 5 times different MLVA types were encountered in these samples. The 17 *S.* Typhimurium and monophasic *S.* Typhimurium isolates originating from carcasses at the re-work station matched with MLVA types isolated at exsanguination from the same animals, except in three instance (Table 6; Animals 149_1, 657_1 and 657_2).

Table 6
All paired occurrences of (monophasic) *S.* Typhimurium typed by MLVA on single carcasses.

Date	Herd_animal	Origin[a]	Serovar[b]	MLVA allele string
14-04-2009	A_1	EE-RE	mSTM	02-06-04-00-02
	C_1	EE-RE-RI	mSTM	02-06-04-00-02
11-05-2009	396_1	EE-FS	mSTM	02-07-06-00-02
	396_2	EE-FS	STM	02-05-05-00-02
	646_2	EE-FS	STM	02-05-05-00-02
25-05-2009	723_1	EE, FS	STM	04-01-17-14-02, 02-05-06-00-03
	787_1	EE-FS	STM	04-01-17-14-02
	787_3	EE-FS-RE	STM	04-01-17-14-02
	900_1	EE-FS	STM	04-01-17-14-02
	900_2	EE-RE	STM	04-01-17-14-02
02-06-2009	826_1	EE, FS	STM	02-03-19-01-02, 02-03-19-14-02
08-06-2009	431_1	EE-FS	STM	02-03-19-14-02
	611_2	EE-FS	STM	02-03-19-14-02
	921_1	EE, FS	STM	02-03-19-14-02, 03-02-04-13-02
	921_2	EE-FS	STM	02-03-19-14-02
	968_1	EE, FS	STM	02-03-19-14-02, 03-02-04-13-02
15-06-2009	532_1	EE-FS	STM	02-03-19-14-02
	921_1	EE-FS	STM	02-03-19-14-02
16-06-2009	662_1	EE-FS	STM	02-05-20-00-02
	662_2	EE-FS	STM	02-05-20-00-02
	657_1	EE-FS, RI	STM	02-05-20-00-02, 03-04-04-22-02
	657_2	EE, RI	STM	02-05-20-00-02, 03-04-04-22-02
	657_3	EE-FS	STM	02-07-09-08-03
	657_4	EE-FS-RE-RI	STM	02-07-09-08-03
13-07-2009	149_1	FS, RE	STM	03-04-04-22-02, 03-08-13-19-02
	921_1	EE-FS-RE	STM	02-03-19-14-02
	921_2	EE, FS	STM	02-03-19-14-02, 06-03-00-00-01

Note.
[a] EE: Exsanguination, exterior; FS: Exsanguination, Rectal swab; RE: Re-work station, exterior; RI: Re-work station, interior.
[b] mSTM: monophasic *S.* Typhimurium; STM: *S.* Typhimurium.

The two *S.* Typhimurium MLVA types detected on belly opener 2 (BO2) were also found on *Salmonella* samples originating from the incoming animals on those sampling days. In addition, both of these MLVA types were identified in samples taken at the re-work station (Table 5).

A selection of the *S.* Derby and *S.* Rissen isolates (n = 96) was genotyped using PGFE. The dendrogram (Fig. 3) shows that the *S.* Rissen

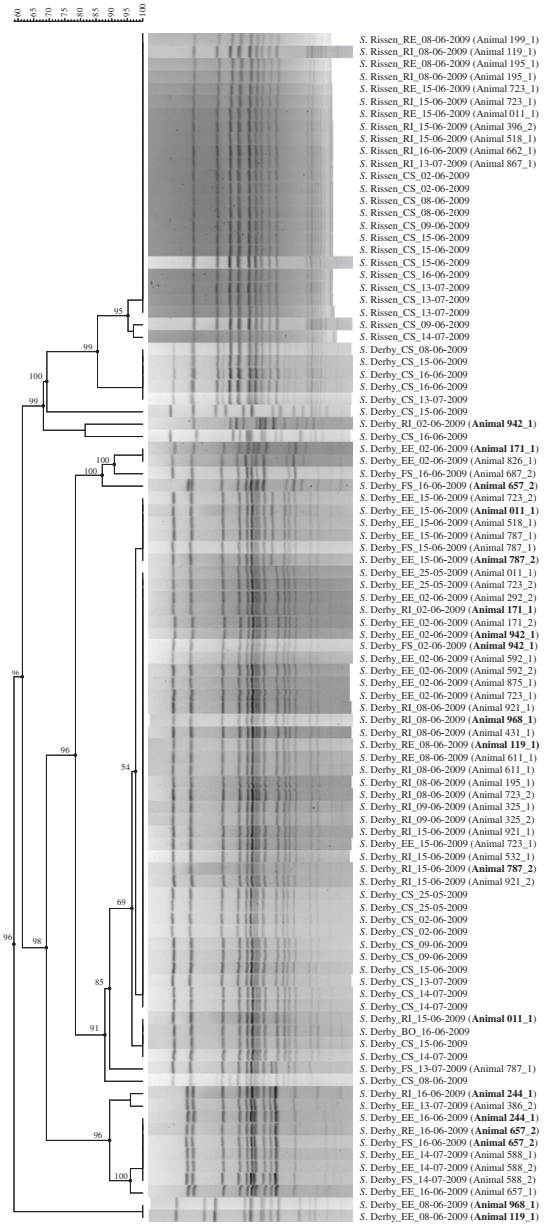


Fig. 3. PFGE dendrogram of *S.* Rissen and *S.* Derby isolates from the slaughter-line and pigs. BO: Belly opener; CS: Carcass splitter; EE: Exsanguination, exterior; FS: Rectal swab; RE: Re-work station, exterior; RI: Re-work station, interior. Sample names in bold indicate *S.* Derby isolates from individual carcasses isolated at different stages of the slaughter-line.

isolates belonged to one indistinguishable type, whereas the PFGE profiles varied among the *S.* Derby isolates analysed, although one particular *S.* Derby genotype clearly dominated the phylogenetic tree. Isolates belonging to this branch originated from various sampling days and all types of samples taken at the slaughterhouse, except the belly opener. In contrast, one branch with a PFGE pattern very similar to the *S.* Rissen profile contained 5 *S.* Derby isolates isolated only from the carcass splitter but at different sampling days.

From several individual carcasses, *S.* Derby was isolated at two or more sampling sites (Table 4). The phylogenetic tree in Fig. 3 includes some of these isolates (in bold). *S.* Derby isolates originating from rectal swabs (FS) and skin samples (EE) showed an identical PFGE pattern in 75% of the cases (n = 4), whereas, only different PFGE profiles were encountered among the exsanguination (EE) and the re-work station isolates of the same animal of this serovar (n = 8).

4. Discussion

The prevalence of *Salmonella* contaminated carcasses started with 96.6% at exsanguination and was 35.9% after slaughtering at the re-work station. The level of contaminated carcasses in this study was relatively high, compared to other studies (Bouvet et al., 2003; de Busser et al., 2011; Swanenburg et al., 2001a). This high level of *Salmonella* positive samples gave the opportunity to get a clear picture of the contamination routes.

At the re-work station, over 35% of the carcasses tested were *Salmonella* positive. In 10.3% of all tested carcasses, *Salmonella* was detected on both the cut section and on the skin, 19.7% of the tested carcasses were only contaminated at the cut section, and 6.0% contained *Salmonella* only on the skin. So the slaughter process reduces the number of skin contaminated carcasses from 96.6 to 16.2%. Cross contamination via the slaughter process was responsible for at least 30% of all carcasses, i.e. the carcasses were contaminated at the interior side. These results correspond to data reported by others (Berends et al., 1997; Botteldoorn et al., 2003). However, this cross contamination percentage might be an underestimated value since they do not take into account the genotypic diversity of *Salmonella* serovars. In the present study on the one hand the same MLVA type was found at exsanguination and re-work station (Table 6), but on the other hand it was clearly shown that genotypically different subtypes of the same *Salmonella* serotype can be present on one carcass at exsanguination and at the re-work station (see Fig. 3 and Table 5).

An excision technique was used as the sampling method for pig skins and carcasses. In many studies (Botteldoorn et al., 2003; EFSA, European Food Safety Authority, 2008; Hald et al., 2003; Oosterom et al., 1985; Swanenburg et al., 2001a, 2001b) dry-wet swabbing was the technique of choice. Comparison of both techniques showed that the excision technique was approximately 10-fold more sensitive, but there seemed to be no linear relationship between the two results (Hutchison et al., 2005; Martínez et al., 2010). In case of low concentrations, swabbing a large area is to be preferred above excision of a small area (Lindblad, 2007), since the excision techniques only samples 5 cm^2 per excision. The concentration data obtained in this study clearly showed that the level of contamination of the sampled carcasses was high enough to use the excision technique.

The average number of *Salmonella* per carcass was almost 2log lower at the end of the slaughter-line. On the skin (12,000 cm^2) a 10 fold lower number was found, i.e. 3.8 to 0.37 *Salmonella* per cm^2. At the cutting area (3000 cm^2), the average MPN of *Salmonella* was 0.48 per cm^2. As a consequence, the average number of *Salmonella* per carcass decreased from 44,050 (prevalence × concentration × surface; 0.966 × 3.8 × 12,000) at exsanguination to 1150 per carcass (0.162 × 0.37 × 12,000 + 0.299 × 0.48 × 3000) at the re-work station. As 37.5% of all salmonellae on carcasses at the re-work station were found on the cutting edges, cross contamination is responsible for more than 35% of all *Salmonella* on pork carcasses based on bacterial counts.

The seven *Salmonella* serovars identified in this study, i.e. *S.* Bredeney, *S.* Brandenburg, *S.* Derby, *S.* Infantis, monophasic *S.* Typhimurium (*Salmonella* 4,5,12:i:–), *S.* Rissen and *S.* Typhimurium were also described by various other authors on pigs at the slaughterhouse stage (Arguello et al., in press; Bouvet et al., 2003; de Busser et al., 2011; Hald et al., 2003; Swanenburg et al., 2001a).

At the re-work station, five different serovars were detected, whereas at exsanguination six *Salmonella* serotypes were characterised (Tables 3 and 4). Two serovars detected at exsanguination, i.e. *S.* Bredeney and *S.* Infantis, were not detected at the re-work station. It might be possible that the contamination level with these serovars was very low and that they disappeared during the slaughter process. In contrast, one serovar, i.e. *S.* Rissen, was not detected at exsanguination but was detected at the re-work station and on one of the carcass splitters. The companies own monitoring program reflected that this slaughterhouse encountered hygiene problems during and after the study (data not shown). The serological patterns (Tables 3 and 4) clearly indicated complicated contamination routes.

The phenomenon of multiple serovars present in individual samples (15%), especially in those taken from carcasses at exsanguination suggested an underestimation of *Salmonella* serotypes in pork, since routinely only one isolate per sample is serotyped.

In order to determine their origin, isolates of the serovars *S.* Rissen, *S.* Derby and (monophasic) *S.* Typhimurium were subtyped. The results differed per serotype. *S.* Rissen was not detected on any of the incoming pigs. Only one PFGE genotype was found on cutting areas of carcasses at the re-work station and on the carcass splitter on various sampling days. This result strongly suggested that resident house flora was a source of carcass contamination.

S. Derby showed the characteristics of a cross contaminator as none of the strains detected on a single carcass at the re-work station was detected on the same carcass at exsanguination. Comparing MLVA types of (monophasic) *S.* Typhimurium isolates on carcasses at exsanguination and re-work station revealed that (monophasic) *S.* Typhimurium can originate from pigs carrying *Salmonella* into the slaughterhouse. The observation that some carcasses at the re-work station contained MLVA types that were not detected on the same carcass at exsanguination, again showed that cross contamination from one carcass to another can also have occurred.

In this study the carcass splitter was identified as an important source of *S.* Rissen contamination. In previous assessments the carcass splitter has been considered an unimportant attributive source of *Salmonella*, because of the high infection status of the pigs entering the slaughterhouse, especially, if the splitter is equipped with automatic disinfection between each carcass and faecal contamination during evisceration is controlled (Berends et al., 1997; Borch et al., 1996). However, other reports showed that a significant *Salmonella* contamination via the slaughterhouse environment was caused by the carcass splitter (Sørensen et al., 1999; Swanenburg et al., 2001a, 2001b). In the present study slaughter equipment apparently contributed also to *Salmonella* on pig carcasses. Despite cleaning and disinfection, one of the robots was repeatedly contaminated with *S.* Rissen. Moreover, once this serovar was even present on this carcass splitter prior to the start of slaughter on that day and over the weekend (Fig. 3; *S.* Rissen; CS, 13-07-2009).

In the slaughterhouse studied, cross-contamination contributed significantly to the carcass contamination. Resident flora was detected throughout the study on one of the slaughter robots. The serovar identified, *S.* Rissen, contributed significantly to the contamination at the end of the slaughter-line, whereas it was not found on any of the incoming carcasses. In addition, serovars on carcasses at the re-work station were many times other types than the ones detected at exsanguination in skin and faeces samples. The data collected, especially the *Salmonella* enumeration results and the sero- as well as genotyping data, gave unique insight in the dynamics of transmission in a slaughter-line.

The sero- and genotyping data will be compared using a variety of statistical tests and implemented in a tracing scheme to predict the source of *Salmonella* on a carcass at the re-work station (Smid et al., submitted for publication).

Acknowledgements

This work was supported by the European Union funded Integrated Research Project BIOTRACER (contract FOOD-2006-CT-036272) under the 6th RTD Framework.

We would like to thank Paul Hengeveld for his help during sampling at the slaughterhouse.

References

Alban, L., Stärk, K.D.C., 2005. Where should the effort be put to reduce the *Salmonella* prevalence in the slaughtered swine carcass effectively? Preventive Veterinary Medicine 68, 63–79.

Arguello, H., Carvajal, A., Collazos, J.A., García-Feliz, C., Rubio, P., *in press*. Prevalence and serovars of *Salmonella enterica* on pig carcasses, slaughtered pigs and the environment of four Spanish slaughterhouses. Food Research International

Baptista, F.M., Dahl, J., Nielsen, L.R., 2010. Factors influencing *Salmonella* carcass prevalence in Danish pig abattoirs. Preventive Veterinary Medicine 95, 231–238.

Berends, B.R., van Knapen, F., Snijders, J.M.A., Mossel, D.A.A., 1997. Identification and quantification of risk factors regarding *Salmonella* spp. on pork samples. International Journal of Food Microbiology 36, 199–206.

Borch, E., Nesbakken, T., Christensen, H., 1996. Hazard identification in swine slaughter respect to foodborne bacteria. International Journal of Food Microbiology 30, 9–25.

Botteldoorn, N., Heyndrickx, M., Rijpens, N., Grijspeerdt, K., Herman, L., 2003. *Salmonella* on pig carcasses: positive pigs and cross contamination in the slaughterhouse. Journal of Applied Microbiology 95, 891–903.

Bouvet, J., Bavai, C., Rossel, R., Le Roux, A., Montet, M.P., Mazuy, C., Vernozy-Rozand, C., 2003. Evolution of pig carcass and slaughterhouse environment contamination by *Salmonella*. Revue de Médecine Vétérinaire 154, 775–779.

De Busser, E.V., Maes, D., Houf, K., Dewulf, J., Imberechts, H., Bertrand, S., De Zutter, L., 2011. Detection and characterization of *Salmonella* in lairage, on pig carcasses and intestines in five slaughterhouses. International Journal of Food Microbiology 145, 279–286.

De Man, J.C., 1983. MPN tables, corrected. European Journal of Applied Microbiology and Biotechnology 17, 301–305.

EFSA, European Food Safety Authority, 2006. Opinion of the scientific panel on biological hazards on "Risk assessment and mitigation options of *Salmonella* in pig production". The EFSA Journal 341, 1–131.

EFSA, European Food Safety Authority, 2008. Report of the task force on zoonoses data collection on the analysis of the baseline survey on the prevalence of *Salmonella* in slaughter pigs, Part A. The EFSA Journal 135, 1–111.

EFSA, European Food Safety Authority, 2010. The community summary report on trends and sources of zoonoses, zoonotic agents and food-borne outbreaks in the European Union in 2008. The EFSA Journal 8, 1496.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 2004. Bayesian Data Analysis. Chapman and Hall/CRC, London, UK.

Grimont, P.A.D., Weill, F.-X., 2007. Antigenic Formulae of the *Salmonella* Serovars (9th ed.). WHO Collaborating Center for Reference and Research on *Salmonella*, Institut Pasteur, Paris.

Hald, T., Wingstrand, A., Swanenburg, M., von Altrock, A., Thorberg, B.-M., 2003. The occurrence and epidemiology of *Salmonella* in European pig slaughterhouses. Epidemiology and Infection 131, 1187–1203.

Hutchison, M.L., Walters, L.D., Avery, S.M., Reid, C.-A., Wilson, D., Howell, M., Johnston, A.M., Buncic, S., 2005. A comparison of wet-dry swabbing and excision sampling methods for microbiological testing of bovine, porcine, and ovine carcasses at red meat slaughterhouses. Journal of Food Protection 68, 2155–2162.

Lim, Y.-H., Hirose, K., Izumiya, H., Arakawa, E., Takahashi, H., Terajima, J., Itoh, K., Tamura, K., Kim, S.-I., Watanabe, H., 2003. Multiplex polymerase chain reaction assay for selective detection of *Salmonella enterica* serovar Typhimurium. Japanese Journal of Infectious Diseases 56, 151–155.

Lindblad, M., 2007. Microbiological sampling of swine carcasses: a comparison of data obtained by swabbing with medical gauze and data collected routinely by excision at Swedish abattoirs. International Journal of Food Microbiology 118, 180–185.

Lindstedt, B.-A., Vardund, T., Aas, L., Kapperud, G., 2004. Multiple-locus variable-number tandem-repeats analysis of *Salmonella enterica* subsp. *enterica* serovar Typhimurium using PCR multiplexing and multicolor capillary electrophoresis. Journal of Microbiological Methods 59, 163–172.

Malorny, B., Paccassoni, E., Fach, P., Bunge, C., Martin, A., Helmuth, R., 2004. Diagnostic real-time PCR for detection of *Salmonella* in food. Applied and Environmental Microbiology 70, 7046–7052.

Martínez, B., Celda, M.F., Anastasio, B., García, I., López-Mendoza, M.C., 2010. Microbiological sampling of carcasses by excision or swabbing with three types of sponge or gauze. Journal of Food Protection 73, 81–87.

Oosterom, J., Dekker, R., de Wilde, G.J., van Kempen-de Troye, F., Engels, G.B., 1985. Prevalence of *Campylobacter jejuni* and *Salmonella* during pig slaughtering. The Veterinary Quarterly 7, 31–34.

Ribot, E.M., Fair, M.A., Gautom, R., Cameron, D.N., Hunter, S.B., Swaminathan, B., Barrett, T.J., 2006. Standardization of pulsed-field gel electrophoresis protocols for the subtyping of *Escherichia coli* O157:H7, *Salmonella*, and *Shigella* for PulseNet. Foodborne Pathogens and Disease 3, 59–67.

Smid, J.H., van Hoek, A. H. A. M., Aarts, H.J.M., Havelaar, A.H., Heres, L., de Jonge, R., Pielaat A., *submitted for publication*. Quantifying the sources of *Salmonella* contamination in a Dutch pig slaughter plant. Food Control.

Sørensen, L.L., Sørensen, R., Klint, K., Nielsen, B., 1999. Persistent environmental strains of *Salmonella infantis* at two Danish slaughterhouses, two case-stories. Proceedings of the 3rd International Symposium on Epidemiology and Control of *Salmonella* in Pork, Washington, DC, 4–7 August, pp. 285–286.

Swanenburg, M., Urlings, H.A.P., Snijders, J.M.A., van Keuzenkamp, D.A., Knapen, F., 2001a. *Salmonella* in slaughter pigs: prevalence, serotypes and critical control points during slaughter in two slaughterhouses. International Journal of Food Microbiology 70, 243–254.

Swanenburg, M., van der Wolf, P.J., Urlings, H.A.P., Snijders, J.M.A., van Knapen, F., 2001b. *Salmonella* in slaughter pigs: the effect of logistic slaughter procedures of pigs on the prevalence of *Salmonella* in pork. International Journal of Food Microbiology 70, 231–242.

Torpdahl, M., Sorensen, G., Lindstedt, B.A., Nielsen, E.M., 2007. Tandem repeat analysis for surveillance of human *Salmonella* Typhimurium infections. Emerging Infectious Diseases 13, 388–395.

Valdezate, S., Vidal, A., Herrera-Leon, S., Pozo, J., Rubio, P., Usera, M.A., Carvajal, A., Echeita, M.A., 2005. *Salmonella* Derby clonal spread from pork. Emerging Infectious Diseases 11, 694–698.

Valkenburgh, S., van Oosterom, R., Stenvers, O., Aalten, M., Braks, M., Schimmer, B., van de Giessen, A., van Pelt, W., Langelaar, M., 2007. Zoonoses and Zoonotic Agents in Humans, Food, Animals and Feed in The Netherlands 2003–2006 http://www.rivm.nl/bibliotheek/rapporten/330152001.pdf.

van Pelt, W., van Giessen, A., van Leeuwen, W., Wannet, W., Henken, A., Evers, E., de Wit, M., van Duynhoven, Y., 2000. Oorsprong, omvang en kosten van humane salmonellose. Infectieziekten Bulletin 11, 4–8.

Visscher, C.F., Klein, G., Verspohl, J., Beyerbach, M., Stratmann-Selke, J., Kamphues, J., 2011. Serodiversity and serological as well as cultural distribution of *Salmonella* on farms and in abattoirs in Lower Saxony, Germany. International Journal of Food Microbiology 146, 44–51.

Vose, D., 2000. Risk Analysis: A Quantitative Guide. John Wiley & Sons Ltd, Chichester, United Kingdom.

Warriner, K., Aldsworth, T.G., Kaur, S., Dodd, C.E.R., 2002. Cross-contamination of carcasses and equipment during pork processing. Journal of Applied Microbiology 93, 169–177.

Chapter 5

A Biotracing Model of *Salmonella* in the Pork Production Chain

J. H. Smid [a,b]* L. Heres [c], A. H. Havelaar [a,b], A. Pielaat [a]

[a] RIVM, A. van Leeuwenhoeklaan 9, 3720 BA Bilthoven, The Netherlands
[b] IRAS, Yalelaan 2, 3584 CM Utrecht, The Netherlands
[c] VION Fresh Meat West, Boseind 10, 5281RM Boxtel, The Netherlands

Journal of Food Protection 2012; 75(2): 270–280
including erratum in Journal of Food Protection 2012; 75(4): 628

Journal of Food Protection, Vol. 75, No. 2, 2012, Pages 270–280
doi:10.4315/0362-028X.JFP-11-281

A Biotracing Model of *Salmonella* in the Pork Production Chain

J. H. SMID,[1,2]* L. HERES,[3] A. H. HAVELAAR,[1,2] AND A. PIELAAT[1]

[1]*National Institute for Public Health and the Environment, Centre for Infectious Disease Control, Antonie van Leeuwenhoeklaan 9, 3721 MA Bilthoven, The Netherlands;* [2]*Institute for Risk Assessment Sciences, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands; and* [3]*VION Fresh Meat West, Boseind 10, 5281RM Boxtel, The Netherlands*

MS 11-281: Received 8 June 2011/Accepted 25 September 2011

ABSTRACT

In biotracing systems, downstream chain information and model-based approaches are used to trace the sources of microbial contamination in a food chain. This article includes the results of a biotracing model for *Salmonella* in the pork slaughter process chain. A Bayesian belief network model was used in which information on the *Salmonella* level at different locations in the slaughterhouse were used in combination with prior knowledge about the dynamics of *Salmonella* throughout the slaughter line. Data collected in a Dutch slaughterhouse were used to specify prior beliefs about the model inputs and to iteratively refine the distributions of the parameters in the model to obtain an optimal description of that specific slaughterhouse. The primary purpose of the model is to trace the sources of contamination for individual *Salmonella*-positive carcasses at the end of the slaughter line. The model results indicated that house flora on or in the carcass splitter was the source of contamination for many carcasses, especially for those that carried contamination on the cutting side. The results also indicated that the parameter values of the model may be subject to temporal variation and can be used as a tool to provide estimates of such trends. This model illustrates the concept of biotracing, gives insight into the dynamics of *Salmonella* in the slaughter line, and indicates the sites in the line where data collection is most effective for biotracing. This biotracing model was implemented as an interactive computer application, which is a step in the process toward an operational biotracing system by which a stakeholder can initiate immediate responses to *Salmonella* contamination and other hazards in the pork slaughterhouse.

Pork is one of the main sources for human salmonellosis, associated with 15 to 25% of the cases in industrialized countries *(12, 14, 39)*. However, it is not clear which steps of the pork processing chain contribute most to these cases, and findings from studies on the source of *Salmonella*-positive finished carcasses have been inconsistent.

Feed is a potential pathway by which *Salmonella* may be introduced to farm livestock *(6, 26, 27)*. Pigs are normally asymptomatic carriers of *Salmonella* and can contaminate and infect each other on the farm via contaminated feces. Transport and lairage induce shedding among carriers due to stress, increasing the proportion of contaminated pigs. In a recent study in The Netherlands, *Salmonella* was detected in 97% of the cork bore skin samples and 62% of the fecal samples of pigs before slaughter *(38)*.

Numerous researchers *(4, 10, 18, 21, 24, 30, 40)* have reported a strong association between the proportion of animals carrying or excreting *Salmonella* before slaughter and the proportion of contaminated carcasses at the end of the slaughter line. In a Canadian study in which the genetic profiles of *Salmonella* isolates were compared using pulsed-field gel electrophoresis, the researchers concluded that most isolates from carcasses were similar to those from animals or the preevisceration environment *(17)*.

In contrast, Swanenburg et al. *(34)* conducted a similar study in The Netherlands and found that *Salmonella* serovars may survive in certain niches in the slaughterhouse environment, become part of the house flora of the slaughterhouse, and contaminate carcasses passing through this environment. Botteldoorn et al. *(8)* concluded that complex contamination cycles can exist in slaughterhouses, resulting in the isolation of various genotypes circulating in the slaughter environment. In several studies, house flora was particularly found to be associated with the carcass splitter *(31, 36)*. In a recent study in The Netherlands *(38)*, one carcass splitter was consistently contaminated. The predominant *Salmonella* serovar found on the splitter also was found on some carcasses, but this serovar was not found on the living pigs. The existence of resident strains in slaughterhouses has been reported *(2, 35, 41)*.

Several potential sources of *Salmonella* contamination of pig carcasses have been reported, and these sources may vary by country, slaughterhouse facility, and slaughtering day. Currently, when increased bacterial levels are found on carcasses, an ad hoc approach is employed to trace the source, which is time inefficient and costly. From a public health perspective, more systematic and hence faster and possibly cheaper approaches are needed to trace sources of contamination in the pork slaughter chain. A model-based

* Author for correspondence. Tel: +31-30-2744284; Fax: +31-30-2744434; E-mail: joost.smid@rivm.nl.

approach is provided by the recently developed biotracing methodology *(3)*. Biotraceability is defined as the ability to use downstream information to point to materials, processes, or actions within a particular food chain that can be identified as the source of undesirable agents. This nontypical application of probabilistic farm-to-fork modeling can be used to recover the origins of foodborne hazards. For this purpose, a model must be able to answer questions in the reverse direction of the chain processing order. This directionality requires that multiple pieces of evidence (observations of variables in the model) be used to update the statistics of the model parameters, which can be done by consistently applying Bayes's rule in a network model. Bayesian belief networks (BBNs) are probabilistic network models that allow for such inferential queries and are an appropriate choice of model for biotracing *(31)*.

This article includes results of the biotracing methodology used to trace the sources of *Salmonella* in the pork slaughter chain. A model we proposed previously *(28)* and the *Salmonella* counts provided by van Hoek et al. *(38)* were used. Multiple sources may contribute to contamination of a carcass in this chain. The dominant source was defined as the source that contributed most to the total number of bacteria on a specific carcass at the time of inspection. The following potential dominant sources were deduced from published reports and considered in the model: contaminated skin surface of pigs at the onset of slaughter (skin), contamination from the gut contents of the pigs (feces), and house flora isolated from the carcass splitter. The purpose of the model was to find the dominant source for each individual carcass that was positive for *Salmonella* on the lard and/or cutting side at the final meat inspection, which is one of the final steps of slaughtering.

MATERIALS AND METHODS

Core model. In our previous article *(28)*, a model was constructed to provide an answer to the following question: ''Given a downstream observed number of salmonellae on the lard side (*exterior*) or on the cutting side (*interior*) of a carcass after meat inspection, can the dominant source of those salmonellae be inferred?'' An overview of this model is shown in Figure 1. The model parameters as previously presented *(28)* described a generic slaughterhouse, and only summary statistics obtained from the literature were used for estimating the parameters values, leaving considerable uncertainty in the model. That partially parameterized model is the core model *(1)*.

The core model was based on a Monte Carlo simulation model for *Salmonella* in slaughter and breeder pigs *(13)*. Stages of slaughtering that were considered relevant to our question were selected from the set of stages described in the European Food Safety Authority (EFSA) model *(13)*. The changes occurring during these stages were modeled using modular process risk model equations *(20)*. These equations were used to construct a BBN on a slightly simplified domain to reduce computational complexity, which is a practical necessity. The resulting model describes the transmission dynamics of *Salmonella* during all stages of slaughtering in a generic slaughter plant and provides the expected prevalence and population of *Salmonella* on a whole carcass after these stages. The *Salmonella* population on the exterior and interior of a carcass was included in the model as a separate variable because the location of the pathogen on the carcass is expected to be indicative of the origin of the contamination, e.g., pathogen cells that are transferred from cutting robots are expected to be found primarily on the carcass interiors. We assumed that the salmonellae were homogeneously spread over these surfaces. The BBN was implemented in the HUGIN software (HUGIN Expert, Aalborg, Denmark). A graphical representation and the equations used for constructing the core model were published previously *(28)*.

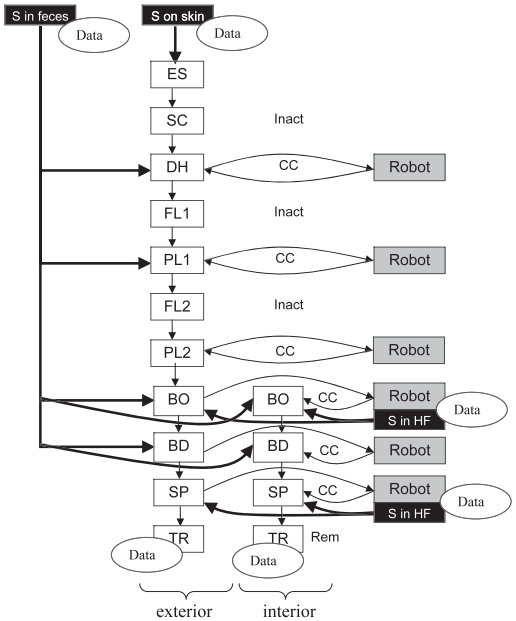


FIGURE 1. *Steps (ES, exsanguination; SC, scalding; DH, dehairing; FL, flaming; PL, polishing; BO, belly opening; BD, bung dropping; SP, splitting; TR, trimming) and transmission dynamics of* Salmonella *included in the model. Black boxes indicate contamination sources:* Salmonella *on the skin or feces of incoming pigs and* Salmonella *from house flora (HF) on the cutting robots. Fat arrows represent direct transfer of* Salmonella *from the sources; thin arrows represent other microbial dynamics, i.e., inactivation (Inact), removal (Rem), and cross-contamination (CC). Ovals indicate where data were collected. Adapted from Smid et al. (28) with permission.*

Model inputs. By including prevalence and population data from one specific slaughterhouse in The Netherlands, the model was trained to describe the dynamics of *Salmonella* in that particular slaughterhouse. The sampling experiment and the data were fully described previously *(38)*. On 11 days from April to July 2009, samples were taken from 118 randomly selected carcasses and analyzed for *Salmonella* prevalence using enrichment followed by PCRs. When positive results were obtained, *Salmonella* populations were determined using most-probable-number (MPN) techniques. Four cork bore samples, each taking 5 cm^2 of skin, and one swab sample from the anal region of each pig were collected immediately after exsanguination, generating the upstream data. The carcasses were marked, and at the end of the slaughter process during meat inspection, another four cork bore samples were collected from the exterior and interior of the same carcasses, generating the downstream data. At the end of

each sampling day, two swab samples were collected from the knives of two belly opening robots and two carcass splitting robots. Samples were collected from small pockets in the robots close to the knives using a flexistem brush. The two belly openers and one carcass splitter were contaminated on only 1 or 2 days. However, the other carcass splitter was consistently contaminated and, hence, considered a potential source of house flora. The data obtained from this robot were the environmental data.

The upstream data were used to generate the variability distributions for the input variables of the model. These inputs were (all directly after exsanguination) (i) the probability of a pig being positive for *Salmonella* on its skin, (ii) the probability of a pig being positive for *Salmonella* in its feces, (iii) the number of *Salmonella* cells on the skin for skin samples that were *Salmonella* positive, and (iv) the number of *Salmonella* cells per gram of feces for fecal samples that were *Salmonella* positive.

For each sampling location, the observed prevalence was treated in the model as the probability of an arbitrary pig being positive for *Salmonella,* expressed by a probability distribution on a binary variable, i.e., a variable with values of 0 or 1. *Salmonella* populations in samples were approximately lognormally distributed *(15)*. The mean and variance of this distribution were estimated using maximum likelihood methods. Samples that were positive based on PCR results but below the detection limit based on the MPN technique were regarded as censored positive and were accounted for in these estimates by using the cumulative lognormal distribution. The number of *Salmonella* cells on a pig's skin was extrapolated to an expected number of *Salmonella* cells on a whole pig by assuming an average body surface of 12,000 cm^2 *(37)*. When correlations in the upstream data were demonstrated at or below the significance level of 0.05, then such interdependencies were included in the BBN.

Model training. The data were used to refine the probability distributions of the parameters in the model (i.e., to train the model) through an algorithm called sequential adaptation *(33)*. During one adaptation step, one subset of data (consisting of upstream and downstream *Salmonella* population data for one carcass or environmental data from one sample from the carcass splitter) was fed to the model, generating posterior probability distributions for the parameters in the model. In the Bayesian context, the posterior (probability) distribution of a variable refers to the distribution of that variable given that any other (sets of) variables are observed, whereas the prior (probability) distribution is the distribution of that variable without taking observations into account. Subsequently, these posterior distributions were used to update the prior distributions for these parameters in the light of the new data and the strength of belief in the correctness of the prior distributions. This belief is expressed by the notion of experience *(33)*. The data were subsequently retracted from the model, concluding one full adaptation step.

These steps were successively repeated with the datasets for all carcasses, gradually reducing the uncertainty about the parameter values of the model (which can be seen by increasing experience) and improving it. Sequential adaptation shows how the distribution of parameter values changes as a function of new data inputs. To correctly represent such changes over real time, the data must be used for adaptation in the same order as they were collected during the experiment. The data sets, consisting of up- and downstream data for one carcass or environmental data for 1 day, were therefore numbered chronologically in the order in which the data were collected during the experiment and used for adaptation in that order.

The environmental data also were used for adaptation. These data are observations of a variable that may be the result of two distinct processes: direct transfer of resident bacteria from the carcass splitter (house flora) and transfer of transient bacteria from the carcass via the robot (cross-contamination). Therefore, an auxiliary variable was added to the model, summing the contributions of both processes. The environmental data were treated as observations of this auxiliary variable.

To verify the sensitivity of the adaptation results to a partial omission of the upstream data, the upstream (carcass and environmental) data for the second half of the sampling experiment were deleted. Next, adaptation was performed. The adaptation results of three variables in the model (the parameter quantifying the transfer of house flora to a carcass and the two report variables; see below) were compared with the results obtained when all data were used. Adaptation also was performed with all upstream data omitted, and the results were compared.

Forward functionality of the model. The trained model was used as a tool to generate baseline predictions, i.e., estimates of the number of *Salmonella* cells on a carcass throughout the slaughter process. Baseline predictions were obtained by computing the marginal probability distributions of the variables in the model representing these numbers. The marginal (probability) distribution of a variable X is obtained by summing or integrating the joint probability distribution over the distribution of all variables minus X.

The correctness of the baseline predictions was investigated by comparing the model predictions with results from similar models and with data from the literature. The model was maximally trained for these comparisons, i.e., all data collected in the slaughterhouse were used to learn the parameters. The capacity of the model to adapt to the data also was investigated, i.e., can the uncertain parameters of the model be learned in such a way that the downstream data are likely samples from the baseline predictions of the model? A chi-square goodness-of-fit test was used in which it was assumed that the (hypothetical) data set tested was equal to the training data set. Theoretically, it would be more accurate to split the data set into two parts, train the model with one data set, and compare the second data set with the baseline predictions of the model. However, the amount of data is limited, and the parameter values of the model may change over time. Therefore, all data were used for training the model.

Backward functionality of the model. *Salmonella*-positive results for samples collected from the exterior and/or interior of the carcass at meat inspection may trigger biotracing. The variables in the model that represent the number of *Salmonella* cells at these sites were therefore designated as the triggers *(30)*: Tot_e is the trigger representing the total number of *Salmonella* cells on the exterior, and Tot_i is the trigger representing the total number of *Salmonella* cells on the interior. The downstream data provide observations of the triggers. Given such observations, causal relations can be inverted in the BBN by a subsequent use of Bayes's rule, allowing for reasoning back in the chain.

The answer to the question ''which was the dominant source of contamination of a positive carcass?'' is provided by the report variables *(28)*, which return the posterior probabilities that each possible source was the dominant source: Rep_e for exterior contamination and Rep_i for interior contamination. These posterior probabilities were registered after each adaptation step, in which sets of upstream and downstream data for one carcass were input to the model.

Biotracing differs from model training, although the same data are used for both applications. Biotracing consists of registering the posterior distributions of the report variables, which are calculated from realizations of the triggers. Model training

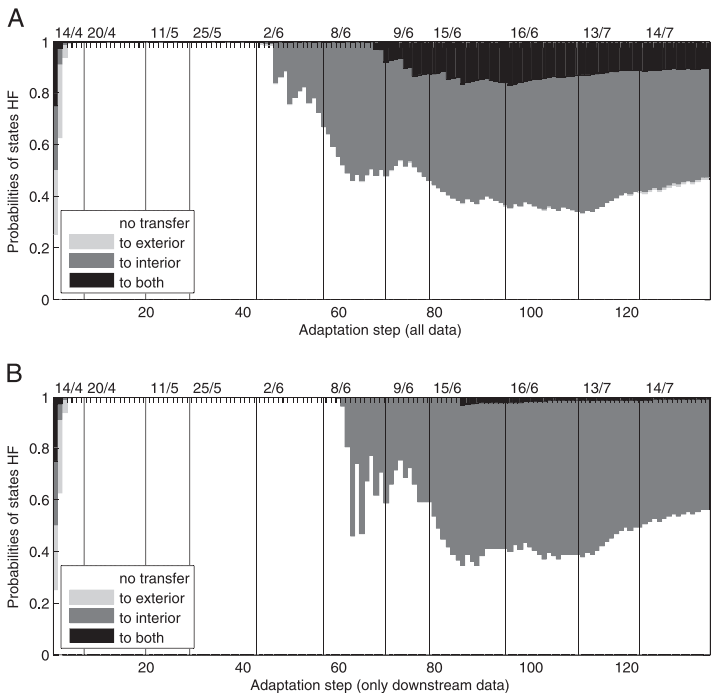


FIGURE 2. *Marginal prior probabilities of the states of parameter HF, representing the probability of transfer of house flora to the carcass exterior, interior, or both before each sequential adaptation step. The upper* x *axis indicates the date (day and month) on which the data were obtained. (A) Results when all data were used for adaptation. (B) Results when only downstream data were used for adaptation.*

consists of updating the prior distributions of parameters, which are calculated using the long-term averages over many realizations of the triggers.

As a second type of analysis, hypothetical downstream data were input to the trained model to (i) verify whether it produces discriminatory biotracing results when only downstream data are available and (ii) explore the predictions of the model as a function of the data.

Sensitivity analysis. Pollino et al. *(23)* proposed two types of sensitivity analysis to evaluate BBNs: sensitivity to findings and sensitivity to parameters. The first analysis is used to investigate whether and to what degree evidence about one variable influences the probability distribution of another variable, which is quantified by their mutual information. High mutual information indicates that evidence about one variable largely reduces uncertainty about another variable, whereas zero mutual information means that the variables are independent (see Korb and Nicholson *(16)* for an excellent technical introduction to this kind of sensitivity analysis). For this study, the mutual information between the report variables and the variables for which data were collected was analyzed. The mutual information between the report variables and the process parameters also was analyzed.

The sensitivity to parameters approach is empirical, based on changing the values of each of the parameters and observing the changes in the (marginal) posterior probability distribution of the report variable. The magnitude of such changes may be expressed by a sensitivity value *(9),* which provides an analytic summary of how sensitive the probabilities for the values of the report variables are to the initial assessments (prior probabilities) of the parameters. The larger the absolute value of a sensitivity value for a parameter, the more carefully its prior distribution should be chosen because it may heavily influence the posterior probability of the report variable (by which the biotracing conclusion of the model is reported). Sensitivity values can be calculated for all possible values of a parameter and all possible values of the report variable. Yet, the overall sensitivity of the report variable to that parameter may be summarized by the minimum, maximum, and mean sensitivity value taken over all combinations of values of the parameter and the report variable. A more detailed introduction to this kind of analysis was provided by Kjaerulff and Madsen *(15)* and, for this particular application, in our previous article *(28).* Automated routines were available in HUGIN to do both types of sensitivity analysis.

RESULTS

Model training. In Figure 2, the data sets (the adaptation steps on the lower *x* axis) were clustered per sampling date (upper *x* axis) on which they were obtained. In Figure 2A, the stacked bars illustrate the prior probabilities of the parameter, indicating the transfer of *Salmonella* from house flora to the carcass (HF) before each adaptation step. All upstream, environmental, and downstream data were used for adaptation. HF is an uncertain but important parameter in the model *(28)*. This uncertainty is expressed by the equal probabilities for each state of HF before adaptation (left stacked bar). After only two adaptation steps, the probability of the state of HF indicating ''no transfer of house flora'' increased to nearly 1. Thus, the model quickly updated its prior belief. However, the belief that house flora will not be transferred to a carcass did not remain close to 1 throughout the experiment. After the 43rd adaptation step (data obtained on 2 June), the probability that *Salmonella* would be transferred to the interior of a

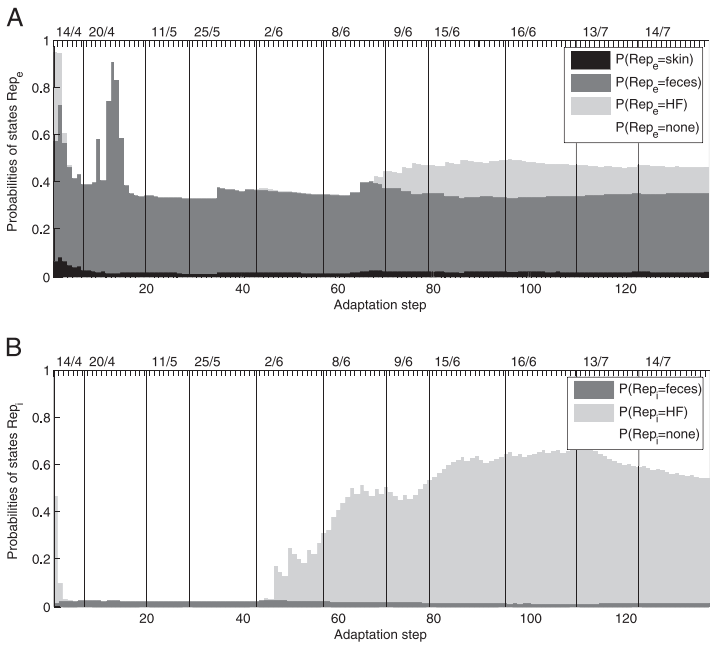


FIGURE 3. *Marginal probabilities of the report variables (A, Rep_e; B, Rep_i) after each sequential adaptation step. The upper* x *axis indicates the date (day and month) on which the data were obtained.*

carcass increased to a maximum of 0.65. After the 65th adaptation step (data obtained on 8 June), the probability that *Salmonella* would be transferred to both the exterior and interior of a carcass increased to a maximum of 0.16 (end of 16 June). At the end of the experiment, the probability of transfer of *Salmonella* to a carcass slightly decreased again; the probability of no transfer was 0.47, the probability of *Salmonella* transfer only to the interior was 0.42, the probability of *Salmonella* transfer only to the exterior was 0.01, and the probability of *Salmonella* transfer to both exterior and interior was 0.10.

Figure 3 shows the marginal prior probabilities of the report variables Rep_e and Rep_i after each full adaptation step. The prior belief of a carcass being contaminated by feces remained remarkably constant during adaptation, around 0.33 for exterior contamination and around 0.02 for interior contamination. However, the prior belief of a carcass being contaminated by house flora increased dramatically after the 45th adaptation step (data obtained on 2 June), to around 0.16 for exterior contamination (end of 15 June), and to around 0.66 for interior contamination (end of). After 16 June, these probabilities decreased again to 0.11 and 0.52, respectively.

These findings indicate that house flora did not exist or did not contaminate carcasses during the first sampling days but then became increasingly important, especially for contamination of the carcass interior, but seemed to decline slightly at the end of the experiment.

When we deleted the second half of the upstream and environmental data and performed sequential adaptation with the reduced data set, we obtained figures practically indistinguishable from Figures 2 and 3 (results not shown), which indicates that the model does not necessarily need all upstream data for a correct adaptation of its parameters. When adaptation was performed with all upstream and environmental data omitted, the probabilities of the states of HF after each full adaptation step were only slightly shifted (see Fig. 2B). This shift also occurred for the adaptation results of Rep_e and Rep_i (results not shown).

Baseline predictions. The baseline predictions of the model are shown in Figure 4. In this analysis, the predicted number of *Salmonella* cells on the exterior was added to the number of *Salmonella* cells on the interior of the carcass to obtain the total number of *Salmonella* cells on a carcass. In general, the following conclusions were reached.

Scalding reduces the expected number of bacteria on a *Salmonella*-positive carcass by more than 3 log units. During dehairing, cross-contamination occurs, especially from fecal matter assumed to be extruded from the carcass by pressure. This contamination increases the prevalence on carcasses to close to 100% and the expected number of bacteria on each carcass to around 2.5 log units. Flaming decreases this number again by around 1 log unit. Polishing causes cross-contamination by fecal matter, which is again assumed to be extruded from the carcass by pressure, increasing the prevalence on carcasses to more than 90%. A second flaming operation reduces the expected number of bacteria on the carcass to less than 1 log unit. Upon entering the clean area, the expected number of bacteria on the carcass remains almost constant, although cross-contamination by fecal matter is limited. However, after splitting, considerable contamination by bacteria from the robot occurs, increasing the expected number of bacteria on the

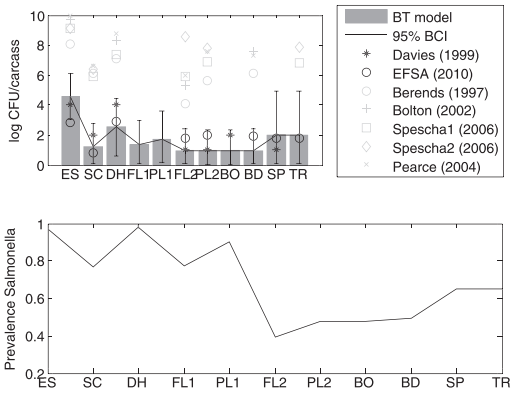


FIGURE 4. *(Top) The bar charts show expected number of* Salmonella *cells (and 95% Bayesian credible intervals) on a positive carcass after each production step, as computed by the fully trained model. The expected numbers are the sums of external and internal contamination. Results were compared with published reports. Solid symbols indicate studies specifying the level of* Salmonella *through the chain; shaded symbols indicate studies specifying the levels of indicator organisms (*Enterobacteriaceae, *total viable bacterial counts [TVC], and total aerobic bacterial counts [TAC]) throughout the chain. (Bottom) Computed* Salmonella *prevalence throughout the chain. ES, exsanguination; SC, scalding; DH, dehairing; FL, flaming; PL, polishing; BO, belly opening; BD, bung dropping; SP, splitting; TR, trimming.*

carcass to 1.7 log units, with a 5% upper bound of the probability distribution of the number of bacteria on the carcass of 4.9 log units.

Other researchers have investigated the dynamics of *Salmonella* or related bacteria (*Enterobacteriaceae,* total viable counts, and total aerobic counts) along the food chain *(5, 7, 11, 13, 22, 32)*. The findings of these studies also are shown in Figure 4, and similar trends in the dynamics of the bacteria were found along the chain. *Enterobacteriaceae* and total viable counts were much higher on the carcasses, but the dynamics seem to reflect those of *Salmonella*.

The baseline prevalences at the end of the slaughter chain, as calculated by the model, were 46.1% for exterior contamination and 53.8% for interior contamination (not shown in Fig. 4). These percentages surpassed the actually observed prevalences in the data. However, the percentages calculated by the model should be interpreted as the fraction of carcasses with zero salmonellae on the whole carcass. Most probably, low numbers of bacteria on the carcass will not be found during screening. Cork bore samples allow evaluation of only 5 cm^2 of carcass surface per sample, so a pool of four cork bore samples, as taken by van Hoek et al. *(38),* allows evaluation of only 20 cm^2. Assuming a PCR recovery rate of 100% in this sample, one *Salmonella* cell per 20 cm^2 could in principle be found. Assuming a carcass surface of 12,000 cm^2 *(37)* and homogeneous bacterial distribution over this surface, the minimum number of bacteria that could be found on this surface would then be 2.78 log CFU. Similarly, assuming that the total surface of the carcass interior is around one-quarter of its exterior surface, the minimum number of bacteria that could be found is 2.18 log CFU. Thus, low numbers of salmonellae (which the model judges as most likely) often go undetected. When samples with numbers of salmonellae below these limits are assumed to be negative for *Salmonella,* then the probability that a carcass will be positive at the meat inspection is 19.1% for exterior contamination and 22.2% for interior contamination. The expected number of bacteria on the exterior of positive carcasses would then be 4.0 log CFU (95% Bayesian credible interval, 2.3 to 6.1 log CFU), and the expected number of bacteria on the interior of positive carcasses would then be 3.7 log (95% Bayesian credible interval, 2.2 to 5.9 log CFU). These results are in closer agreement with the data. Using a chi-square goodness-of-fit test, the null hypothesis that the data are likely samples from the probability distributions of Tot_e and Tot_i cannot be rejected at a 5% significance level ($P = 0.45$ for Tot_e, and $P = 0.19$ for Tot_i).

Biotracing results. Each stacked bar chart in Figure 5A and 5B represents the posterior (marginal) distribution of the report variable Rep_e or Rep_i for carcasses with contaminated exteriors and interiors, respectively, after having observed up- and downstream data for a particular carcass. The shaded fractions of each bar denote the posterior probabilities that each source was the dominant one for that carcass. *Salmonella*-positive carcasses were clustered by sampling date (*x* axis).

The most likely dominant source of the first five carcasses with contaminated exteriors in the study (sampled on 14 April, 25 May, and 8 June) was pig's feces, according to the model. Most carcasses with contaminated exteriors sampled on 8, 9, and 15 June were highly contaminated *(38),* which the model conceptualized as salmonellae detached from the house flora. Exceptions were the third contaminated carcass sampled on 8 June and the second contaminated carcass sampled on 9 June, which carried *Salmonella* below the detection limit by MPN *(38)*. For these carcasses (except for that sampled on 9 June) it was therefore also considered less likely by the model that their dominant source had been house flora. At the end of the sampling period (July), *Salmonella* was again found with numbers below the detection limit by MPN. These cases were predominantly attributed to contamination by feces.

For almost all carcasses carrying contamination in the interior, the most likely dominant source was house flora. Exceptions were the first two carcasses that carried *Salmonella* on the interior. However, this result may be an artifact of the model; these carcasses were sampled in the beginning of the experiment when the parameter values quantifying the transfer rates from house flora were still poorly estimated.

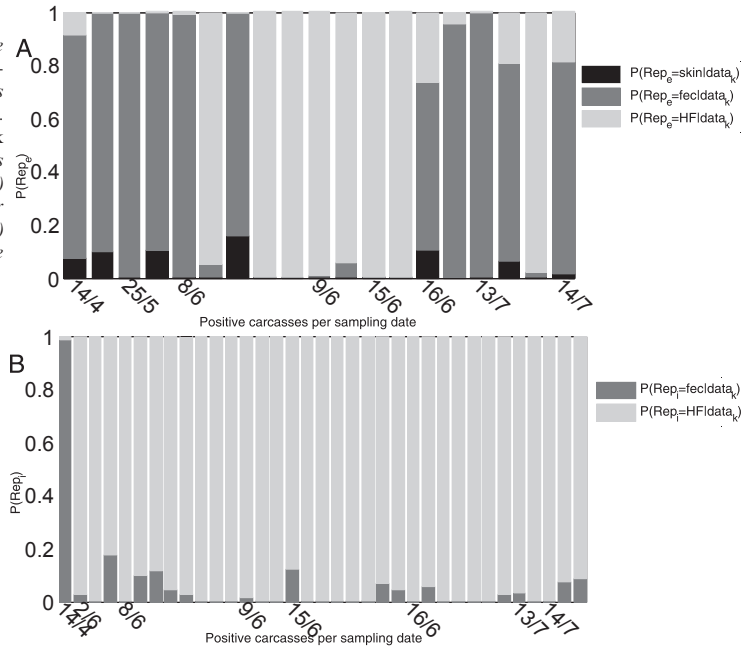


FIGURE 5. *For each potential source, the estimated probabilities of being the dominant source of contamination of a carcass are indicated by fractions of a bar chart. The* k*th data set corresponds to carcass* k *(carcass number not shown), which was sampled on the dates (day and month) shown on the* x *axis. (A) Results for contamination of the carcass exterior. (B) Results for contamination of the carcass interior.*

After training the model, it was assumed that new (for now, hypothetical) downstream data were available, consisting of observed numbers of salmonellae on the exterior and interior of the carcass after meat inspection. These data points were propagated case by case through the model, and the posterior marginals of the report variables were plotted in Figure 6. This analysis using the fully trained model is different from that described in the previous section, where the model was not yet fully trained and where upstream and environmental data also were used.

When low numbers of *Salmonella* cells (<2 log units) are on the exterior of the carcass at the time of meat inspection, then the chance is highest that the dominant source is feces, although the dominant source also can be house flora. When numbers are between 2 and 4 log CFU, the chance is highest that the dominant source is house flora, although it also could be feces. For higher numbers, the source is most probably house flora. Skin can act as a source but only for low numbers of *Salmonella*.

The most probable source for almost any number of *Salmonella* cells on the interior of a carcass after meat inspection is house flora. Only for very low numbers (between 0 and 2 log units) is the source likely to be feces, but these low values are hard to detect when the contamination is spread homogeneously across the carcass. Skin is never assigned as a source for *Salmonella* on the carcass interior because of the modeling assumption that the salmonellae that are left on the skin after polishing are irreversibly attached *(19)* and consequently cannot cross-contaminate the carcass interior.

Sensitivity analysis. Table 1 shows the results of a sensitivity analysis to findings for the report variables Rep_i and Rep_e in the fully trained model. Results indicated that the report variables are most sensitive to observations of downstream variables and less sensitive to observations of process parameters. For future use in biotracing studies, it will be most informative to monitor the number of *Salmonella* cells on the carcass after slaughtering and on

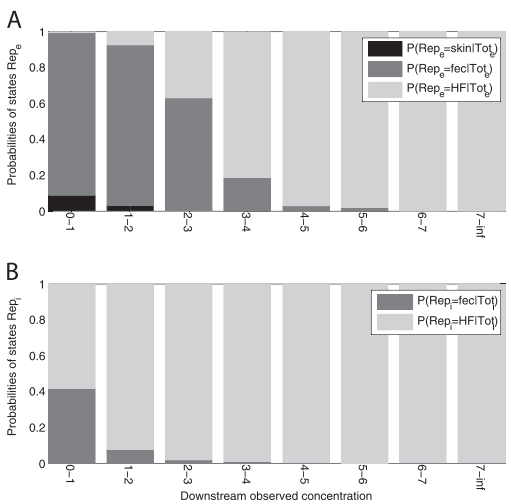


FIGURE 6. *Stacked probabilities of the report variables (A,* Rep_e*; B,* Rep_i*) given (hypothetical) observations of the triggers (*Tot_e *and* Tot_i*). To obtain these results, the model was fully trained by the data.*

TABLE 1. *Sensitivity to findings analysis for the report variables Rep_e and Rep_i*

Variable[a]	Mutual information for Rep_e[b]	Variable[a]	Mutual information for Rep_i[b]
Total no. of salmonellae on skin after trimming	0.9	Total no. of salmonellae on cut section after trimming	0.7
Total no. of salmonellae on carcass splitter	0.5	Total no. of salmonellae on carcass splitter	0.6
Total no. of salmonellae on cut section after trimming	0.5	Total no. of salmonellae on skin after trimming	0.4
No. of salmonellae per gram of feces	0.2	No. of salmonellae per gram of feces	8.3×10^{-4}
No. of salmonellae on skin of incoming pig	1.8×10^{-2}	Probability of cutting colon during bung removal	3.3×10^{-4}
Rate of *Salmonella* transfer from polishing robot to pig	4.3×10^{-3}	Remaining measurable variables	$<1.0 \times 10^{-5}$
Rate of *Salmonella* transfer from dehairing robot to pig	8.0×10^{-4}		
Temp of scalding water	4.2×10^{-4}		
Remaining measurable variables	$<1.0 \times 10^{-5}$		

[a] Only those variables that can be measured theoretically in a laboratory or practically in the slaughterhouse were considered. Variables for which the mutual information is smaller than 10^{-5} are not shown.
[b] Higher mutual information means greater sensitivity of the report variable to observations of considered variables.

the machines instead of monitoring process variables or upstream measurements.

Table 2 shows the results of a parameter sensitivity analysis. Variable Rep_i is most sensitive to the parameters indicating the probability of detachment of *Salmonella* from house flora and the fraction of *Salmonella* from house flora that is transferred to the carcass interior. Limited sensitivity was found to the parameter indicating the probability of cutting through the intestines during bung dropping. Variable Rep_e also was sensitive to the parameter indicating the probability of detachment of *Salmonella* from house flora, the parameter quantifying the number of *Salmonella* cells on the exterior of a positive pig, and the parameter indicating the transfer and inactivation rates of *Salmonella* during dehairing, flaming, and polishing.

TABLE 2. *Parameter sensitivity analysis for the report variables Rep_e and Rep_i*

	Sensitivity value[a]					
	Rep_i			Rep_e		
Parameter	Minimum	Maximum	Mean	Minimum	Maximum	Mean
Probability that incoming pig has *Salmonella* on skin	0	0	0	-8.4×10^{-3}	8.4×10^{-3}	4.4×10^{-3}
No. of salmonellae on positive incoming pig	0	0	0	-4.9×10^{-2}	6.5×10^{-2}	2.7×10^{-3}
Temp of scalding (°C)	0	0	0	-9.4×10^{-2}	1.0×10^{-1}	3.9×10^{-2}
Probability that incoming pig has *Salmonella* in feces	-1.1×10^{-2}	1.1×10^{-2}	8.1×10^{-3}	-3.0×10^{-1}	3.0×10^{-1}	2.0×10^{-1}
No. of salmonellae per gram in feces-positive pig	-1.3×10^{-2}	1.8×10^{-2}	5.7×10^{-3}	-3.0×10^{-1}	3.0×10^{-1}	2.0×10^{-1}
Rate of *Salmonella* transfer from dehairing machine to pig	-5.4×10^{-5}	4.4×10^{-5}	1.4×10^{-5}	-3.0×10^{-1}	2.0×10^{-1}	9.5×10^{-2}
Rate of inactivation during flaming	-3.7×10^{-5}	4.1×10^{-5}	2.0×10^{-5}	-2.0×10^{-1}	2.0×10^{-1}	1.0×10^{-1}
Rate of *Salmonella* transfer from polishing machine to pig	-6.3×10^{-4}	5.6×10^{-4}	2.7×10^{-4}	-4.0×10^{-1}	5.0×10^{-1}	3.0×10^{-1}
Temp reached during decontamination (°C)	-3.6×10^{-3}	3.5×10^{-3}	1.9×10^{-3}	0	0	0
Probability of cutting colon	-9.1×10^{-2}	9.9×10^{-2}	1.9×10^{-2}	-1.8×10^{-3}	2.8×10^{-3}	6.0×10^{-4}
Amt of feces released when colon is cut	-7.9×10^{-4}	7.2×10^{-4}	3.9×10^{-4}	-4.3×10^{-4}	4.3×10^{-4}	1.8×10^{-4}
Temp reached during bung removal (°C)	-1.2×10^{-2}	1.2×10^{-2}	5.9×10^{-3}	0	0	0
Fraction of feces transferred to exterior when colon is cut	-6.6×10^{-3}	6.1×10^{-3}	2.9×10^{-3}	-4.6×10^{-4}	7.0×10^{-4}	3.4×10^{-4}
Temp reached during decontamination (°C)	-3.8×10^{-3}	3.7×10^{-3}	1.9×10^{-3}	0	0	0
Multiplication factor amt of feces	-4.5×10^{-3}	5.0×10^{-3}	1.3×10^{-3}	-1.1×10^{-3}	1.3×10^{-3}	3.3×10^{-4}
Probability that particles are released from house flora	-4.3×10^{-2}	4.0×10^{-2}	3.1×10^{-4}	-5.8×10^{-2}	3.6×10^{-2}	1.6×10^{-4}
Expected log no. of salmonellae transferred from biofilm to carcass exterior	-1.0×10^{-0}	1.0×10^{-0}	8.0×10^{-1}	-9.0×10^{-1}	9.0×10^{-1}	6.0×10^{-1}
Expected log no. of salmonellae transferred from biofilm to carcass interior	-7.7×10^{-2}	7.3×10^{-2}	1.7×10^{-2}	-1.0×10^{-1}	7.1×10^{-2}	2.1×10^{-2}

[a] Minimum, maximum, and mean sensitivity value is taken over all states of the parameter and all states of the report value.

DISCUSSION

We built a model by which observed *Salmonella* data for carcasses at meat inspection, one of the final stages of the slaughtering process, can be used for biotracing. For this model it was assumed that pig carcass contamination might result from exterior or intestinal carriage of *Salmonella* on or in pigs or from direct transfer of house flora in the slaughterhouse environment, which is based on the conclusions of many other researchers. Meat inspection was chosen as reference point for model validation because (i) this stage is the last step in the chain where *Salmonella* was found and (ii) at this stage the *Salmonella* status of carcasses also is assessed in existing monitoring programs.

Most cases of contamination in which *Salmonella* has been found on the carcass interior have the highest probability of being caused by house flora from the carcass splitter. These findings support those of Swanenburg et al. *(34),* who used an epidemiological approach. *Salmonella* that was found on the exterior of the carcasses in high numbers also was likely to have originated from house flora. Lower numbers of *Salmonella* were more likely to have originated from feces, which may have spread to the carcass through cross-contamination. Other researchers have used epidemiological approaches to show that cross-contamination is an important process during slaughtering and is responsible for many cases of contamination *(25, 41)*.

Epidemiological approaches to finding the dominant source of contamination in a slaughter process are based on averaging findings over many carcasses examined. The clear advantage of this model is that, once trained, it can make an inference statement associated with a specific observation for a specific carcass. Simple examples are given in Figure 6. This model naturally incorporates temporal variation in the inputs or the parameter values into the model. An example of temporal variation in the input variable quantifying the transfer of *Salmonella* from house flora is given in Figure 2. In epidemiological methods, this approach is less convenient, and often such studies are done only occasionally to account for temporal variation.

Once the causal relations have been implemented in the BBN, the user can input all data known for a specific slaughterhouse, including real-time measurements of process parameters and bacterial numbers on a carcass at several stages of the chain. The model includes these observations in the form of prior uncertainty distributions of the parameters, including experience, and refines these prior distributions iteratively with new data, which reduces the total uncertainty in the model. The longer the model is in use, the better its predictions become because it better describes the specific slaughterhouse through training. The model also can be used for risk assessment in a specific slaughterhouse because it generates estimates of the number of *Salmonella* cells on a carcass as a function of the process parameters in that slaughterhouse. In theory, specific slaughterhouses also could be modeled using standard Monte Carlo techniques. However, inference of unknown values of process parameters using downstream data is not straightforward using those techniques *(29)*. Therefore, most existing risk assessments based on Monte Carlo sampling, such as the EFSA model *(13),* describe a generic (nonspecific) slaughterhouse.

The proposed system is a systematic approach to finding the source of contamination in a slaughterhouse. Once a reasonable basic understanding of the parameter values governing the dynamics of *Salmonella* in the slaughterhouse is obtained, the model can make use of historic data for biotracing. Thus, ultimately few or no new data may be needed for biotracing; the model can make rapid conclusions. Compared with the current ad hoc approach to tracing the source of microbial contamination on carcasses in most slaughterhouses, this model provides results more quickly. To obtain a basic understanding of the parameter values, initially some data should be collected. After this initial data collection period, only downstream data are required for biotracing. Thus, this approach may eventually be cheaper. However, no comparison of the costs of both approaches was made.

Uncertainties in the model. The purpose of this model was to answer the question of which source contaminated a specific *Salmonella*-positive carcasses in a particular slaughterhouse, allowing plant operators to prioritize decontamination measures. This question may be difficult to answer when the dynamics of *Salmonella* in the slaughterhouse are poorly understood. To diminish the uncertainty, many samples were collected, but in a large number of samples the *Salmonella* level was below the detection limit. Indicator organisms were considered as a way to increase the share of countable samples. However, although *Salmonella* present in feces is assumed to be correlated with fecal indicator organisms such as *Enterobacteriaceae* and *Escherichia coli,* this correlation probably is weak for *Salmonella* potentially present in the house flora of the slaughterhouse. Hence, we used only *Salmonella* data for this model.

As more data were used for adapting the model, the estimated distributions of the model parameters became more certain. Therefore, the model predictions in Figure 5, generated by the later data sets, are less uncertain than those generated by earlier data. However, even after full sequential adaptation, the parameters describing the probability and degree of transfer of *Salmonella* from the house flora remained uncertain. Parameter sensitivity analysis indicated that these parameters were sensitive to the number of bacteria on the interior of carcasses. To improve the model, we recommend that efforts target an understanding of the dynamics of *Salmonella* in the house flora on slaughterhouse robots.

Like all models, this model is a reduced explanation of the real world. The more sophisticated the model, the more precisely the world will be explained, but also the more data will be needed and the less transparent the results may be *(42)*. The biotracing model presented here is the result of the accumulated data and the knowledge of experts. Therefore, any conclusion derived from this model is dependent of the current state of information.

Operational biotracing system. The inferential power of the model will increase when the model is better trained.

Therefore, the model can form a basis for an operational biotracing system when data from systematic monitoring in a slaughterhouse are used. How this works was demonstrated with a limited data set for one particular slaughterhouse in The Netherlands; therefore, the results are valid only for the period of the sampling experiment and only for this slaughterhouse. However, the potential sources are generic and the slaughter processes are similar in many slaughterhouses, so with slight modifications of the model targeted at another slaughterhouse and a data set collected in that slaughterhouse, results can be obtained for any slaughterhouse with this proposed methodology.

We implemented a BBN as an interactive computer application, which is a quantitative representation of the model equations. This application provides a tool for monitoring in real-time the process of contamination and for tracing of the sources of that contamination. This study is a first attempt to describe a biotracing model of *Salmonella* in the pork slaughter chain and can be seen as a step in the process toward an operational biotracing system in which a stakeholder can implement immediate responses to hazards in the pork slaughterhouse.

ACKNOWLEDGMENT

This work was supported by the Integrated Project BIOTRACER (contract 036272) funded by the European Union under the 6th RTD Framework.

REFERENCES

1. Albert, I., E. Grenier, J. B. Denis, and J. Rousseau. 2008. Quantitative risk assessment from farm to fork and beyond: a global Bayesian approach concerning food-borne diseases. *Risk Anal.* 28:557–571.
2. Baptista, F. M., J. Dahl, and L. R. Nielsen. 2010. Factors influencing *Salmonella* carcass prevalence in Danish pig abattoirs. *Prev. Vet. Med.* 95(3–4):231–238.
3. Barker, G. C., N. Gomez, and J. Smid. 2009. An introduction to biotracing in food chain systems. *Trends Food Sci. Technol*. 20:220–226.
4. Berends, B. R., F. Van Knapen, D. A. A. Mossel, S. A. Burt, and J. M. A. Snijders. 1998. *Salmonella* spp. on pork at cutting plants and at the retail level and the influence of particular risk factors. *Int. J. Food Microbiol.* 44:207–217.
5. Berends, B. R., F. Van Knapen, J. M. A. Snijders, and D. A. A. Mossel. 1997. Identification and quantification of risk factors regarding *Salmonella* spp. on pork carcasses. *Int. J. Food Microbiol.* 36:199–206.
6. Blackman, J., T. Bowman, J. Chambers, J. Kisilenko, P. Parr, A. M. St.-Laurent, and J. Thompson. 1992. Controlling *Salmonella* in livestock and poultry feeds. Plant Products Division, Centre for Food and Animal Research of Agriculture, London, Ontario, Canada.
7. Bolton, D. J., R. A. Pearce, J. J. Sheridan, I. S. Blair, D. A. McDowell, and D. Harrington. 2002. Washing and chilling as critical control points in pork slaughter hazard analysis and critical control point (HACCP) systems. *J. Appl. Microbiol.* 92:893–902.
8. Botteldoorn, N., L. Herman, N. Rijpens, and M. Heyndrickx. 2004. Phenotypic and molecular typing of *Salmonella* strains reveals different contamination sources in two commercial pig slaughterhouses. *Appl. Environ. Microbiol.* 70:5305–5314.
9. Coupé, V. M. H., and L. C. van der Gaag. 2002. Properties of sensitivity analysis of Bayesian belief networks. *Ann. Math. Artif. Intel.* 36:323–356.
10. Davies, R. H., R. Dalziel, J. C. Gibbens, J. W. Wilesmith, J. M. Ryan, S. J. Evans, C. Byrne, G. A. Paiba, S. J. Pascoe, and C. J. Teale. 2004. National survey for *Salmonella* in pigs, cattle and sheep at slaughter in Great Britain (1999–2000). *J. Appl. Microbiol.* 96:750–760.
11. Davies, R. H., I. M. MacLaren, and S. Bedford. 1999. Distribution of *Salmonella* contamination in two pig abattoirs, p. 891–903. *In* Proceedings of the 3rd International Symposium on the Epidemiology and Control of *Salmonella* in Pork, Washington, DC.
12. European Food Safety Authority. 2006. Opinion of the Scientific Panel on Biological Hazards on the request from the Commission related to ''Risk assessment and mitigation options of *Salmonella* in pig production.'' *EFSA J*. 341:1–131.
13. European Food Safety Authority. 2010. Quantitative microbial risk assessment on *Salmonella* in slaughter and breeder pigs: final report. Veterinary Laboratories Agency, Weybridge, UK
14. Hald, T., D. Vose, H. C. Wegener, and T. Koupeev. 2004. A Bayesian approach to quantify the contribution of animal-food sources to human salmonellosis. *Risk Anal.* 24:255–269.
15. Kjaerulff, U. B., and A. L. Madsen. 2008. Bayesian networks and influence diagrams. Springer, New York.
16. Korb, K. B., and A. E. Nicholson. 2004. Bayesian artificial intelligence. Chapman and Hall, CRC Press, London.
17. Letellier, A., G. Beauchamp, E. Guevremont, S. D'Allaire, D. Hurnik, and S. Quessy. 2009. Risk factors at slaughter associated with presence of *Salmonella* on hog carcasses in Canada. *J. Food Prot.* 72:2326–2331.
18. Morgan, I. R., F. L. Krautil, and J. A. Craven. 1987. Effect of time in lairage on caecal and carcass *Salmonella* contamination of slaughter pigs. *Epidemiol. Infect.* 98:323–330.
19. Namvar, A., and K. Warriner. 2005. Attachment strength to pork skin and resistance to quaternary ammonium salt and heat of *Escherichia coli* isolates recovered from a pork slaughter line. *J. Food Prot.* 68: 2447–2450.
20. Nauta, M. 2008. The modular process risk model (MPRM): a structured approach to food chain exposure assessment, p. 99–136. *In* D. Schaffner (ed.), Microbial risk analysis of foods. ASM Press, Washington, DC.
21. Oosterom, J., R. Dekker, G. J. de Wilde, F. van Kempen-de Troye, and G. B. Engels. 1985. Prevalence of *Campylobacter jejuni* and *Salmonella* during pig slaughtering. *Vet. Q.* 7:31–34.
22. Pearce, R. A., D. J. Bolton, J. J. Sheridan, D. A. McDowell, I. S. Blair, and D. Harrington. 2004. Studies to determine the critical control points in pork slaughter hazard analysis and critical control point systems. *Int. J. Food Microbiol.* 90:331–339.
23. Pollino, C. A., O. Woodberry, A. E. Nicholson, K. B. Korb, and B. T. Hart. 2007. Parameterisation and evaluation of a Bayesian network for use in an ecological risk assessment. *Environ. Model. Softw.* 22: 1140–1152.
24. Quirke, A. M., N. Leonard, G. Kelly, J. Egan, P. B. Lynch, T. Rowe, and P. J. Quinn. 2001. Prevalence of *Salmonella* serotypes on pig carcasses from high- and low-risk herds slaughtered in three abattoirs. *Berl. Muench. Tierarztl. Wochenschr.* 114:360–362.
25. Rivas, T., J. A. Vizcaino, and F. J. Herrera. 2000. Microbial contamination of carcasses and equipment from an Iberian pig slaughterhouse. *J. Food Prot.* 63:1670–1675.
26. Sauli, I., J. Danuser, A. H. Geeraerd, J. F. Van Impe, J. Rufenacht, B. Bissig Choisat, C. Wenk, and K. D. C. Stark. 2005. Estimating the probability and level of contamination with *Salmonella* of feed for finishing pigs produced in Switzerland—the impact of the production pathway. *Int. J. Food Microbiol.* 100:289–310.
27. Shapcott, R. 1984. Practical aspects of *Salmonella* control: progress report on a programme in a large broiler integration. Presented at the International Symposium on *Salmonella*. American Association of Avian Pathology, New Orleans, LA.
28. Smid, J. H., A. N. Swart, A. H. Havelaar, and A. Pielaat. 2011. A practical framework for the construction of a biotracing model: application to *Salmonella* in the pork slaughter chain. *Risk Anal.* 31: 1434–1450. doi:10.1111/j.1539-6924.2011.01591.x.
29. Smid, J. H., D. Verloo, G. C. Barker, and A. H. Havelaar. 2010. Strengths and weaknesses of Monte Carlo simulation models and Bayesian belief networks in microbial risk assessment. *Int. J. Food Microbiol.* 139:57–63.

30. Sorensen, L. L., L. Alban, B. Nielsen, and J. Dahl. 2004. The correlation between Salmonella serology and isolation of Salmonella in Danish pigs at slaughter. *Vet. Microbiol.* 101:131–41.
31. Sorensen, L. L., R. Sorensen, K. Klint, and B. Nielsen. 1999. Persistent environmental strains of *Salmonella* Infantis at two Danish slaughterhouses, two case-stories, p. 285–286. *In* Proceedings of the 3rd International Symposium on the Epidemiology and Control of *Salmonella* in Pork, Washington, DC.
32. Spescha, C., R. Stephan, and C. Zweifel. 2006. Microbiological contamination of pig carcasses at different stages of slaughter in two European Union–approved abattoirs. *J. Food Prot.* 69:2568–2575.
33. Spiegelhalter, D., and S. L. Lauritzen. 1990. Sequential updating of conditional probabilities on directed graphical structures. *Networks* 20:579–605.
34. Swanenburg, M., B. R. Berends, H. A. P. Urlings, J. M. A. Snijders, and F. van Knapen. 2001. Epidemiological investigations into the sources of *Salmonella* contamination of pork. *Berl. Muench. Tierarztl. Wochenschr.* 114:356–359.
35. Swanenburg, M., H. A. P. Urlings, J. M. A. Snijders, D. A. Keuzenkamp, and F. van Knapen. 2001. *Salmonella* in slaughter pigs: prevalence, serotypes and critical control points during slaughter in two slaughterhouses. *Int. J. Food Microbiol.* 70:243–254.
36. Swanenburg, M., P. J. van der Wolf, H. A. Urlings, J. M. Snijders, and F. van Knapen. 2001. *Salmonella* in slaughter pigs: the effect of logistic slaughter procedures of pigs on the prevalence of *Salmonella* in pork. *Int. J. Food Microbiol.* 70:231–242.
37. Titus, S. 2007. A novel model developed for quantitative microbial risk assessment in the pork food chain. Ph.D. dissertation. Massey University, Palmerston North, New Zealand.
38. van Hoek, A. H. A. M., R. de Jonge, W. M. van Overbeek, E. Bouw, A. Pielaat, J. H. Smid, C. Löfström, B. Malorny, E. Junker, K. Pedersen, H. J. M. Aarts, and L. Heres. A quantitative approach toward a better understanding of the dynamics of *Salmonella* spp. in a pork slaughter line. *Int. J. Food Microbiol.,* in press.
39. van Pelt, W., D. Notermans, D. J. Mevius, H. Vennema, M. P. G. Koopmans, and Y. T. H. P. van Duynhoven. 2008. Trends in gastro-enteritis van 1996–2006: Verdere toename van ziekenhuisopnames, maar stabiliserende sterfte. *Infect. Bull.* 19:24–31.
40. Vieira-Pinto, M. M., and C. Martins. 2003. Detection of *Salmonella* in pork at the slaughterhouse, p. 237. *In* Proceedings of the National Congress of Microbiology, Tomar, Portugal.
41. Warriner, K., T. G. Aldsworth, S. Kaur, and C. E. Dodd. 2002. Cross-contamination of carcasses and equipment during pork processing. *J. Appl. Microbiol.* 93:169–177.
42. Zwietering, M. H. 2009. Quantitative risk assessment: is more complex always better? Simple is not stupid and complex is not always more correct. *Int. J. Food Microbiol.* 134:57–62.

Chapter 6

Variability and uncertainty analysis of the cross-contamination ratios of *Salmonella* during pork cutting.

J.H. Smid [a,b], R. de Jonge [a], A. Swart [a], A.H. Havelaar [a,b] , A. Pielaat [a]

[a] RIVM, A. van Leeuwenhoeklaan 9, 3720 BA Bilthoven, The Netherlands

[b] IRAS, Yalelaan 2, 3584 CM Utrecht, The Netherlands

Submitted for publication

ABSTRACT

The transfer ratio of bacteria from one surface to another is often estimated from laboratory experiments and quantified by dividing the expected number of bacteria on the recipient surface by the expected number of bacteria on the donor surface. Yet, the expected number of bacteria on each surface is uncertain due to the limited number of colonies that is counted and/or samples that can be analyzed. The expected transfer ratio is, therefore, also uncertain and its estimate may exceed 1 if real transfer is close to 100%. In addition, the transferred fractions vary over experiments but it is unclear, using this approach, how to combine uncertainty and variability into one estimate for the transfer ratio. A Bayesian Network model was proposed which allows the combination of uncertainty within one experiment and variability over multiple experiments and prevents inappropriate values for the transfer ratio. Model functionality was shown using data from a laboratory experiment in which the transfer of *Salmonella* was determined from contaminated pork meat to a butcher's knife, and vice versa. Recovery efficiency of bacteria from both surfaces was also determined and accounted for in the analysis. Transfer ratio probability distributions showed a large variability, with a mean value of 0.19 for the transfer of *Salmonella* from pork meat to the knife and 0.58 for the transfer of *Salmonella* from the knife to pork meat. The proposed Bayesian model can be used for analyzing data from similar study designs in which uncertainty should be combined with variability.

1. INTRODUCTION

Cross-contamination is an important process in food safety and it is, therefore, included in many quantitative microbial risk assessment (QMRA) models describing the dynamics of pathogens in food chains. Cross-contamination can be the direct transfer of cells from one product unit to another, e.g. by the physical contact between them, or the indirect transfer, e.g. via the processing environment (Nauta, 2008). Typically, transfer ratios are experimentally determined in laboratory studies or in real environments such as factories (Perez-Rodriguez et al., 2008). The first step in these experiments consists of counting the number of bacteria on the contaminated *donor surface*. The surface, or a section thereof, is subsequently pressed against another surface, the *recipient surface*. Next, the transferred bacteria are recovered from the recipient surface. This is often done by swabbing, which consists of stroking a surface using a cotton or cellulose swab and then inserting the swab into a sterile extracting solution for detection and / or enumeration. Quantifying the bacteria in the solution is typically done by plate counting.

In spite of its popularity, plate counting poses certain methodological and theoretical drawbacks for determining transfer ratios. First, the number of bacteria on a surface is determined by swabbing a section of that surface or, in case of the donor surface, by counting the number of bacteria in the cell suspension used for contaminating that surface. Yet, the bacteria that are physically involved in the transmission experiment are in all cases a different subset of the population than the subset of bacteria that is counted. The size of the subpopulation from which bacterial cells are transferred is, therefore, known with limited precision.

Next, plate counts vary according to a Poisson distribution if a homogeneous distribution of the bacteria in the plated volume may be assumed. Jongenburger *et al.* (2010) showed that the Poisson distributed count variability greatly determines the uncertainty of the estimated number of bacteria in the original suspension. In particular, this error increases significantly for lower plate counts. The counted number of colony forming units (CFU) on the plate may lead to an inaccurate estimate of the Poisson parameter due to variability between counts, and this inaccuracy is multiplied by the dilution factor for the estimated number of bacteria in the original suspension.

Computational difficulties have been avoided in published transfer studies by approximating the real number of bacteria on a surface by the *expected* number of bacteria on that surface. Uncertainty around the true value of this parameter is often disregarded (Chen et al., 2001; Kusumaningrum et al., 2003; Moller et al., 2012; Moore et al., 2003; Nauta et al., 2005). Yet, this uncertainty may be considerable because of the reasons mentioned above.

If uncertainty is taken into account, subtle technical difficulties in determining the transfer ratio may arise. The transfer ratio (TR) is often defined as proposed by Chen *et al*. (2001):

$$TR = (\text{bacteria on surface } B / \text{bacteria on surface } A),$$

where A and B represent the donor and recipient surfaces, respectively. Approximations of the real

numbers of bacteria by expected numbers results in an *expected* transfer ratio, its probability distribution representing uncertainty. However, it has been acknowledged by several authors that transfer ratios may vary over experiments. It is unclear how the uncertainty distributions of expected transfer ratios fit into this variability model. Therefore, uncertainty was in many transfer studies neglected in the description of the variability of the transfer ratio over experiments. In the remainder of the text, we refer to this, commonly used, approach as the *classical model*.

In this paper we propose an alternative *Bayesian network model* to determine transfer ratio probability distributions. The probability distributions describing uncertainty about the true number of bacteria on both surfaces in each individual experiment and the probability distribution describing the variability of the transfer ratio over experiments were combined in this one model. We show the deficiencies of the classical model and compare its results to the results of the proposed network model using data from transfer experiments of *Salmonella* from a raw pork meat slab to the blade of a butcher's knife, and vice versa. Salmonellae were quantified by plate counting.

The transfer experiments that were done in this paper mimic carcass splitting in a pig slaughterhouse. During carcass splitting, a carcass may have a contaminated skin, from which salmonellae are transferred to the blade of the carcass splitter. The cutting section of carcasses processed thereafter may, subsequently, be cross-contaminated by this blade. These specific transfer ratios are so far not known and approximations have been made in published QMRA modeling studies to the true distributions of values of the transfer ratios, e.g (EFSA, 2010; Smid et al., 2011). The bacterial recovery ratio using swab samples was also examined, and the impact of this recovery on the reported transfer ratios was discussed.

2. MATERIALS AND METHODS

The M&M is subdivided into two sections. The first section describes the theoretical construction of the models used for determining the transfer and recovery ratios. The second section provides details of the experimental set-up by which the data were generated to determine these transfer ratios for *Salmonella* between raw pork meat and a butcher's knife and the recovery ratios of *Salmonella* from these surfaces.

2.1 Model Construction

2.1.1 Notation

Subscript i refers to the i^{th} experiment to determine a transfer ratio. Subscript j refers to the j^{th} repetitive counting of CFU on the plate within one experiment. Unless otherwise specified, in the notation below we refer to experiment i. Random variables are indicated by capital letters. Realizations of random variables and fixed values in the model are indicated by lowercase letters. Subscripts S refers to (swab samples from) a general surface area. Subscripts A and B specify this notion and refer to (swab samples from) the donor surface area (A) and the recipient surface area (B).

2.1.2 Expected number of bacteria on a surface

The first dilution factor for which CFU can be counted on the plate is notated as $f_{S,i} = 10^n$ ($n \in \mathbb{N}$). Successive 10-fold dilutions from this first dilution factor are notated as $d_{S,i,j} \in \{10^0, 10^1, 10^2\}$,

Table I. Overview of the variables in both models [a]

Variable		Description	(Likelihood) function / distribution	
			Classical model	Bayesian model.
1	$\Lambda_{S,i}$	Expected bacterial number on surface area *S*.	$Gamma(k_S, \theta_S)$	
2	$f_{S,i}$	First dilution factor providing countable bacterial numbers on the plate from a surface area *S* sample.	value	
3	$E_{S,i}$	Expected bacterial number in a $f_{S,i}$-diluted surface area *S* sample.	$\Lambda_{S,i}/f_{S,i}$	
4	$d_{S,i,j}$	Additional dilution factor used for counting the bacterial number on the plate from the $f_{S,i}$-diluted surface area *S* sample.	value	
5	$D_{S,i,j}$	Expected bacterial number in a $(f_{S,i} \times d_{S,i,j})$-diluted surface area *S* sample.	$E_{S,i}/d_{S,i,j}$	
6	$X_{S,i,j}$	Number of CFU on the plate from a $d_{S,i,j}$-diluted sample from surface area *S*.	$Poisson(D_{S,i,j})$	
7	$N_{A,i}$	Total bacterial number on surface area *A*.	not in model	$Poisson(\Lambda_{A,i})$
8	$N_{B,i}$	Total that is transferred via cross-contamination from area *A* to *B*.	not in model	$Bin(N_{A,i}, TR_{AB})$
9	$\Lambda_{B,i}$	Expected bacterial number on surface area *B*.	$Gamma(k_B, \theta_B)$	$Gamma(k_B + N_{B,i}, \frac{\theta_B}{\theta_B+1})$
10	$f_{B,i}$	First dilution factor providing countable bacterial numbers on the plate from a surface area *B* sample.	value	
11	$E_{B,i}$	Expected bacterial number in a $f_{B,i}$-diluted surface area *B* sample.	$\Lambda_{B,i}/f_{B,i}$	
12	$d_{B,i,j}$	Additional dilution factor used for counting the bacterial number on the plate from the $f_{B,i}$-diluted surface area *B* sample.	value	
13	$D_{B,i,j}$	Expected bacterial number in a $(f_{B,i} + d_{B,i,j})$-diluted surface area *B* sample.	$E_{B,i}/d_{B,i,j}$	
14	$X_{B,i,j}$	Number of CFU on the plate from a $d_{B,i,j}$-diluted sample from surface area *B*.	$Poisson(D_{B,i,j})$	
15	$T_{AB,i}$	Expected transfer ratio in experiment *i*	Empirical distribution: $min(\frac{\lambda_{A,i}}{\lambda_{B,i}}, 1)$	not in model
16	TR_{AB}	Probability for one bacterial cell to be transferred from area *A* to *B*.	Distribution fitted to all $E[T_{AB,i}]$, with shape: $Beta(\hat{p}, \hat{q})$	$Beta(P, Q)$
17	P	Beta distribution parameter	not in model	$Unif[p_1, p_2]$
18	Q	Beta distribution parameter	not in model	$Unif[q_1, q_2]$

[a] *In the notation for the variables in the model, subscript S refers to a general surface, A refers to the donor surface, and subscript B refers to the recipient surface.*

where $d_{S,i,j} = 10^0$ represents, de facto, no further dilution. The total dilution factor used for counting the number of bacteria on plate j (in transfer experiment i) is $f_{S,i} \times d_{S,i,j}$. For each total dilution, the number of CFU is counted. The expected bacterial number in the original suspension ($\Lambda_{S,i}$) is calculated from the plate counts, the dilution factor and the plate volume by using Bayesian inversion (Vose, 2000).

A prior [1] $Gamma(k_S, \theta_S)$ distribution was chosen for $\Lambda_{S,i}$. This distribution is the natural conju-

1 In the Bayesian context, the prior (probability) distribution of a variable refers to the distribution of that variable without taking observations into account. Its posterior (probability) distribution is the distribution of that variable given that other (sets of) variables are observed.

gate of a Poisson distribution (Gelman et al., 2004) and is a suitable choice for $\Lambda_{S,i}$ for reasons that will become clear in Section 2.2.4. The parameters k_S and θ_S were chosen as fixed numbers in such a manner that this prior was slightly informative, with the mean value for the expected number of CFU counted after the first dilution ($E_{S,i} = \Lambda_{S,i}/f_{S,i}$) equaling 100 and most probability mass in the range from 0 to 300, which covers all potentially countable numbers of bacteria on the plate. A division of $\Lambda_{S,i}$ by the total dilution factor $f_{S,i} \times d_{S,i,j}$ gives $D_{S,i,j}$, the expected number of bacteria on the plate. A homogeneous distribution of bacteria in the plated volume was assumed, therefore the likelihood function for the number of CFU on the plate ($X_{S,i,j}$) is:

$$X_{S,i,j} \sim Poisson(D_{S,i,j}) \tag{1}$$

$x_{S,i,j}$ is the actually counted number of CFU. Bayesian inversion was used to obtain the posterior distribution $D_{S,i}|X_{S,i,j} = x_{S,i,j}$. Multiplication of $D_{S,i,j}$ by the total dilution factor provides the posterior distribution of $\Lambda_{S,i}$. See variables 1-6 in Table I. The leftmost circled part of Fig. 1 shows these relations as a network graph.

For Poisson distributed numbers it should hold that their variance equals their mean. Therefore, to examine whether the sets of counts for one experiment, one dilution factor and for one surface area are indeed samples from a Poisson distribution, the variance divided by the mean was computed for each such dataset. A one-sided sign test was used to test the null-hypothesis that the median of the population of these fractions is equal to1. The alternative hypothesis is that the median exceeds one, which indicates overdispersion of the counts.

2.1.3 Classical model for assessing the transfer ratio probability distribution

In the classical model (Table I), Bayesian inference was only used to obtain the posterior distributions of the expected number of bacteria on the donor and recipient surfaces A and B ($\Lambda_{A,i}$ and $\Lambda_{B,i}$, respectively). This was done such as described in the previous section for a general surface area S . Variables $\Lambda_{A,i}$ and $\Lambda_{B,i}$ estimate the real number of bacteria on both surfaces in this model (see Introduction). The empirical distribution of expected transfer ratios of bacteria from surface area *A* to *B* ($T_{AB,i}$) was obtained by simulating numbers $\lambda_{A,i}$ and $\lambda_{B,i}$ from $\Lambda_{A,i}$ and $\Lambda_{B,i}$ and computing

$$\tau_{AB,i} = min(\frac{\lambda_{B,i}}{\lambda_{A,i}}, 1) \tag{2}$$

A maximum of 1 was set to each $\tau_{AB,i}$ to avoid expected transfer ratios exceeding one, which may be observed if the true transfer is high. A simple division of $\lambda_{B,i}$ by $\lambda_{A,i}$ is allowed because $\Lambda_{A,i}$ and $\Lambda_{B,i}$ have distributions representing uncertainty, i.e. each pair of realizations of these two distributions is independent.

In the next step of the classical model, the expected values of the empirical distributions of $T_{AB,i}$ (for different *i*) were used to assess the variability distribution of expected transfer ratios over all experiments. Chen *et al* (2001) found that a Beta distribution often provides the best description of the vari-

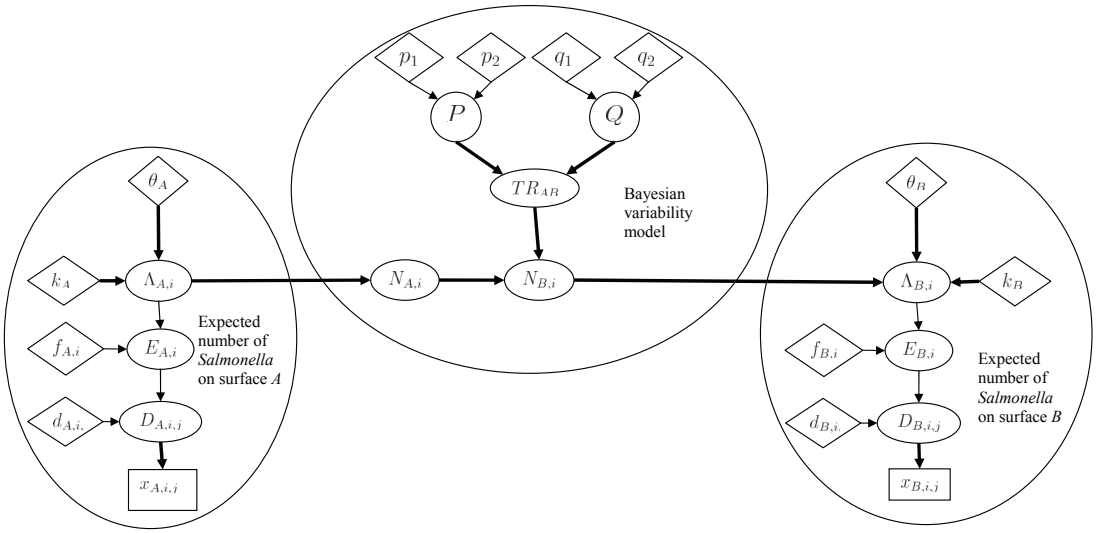


Fig. 1: *Network description of the different parts of the BBN. The network parts "Expected bacterial number on surfaces A and B" were used for inferring the uncertainty distributions for the expected bacterial numbers on the donor and recipient surfaces, respectively. These parts were also used in the classical model. The network part "Bayesian variability model" was used for inferring the variability distribution of the transfer rates over different experiments and was used in the Bayesian model exclusively. Random variables (RV) are circled. Realizations of RV and fixed values in the model are indicated by diamond symbols. Thick arrows indicate a probabilistic relationship between variables, thin arrows a functional relationship.*

ability of the transfer ratios over experiments, as it can take many shapes. We estimated the parameters of the Beta distribution that provided the best fit to the point values $E[T_{AB,i}]$ over experiments i. The Log-likelihood of the fit was used to assess which distribution fitted the data best. MATLAB's dfittool (The MathWorks Inc., Natick, MA, 2010), based on maximum likelihood estimation, was used for fitting. The transfer ratio TR_{AB} was defined as this fitted variability distribution.

2.1.4 Bayesian model for assessing the transfer ratio probability distribution

In the full Bayesian network model (Table I and Fig. 1), bacteria on the donor surface were again assumed to be homogeneously distributed. Therefore,

$$N_{A,i} \sim Poisson(\Lambda_{A,i}) \tag{3}$$

where the posterior distribution for $\Lambda_{A,i}$ was obtained as described in 2.1.2. The transfer ratio TR_{AB} was in this model defined as the probability for one bacterial cell to be transferred from the donor surface area to the recipient surface (Perez-Rodriguez et al., 2008). The number of bacteria on the recipient surface is then described by a binomial distribution:

$$N_{B,i} \sim Bin(N_{A,i}, TR_{AB}) \tag{4}$$

In addition, it was assumed that bacteria on the recipient surface were homogeneously distributed. This

means that the probability distribution for $N_{B,i}$ has an alternative description via a $Poisson(\Lambda_{B,i})$ likelihood function. Let the prior distribution for $\Lambda_{B,i}$ be $Gamma(k_B, \theta_B)$. Then the causal direction of the Poisson likelihood can be inverted by conjugacy (Gelman et al., 2004) with the Poisson distribution so that

$$\Lambda_{B,i} \sim Gamma(k_B + N_{B,i}, \tfrac{\theta_B}{\theta_B+1}) \tag{5}$$

This Gamma distribution is further updated by plate count data from the recipient surface $x_{B,i,j}$, as described before. Again, parameters k_B and θ_B were chosen as fixed numbers in such a manner that the Gamma prior was slightly informative on an appropriate interval.[2] The variability of TR_{AB} over all experiments was again defined by a Beta distribution:

$$TR_{AB} \sim Beta(P, Q) \tag{6}$$

The posterior distributions of the parameters P and Q were estimated by using the count data and Bayesian inference in the network model. Hereto, P and Q were chosen to have a non-informative (uniform) prior distribution on an appropriate interval. The count data (i.e., the observations of the variables $x_{A,i,j}$ and $x_{B,i,j}$) were used to refine the prior distributions of parameters P and Q through a Bayesian algorithm called *sequential adaptation* (Spiegelhalter & Lauritzen, 1990). During one adaptation step, the count data for one experiment *i* was fed to the model, generating posterior probability distributions for these parameters. Subsequently, the posterior distributions were used to update the prior distributions for the parameters in the light of the new data and of the strength of belief about the correctness of the prior distributions. This belief is expressed by the notion of *experience* (Spiegelhalter & Lauritzen, 1990). The data were subsequently retracted from the model. Such steps were successively repeated with the count data for all experiments, which reduces the uncertainty about the parameter values of the model (which can be seen as increasing experience).

The model was constructed as a Bayesian network in HUGIN software (HUGIN Expert, Aalborg, Denmark). In this software, the domain of continuous variables is to be discretized. The general guideline that was used for an optimal discretization was to have an equal probability mass per bin. This was done by a trial and error approach.

2.1.5 Transfer ratios corrected for limited recovery

The equations describing recovery are similar to equations describing transfer, because recovery can be seen as the event in which bacteria are transferred from a surface to a swab sample. Therefore, both models were also used to assess recovery ratios. To indicate that we refer to the recovered fraction of bacteria from surface area S, we use the superscript ^{rec} in the notation of the variables, e.g. $\Lambda^{rec}_{S,i}$ is the expected number of recovered bacteria from surface area *S* in experiment *i*. The recovery analogues of $T_{AB,i}$ and TR_{AB} were defined as $R_{S,i}$ and Rec_S, respectively.

2 The criterion for which interval was "appropriate" was that the probability mass of the posterior distribution was well in the centre between the lower and upper bounds of the interval

In the methods to determine the transfer ratio distributions, such as described above, limited recovery was not accounted for. If the researcher wants to correct the experimentally determined transfer ratios for limited recovery, the bold assumption needs to be made that these variables are independent of each other (Perez-Rodriguez et al., 2008). In the Discussion we will elaborate on why this is a bold assumption. The empirical distribution of the real transfer ratio p_{AB}, defined as the probability for a bacterial cell to be transferred from surface area *A* to surface area *B*, may then be computed as

$$p_{AB} = TR_{AB}\frac{Rec_A}{Rec_B} \tag{7}$$

Yet, this may result in values for p_{AB} exceeding 1. The distribution of p_{AB} was truncated at 1 to prevent such values. Although limited recovery is an important issue in transfer experiments, a detailed analysis of this phenomenon was beyond the scope of this study. In most of the comparative analyses of the classical and the Bayesian transfer models the experimental transfer ratios were, therefore, not corrected for limited recovery. The effect of limited recovery for the transfer rations of *Salmonella* between raw pork meat and a butcher's knife will be shown briefly, using the Bayesian model.

2.2 Experimental Set-up

The functionality of both transfer models was shown using experimental data for the transfer of *Salmonella* between raw pork meat (M) and a butcher's knife (K) and for the recovery of *Salmonella* from these surfaces. The performed transfer experiments mimic cross-contamination in a pig slaughterhouse from the carcass skin, via the carcass splitter to the cutting section of carcasses processed thereafter. Therefore, the rind side of the meat slab was designated as the donor surface and the cutting section as the recipient surface. Specifically, the following transfer and recovery ratios were determined:

- The experimental transfer ratio of *Salmonella* from the contaminated rind side of pork meat to the knife (TR_{MK})
- The experimental transfer ratio of *Salmonella* from the contaminated blade of the knife to the cutting side of pork meat (TR_{KM})
- The experimental transfer ratio of *Salmonella* from the contaminated rind side of pork meat via the blade of the knife to the cutting side of uncontaminated pork meat (TR_{MKM})
- The recovery ratio of *Salmonella* from the blade of the knife by swabbing (Rec_K)
- The recovery ratio of *Salmonella* from the cutting side of pork meat by swabbing (Rec_M)

2.2.1 Culture preparation

Salmonella Derby was the predominant serotype in a Dutch pork slaughterhouse (Hoek et al., 2012). A strain of *S.* Derby isolated from pigs was chosen for the experiments and grown overnight at 37ºC in brain heart infusion (BHI) broth. For additional selectivity, the isolate was selected to be tetra-cycline resistant. To study the effect of the inoculum's concentration, one half of this culture was diluted by a factor 100 in phosphate buffered saline (PBS). Directly before the experiment, 6g of pig feces was homogenized with 3ml PBS. For experiments in which a high initial bacterial concentration was used, 1ml of the undiluted cell culture was added to this matrix. This mix is referred to as the "concentrated fecal mixture". For other experiments, a lower initial bacterial concentration was used. In this case,

1ml of the diluted cell culture was added to the matrix (the "diluted fecal mixure").

Fig. 2: *Set-up of the transmission experiment.*

2.2.2 Transfer experiment

Standard pork bellies, having one rind side, were used which were cut into parts of approximately 30x10cm^2. Tests were done to verify whether the rind side of the meat slabs did not already contain tetra-cycline resistant *S.* Derby prior to the experiment. These tests were in all cases negative. For the transfer experiments, meat slabs were held upright with the rind side facing forward (Fig. 2). The knife used in this experiment was a butcher's knife (Wüsthof 4685, X50 Cr Mo V15, Solingen, Germany) with the blade having a length of 188mm and cross section of 4mm. The blade was made of stainless steel. Before the transfer experiments, the knife was sterilized using alcohol, soap and hot water.

In order to calculate the actual number of transferred bacteria, the contact surface of the knife with the rind side of the meat slab needs to be known. This surface was assessed by putting colored lipstick on the blade of the knife and subsequently cutting through the rind. By measuring the width (0.1-0.15cm on both sides of the cut) and length of the cut (10cm), the contact surface of the knife with the rind side was assessed to be approximately 2-3cm^2. This assessment was only done once.

- To determine TR_{MK}, 1g of the fecal mixture was spread homogeneously onto the rind side of the meat slab. Subsequently, the sterile knife was used to cut through the contaminated meat slab. The entire surface of the knife's blade was swabbed with a sponge (3M Dry-sponge) soaked in 10ml of buffered peptone water (BPW). Six repeats were done with the concentrated and six with the diluted fecal mixture.
- To determine TR_{KM}, 0.5g of the fecal mixture was spread homogeneously onto the blade of the knife. Subsequently, the contaminated knife was used to cut through the uncontaminated meat slab. The cutting side of the meat slab was swabbed with a sponge soaked in 10ml of BPW. Six repeats were done with the concentrated and six with the diluted fecal mixture.
- To determine TR_{MKM}, 1g of the fecal mixture was spread homogeneously onto the rind side of the meat slab. Subsequently, a sterile knife was used to cut through the contaminated meat slab. A second meat slab was cut through with the same knife. The cutting side of the second meat slab was swabbed with a sponge soaked in 10ml of BPW. Five repeats were done with the concentrated and five with the diluted fecal mixture.
- To determine Rec_K, 0.5g of the fecal mixture was brought onto the blade of the knife. Subsequently, the knife's blade was swabbed with a sponge soaked in 9ml of BPW. Six repeats were done with the concentrated fecal mixture.
- To determine Rec_M, 1g of the fecal mixture was brought onto the cutting side of the meat slab. Subsequently, the cutting side of the meat slab was swabbed with a sponge soaked in 10ml of BPW. Four repeats were done with the concentrated and two with the diluted fecal mixture.

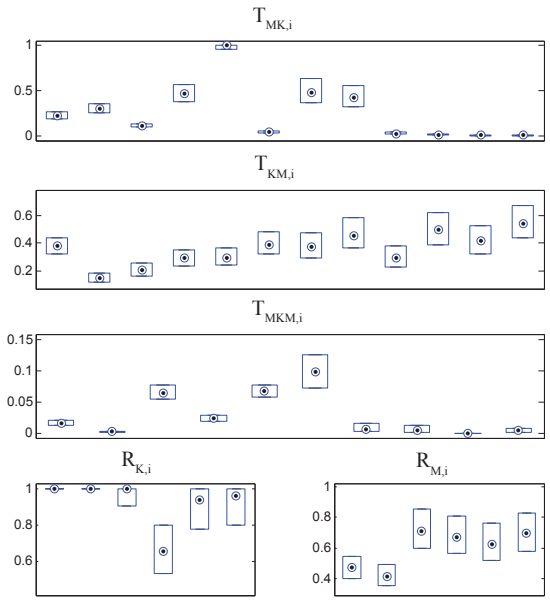


Fig. 3: *Expected transfer ($T_{AB,i}$) and recovery ($R_{S,i}$) rate distributions per experiment, as computed by the classical model. The central black dot of each box is the median, and the edges of each box are the 2.5% and the 97.5% percentiles.*

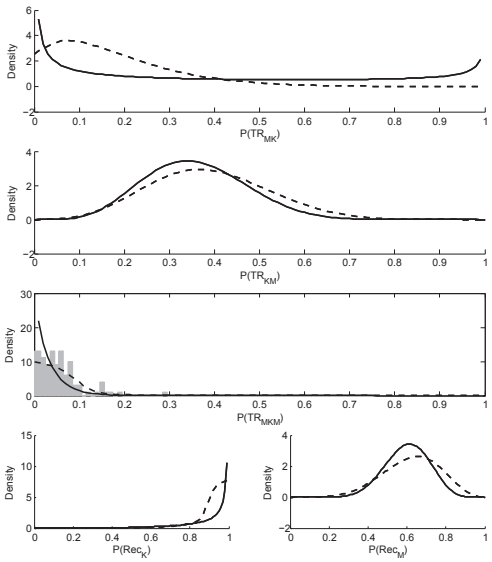


Fig. 4: *Density plots of transfer and recovery ratio distributions, as inferred by the different models. For the classical model (solid line), the density plots only represent variability over experiments. For the Bayesian model (dashed line), the density plots combine uncertainty within one experiment with variability over experiments. The histogram in the third window shows random samples from the distribution of T_{MK} multiplied with random samples from the distribution of T_{KM}*

The sponges were put into sterile plastic bags to which another 10ml of BPW was added.

2.2.3 Cell enumeration

In all experiments, the number of *Salmonella* in the fecal mixtures and in the sponge samples was assessed by spread plating appropriate 10-fold dilutions in peptone physiological salt on brilliant green agar (BGA) plates supplemented with 50µg/ml tetracycline. Plates were incubated for 20±2h at 37°C. For each experiment, this was done in two or three fold and up to three successive 10-fold dilutions from the first dilution factor which provided countable numbers on the plate (CFU$\leq$ 200). The swab samples from the knife and from the meat were also enumerated using appropriate 10-fold dilutions followed by plating, incubation and direct plate counting.

3. RESULTS

3.1 Expected Number of Salmonella

For all surface areas, the counted number of *Salmonella* could not be shown to be draws from a more dispersed distribution than Poisson ($p<0.05$).

Table II: Statistical analysis results for transfer and recovery ratio distribution fitting[a]

Type of transfer	Parameter values Beta distribution		Log-likeli-hood	Statistics TR		Percentiles TR				
	$\hat{p}$	$\hat{q}$		mean	SD	2.5%	25%	50%	75%	97.5%
TR_{MK}	0.33	0.53	*7.24*	0.38	0.36	0.00	0.04	0.28	0.72	1.00
TR_{KM}	6.26	11.20	*9.41*	0.36	0.11	0.16	0.28	0.35	0.43	0.59
TR_{MM}	0.85	25.34	*21.97*	0.03	0.03	0.00	0.01	0.02	0.05	0.12
Rec_K	2.91	0.22	*19.12*	0.93	0.13	0.52	0.92	0.99	1.00	1.00
Rec_M	10.83	7.20	*4.64*	0.60	0.11	0.37	0.52	0.61	0.68	0.81

[a] Only the expected values of the transfer rates were used, so Poisson uncertainty of the count data is not accounted for.

3.2 Classical model

3.2.1 Uncertainty

The expected transfer ratios, directly derived from the expected total number of *Salmonella* on each surface by the classical model, are shown as a box plot in Fig. 3. The expected transfer ratio from meat to the knife was found to be highly variable, with $E[T_{MK,i}]$ ranging from less than 0.01 to 1. The expected transfer ratio from the knife to meat was less variable, with $E[T_{KM,i}]$ ranging between 0.15 and 0.54. Expected transfer from a meat slab, via the knife to another meat slab ranged between 0 and 0.10. The uncertainty about the expected transfer ratios, originating from the uncertainty about the Poisson parameters, was considerable. Yet the non-overlapping 95% uncertainty intervals of the expected transfer ratios indicated that their variability was in most cases larger. For the recovery of *Salmonella* (Fig. 3) from the knife, almost all expected recovery ratios were 1 or close to 1, with the exception of $R_{K,4}$ which has a 95% uncertainty interval between 0.53 and 0.81. The expected recovery ratios from meat were significantly lower, with $E[R_{M,i}]$ ranging between 0.41 and 0.70 and the 95% uncertainty intervals ranging between 0.35 and 0.86.

3.2.2 Variability

The results obtained by starting experiments with concentrated and diluted fecal mixtures were combined to fit Beta probability distributions describing variability over experiments. The parameters of the best-fitting Beta distributions and the log-likelihood of the fits are shown in Table II. Also shown are the summarizing statistics of these distributions. Fig. 4 shows the density plots of the best-fitting Beta distributions. For all types of transfer, the null-hypothesis that the calculated expected transfer and recovery ratios were draws from the fitted Beta distributions was not rejected by a two-sample Kolmogorov-Smirnov test.

3.3 Bayesian Model

Table III shows the mean values and 95% credible intervals of the posterior parameter distributions for the transfer and recovery ratios, as inferred by the Bayesian model. The majority of these mean values were close to the point estimates by the classical model (Table II). The results for TR_{MK} and Rec_K were exceptions as its estimated parameters by the classical model were close to (Rec_K) and outside (TR_{MK}) the 95% credible bounds. Table III also shows the summarizing statistics of the posterior transfer and recovery ratio distributions. Their density plots are shown in Fig. 4. It can be seen that all

Table III: Distributions of transfer and recovery ratios inferred by the Bayesian model

Type of transfer	Parameter values Beta distribution		Statistics TR		Percentiles TR				
	P [95% CI]	Q [95% CI]	mean	SD	2.5%	25%	50%	75%	97.5%
TR_{MK}	1.8 [1.1,2.9]	8.89 [4.9,14.0]	0.17	0.13	0.01	0.07	0.14	0.24	0.49
TR_{KM}	7.43 [5.1,9.9]	11.62 [7.2,18.0]	0.39	0.13	0.16	0.30	0.38	0.47	0.66
TR_{MM}	0.61 [0.00,1.9]	20.94 [7.0,45.0]	0.03	0.05	0.00	0.00	0.01	0.04	0.18
Rec_K	15.53 [2.7,45]	0.73 [0.0,2.5]	0.93	0.11	0.62	0.91	0.97	0.99	0.99
Rec_M	13.94 [7.2,23.0]	8.16 [3.6,14.0]	0.62	0.14	0.32	0.53	0.63	0.73	0.87

[a] CI stands for credible interval (a.k.a. Bayesian confidence interval).

distributions were less peaked when using the Bayesian model, compared to when using the classical model. This means that the probability mass of these ratios was more spread over the interval $[0, 1]$. This is logical as the ratio distributions in the Bayesian model combined variability over experiments with uncertainty within one experiment, instead of only variability as in the classical model. The results for TR_{MK} between the classical and the Bayesian model were most markedly different. These differences were caused by multiple findings of an expected transfer ratio below 0.5 and one close to 1, using the classical model (Fig. 3). This forced the fitting curve of TR_{MK} to its current shape. A more right-skewed variability distribution was obtained by the Bayesian model because that model accounted for uncertainty of the count data (i.e. it smoothens out the extreme finding of the expected transfer ratio close to 1). The effect of uncertainty in relation to variability is also visualized in Fig. 5. For each type of transfer and recovery, this spaghetti plot shows 25 cumulative density functions (CDF) representing possible variability distributions of these ratios inferred by the Bayesian model. The estimated transfer and recovery ratio distributions inferred by the classical model are also shown. The effect of including uncertainty in the analyses was considerable, especially if the number of experiments per type of transfer was low (such as for Rec_M).

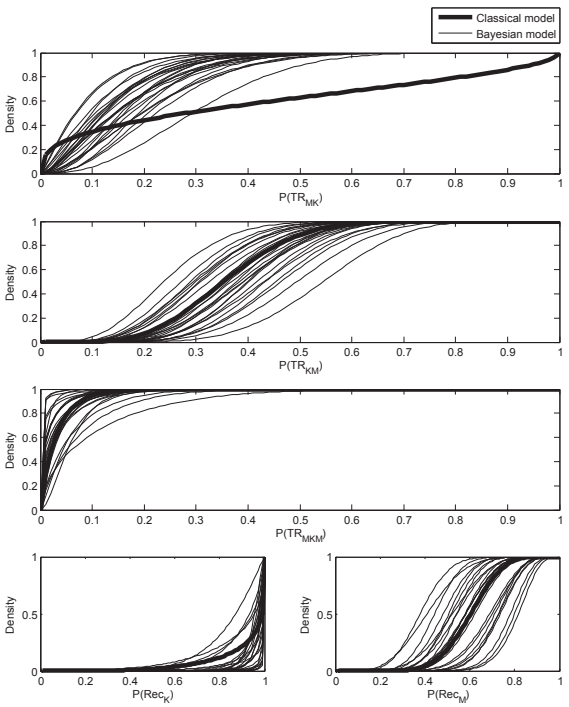


Fig. 5*: Spaghetti plot showing uncertainty and variability of the transfer ratio, as inferred by the Bayesian model. Each cumulative distribution function (CDF) represents a possible transfer ratio variability distribution. The CDF in bold represents the transfer ratio distribution as inferred by the classical model.*

The histogram in the third window of Fig. 4 shows random samples from the posterior distribution of TR_{MK} multiplied with random samples from the posterior distribution of TR_{KM}. This histogram largely overlaps the probability density function of TR_{MKM} . This indicates an absence of correlation between TR_{MK} and TR_{KM} suggesting that the transfer of *Salmonella* from meat to the knife and the transfer of *Salmonella* from the knife back to meat can be modeled as independent events.

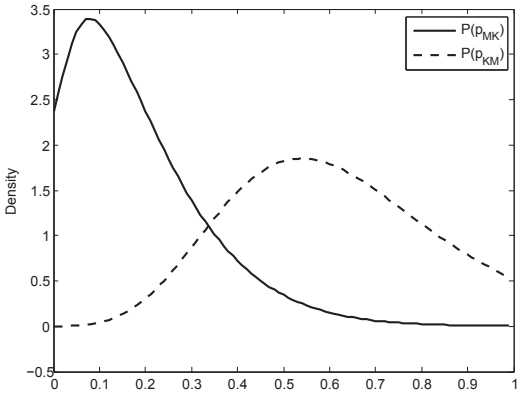


Fig. 6: *Transfer ratio distributions corrected for recovery of the swab procedures.*

3.4 Transfer Ratios Corrected for Limited Recovery

In our experiments, the *Salmonella* population size in the fecal mixture was assessed before it was brought on the donor surface; therefore Equation 7 simplifies to $p_{AB} = TR_{AB}/Rec_B$. The mean transfer ratios corrected for recovery were 0.19 for the transfer of *Salmonella* from pork meat to the knife and 0.58 for the transfer of *Salmonella* from the knife to pork meat (Fig.6). The variance of the distributions of the corrected transfer ratios was larger than the variance of the measured transfer ratio distributions, which is explained by the low precision of the recovery ratios.

4. DISCUSSION

The simple approach of computing the transfer ratio by taking the fraction of the expected number of *Salmonella* on the recipient surface to the expected number of *Salmonella* on the donor surface may result in transfer ratios exceeding 1. E.g., $E[\Lambda_{M,5}]$ exceeded $E[\Lambda_{K,5}]$ in the meat to knife transfer experiment. Such inconsistencies can be explained by uncertainty about the expected number of bacteria on both surfaces as a result of limited counting data, in particular if the transfer and recovery ratios are close to 1. The strategy used in the classical model to prevent such inconsistencies was setting an upper bound of1to the calculated transfer ratio. The approach of considering the full joint probability of the model domain including uncertainty, and regarding the transfer ratio as a probability, as was taken in the Bayesian model, is a less arbitrary approach to prevent such inconsistencies.

A researcher's knowledge about transfer and recovery is imperfect because only a limited number of experiments can be done. The classical model simplifies the analysis of the transfer / recovery ratios by neglecting uncertainty in the description of the variability of these ratios. By using a small dataset we have provided a proof-of-principle that this approximation is not always justifiable. Since the Bayesian model explicitly quantifies uncertainty instead of using point values for the parameters, it spreads the probability mass of the transfer ratio distributions more evenly over the interval $[0, 1]$ (Fig. 4). If uncertainty is disregarded, the estimated transfer ratio variability distribution may be markedly different from the true transfer ratio variability distribution (which may be one of the curves in Fig. 5).

The process generating the transfer ratios is complex and variability of these ratios can, therefore, not be perfectly summarized by one mathematical probability distribution (Perez-Rodriguez et al., 2008). The number of experiments that are done to analyze one transfer event is typically low, which precludes a reliable analysis to find a best-fitting probability distribution. A Beta distribution can take many shapes and was therefore chosen for describing variability of the transfer and recovery ratios (Chen et al., 2001). In several other studies (Chen et al., 2001; Perez-Rodriguez et al., 2008) a log -normal distribution has also been proposed as an adequate distribution to describe such variability. Yet, a log-normal distribution does not have an upper bound, which allows potential transfer / recovery ratios exceeding 1. A log-normal distribution may, therefore, not be suitable if the expected transfer or recovery is close to one, such as was the case for the recovery ratio from the knife. Results for our dataset (not shown) showed that the fit using log-normal variability distributions was worse in the classical model for almost all types of transfer and recovery studied in our experiments (for TR_{MM} the fit was equal).

The measured transfer ratios were recalculated while accounting for the limited recovery efficiency of bacteria on the donor and recipient surfaces by assuming independency between these ratios. Although this approach was taken in many similar studies (Perez-Rodriguez et al., 2008), it is likely that the recovery efficiency and the true transfer ratio are correlated. A high donor surface recovery ratio indicates a weak attachment of bacteria to the donor surface. Weakly attached bacteria are more likely to be transferred to the recipient surface. Similarly, a low recipient surface recovery ratio indicates a strong attachment of bacteria to the recipient surface which may imply a high transfer ratio. Yet, the extent of this correlation cannot be computed from the data because of the researcher's inability to measure the transfer ratios and the recovery ratios in one experiment.

In this article we have given a proof-of-principle that most currently used approaches to describing transfer ratios, in which uncertainty originating from limited count data is often neglected, may lead to inconsistencies and to an underestimation of the total uncertainty in the model. We have proposed a Bayesian model which includes all sources of uncertainty and prevents these inconsistencies. By doing laboratory cross-contamination experiments, the functionality of the model was shown, and more insight was gained into the transfer probabilities of *Salmonella* from raw pork to the knife, and back to pork. The variability in all of these transfer events was shown to be large. Yet, uncertainty originating from limited count data was also shown to be considerable. In addition, the recovery ratios of *Salmonella* from stainless steel and from meat were determined. Improved insight in these biological parameters and a correct representation of their uncertainty can be used for the purpose of better QMRA models.

Acknowledgements

The authors would like to gratefully acknowledge Alice van der Meij-Florijn, Ellen Delfgou-van Asch, and Christiaan Veenman for their dedicated laboratory work.

REFERENCES

Chen, Y., Jackson, K. M., Chea, F. P., & Schaffner, D. W. (2001). Quantification and variability analysis of bacterial cross-contamination rates in common food service tasks. J Food Prot, 64(1), 72-80.

EFSA. (2010). Quantitive Microbial Risk Assessment on Salmonella in Slaughter and Breeder pigs: Final Report.

Gelman, A., Carlin, J., Stern, H., & Rubin, D. (2004). Bayesian Data Analysis. Chapman and Hall/CRC.

Hoek, A. H. A. M. v., Jonge, R. d., Overbeek, W. M. v., Bouw, E., Pielaat, A., Smid, J. H., Löfström, C., Malorny, B., Junker, E., Löfström, C., Pedersen, K., Aarts, H. J. M., & Heres, L. (2012). A quantitative approach toward a better understanding of the dynamics of Salmonella spp. in a pork slaughter-line. Int J Food Microbiol, 145(1-2), 45-52.

Jongenburger, I., Reij, M. W., Boer, E. P., Gorris, L. G., & Zwietering, M. H. (2010). Factors influencing the accuracy of the plating method used to enumerate low numbers of viable micro-organisms in food. Int J Food Microbiol, 143(1-2), 32-40.

Kusumaningrum, H. D., Riboldi, G., Hazeleger, W. C., & Beumer, R. R. (2003). Survival of foodborne pathogens on stainless steel surfaces and cross-contamination to foods. Int J Food Microbiol, 85(3), 227-236.

Moller, C. O., Nauta, M. J., Christensen, B. B., Dalgaard, P., & Hansen, T. B. (2012). Modelling transfer of Salmonella Typhimurium DT104 during simulation of grinding of pork. J Appl Microbiol, 112(1), 90-98.

Moore, C. M., Sheldon, B. W., & Jaykus, L. A. (2003). Transfer of Salmonella and Campylobacter from stainless steel to romaine lettuce. J Food Prot, 66(12), 2231-2236.

Nauta, M. (2008). The Modular Process Risk Model (MPRM): a Structured Approach to Food Chain Exposure Assessment. In: Microbial Risk Analysis of Foods (pp. 99-136): ASM Press.

Nauta, M., van der Fels-Klerx, I., & Havelaar, A. (2005). A poultry-processing model for quantitative microbiological risk assessment. Risk Anal, 25(1), 85-98.

Perez-Rodriguez, F., Valero, A., Carrasco, E., Garcia, R. M., & Zurera, G. (2008). Understanding and modelling bacterial transfer to foods: a review. Tr Food Sci Tech, 19, 131-144.

Smid, J. H., Swart, A. N., Havelaar, A. H., & Pielaat, A. (2011). A practical framework for the construction of a biotracing model: application to Salmonella in the pork slaughter chain. Risk Anal, 31(9), 1434-1450.

Spiegelhalter, D., & Lauritzen, S. L. (1990). Sequential updating of conditional probabilities on directed graphical structures. Networks, 20.

Vose, D. (2000). Risk Analysis, A Quantitative Guide. Chichester. John Wiley and Sons.

Chapter 7

Quantifying the sources of *Salmonella* contamination in a Dutch pig slaughter plant

J.H. Smid [a,b], A.H.A.M. van Hoek [a], H.J.M. Aarts [a], A.H. Havelaar [a,b], L. Heres [c], R. de Jonge [a], A. Pielaat [a]

[a] RIVM, A. van Leeuwenhoeklaan 9, 3720 BA Bilthoven, The Netherlands

[b] IRAS, Yalelaan 2, 3584 CM Utrecht, The Netherlands

[c] VION Fresh Meat West, Boseind 10, 5281RM Boxtel, The Netherlands.

Submitted for publication

ABSTRACT

The purpose of this study was to investigate the relevance of different sources and pathways of *Salmonella* for the contamination of pig carcasses during slaughtering. Supplementary to previous work, the role of house flora on the equipment versus cross-contamination via the equipment was studied, which resulted in new, quantitative estimates of their importance. Salmonella serotyping data from samples of carcasses and from the cutting robots in a pig slaughterhouse were obtained from the study by Van Hoek *et al*, 2012. The serovars obtained from carcasses at the re-work station were compared to serovars obtained from the possible sources of contamination, using a variety of statistical tests and a newly proposed tracing scheme. It was concluded that house flora was an important source of *Salmonella*, having contaminated approximately two-third of carcasses that carried *Salmonella* on their cutting side at the re-work station. For carcasses that carried *Salmonella* on the lard side, *Salmonella* that were already on the skin or in the feces of the pigs at the moment of exsanguination were more important sources. The relevance of the different sources of *Salmonella* varied within one sampling day and also between different sampling days. The proposed tracing scheme provides a valuable contribution to decision making in existing quality control systems at the slaughter plant level. Close cooperation with quality managers enhanced the awareness of the status of the dynamics of *Salmonella* in the pork production chain and, with that, created new insights in the development of food safety control measures.

1. INTRODUCTION

Salmonella causes around 33,000 human cases of illness per year in the Netherlands (de Wit, et al., 2001) and it was estimated that between 15 and 25% of these cases were caused by pigs and the consumption of pork and pork products (Van Pelt et al., 2008). Several studies have found that slaughtering plays an important role in the contamination of pork, accounting for between 69% (Duggan et al., 2010) and 80% (Swanenburg et al., 2001a) of the contaminated carcasses at the end of the slaughter line.

The slaughtering of pigs becomes more and more concentrated. Fewer companies are involved and pig slaughterhouse process higher number of pigs than in the past. Automation of the slaughter process is introduced step by step. With the introduction of large, centralized breeding farms and subsequent transportation to centralized slaughterhouses, traceability is required for disease control in modern animal breeding. The meat production and processing chain, including control measures, is directly affected by the desire of consumers to know more about the origin and health status of the pork meat they eat. Slaughter companies want to improve the food safety of their product, and the EU wants to set targets to reduce *Salmonella* exposure of the consumers. With respect to *Salmonella* in pigs, control measures in primary production have decreased human disease incidence (Pielaat, 2011). EU wide control of *Salmonella*, however, requires long-term policies and programs.

Salmonella infected or contaminated pigs at the farm are estimated to be an important source of contamination, because these animals can carry *Salmonella* in their feces, on their skin and/or in their lymphoid tissues and introduce the pathogen into the slaughter process. During the slaughter process, *Salmonella* may cross-contaminate other carcasses and/or the equipment in the slaughterhouse (Duggan, et al., 2010; Warriner et al., 2002). At the end of every slaughter day the slaughterhouse environment is sanitized. Yet, results from several studies show that *Salmonella* can persist and multiply in the slaughterhouse environment over many days and act as a separate source of contamination, potentially transferring *Salmonella* to carcasses being processed in this environment (Botteldoorn et al., 2001a). Swanenburg *et al* (2011a) observed that the existence of such *house flora* is particularly associated to the cutting robots, i.e. the belly opener and the carcass splitter, compared to the other robots in the slaughter process. The relative importance of these *Salmonella* sources and their modes of transmission in the slaughterhouse is unclear. Yet, understanding of the microbiological behavior of Salmonella in the environment as well as its dynamics is important to build scientific based quality assurance and control programs.

Van Hoek et al (2012) described a sampling experiment in a Dutch slaughterhouse, in which skin and rectal samples were collected from carcasses directly after exsanguination, from the same carcasses at the rework station and from the cutting robots in the slaughterhouse. *Salmonella* isolates from these samples were enumerated and serotyped. By comparing the data from the different sampling locations it was concluded that the cutting robots were relevant for the contamination of carcasses at the rework station. In particular, one carcass splitter in this slaughterhouse was repetitively contaminated with *Salmonella*, which reflected that this robot encountered hygiene problems during the sampling study. Yet, the role of house flora on the cutting robots versus cross-contamination via the cutting robots was

not studied and a quantitative estimate of the relative importance of the different *Salmonella* sources was not given.

The purpose of this study was to generate such estimates and to analyze how the importance of the different sources varies over time (per day and per month). In addition, the importance of cross-contamination versus house flora during slaughtering was explored. This paper presents a structured statistical scheme to increase source traceability of *Salmonella* in the pork production chain. *Salmonella* serovars found at the end of the slaughtering line were matched with their most likely source during processing based on frequency distributions of occurring serovars in the chain. Hereto, a modified version of the algorithm described by Swanenburg *et al*. (2001a) to compare the *Salmonella* genotypes and to include the effects of cross-contamination has been used. Our approach can help to identify the spots in the chain where the major risks are and help in decision making for appropriate control measures.

2. MATERIALS AND METHODS

The slaughterhouse characteristics, the sampling experiment and the data that were obtained are fully described by Van Hoek et al. (2012). A brief summary of these is provided for general understanding of further analyses.

2.1 Slaughterhouse characteristics

In the Dutch slaughterhouse investigated for this study, between 5,000 and 6,000 animals were slaughtered per day. Pigs were electrically stunned, sticked on a table, scalded in a tank, dehaired, flamed and wet polished. The slaughterhouse was partly automated with robots for belly opening, rectum drilling, splitting, leaf lard removal, neck cutting and marking. There were two robots for belly opening and carcass splitting on parallel lines, each line used for processing half of the carcasses. The head was not cut by the carcass splitter. Meat inspection was visual only, i.e. there was no palpatation and the mandibular lymph nodes were not incised.

2.2 Sampling set-up

Samples were collected in the slaughterhouse from April – July 2009, on 11 days which were roughly evenly spread within the sampling period. Skin samples (obtained by pooling 4 cork borer samples of the shoulder) and rectal swab samples were collected from randomly selected pigs directly after exsanguination. In the remainder of this article, these two types of samples are referred to as the **upstream** samples, as these samples were taken from the incoming pigs. The number of upstream sampled pigs per day ranged between 6 and 12, with a total of 118 sampled pigs throughout the study. In addition, one or two samples were taken from the knives and from small pockets in the robot close to the knives of two belly opening robots (BO1 and BO2) and from two carcass splitting robots (SP1 and SP2), using swabs and a flexistem brush. This was done at the beginning (around 6 AM) and at the end (around 3 PM) of each slaughtering day. The samples taken from the cutting robots are referred to as the **environmental** samples. The upstream sampled pigs were tracked in the slaughter line and at the end of the slaughter process, at the re-work station, pooled samples (using 4 cork borer samples) were obtained from the exterior (back, jowl, ham and belly) and the interior (ham, back before pelvis, sternum and shoulder muscle) of these carcasses (the **downstream** samples).

All samples were analyzed for *Salmonella* using enrichment by BPW followed by PCR. Original samples of PCR-positive results were further analyzed to obtain *Salmonella* concentrations by culture using the MPN method (de Man, 1983). Depending on the *Salmonella* concentration, up to five (representative) isolates from each sample were randomly selected. These isolates were subsequently serotyped by slide and tube agglutination following the Kauffmann–White scheme.

2.3 Considered Salmonella sources

The following potential sources of *Salmonella* contamination were considered important for the Dutch slaughterhouse analyzed for this study, and deduced from the dynamics of *Salmonella* in pork during slaughtering (Introduction) and the description of the slaughterhouse and the sampling set-up. Also see the Discussion.

- *Salmonella* on the skin of pigs at exsanguination (**skin**),
- *Salmonella* in the intestinal tract of pigs at exsanguination (**feces**),
- Resident *Salmonella* on the cutting robots (**house flora**).

2.4 Considered transmission routes

Salmonella can be found at the lard side (**exterior**) or at the cutting side (**interior**) of a carcass at the re-work station and originate from either one of the three proposed potential sources through a multitude of pathways. The first stages of the slaughter process, in particular scalding, flaming, polishing and water hosing, result in a massive reduction of bacteria on the carcass' skin due to bacterial removal and inactivation. Nevertheless, *Salmonella* in the recta or in skinfolds of carcasses may not be exposed to these stages and persist (Spescha et al., 2006). It is thought that, if inactivation was insufficient, of all salmonellae that were on the skin of pigs at exsanguination only the irreversibly attached ones may be left on the skin after polishing (EFSA, 2010). This may lead to a *Salmonella*-positive carcass at the re-work station, but it is unlikely that such salmonellae cross-contaminate other carcasses during slaughtering due to their strong attachment to the carcass surface. However, pressure of the dehairing and polishing robots may result in fecal leakage (Morgan et al., 1989). This may also occur if the pig's colon is punctured during belly opening, bung dropping or pluck removal. If leaking feces contains salmonellae, these may be transferred to the carcass and/or to the slaughter environment, e.g. the cutting robots. *Salmonella* in the slaughter environment may subsequently cross-contaminate the interior or the exterior of other carcasses. In general, all features in the slaughterhouse environment that are in direct contact with carcasses (e.g. hands of personnel, conveyor belts, robots) constitute pathways by which salmonellae may be transferred to other carcasses. The knives of the robots are routinely sanitized upon cutting, but it is unlikely that this sanitizing program excludes the possibility of cross-contamination (Alban & Stark, 2002). Moreover, resident salmonellae may persist as house flora and multiply in small nooks of the robots which are difficult to clean (Joseph et al., 2001). Such strains may be transferred with drip water to the exterior or, in case of cutting robots, to the interior of the carcasses.

2.5 Analysis of contamination routes

2.5.1 Within day prevalence comparison

First, the per-day prevalences of *Salmonella* at the upstream and downstream sampling sites were assessed. The prevalences on different days at one site were regarded as elements of a vector. The (Pear-

son's product moment) correlation between vectors representing different sites was calculated, and it was investigated whether correlations were significant ($p<0.05$). Significant correlations between the prevalence vectors for two sampling sites could be indications of transmission pathways between these sites. Vectors representing daily prevalences at environmental sampling sites were not constructed, as the number of samples taken per day from those sites was too low for statistical calculations.

2.5.2 Between day prevalence comparison

An additional analysis to identify possible transmission pathways was performed by comparing the variability in prevalence over different days at a given sampling site to the variability in prevalence over those days at other sampling sites. For each sampling site it was assessed between which days statistically significant different per-day prevalences existed ($p<0.05$). The days between which such differences existed were compared to the days between which such differences existed for another sampling site. Similarity between such (pairs of) days for two sampling sites may indicate that the *Salmonella* prevalence at these sites are related. Precisely, if a pathway exists from an upstream sampling site U to a downstream sampling site D, then an increased *Salmonella* prevalence at U should lead to an increased *Salmonella* prevalence at a D. Conversely, if such a pathway does not exist then the *Salmonella* prevalence at U and D are unrelated. For each sampling site, 2 x 2 contingency tables and Fisher's exact test were used to test whether the *Salmonella* per-day prevalences showed significant differences from prevalences on other sampling days.

2.5.3 Serovars comparison

Similarity in serovars across an upstream sampling site U and a downstream sampling site D also indicate a potential transmission route from U to D. McNemar's test, using a 2 x 2 contingency table, was used to infer whether such transmission routes exist. The test requires the same carcasses to be included in the upstream and downstream measurements (matched pairs). The purpose of the test was to examine whether the null-hypothesis, that this serovar occurs in equal proportions in U and in D, can be rejected. This was done for any pair of upstream and downstream sampling sites and for all individual serovars. Only cases in which both sampling sites were found positive for *Salmonella* were included in this analysis to account for differences in *Salmonella* prevalence at the different sites. Environmental samples could not be considered by McNemar's test because they are not matched pairs with the downstream samples taken from the carcasses.

2.6 Source tracing scheme

The serovars isolated from the downstream samples were compared to the serovars from upstream and environmental samples to establish potential matches, using the scheme proposed hereafter. The basic idea of the scheme is to estimate the probabilities for each source of having contaminated the carcasses. The assumption in these calculations is that one case of contamination is caused by one source (see Discussion). The calculations proceed according to a number of main principles, which are schematically represented in Fig. 1 and described in more detail below. Roman numbers in the text below refer to those in Fig. 1.

Pairwise comparison of the downstream and upstream data for one carcass allows to establish whether the serovar found in a source matches the serovar in or on the carcass. However, a carcass can also

be contaminated via cross-contamination, so that the serovar on a carcass may originate from another pig, e.g. a pig from the same herd or a pig that is slaughtered on the same day. Because of reasons mentioned earlier, it was assumed that cross-contamination through contact with the cutting equipment is only relevant for salmonellae in feces, originating from the pig's intestinal tract (I). Yet, it is unknown between which carcasses cross-contamination may occur. Two extreme scenarios were studied, between which the true situation is assumed to be in between (II):

- **Maximal cross-contamination:** cross-contamination of salmonellae in feces may occur between all carcasses processed on the same day. To establish a match, the *Salmonella* serovars of the downstream samples were compared to the serovars from the feces of all pigs slaughtered on the same day.
- **No cross-contamination:** cross-contamination of salmonellae in feces between carcasses does not occur. To establish a match, only the *Salmonella* serovars of the downstream samples were compared to the serovars from the feces of that same pig.

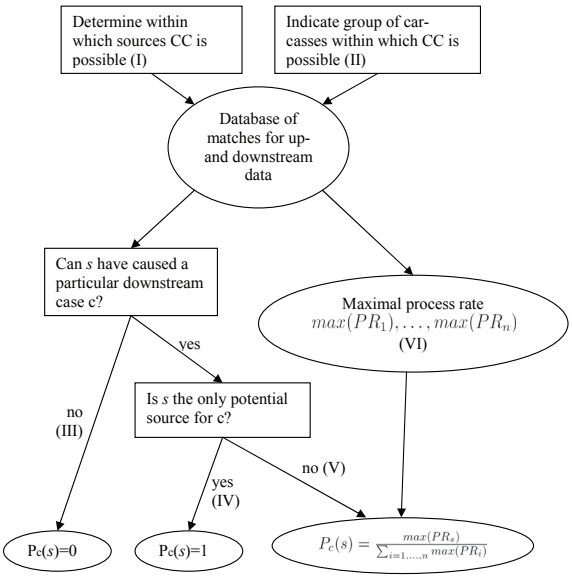


Fig. 1: *Proposed scheme for computing the probability that a potential source s is the active source of contamination of a downstream observed case of contamination c, i.e. $P_c(s)$. This scheme is described in words in the text in the section 'Source tracing scheme'. CC stands for cross-contamination.*

The *Salmonella* serovars of the downstream samples were also compared to serovars of the environmental samples. However, pairwise comparisons of the serovars obtained from the downstream samples and from the environmental samples was not possible, since only environmental samples taken at the beginning and end of each slaughtering day were available. To circumvent this problem it was assumed that the serovars, obtained from the robots, had been present on those robots during the whole day. So a match of an environmental sample and a downstream sample collected on the same day implies a potential case of transmission. A mismatch means that in that case the robot has not transmitted *Salmonella* to the carcass. Note that *Salmonella* serovars on the robots may also have originated from the other sources through cross-contamination. This may be true if, within one slaughtering day, the same types are found on the robots as in a source where cross-contamination is assumed to be a relevant process. The likelihood of this was investigated by comparing the serovars identified in the upstream and environmental samples.

Next, the number of occasions that a serovar on a carcass *c* (or, via cross-contamination, on other related carcasses) matched the serovar in source *s* was determined. This number may differ per scenario 1 or 2. Based on the number of matches, the probability for each source of having contaminated a carcass *c* was obtained as follows:

If there is no match between the serovars on *c* (or, in scenario 1, on related carcasses of *c*) and in *s* then it was assumed that *s* could not have contaminated *c* (III). Else, if the serovars on *c* (or, in scenario 1, on related carcasses of *c*) match with serovars in source *s* and not with serovars in sources other than *s*, then it was assumed that *s* has contaminated *c* with probability 1 (IV). Else, if the serovars on *c* (or, in scenario 1, on related carcasses of *c*) match serovars in multiple (say *n*) sources among which source *s* (V), then the relevance of those sources was scaled. First, the *process rate* of salmonellae from *s* (notation PR_s) was defined as the ratio of the frequency that salmonellae, transferred from *s*, end up on a carcass, divided by the total frequency that salmonellae are transferred from *s*. This ratio is indicative for the source-dependent property to what extent the salmonellae, transferred from that source, survive the chain process. Naturally, the *exact* number of times that an individual positive source has transferred salmonellae to a carcass can often not be inferred from the data, because more than one source may display serovars that match those on the carcass. Consequently, the process rate of salmonellae from a source *s* cannot be estimated from the data. However, its upper limit equals the number of occasions in which the *Salmonella* serovars in *s* matched the serovar found on the carcasses $c = 1, 2, 3, \ldots$ divided by the number of occasions in which any *Salmonella* serovar was found in *s*:

$$max(PR_s) = N_{s \wedge c}/N_s \tag{1}$$

This limit will be referred to as the *maximal process rate* (VI) and can be seen as the maximal probability that source *s* is responsible for the observed cases of contamination, given the observed number of matches and mismatches of serovars. Such probabilities were calculated for the *n* sources that could have contaminated the carcasses. It was assumed that $max(PR_{hf})$, where hf stands for house flora as a source of *Salmonella*, equals one. This implies that, if house flora is transferred from the cutting robots to a carcass, it will certainly remain on the carcass until meat inspection at the re-work station. This assumption is reasonable because the point of introduction of house flora in the chain is close to the re-work station and there are no salmonellae-reducing steps in between. Logically, the summation $\sum_{i=1,\ldots,n} max(PR_i)$ is greater than one because for all sources the upper limit of the process rate was taken. The maximal process rates of different sources were scaled to compute the probability that a source *s* has contaminated a carcass *c*:

$$P_c(s) = \frac{max(PR_s)}{\sum_{i=1,\ldots,n} max(PR_i)} \tag{2}$$

The source attributable proportions for positive carcasses over the whole period (F_s) were computed by averaging the source probabilities per contaminated carcass over all carcasses: $F_s = \sum_c P_c(s)$.

3. RESULTS

3.1 Data

The per-day prevalence of *Salmonella* at all sampling sites and the strain diversity of the isolates obtained from the upstream, downstream and environmental samples are fully provided in the article

by Van Hoek et al. (2012). A numerical summary of these data is presented in Table I. Briefly, the skin of almost all incoming pigs was contaminated with *Salmonella*. *S.* Derby, *S.* Typhimurium and *S.* Brandenburg were the predominant serovars found in these samples. Nearly two-third of the rectal samples was contaminated with *Salmonella*. The same *Salmonella* serovars were found as in the skin samples, but in slightly different proportions. Of the cutting robots, only carcass splitter 2 was frequently contaminated: 69% of these samples were *Salmonella*-positive. *S.* Derby and *S.* Rissen were the only serovars found in these samples. Remarkably, *S.* Rissen was never found in the samples taken from the pigs at exsanguination. All samples taken from carcass splitter 1 tested negative for *Salmonella*.

Table I: Prevalence of *Salmonella* and distribution of serovars per sampling site.

	Skin pigs	Faeces pigs	BO1	BO2	SP1	SP2	Exterior carcasses	Interior carcasses
Pos. samples/ All samples (prevalence)	114/118 (0.97)	70/112 (0.63)	1/31 (0.03)	2/31 (0.06)	0/35 (0.00)	24/35 (0.69)	19/117 (0.16)	35/117 (0.30)
Serovar (% of isolates)								
Bredeney	7 (2)	2 (2)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Brandenburg	63 (20)	17 (13)	0 (0)	0 (0)	0 (0)	0 (0)	3 (10)	1 (2)
Derby	118 (38)	47 (36)	1 (100)	0 (0)	0 (0)	47 (53)	5 (17)	37 (62)
Infantis	6 (2)	1 (1)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)
Rissen	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	41 (47)	10 (34)	12 (20)
monophasic Typhimurium	17 (6)	2 (2)	0 (0)	0 (0)	0 (0)	0 (0)	3 (10)	1 (2)
Typhimurium	97 (31)	60 (47)	0 (0)	6 (100)	0 (0)	0 (0)	8 (28)	8 (13)
Unknown	1	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	1
Total # isolates	*309*	*129*	*1*	*6*	*0*	*88*	*29*	*60*

Reprinted from Hoek *et al.*(2012) with permission. BO: belly opener, SP: carcass splitter.

The prevalence of *Salmonella* on the skin of carcasses had decreased markedly after slaughter to 16%. However, in 30% of the interior samples of the carcasses *Salmonella* was found at the re-work station. In the downstream samples, *S.* Rissen, *S.* Typhimurium and *S.* Derby were found most often. The prevalence of *Salmonella* and the distribution of *Salmonella* serovars are further specified in Table 2. Per-day *Salmonella* prevalences are shown and per carcass, per sampling site and per day is indicated if, and which, *Salmonella* serovars were found. Also the strain diversity of the environmental samples per sampling day is shown. The results show that *S.* Derby and *S.* Rissen were consistently found in the afternoon samples taken from carcass splitter

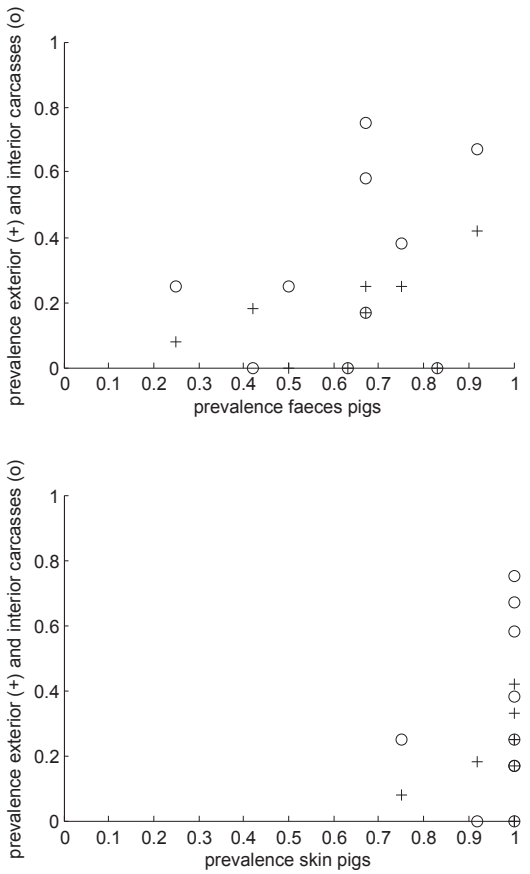


Fig. 2: *Per-day prevalences obtained from the downstream samples (interior: o, exterior: +) plotted against the per-day prevalences from the upstream samples.*

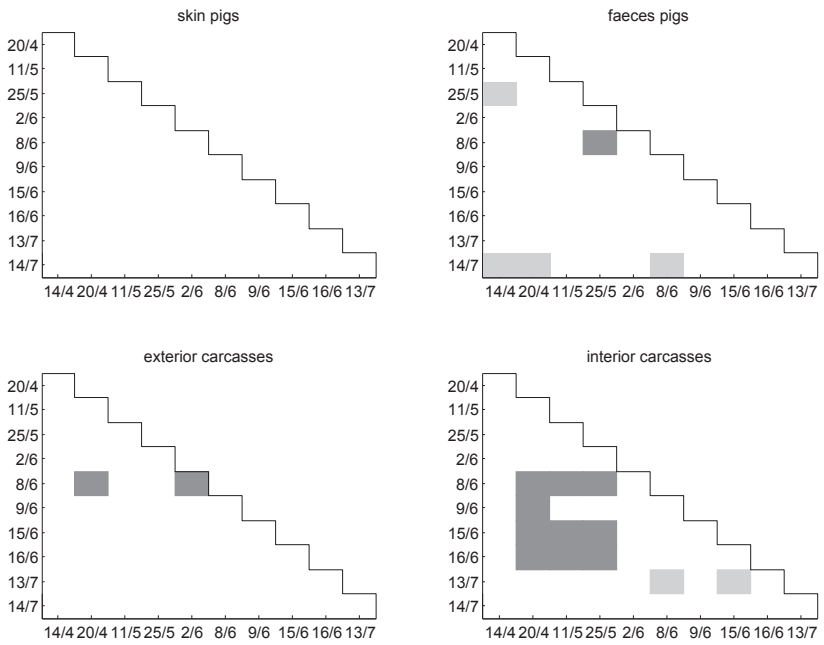


Fig. 3: *Per-day Salmonella prevalence comparison. Grey boxes indicate that the prevalence of Salmonella (at a given sampling site) was significantly different on two days. Dark grey boxes indicate that the Salmonella prevalence was significantly higher on the day on the y-axis compared to the day on the x-axis. Light grey boxes indicate that the Salmonella prevalence was significantly lower on the day on the y-axis compared to the day on the x-axis. All dates are shown as Day/Month.*

2 after the third (May 25) and the fourth sampling day (June 2), respectively. On two different days (July 13 and 14), *S.* Rissen and *S.* Derby were found in the morning samples of this robot. Belly opener 1 was found to be contaminated with *S.* Derby on one occasion (June 16), and belly opener 2 was found to be contaminated with *S.* Typhimurium on two occasions (June 9 and 16).

3.2 Transmission routes

In Fig. 2, the per-day *Salmonella* prevalences obtained from the downstream samples of the interior and exterior of the carcasses were plotted against the per-day *Salmonella* prevalences obtained from the upstream fecal and skin samples. The correlation coefficients between the upstream and downstream prevalences were not statistically significant ($p>0.05$), so no indications of potential routes of contamination through the chain could be found.

Fig. 3 shows, for each sampling site and each pair of sampling days, whether the observed prevalence on these days was significantly different ($p<0.05$). No significant differences in prevalence were found between sampling days for the skin of pigs. The majority of statistically significant differences were found for the interior samples taken from the carcasses at the re-work station. In particular, differences were observed between the prevalence found on sampling days in the months April-May and June and between the prevalence on sampling days in June and July, due to a significantly higher prevalence of internally contaminated carcasses in June compared to April and May and a significantly lower prevalence in July compared to June. The pairs of days on which such significant differences were observed for the interior samples did, in general, not coincide with days for which significantly different *Salmonella* prevalences on the pig's skin or in the pig's feces were found. Only on June 8, a significantly higher prevalence of *Salmonella* in the interior samples compared to that on May 25 corresponded with

a significantly higher prevalence of *Salmonella* in the feces of pigs on June 8 compared to May 25.

For the carcass splitters and belly openers, no per-day prevalence could be assessed, due to the limited number of samples taken per day at these sites. Yet, the repetitive finding of *Salmonella* on carcass splitter 2 from the end of May onwards (Table II, Appendix) coincided with the increased prevalence of *Salmonella* in the interior samples of carcasses in June, indicating a potential role of the carcass splitter in contamination of the interior of the carcass at the re-work station from June to mid-July.

Using McNemar's test, significant differences could be established between the upstream and downstream *Salmonella* serovar distributions. These differences were all on the account of *S.* Derby, *S.* Typhimurium and *S.* Rissen. *S.* Typhimurium was significantly more often found in isolates from the skin of pigs and *S.* Rissen was significantly more often found in isolates from the carcass interior compared to the other strains (with $p=0.006$ and $p=0.041$, respectively). *S.* Typhimurium was also significantly more often found in isolates from the pig's feces than from the carcass interior ($p=0.006$). *S.* Derby was significantly more often found in isolates from the interior samples than from the rectal samples ($p=0.009$). It was noticed (Table II, Appendix) that the *Salmonella* serovars that were found on the skin of pigs were also found in the feces of the same pigs. Yet, *S.* Derby was significantly more often found in isolates from the skin than from the feces of pigs ($p=0.0095$). Therefore, the skin and the feces of pigs were considered as distinct potential sources of contamination of the carcass in the source tracing analysis hereafter.

3.3 Source tracing

Salmonella was only sporadically found on the belly openers and not on carcass splitter 1 and, therefore, these sites were not considered as potential sources of house flora. Consequently, only the up- and downstream data obtained from the carcasses and the environmental data obtained from carcass splitter 2 were used in the following analysis. The data obtained on the first day of sampling (April 14) were not considered because on that date no rectal samples were taken. Furthermore, as a first step, it was assumed that cross-contamination of fecal contamination may occur between pigs that are slaughtered on the same day (scenario 1). The tracing results in scenarios 2 and 3 will be discussed at a later stage.

Based on the scheme in Fig. 1, the probabilities for each source of having contaminated the exterior or interior of the individual carcasses at the re-work station were calculated. First the maximal process rates for salmonellae from the sources to both downstream sampling sites were calculated by comparing the detected serovars in the sources and on the carcass (Table 2). The maximal process rate of salmonellae from feces to the exterior of the carcasses was higher (0.17) than that of salmonellae from the skin of pigs to the carcass exterior (0.08). The maximal process rate of salmonellae from feces to the interior of the carcasses was also higher (0.27) than that of salmonellae from the skin of pigs to the carcass interior (0.12). This indicates that salmonellae from the feces of pigs have a higher probability of ending up on either the exterior or the interior of the carcass at the re-work station compared to salmonellae from the skin of pigs.

Fig. 4 shows the estimated probabilities for the potential sources of having contaminated the exterior and the interior of each carcass at the re-work station (carcass number on the horizontal axis). Carcasses

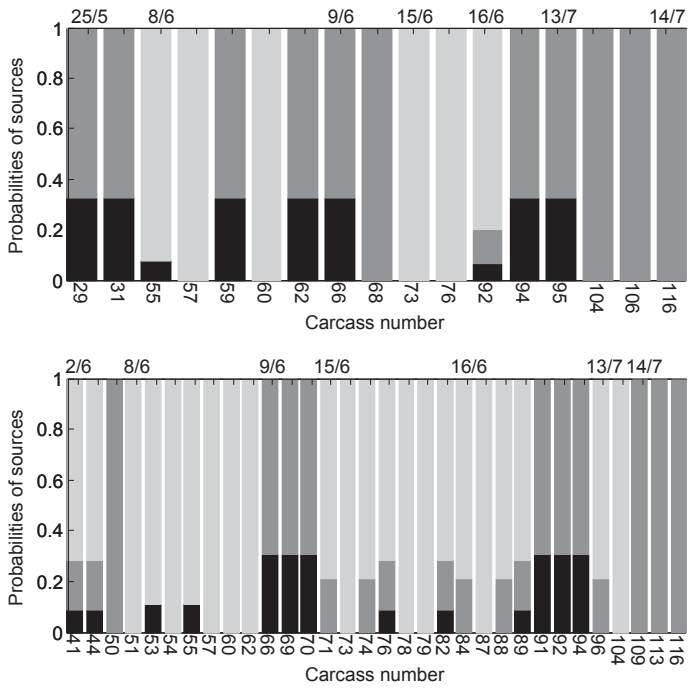


Fig. 4: *For each potential source (black: pig skin; dark grey: feces; light grey: house flora), the estimated probabilities of being the primal source of contamination of a carcass are indicated by fractions of a bar chart. Upper chart: Results for contamination of the carcass exterior. Lower chart: Results for contamination of the carcass interior. The carcass numbers are given on the lower horizontal axes and the dates on which these carcasses were sampled is given on the upper horizontal axes (Day/Month). Postive carcasses sampled on April 14 were not considered because no fecal data were available for that date.*

of which the samples were negative by PCR or not serotyped were omitted in Fig. 4. Roughly half of the externally positive carcasses that were sampled in June were most likely contaminated by house flora whereas those slaughtered in July were more likely to be contaminated by salmonellae from the skin or feces. The majority of the carcasses of which the interiors were *Salmonella*-positive were a result of samples taken in June. The most likely source of contamination of most (22 out of 32) of these carcasses was house flora. Carcasses with positive interiors slaughtered on June 9 and July 14 and some carcasses slaughtered during the second half of the day on June 2 and June 16 were more likely to be contaminated by salmonellae from the feces of pigs.

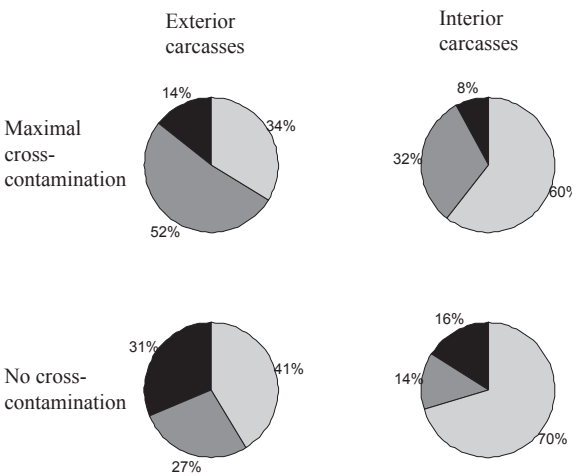


Fig. 5: *For each potential source (black: pig skin; dark grey: feces; light grey: house flora), the estimated probabilities of being the primal source of contamination are indicated by fractions in the pie chart, as computed over all carcasses. Results for contamination of the carcass exterior, the carcass interior and the two considered scenarios about cross-contamination (CC).*

Fig. 5 shows the source attributable proportions for positive carcasses over the whole sampling period in the different cross-contamination scenarios. For carcasses with a contaminated exterior, all three sources of *Salmonella* are important. The importance of the source, however, depends on the cross-contamination

scenario. In the scenario of maximal cross-contamination, where cross-contamination potentially takes place between all carcasses slaughtered on one day, *Salmonella* in the feces of pigs was the most important source, causing over 50% of cases. In the scenario of no cross-contamination, *Salmonella* in the feces of pigs was a less important source. For interior contamination, house flora was distinctively the most relevant source of contamination, having caused 60 – 70 % of the carcass contamination, depending on the cross-contamination scenario. The effect of the type of cross-contamination scenario was small (Fig. 5).

4. DISCUSSION

The *Salmonella* prevalence among the pigs entering the sampled slaughterhouse was high, with 97% of pigs being contaminated on the skin, and 62% carrying contaminated feces in their intestinal tract. The comparison of the between-day *Salmonella* prevalence at different sampling sites and, in particular, the comparison of *Salmonella* serovars at exsanguination and at the re-work station indicated that some salmonellae, originating from the pig, remained on the carcass during slaughtering or were transmitted to other carcasses. In addition, our results indicated that it is unlikely that bacteria on or inside the pig are the only source of contamination for carcasses at the re-work station.

The results showed that the *Salmonella* subtype distributions as observed at the upstream sampling sites (skin and feces) were modified along the chain. Possible explanantions are: (1) different serovars may have a different survival strategy so that some serovars persist along the chain whereas other do not, (2) the serovars transmitted by the different upstream sources are mingled in the chain and (3) the existence of another source in the slaughterhouse environment, i.e. house flora. Any combination of these causes may have led to the differences between the up- and downstream serovar distributions. Yet, the marked absence of *S.* Rissen in the upstream samples and its occurrence in the downstream samples suggests the relevance of house flora in the slaughterhouse environment for the contamination of carcasses. This was confirmed by the persistent finding of *S.* Rissen on carcass splitter 2. It was shown that house flora is particularly relevant for contamination of the interior of the carcasses (Fig. 5). This was expected, as the carcass splitter contacts the interior of the carcass more intensively than the carcass exterior. Yet, also for the carcass exterior, house flora may be an important source of contamination, causing approximately one third of all cases (Fig. 5). This could be caused by meat juice dripping from the carcasses into contaminated parts of the machine, and back to carcasses that are subsequently slaughtered. The potential role of house flora in contaminating carcasses has already been addressed by various studies (Baptista et al., 2010; Swanenburg et al., 2001b; Swanenburg et al., 2001c; Warriner et al., 2002). The results of this study provide additional evidence for their relevance in carcass contamination.

Little is known about the survival mechanisms of *Salmonella* in the slaughterhouse environment. (Kusumaningrum et al., 2003) showed that high numbers (10^5 CFU/cm^2) of free *S.* Enteritidis cells may survive on stainless steel surfaces for multiple days. After each day, the slaughtering machines are washed and decontaminated, which makes the existence of high concentrations of free *Salmonella* cells on the larger exposed surfaces of the equipment unlikely. Yet, *Salmonella* populations may persist in parts of the equipment that are difficult to clean, such as small pockets of the cutting machines.

Moreover, bacteria are able to attach to environmental surfaces by producing biofilms (Donlan & Costerton, 2002), which are surface-associated community forming micro-colonies surrounded by a matrix of exopolymers that trap other bacteria, nutrients and debris (Chavant et al., 2002). Meat and meat juices provide rich substrates for biofilms, potentially leading to overnight growth of the resident *Salmonella* population. The matrix enhances the resistance of the bacterial cells to antimicrobial agents (Anderl et al., 2000; Stewart & Costerton, 2001). This impedes a proper sanitation of the slaughtering machines. The persistence of only two strains on carcass splitter 2 throughout the last six weeks of the experiment, opposed to the diversity of strains found in the pigs at the onset of slaughter during the same period, suggests an improved capacity of these strains to attach to the equipment surface, or even to form biofilms. Bacteria growing in biofilms are known to release cells that can migrate to other locations (Costerton et al., 1995), thereby contaminating carcasses passing through a contaminated machine.

By our knowledge, no comprehensive study has been done to elucidate either the attachment strengths or the biofilm-forming capacities of different *Salmonella* serovars. (Kim & Wei, 2007) found that *S.* Typhimurium DT104 formed stronger biofilms than *S.* Heidelberg and *S.* Agona and even well-known biofilm formers like *Listeria monocytogenes* and *Pseudomonas aeruginosa*. However, in this study *S.* Typhimurium was not found on the carcass splitters, and in lower proportions on the carcasses than would be expected from the data from pigs after exsanguination. In previous studies in Dutch slaughterhouses, persistent contamination of *S.* Infantis and *S.* Brandenburg was found on the carcass splitter (Swanenburg et al., 2001c). Also in Denmark, *S.* Infantis was found as house flora in two slaughterhouses (Sorensen et al., 1999). Little is known about the biofilm-forming capacities of other *Salmonella* serovars, including *S.* Derby and *S.* Rissen.

Only three potential sources of *Salmonella* contamination were considered in this study. Yet, in existing literature, more sources where *Salmonella* may enter the pig slaughter line have been described. In the traditional European system of meat inspection, the mandibular lymph nodes of the pigs are deliberately incised. If these lymph nodes contain Salmonella then these may be transmitted to carcasses (Vieira-Pinto & Martins, 2003). Also, handling by workers may be an additional source of contamination (Borch et al., 1996). Another potential source of contamination is house flora from other equipment than the cutting robots. For this study, a careful prioritization selection of sources was made because of limited sampling capacity. In the studied slaughterhouse, the lymph nodes of the carcasses were not incised. Also, carcasses from which samples were taken in this study were marked and not touched by slaughterhouse personnel. Environmental samples were taken from the belly opener and the carcass splitter only because of the findings of Swanenburg *et al* (2001a), who observed that the existence of house flora is particularly associated to these robots. Recognizing discrepancies in comparing the status of *Salmonella* serovars in the food chain has taken the discussions a step further in awareness of the need for uniform sampling and sample analysis schemes.

The study of Swanenburg *et al* (2001a) showed that the slaughter line was the most important source of *Salmonella* contamination of carcasses, but the importance of cross-contamination versus house flora was not discussed. The study described here looked in more detail into the sources of contamination within the pig slaughter chain and in particular the role of house flora. Whether the importance of

these potential sources may alter in time was also analyzed. It was seen that the sources of contamination may vary per month with, for this study, an increased probability of contamination by house flora in June 2009 (Fig. 3 and 5). After feedback of the data to the quality manager of the slaughterhouse and intervening actions of a stricter cleaning regime in July, the role of house flora seemed to decline again. The source tracing results also suggest an increased probability of contamination through cross-contamination during the second half of the day. This seems plausible because the *Salmonella* load transferred to the slaughter machines may build up in time during the day. This hypothesis is in line with the findings of (Delhalle et al., 2008) showing that extensive cleaning of the splitting machine several times a day is beneficial for reducing the *Salmonella* prevalence.

The results of this study can also be compared to the model described in a recent paper (Smid et al., 2012). The same dataset, yet another approach, was used to trace the sources of contamination of individual contaminated carcasses. The bacterial concentrations of *Salmonella* on the carcasses were used in that model for tracing (the *concentration model*). Instead of using a statistical model for tracing the sources of cases of contamination, such as done here, the concentration model explicitly describes the dynamics of *Salmonella* through the chain and can, therefore, be used as a tool for monitoring and risk assessment. In addition, the concentration model is implemented as a Bayesian Belief Net (Kjaerulff & Madsen, 2008), in which uncertainty about the values of process parameters is considered for tracing. The main conclusions of the concentration model are in agreement with the conclusions of this paper. It was also found that cases of interior *Salmonella* contamination and approximately half of the cases of exterior contamination in June were caused by house flora from the carcass splitter. The remaining cases were, in both approaches, found to be mainly caused by either *Salmonella* in the feces of pigs (most often), or by *Salmonella* on the skin of pigs. Nevertheless, in some cases different tracing results were found, caused by the different approaches taken and the different assumptions made in both models. For example, in this approach it was assumed that each contaminated carcass is exclusively contaminated by one source. Yet, in practice, multiple sources may contribute to one case of contamination. The assumption of a unique source was not made in the concentration model.

Yet, the algorithm proposed here is much simpler than the concentration model. This makes it intuitive to understand, easy to use and applicable to any slaughterhouse for which similar data is available. The concentration model is built for one specific slaughterhouse. It will only be of use to quality managers of other slaughterhouses where the stages of slaughtering are exactly identical to the one described here. In most cases, the adaptations of the model to fit a new slaughterhouse will only be minor, but, nevertheless, an expert modeler is needed to do this. Ultimately, the choice for one of these models should depend on the available data (serotypes vs. concentrations) and the purpose for which it is used (strictly source tracing for the approach described in this paper, or a combination of monitoring, risk assessment and source tracing for the concentration approach). Yet, both models have a complementary value for long-time monitoring in the chain to arrive at operational biotracing.

It can be concluded that house flora may play a significant role in carcass contamination, especially for contamination found on the interior of the carcass. This conclusion gets additional support from several other studies (Botteldoorn et al., 2003; Delhalle, et al., 2008; Swanenburg et al., 2001a). It is, therefore, obvious that slaughterhouse hygiene is a determinative factor for managing carcass contami-

nation. Slaughter companies that take into account that contamination in the slaughter line can, in some situation, contribute significantly to contamination levels on carcasses can strengthen their efforts to clean the equipment even more thoroughly, being aware that biofilm formation or colonization of the equipment can happen. Industry can interpret from collected data that they can do much by slaughter hygiene to reach performance goals at carcass level at the end of the slaughter line, which is a relevant issue in the EU discussion on setting targets for Salmonella in the supply chain. A future step would be to implement the developed algorithm into existing quality control systems at several slaughter plants over Europe. This would be a valuable contribution to a uniform decision making at the slaughter plant level.

APPENDIX

		Data carcasses				Environmental data			
		……upstream		downstream					
day	**carcass number**	**skin**	**faeces**	**exterior carcass**	**interior carcass**	**BO1**	**BO2**	**SP1**	**SP2**
14/4	**1**	F_2G_1	ND	F_1	-	ND	ND	ND	ND
	2	C_4G_1	ND	-	-				
	3	U	ND	-	-				
	4	D_1F_4	ND	-	-				
	5	D_2F_3	ND	F_1	F_1				
	6	$D_3F_1G_1$	ND	-	-	-	-	-	-
	Prev	**6/6**	**ND**	**2/6**	**1/6**	**0/1**	**0/1**	**0/1**	**0/1**
20/4	**7**	C_5	C_5	-	-	-	-	-	-
	8	C_5	C_5	-	-				
	9	C_2F_2	C_1	-	-				
	10	C_5	C_2	-	-				
	11	C_5	C_5	-	-				
	12	C_1	D_1	-	-				
	13	C_4	C_1	-	-				
	14	C_1	-	-	-				
	15	C_1	-	-	-				
	16	C_1	G_1	-	-				
	17	C_2	C_1	-	-				
	18	C_1	C_2	-	-	-	-	-	-
	Prev	**12/12**	**10/12**	**0/12**	**0/12**	**0/2**	**0/2**	**0/2**	**0/2**
11/5	**19**	G_1	-	-	-	-	-	-	-
	20	G_1	-	-	-				
	21	G_1	-	-	-				
	22	G_1	G_1	-	-				
	23	C_5	C_5	-	-				
	24	C_5	C_5	-	-				
	25	F_4	F_1	-	-	-	-	-	-
	26	F_1G_2	G_1	-	-	-	-	-	-
	Prev	**8/8**	**5/8**	**0/8**	**0/8**	**0/3**	**0/3**	**0/3**	**0/3**
25/5	**27**	G_1	G_1	-	-	-	-	-	-
	28	G_1	-	-	-				
	29	G_1	G_1	G_1	-				
	30	G_1	G_1	-	-				
	31	G_1	-	G_1	-				
	32	G_1	-	-	-				
	33	-	-	-	-				
	34	G_1	-	ND	ND				
	35	C_1	-	-	-				

Table II: Strain diversity per sampling date at the 8 different sampling sites.

Salmonella serovars found were coded as follows: A: S. Bredeney, B: S. Brandenburg, C: S. Derby, D: S. Infantis, E: S. Rissen, F: monophasic S. Typhimurium, G: S. Typhimurium, U: unknown, -: negative sample. Subscripts denote the number of isolates of one serotype that were found per sample. Per-day prevalences per sampling site are given below each sampling date. Strain diversity within one sampling date is subdivided by a dashed line, indicating carcasses sampled before and after noon and environmental samples collected pre- and post-slaughter (upper and lower blocks, respectively). ND: not determined.

	36	G_1	-	-	-				
	37	G_1	G_1	-	-	-	-	-	C_1
	38	C_1	G_1	-	-	-	-	-	C_1
	Prev	**11/12**	**5/12**	**2/11**	**0/11**	**0/3**	**0/3**	**0/3**	**2/3**
2/6	**39**	G_1	C_1	-	-	-	-	-	-
	40	C_1	-	-	-				
	41	C_5	-	-	C_1				
	42	C_4G_1	-	-	-				
	43	C_4G_1	G_4	-	-				
	44	C_5	C_1G_1	-	C_1				
	45	C_2	F_1	-	-				
	46	C_1	-	-	-				
	47	C_1	G_1	-	-				
	48	C_1	-	-	-				
	49	C_1	G_1	-	-	-	-	-	C_2E_2
	50	C_2	-	-	G_1	-	-	-	C_2E_3
	Prev	**12/12**	**6/12**	**0/12**	**3/12**	**0/3**	**0/3**	**0/3**	**2/3**
8/6	**51**	G_1	G_1	-	C_5	-	-	-	-
	52	G_1	G_1	-	U				
	53	$A_2C_1G_2$	G_1	-	C_1				
	54	G_4	G_5	-	C_1				
	55	C_1	G_1	C_2E_1	C_3				
	56	A_1	G_1	-	-				
	57	A_1G_3	-	C_2	C_1				
	58	G_1	G_3	-	-				
	59	B_5	A_1	B_1	-				
	60	B_5	B_1	E_4	C_3E_1	-	-	-	C_3E_2
	61	B_1	G_1	-	-	-	-	-	C_2E_3
	62	B_1	B_1	B_1	C_5	ND	ND	ND	C_3E_2
	Prev	**12/12**	**11/12**	**5/12**	**8/12**	**0/3**	**0/3**	**0/3**	**3/4**
9/6	**63**	G_1	-	-	-	-	-	-	-
	64	B_1	-	-	-				
	65	B_4	B_4	-	-				
	66	B_5	B_2	B_1	B_1				
	67	B_2G_1	U	-	-				
	68	B_5	B_1	G_1	-				
	69	B_2G_3	B_1	-	B_4	-	G_5	-	C_2 E_3
	70	B_3 G_2	G_1	-	B_2	-	-	-	C_4 E_1
	Prev	**8/8**	**6/8**	**2/8**	**3/8**	**0/3**	**1/3**	**0/3**	**1/3**
15/6	**71**	G_1	G_5	-	C_1	-	-	-	-
	72	G_1	A_1	-	U				
	73	C_1	-	E_1	E_1				
	74	G_3	G_1	-	C_3				
	75	C_1	C_1	-	-				
	76	B_3C_1	B_1	E_4	C_1				
	77	B_1	-	-	-				
	78	C_4	B_1	-	E_5				
	79	C_5	C_1	-	E_1	-	-	-	C_4E_1
	80	C_1	-	-	-	-	-	-	C_3E_2
	81	C_3	C_1	-	-	ND	ND	-	C_1E_4
	82	C_4G_1	-	-	C_1	ND	ND	-	C_4E_1
	Prev	**12/12**	**8/12**	**2/12**	**8/12**	**0/3**	**0/3**	**0/5**	**4/5**
16/6	**83**	U	-	-	-	ND	ND	ND	ND
	84	G_1	-	-	C_1				
	85	G_1	-	-	-				
	86	G_2	C_2	-	-				
	87	G_1	G_1	-	E_1				
	88	G_4	G_1	-	C_1				
	89	C_1G_2	C_1	-	C_1				
	90	G_1	-	-	-				
	91	C_4G_1	G_1	-	G_1				
	92	C_3G_2	C_1	C_1	G_2	C_1	G_1	-	C_2E_1
	93	G_5	G_5	-	-	-	-	-	C_4E_1
	94	C_2G_3	G_5	G_1	G_1	ND	ND	ND	C_2E_3
	Prev	**12/12**	**8/12**	**2/12**	**7/12**	**1/3**	**1/3**	**0/3**	**3/4**
13/7	**95**	G_5	G_5	G_1	-	-	-	-	E_5
	96	A_3G_2	G_5	-	C_1	-	-	-	E_1
	97	G_1	C_1	-	-				
	98	G_1	-	-	-				
	99	G_1	-	-	-				
	100	B_5	B_1	-	-				
	101	C_1	-	-	-				
	102	G_1	-	-	-				
	103	B_5	B_1	-	-				
	104	B_5	B_1	G_1	E_3				
	105	B_5	B_1	-	-	-	-	-	C_1
	106	B_5	B_1	G_1	-	-	-	-	C_1E_4
	Prev	**12/12**	**8/12**	**3/12**	**2/12**	**0/4**	**0/4**	**0/4**	**4/4**
14/7	**107**	G_1	-	-	-	-	-	-	C_1
	108	G_1	-	-	-	-	-	-	C_1
	109	C_5	C_1	-	G_1				
	110	C_4G_1	C_4	-	-				
	111	G_3	-	-	-				
	112	G_1	-	-	-				
	113	-	U	-	G_1				
	114	G_2	-	-	-				
	115	-	-	-	-				
	116	-	G_1	G_1	G_1				
	117	G_4	-	-	-	-	-	-	C_1
	118	G_4	-	-	-	-	-	-	C_2E_2
	Prev	**9/12**	**4/12**	**1/12**	**3/12**	**0/4**	**0/4**	**0/4**	**4/4**
Overall prev.		**114/118**	**70/112**	**19/117**	**35/117**	**1/31**	**2/31**	**0/35**	**24/35**

REFERENCES

Alban, L., & Stark, K. D. C. (2002). Simulating salmonella prevalence from the growing pig to the slaughtered carcass: Where should the effort be put to increase food safety? *Prev Vet Med, 68(1),* 98-110.

Anderl, J. N., Franklin, M. J., & Stewart, P. S. (2000). Role of antibiotic penetration limitation in Klebsiella pneumoniae biofilm resistance to ampicillin and ciprofloxacin. *Antimicrob Agents Ch, 44*(7), 1818-1824.

Baptista, F. M., Dahl, J., & Nielsen, L. R. (2010). Factors influencing Salmonella carcass prevalence in Danish pig abattoirs. *Prev Vet Med, 95(3-4)*, 231-238.

Borch, E., Nesbakken, T., & Christensen, H. (1996). Hazard identification in swine slaughter with respect to foodborne bacteria. *Int J Food Microb, 30*(1-2), 9-25.

Botteldoorn, N., Herman, L., Rijpens, N., & Heyndrickx, M. (2004). Phenotypic and molecular typing of Salmonella strains reveals different contamination sources in two commercial pig slaughterhouses. *Appl Env Microbiol, 70*(9), 5305-5314.

Botteldoorn, N., Heyndrickx, M., Rijpens, N., Grijspeerdt, K., & Herman, L. (2003). Salmonella on pig carcasses: positive pigs and cross contamination in the slaughterhouse. *J Appl Microbiol, 95*(5), 891-903.

Chavant, P., Martinie, B., Meylheuc, T., Bellon-Fontaine, M. N., & Hebraud, M. (2002). Listeria monocytogenes LO28: surface physicochemical properties and ability to form biofilms at different temperatures and growth phases. *Appl Env Microbiol, 68*(2), 728-737.

Costerton, J. W., Lewandowski, Z., Caldwell, D. E., Korber, D. R., & Lappin-Scott, H. M. (1995). Microbial biofilms. *Annu Rev Microbiol, 49*, 711-745.

De Man, J. C. (1983). MPN Tables, Corrected. *Eur J Appl Microbiol, 17*, 301-305.

De Wit, M. A., Koopmans, M. P., Kortbeek, L. M., Wannet, W. J., Vinje, J., van Leusden, F., Bartelds, A. I., & van Duynhoven, Y. T. (2001). Sensor, a population-based cohort study on gastroenteritis in the Netherlands: incidence and etiology. *Am J Epidemiol, 154*(7), 666-674.

Delhalle, L., De Sadeleer, L., Bollaerts, K., Farnir, F., Saegerman, C., Korsak, N., Dewulf, J., De Zutter, L., & Daube, G. (2008). Risk factors for Salmonella and hygiene indicators in the 10 largest Belgian pig slaughterhouses. *J Food Prot, 71*(7), 1320-1329.

Donlan, R. M., & Costerton, J. W. (2002). Biofilms: survival mechanisms of clinically relevant microorganisms. *Clin Microbiol Rev, 15*(2), 167-193.

Duggan, S. J., Mannion, C., Prendergast, D. M., Leonard, N., Fanning, S., Gonzales-Barron, U., Egan, J., Butler, F., & Duffy, G. (2010). Tracking the Salmonella Status of Pigs and Pork from Lairage through the Slaughter Process in the Republic of Ireland. *J Food Prot, 73*(12), 2148-2160.

EFSA. (2010). EFSA Quantitive Microbial Risk Assessment on *Salmonella* in Slaughter and Breeder pigs: Final Report.

Joseph, B., Otta, S. K., Karunasagar, I., & Karunasagar, I. (2001). Biofilm formation by salmonella spp. on food contact surfaces and their sensitivity to sanitizers. *Int J Food Microb, 64*(3), 367-372.

Kim, S. H., & Wei, C. I. (2007). Biofilm formation by multidrug-resistant Salmonella enterica serotype typhimurium phage type DT104 and other pathogens. *J Food Prot, 70*(1), 22-29.

Kjaerulff, U. B., & Madsen, A. L. (2008). *Bayesian Networks and Influence Diagrams*. New York: Springer.

Kusumaningrum, H. D., Riboldi, G., Hazeleger, W. C., & Beumer, R. R. (2003). Survival of foodborne pathogens on stainless steel surfaces and cross-contamination to foods. *Int J Food Microb, 85*(3), 227-236.

Morgan, I. R., Krautil, F. L., & Cravel, J. A. (1989). Bacterial-populations on dressed pig carcasses. *Epidemiol Inf, 1*(15-24).

Pielaat, A. (2011). The data supply chain for tracing Salmonella in pork production. *Int J Food Microb, 145 Suppl 1*, S66-67.

Smid, J. H., Heres, L., Havelaar, A. H., & Pielaat, A. A Biotracing Model of Salmonella in the pork production chain. Accepted for publication in *J Food Prot*.

Sorensen, L. L., Sorensen, R., Klint, K., & Nielsen, B. (1999). Persistent environmental strains of Salmonella infantis at two Danish slaughterhouses, two case-stories. In *Proceedings of the 3rd Int. Symp. on the Epidemiology and Control of Salmonella in Pork* (pp. 285-286). Washington D.C.

Spescha, C., Stephan, R., & Zweifel, C. (2006). Microbiological contamination of pig carcasses at different stages of slaughter in two European Union-approved abattoirs. *J Food Prot, 69*(11), 2568-2575.

Stewart, P. S., & Costerton, J. W. (2001). Antibiotic resistance of bacteria in biofilms. *Lancet, 358*(9276), 135-138.

Swanenburg, M., Berends, B.R., Urlings, H.A.P., Snijders, J.M.A., van Knapen, F. (2001a). Epidemiological investigations into the sources of Salmonella contamination of pork. *Berl Muench Tieraerztl, 114*, 356-359.

Swanenburg, M., Urlings, H. A. P., Snijders, J. M. A., Keuzenkamp, D. A., & van Knapen, F. (2001b). Salmonella in slaughter pigs: prevalence, serotypes and critical control points during slaughter in two slaughterhouses. *Int J Food Microb, 70*(3), 243-254.

Swanenburg, M., van der Wolf, P. J., Urlings, H. A., Snijders, J. M., & van Knapen, F. (2001c). Salmonella in slaughter pigs: the effect of logistic slaughter procedures of pigs on the prevalence of Salmonella in pork. *Int J Food Microb, 70*(3), 231-242.

Van Hoek, A. H. A. M., de Jonge, R., van Overbeek, W. M., Bouw, E., Pielaat, A., Smid, J. H., Löfström, C., Malorny, B., Junker, E., Löfström, C., Pedersen, K., Aarts, H. J. M., & Heres, L. (2012). A quantitative approach toward a better understanding of the dynamics of Salmonella spp. in a pork slaughter-line. *Int J Food Microb, 153*, 45-52.

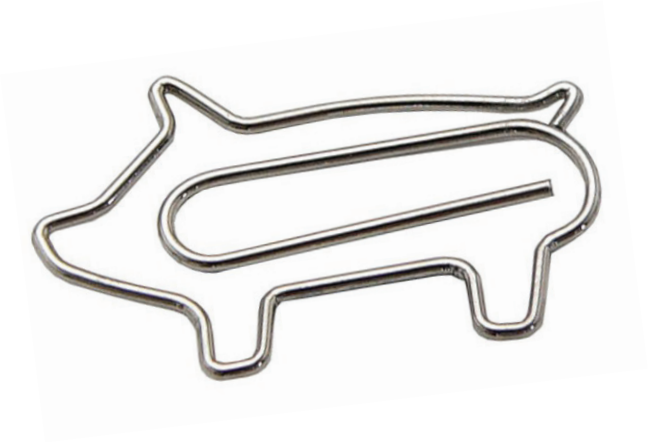
Van Pelt, W., Notermans, D., Mevius, D. J., Vennema, H., Koopmans, M. P. G., & Duynhoven, Y. T. H. P. v. (2008). Trends in gastro-enteritis van 1996 – 2006: Verdere toename van ziekenhuisopnames, maar stabiliserende sterfte. *Infectieziektenbulletin, 19*, 24-31.

Vieira-Pinto, M. M., & Martins, C. (2003). Detection of Salmonella in pork at the slaughterhouse. In *Proceedings of the National Congress of Microbiology* (pp. 237). Tomar, Portugal.

Warriner, K., Aldsworth, T. G., Kaur, S., & Dodd, C. E. (2002). Cross-contamination of carcasses and equipment during pork processing. *J Appl Microbiol, 93*(1), 169-177

Chapter 8

General discussion



1. OUTLINE

In this thesis, biotracing systems for *Salmonella* in the pork slaughter chain were developed. This entailed modeling-work combined with targeted collection of data that were used as inputs for the model. The knowledge obtained in this study is the first step towards the development of an operational biotracing system which should guide evidence-based decision making on prevention and mitigation strategies for *Salmonella* in the pork slaughter chain. In this general discussion, three main issues will be addressed. First, the methodology developed for biotracing *Salmonella* in the pork slaughter chain will be evaluated. The structure of the concentration model (Chapters 3 and 5), as developed for one pig slaughterhouse in The Netherlands, will be discussed and related to the possible application of this structure for other pig slaughterhouses. Possible refinements of the model will be proposed and a theoretical comparison will be made between this model and the serotyping model (Chapter 7). Second, the results obtained in the different chapters of this thesis study will be discussed jointly. In particular, the important role of house flora for the contamination of carcasses, which was a major finding of this study, will be addressed. Third, general conclusions will be given and a perspective will be sketched of the future challenges to make biotracing *Salmonella* in the pork slaughter chain operational.

2. DISCUSSION OF METHODS

2.1 Structure of the concentration model

2.1.1 Modules

The structure of the concentration model (Chapter 3) is such that it describes all relevant processing steps and sources of *Salmonella* in the particular slaughterhouse that was studied for this thesis. This implies that it cannot directly be used in other slaughterhouses where not all, or different, steps are imposed, where the steps are imposed in another order, or where different sources of *Salmonella* may be present. Nevertheless, the general design of individual operations, and their sequence, is very uniform in industrial high-throughput slaughterhouses and has not changed significantly for decades (EFSA, 2006). This implies that minor modifications of the model (e.g. only one flaming/polishing step instead of two) may make it applicable to many slaughterhouses. In practice this means that, at most, a few modules (each one describing the dynamics of *Salmonella* during one processing step) may be added to or removed from the model. The practical framework proposed in Chapter 3 can be used to model the bacterial dynamics during one processing step and, thereby, to compose these modules.

2.1.2 Sources

Potential sources of *Salmonella* contamination for carcasses at the re-work station were predefined in the model structure. Considered sources were: (i) contaminated skin surfaces of pigs at the onset of slaughter, (ii) contamination from the gut contents of the pigs and (iii) house flora. In the slaughterhouse under study, the heads of the pigs were separated from the carcass as a whole at the end of the slaughtering line and not split at the carcass splitting stage as occurs in some slaughterhouses. In addition, the mandibular lymph nodes were not palpated nor incised during meat inspection. These measures prevent the introduction of *Salmonella* residing in the pig's oral cavity or mandibular lymph nodes to carcasses, as was also recommended by the EFSA panel on Biological Hazards (Anonymous, 2011). The pig's oral cavity and the mandibular lymph nodes of pigs were, therefore, not included in the model structure as potential sources of *Salmonella* contamination. If, in another slaughterhouse,

the pig's heads are split or incision of the lymph nodes during meat inspection are still made then these should be included in the model as additional potential sources of contamination together with a description of the dynamics of *Salmonella* that are potentially transferred from these sources. Again, the practical framework proposed in Chapter 3 can be used to provide this description.

At the beginning of the project it was unknown whether house flora was, indeed, a relevant source of *Salmonella* for carcasses in the studied slaughterhouse. Sampling is difficult in small niches in equipment of the slaughterhouse, where *Salmonella* may survive and become part of the house flora of the slaughterhouse. In addition, positive samples taken from the equipment do not directly imply the existence of house flora but may also indicate transient bacteria (through cross-contamination). The initial model assumption was made that house flora can act as a source of contamination that explains the fraction of the *Salmonella* population on the end product that cannot be explained by the contributions of the other sources (Chapter 3).

The belly opener and the carcass splitter were the only slaughter robots from which swab samples were taken because of the findings of Swanenburg *et al.* (2001) and Lo Fo Wong & Hald (2000), who observed that house flora is particularly associated to the cutting robots. By analyzing the samples from these robots, it was found that only one carcass splitter (SP2) was consistently contaminated throughout the sampling period (Chapter 4). Therefore, the general definition of house flora in Chapter 3 was further specified; for the purpose of the particular slaughter house and period studied, house flora was defined as resident *Salmonella* on carcass splitter SP2 which may be transferred to passing carcasses (Chapter 5).

In general (in other slaughterhouses or for other periods than were studied here), house flora may be located at several different slaughter robots. The robots may even act simultaneously as sources of house flora and their relative importance may change over time. The contribution of house flora can be further specified in the model (e.g. house flora per slaughter robot) if data are available over a more extensive period than the studied sampling period. The procedure for determining which sources are relevant, as sketched above, can then be used. First, sampling data from these robots may be used to define which robots should be included in the model as potential sources of house flora. After adapting the model structure based on such an assessment, downstream data from the carcass samples together with historical 'experience' can be used for biotracing.

2.1.3 Recursions

Cross-contamination was expected to be an important factor in the bacterial dynamics for *Salmonella* in the pork slaughter chain (Botteldoorn et al. 2003; Olsen et al. 2003; Warriner et al. 2002), and was modeled using a system of recursive equations. In the implementation of the BBN, the explicit equations defining the dynamics based on individual carcasses were replaced by stochastic relations defining the dynamics based on random carcasses (Chapter 3). In addition, the prior distributions of the cross-contamination parameters were assimilated in the likelihood of their child variables (typically, the variables representing the number of *Salmonella* on the carcass after the cross-contamination step), instead of implementing them as independent parameters in the BBN, such as all other parameters. The choices for this particular model structure were made to avoid the computational complexity of

a dynamic Bayesian belief network and the Expectation-Maximization (EM) algorithm (Spiegelhalter & Lauritzen, 1990) was used for this purpose. A drawback of this modeling choice was that the prior distributions of the cross-contamination parameters could not be learned from data. This implied that the true values of the cross-contamination parameters remained considerably uncertain, even though much upstream and downstream data specifying microbial levels on carcasses were available. Therefore, an additional laboratory experiment was done to obtain more insight in the values of the parameters specifying cross-contamination via the cutting robots (Chapter 6). The results of this experiment were used for an adjusted version of the concentration model. The updated results of this model are shown in section 3.2.

2.2 Possible refinements

The main reason for the approximations to the structure of the concentration model, described in section 2.1.3, was the limited available amount of physical computer memory (RAM). More specifically, large BBNs (e.g. dynamic BBNs) typically have complex *clique* structures (Kjaerulff & Madsen, 2008), which put a heavy demand on physical computer memory. Such practical limitations have lead to specific choices for the model structure in this study. If such limitations can be overcome in the future, it is expected that improvements of the model are obtained by more sophisticated model structures. Some suggestions will be made in this section.

2.2.1 Recursions

A direct implementation of recursive relations describing cross-contamination in a dynamic BBN allows the inclusion of additional information into the model, namely, the carcass number or the time of sampling. This is due to the fact that, using a dynamic BBN, data from a carcass is no longer considered as data from a random product but, instead, as data from a specific product (which has a specific ordering in relation to other products). Because cross-contamination is reported to be of higher importance during the afternoon than in the morning (Chapter 7, Delhalle et al., 2008), it is expected that additional information about the processing order may lead to an improved distinctiveness between potential sources. In addition, the direct implementation of recursive relations enables the explicit inclusion of the cross-contamination parameters as independent parameters in the model. Consequently, their explicit inclusion allows that their values may be learned from longitudinal downstream data.

2.2.2 Different carcass surfaces

In the concentration model, the cutting section and the skin of the carcass were considered as separate sites where *Salmonella* could be detected after slaughter. Each of these sites has its own distribution of *Salmonella*, which were assumed to be homogeneous per surface area. This assumption is unlikely to be true (Lo Fo Wong & Hald, 2000; Richards & Dodd). A more complex model, in which more sites are separately considered, would better represent the inhomogeneous distribution of *Salmonella* over a carcass. Such a model could, therefore, produce more accurate and specific biotracing results. Yet, the complexity of the model structure would then increase dramatically to allow interactions between *Salmonella* on the different carcass surfaces and the model would need much more data to estimate the model parameters than were available in the current study.

2.2.3 Combining data

The *Salmonella* population data (Chapter 5) and serotyping data (Chapter 7) were considered separately for biotracing. Initially, it was planned for these different types of data to be combined in one model. The main problem that was encountered in the attempts to assimilate the serotyping data into the concentration model (Chapter 3) was the necessity to model the dynamics of *Salmonella* of each serotype separately, and combine these sub-models into one big model to describe the dynamics on population level. This would lead to a large BBN with similar parallel yet interacting chain structures, each describing the dynamics of a particular *Salmonella* serotype. The number of parameters in such a model would increase because certain parameters may be serotype-specific, e.g. the inactivation rate of one *Salmonella* serotype may differ from the inactivation rate of another *Salmonella* serotype. Therefore, the complexity of the model structure would, again, increase and much more data would be needed to learn the increased number of parameters.

2.3 Theoretical comparison of the concentration model and the serotyping model

The concentration model (Chapters 3 and 5) and the serotyping model (Chapter 7) could both be used for biotracing but they differ in comprehensiveness as well as in type of data inputs. The concentration model needed prevalence and quantitative bacterial population data. The serotyping model needed upstream and downstream serotyping and prevalence data from carcasses and from suspect slaughter robots. In this section we discuss theoretical differences, and advantages and disadvantages of both approaches. A comparison of the results of both approaches is provided in section 3.1.

2.3.1 Comprehensiveness

The biotracing approach is the most comprehensive. The concentration model was trained with data obtained from contamination events that happened in the past. In principle, routine registrations in the slaughterhouse such as time / temperature registrations of bactericidal procedures can also be used for learning the model parameters. This was not done in the current study because these data were not available. Yet, in general, such measurement may contain valuable information for biotracing (Barker, Gomez, & Smid, 2009), especially if the source of contamination is less distinctly active or more variable than in the current study. The dynamics of *Salmonella* through the chain were preserved in the model as a joint probability distribution, which was used to generate tracing estimates of the sources of contamination of individual carcasses. However, such dynamics may change over time, which was reflected in the joint probability distribution changing over time. For example, it was seen in Chapter 5 that the importance of house flora for the contamination of carcasses varied over the period of sampling. Because these dynamical changes are incorporated in the joint probability distribution, the model can be used to predict sources of potential contamination events happening in the near future. In addition, the concentration model can be used as a predictive tool to estimate the bacterial load on a product after every production step (Chapter 5) and, therefore, allows for a calculation of human health risk reductions due to interventions in the chain.

The serotyping model is less comprehensive and only establishes time averaged association between cases and sources instead of direct causation (Chapter 7, Fig. 1). In particular, the model does not include parameters describing the *Salmonella* dynamics through the chain, the values of which may change over time. Therefore, the sources of observed contaminated carcasses at the end of the chain

could be assessed, but the prediction of sources of future contamination events is less accurate. Also, this approach cannot be used to calculate the consequences of interventions in the chain. Yet, the simplicity of the serotyping model makes it easy to understand and implement, and applicable to any slaughterhouse for which similar data is available.

2.3.2 Data

As the amount and type of data that the models need determine the efforts in taking microbiological samples in the chain and the time and money needed to analyze the samples, the choice of model may also depend on the different data requirements of the models. In this study, the data were obtained by traditional bacteriological methods: the most-probable-number (MPN) test (de Man, 1983) was used for quantifying Salmonella in a sample and the Kaufman-White scheme (Grimont & Weill, 2007) was used for serotyping the Salmonella strain. When using these methods, isolation of Salmonella colonies was not needed for quantification but it was needed for serotyping. Consequently, the time to obtain *Salmonella* concentration data (3 days) was shorter than the time to obtain *Salmonella* serotyping data (4-5 days) from one sample. Yet, quantification can be labor-intensive and expensive on laboratory materials depending on the number of serial dilutions and tubes per dilution used in the MPN test. On the other hand, quantification can be done in standard laboratories whereas specialized laboratories are needed for classical serotyping which may increase the costs. Novel high-throughput methods, such as real-time PCR (Malorny et al. 2008) for quantification and PremiTest (Wattiau et al., 2008) for serotyping are more rapid methods which may, in the future, be more cost-effective in providing data for biotracing. Yet, their cost-benefit analysis was beyond the scope of this study.

The following observations in this study provide indications to the amount of data that were needed by both models. It was seen that removal of the second half of the upstream and environmental data did not influence the adaptation results of the concentration model parameters. Removal of all upstream and environmental data did only modify the adaptation results to a limited extent. This means that the model did not need all data for biotracing: only the downstream data and an initial, limited number of upstream and environmental data points were sufficient for training the model in this case. The serotyping model, on the other hand, needed the upstream and downstream data for each carcass and environmental data for each day to trace the sources of *Salmonella* for individual carcasses.

These observations suggest an advantage of the biotracing approach over the serotyping approach because:

- less samples need to be taken, the time-to-results is shorter,
- taking downstream samples at the re-work station is already current practice in the slaughterhouse and the additional analyses of these samples can therefore be incorporated more easily in existing quality control and monitoring programs, and
- taking environmental samples from the cutting equipment is not straightforward. It is a hazardous task because of the sharp blades of these machines so it can only be done if the machines are not running.

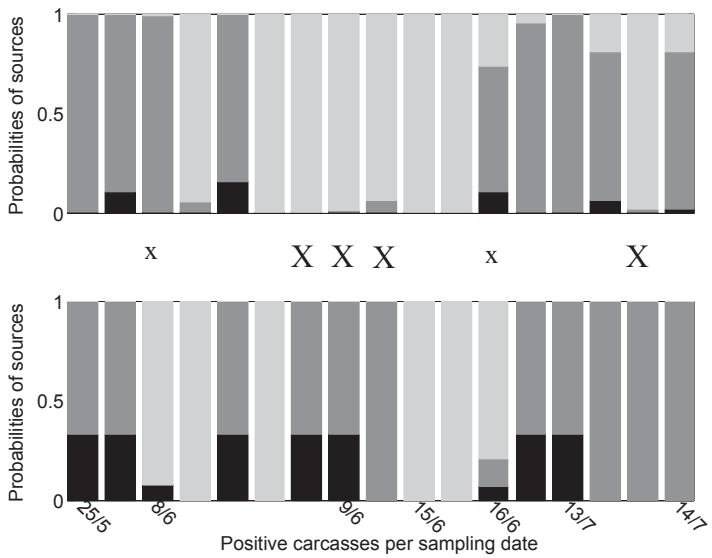


Fig. 1. *Combined results of Chapter 5, Fig. 5A and Chapter 7, Fig. 4A. For each potential source (black: pig skin; dark grey: feces; light grey: house flora), the estimated probabilities of being the primal source of contamination of the carcass exterior are indicated by fractions of a bar chart.* Upper chart: *Results by the concentration model (Chapter 5).* Lower chart: *Results by the serotyping model (Chapter 7). Carcasses for which the models predicted a different primal source of contamination are indicated with a cross. Big cross: see main text.*

3. INTEGRATION AND DISCUSSION OF MAIN RESULTS

3.1 Comparison results of the concentration model and the serotyping model

Chapters 5 and 7 describe different approaches in finding the most likely source of *Salmonella* contamination for pork at the re-work station. It is difficult to judge which approach has a superior performance in tracing the sources of contamination because there is no golden standard. The results of both models were similar for most *Salmonella*-positive carcasses at the re-work station (compare Chapter 5, Fig. 5 to Chapter 7, Fig. 4). By both approaches it was concluded that house flora from the second carcass splitter, SP2, was a relevant source of exterior contamination for carcasses slaughtered in June, and of interior contamination for carcasses slaughtered during the whole studied period.

Nevertheless, differences could be observed: some cases of contamination were attributed to house flora from SP2 by the concentration model but to feces by the serotyping model, or converse. This is shown for cases of external contamination in Fig. 1, but the same observation could be made for some cases of internal contamination (not shown). The downstream samples of most of these cases (these cases are indicated with a big cross in Fig. 1) contained a relatively high *Salmonella* concentration and *Salmonella* serotypes that were similar to serotypes that had only been found in samples from the incoming pigs, in particular *S.* Typhimurium and *S.* Brandenburg. This combination of data may explain the different outcomes of both models for these cases. It was computed by the concentration model that moderate numbers of *Salmonella* on the incoming pigs reduce to zero or to, at most, low numbers on (the exterior or the interior of) carcasses at the re-work station. Therefore, if relatively high bacterial numbers were found in the downstream samples, then, in general, the concentration model

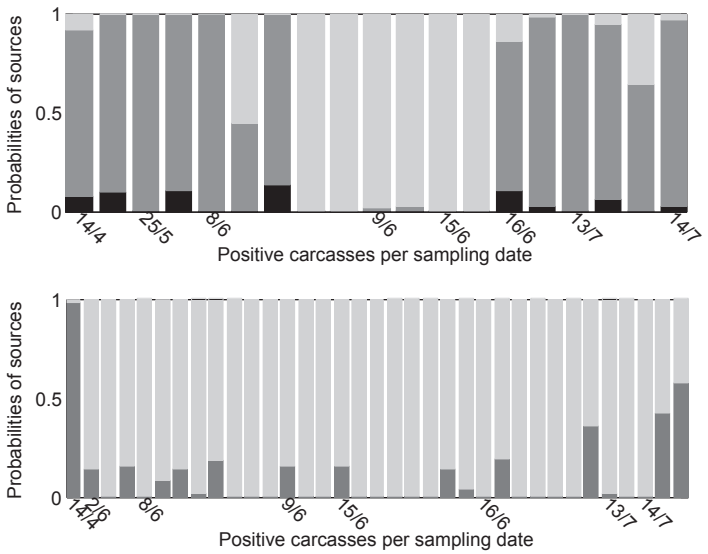


Fig. 2. *Biotracing results of concentration model with fading factor 0.9. For each potential source (black: pig skin; dark grey: feces; light grey: house flora), the estimated probabilities of being the dominant source of contamination of a carcass are indicated by fractions of a bar chart. Upper chart: Results for contamination of the carcass exterior. Lower chart: Results for contamination of the carcass interior.*

attributed these cases of contamination were attributed to house flora. Yet, serovars *S.* Typhimurium and *S.* Brandenburg were attributed to the incoming pigs by the serotyping model, as these serovars had not been found on the equipment but only on incoming pigs.

The apparent discrepancies between the concentration and serotyping data for these cases may be due to the limited number of bacterial colonies per sample that could be serotyped. Only up to five bacterial colonies per sample were serotyped in this study, and in many cases only one. In many instances where multiple bacterial colonies per sample were serotyped, several serovars were found per individual sample. This makes clear that there is a chance that some serovars were missed when serotyping one or few colonies per sample (such as is routinely done). This may lead to an underestimation of *Salmonella* serovars in pork and to incorrect biotracing results of the serotyping model.

Remaining uncertainties about the true values of parameters in the concentration model is another explanation for apparently adverse findings of both models. In particular, the role of cross-contamination was uncertain because the cross-contamination parameters could not be learned from downstream data (section 2.1.3). In addition, learning the transfer rate of house flora may not have been perfect because positive samples taken from the equipment may indicate house flora as well as cross-contamination. Therefore, it was possible that the true importance of house flora differed slightly from the initial estimates of the concentration model. Updated estimates, in which the data from the cross-contamination experiment described in Chapter 6 were used, are shown in section 3.2.

The choice of which experience count is assigned to the prior distributions of the model parameters was another, subjective, choice which could have influenced the results of the concentration model (Chapter

3). We have chosen to assign initial experience counts of 10^{-6} to uncertain parameters. This low count expressed a low initial degree of certainty about the true values of parameters. The uncertainty about the parameter values of the model was gradually reduced by training the model with data, which was seen by increasing experience. However, it is arguable that, if used over an extended period of time, the model should also be able to *loose* experience if the data by which this experience was obtained becomes outdated. This process, called (sequential) *fading*, is indeed possible in BBNs and works in a similar manner as sequential adaptation (Chapter 3). How fast experience is lost is quantified by the fading factor. Fading was not used in the current model because of the limited number of data available for training the model. Yet, this is another subjective choice which may have influenced the results of the concentration model. Fig. 2 shows the results if a fading factor of 0.9 was used. A comparison of Chapter 5, Fig. 5 and this figure shows that the inclusion of a moderate fading factor had only a minor impact on the biotracing results.

3.2 Integration of results of Chapters 3, 5 and 6

It was explained in section 2.1.3 that all parameters of the concentration model could be learned from data (using sequential adaptation, Chapter 3), except those expressing cross-contamination. Chapter 6 of this thesis describes a laboratory experiment that was done to reduce uncertainty about the values of the parameters specifying cross-contamination via the cutting robots (Chapter 6).

The mean ratio of *Salmonella* that was transferred from pork to stainless steel (0.19) as experimentally determined (Chapter 6) was found to be close to the value for this parameter used in the initial model (0.21). The mean ratio of *Salmonella* that was transferred from stainless steel to pork as experimentally determined (0.58) was found to be lower than the value for this parameter used in the initial model (0.94). Yet, the variability of these ratios over different experiments was high and the uncertainty resulting from limited colony counts in one experiment was also high. The updated distributions of transfer ratios (Chapter 6) were used in an adjusted version of the concentration model. Updated biotracing results are shown in Fig. 3. It can be seen that they are almost identical to the results shown in Chapter 5. Fecal contamination seems to have gained a little more importance for the exteriors of the second carcass slaughtered on 8 June and the third carcass slaughtered on 13 July, which carried low numbers of *Salmonella* on their skin. The different results for the 2nd, 3rd and 4th carcasses with contaminated interiors may be artifacts of the model predictions due to poor estimates of the parameters quantifying the transfer rates of *Salmonella* from house flora. The similarity in results indicate that cross-contamination via the cutting knives has not been an important phenomenon for the contamination of carcasses.

3.3 Relevance of house flora for carcass contamination

Many published studies which aim to determine the source of *Salmonella* contamination of carcasses debate the question whether resources should be devoted to on-farm control efforts or to slaughterhouse efforts such as carcass decontamination or increased slaughter hygiene. Different answers to this question were found. Letellier et al (2009) concluded that the bacterial status of the pigs before slaughter was an important risk factor for the bacterial status of carcasses and Duggan et al (2010) found lairage to be more important for carcass contamination than slaughter. Berends et al. (1997) concluded that all carcass contamination comes from the animals themselves (70% because the animals were carriers and

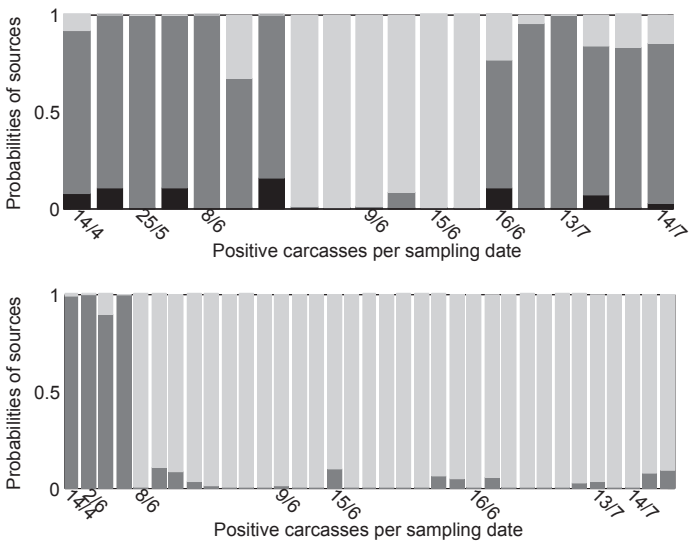


Fig. 3: *Biotracing results of the concentration model with updated cross-contamination rates (Chapter 6). For each potential source (black: pig skin; dark grey: feces; light grey: house flora), the estimated probabilities of being the dominant source of contamination of a carcass are indicated by fractions of a bar chart. Upper chart: Results for contamination of the carcass exterior. Lower chart: Results for contamination of the carcass interior.*

30% because of cross-contamination) and that interim cleaning and disinfection of surfaces will only prevent 10% of all cross-contamination (Berends et al. 1998a, 1998b). In the EFSA QMRA on *Salmonella* in Slaughter and Breeder pigs (EFSA, 2010) it was also found that certain on-farm interventions would be effective in reducing the number of human cases of salmonellosis. Yet, Hurd et al. (2008), in a Danish modeling study, found that even the most stringent on-farm *Salmonella* program (such as the Danish pork *Salmonella* program) cannot completely prevent *Salmonella* infected pigs (even not in herds in the low seroprevalence catergories) and that interventions in the primary production alone have a limited effect and will not markedly improve public health due to cross-contamination and re-contamination of carcasses during subsequent slaughtering. Bollaerts et al. (2010) also found that the most effective strategies were found at the end of the slaughterline, suggesting that re-contamination at later stages of the slaughter process might undo reductions in carcass contamination achieved at earlier stages.

Conclusions about the efficiency of on-farm and slaughterhouse *Salmonella* control efforts probably vary because of the large variability of pig slaughterhouses in respect to the microbiological status of carcasses (Borch et al., 2008; Small et al., 2006), which is indicative of the variability of the technical feasibility of minimizing contamination among slaughterhouses (Anonymous, 2011). Achieving *Salmonella* targets for chilled carcasses at the slaughterhouse, therefore, depends on a) process hygiene in the slaughterhouse and b) *Salmonella* levels in incoming pigs, according to the EFSA panel on Biological Hazards (Anonymous, 2011).

Several studies have found that complex contamination cycles exist in a slaughterhouse (Botteldoorn et al., 2004; Swanenburg, 2000; Vieira-Pinto et al., 2006), yet cross-contamination via the equipment and

re-contamination from house flora on the equipment were in most studies not considered separately. Supplementary to previous work, we have provided quantitative estimates of these distinct phenomena. A major finding of this study was that *Salmonella* can be transferred from sites harboring house flora to the carcasses and can contaminate many carcasses, and be detectable at the re-work station. Roughly half of the externally contaminated carcasses were most likely contaminated by this source and almost all internally contaminated carcasses were deemed to be contaminated by this source. This house flora was associated to one carcass splitter in the slaughterhouse. In the microbiological samples taken from this equipment, only *S.* Derby and *S.* Rissen were found. Further genotyping using PFGE showed that the *S.* Rissen isolates were identical and that there was a clear domination of one particular *S.* Derby genotype, clearly indicating the persistence of these strains. Viera-Pinto et al. (2006) also characterized *S.* Rissen as a house flora strain. Persistent contamination of the carcass splitter by S. Infantis has also been found in an earlier study (Swanenburg et al., 2001). Next, results of the SALINPORK study (Lo Fo Wong & Hald, 2000) also suggested that house flora may be an important source of contamination for carcasses.

Although cross-contamination between *Salmonella* positive pigs slaughtered on the same day has been taken into account in the majority of published QMRA models for *Salmonella* in the pork slaughter chain, the role of persistent contamination of the equipment has not often been addressed or assumed to be of minor importance (Bollaerts et al., 2010; EFSA, 2010; Hill et al., 2003). It is indeed difficult to obtain data that directly quantify the transfer of *Salmonella* from house flora. We have shown how to infer such transfer rates using a modeling approach. The existence of house flora should be an important aspect in discussions regarding the efficiency of on-farm control measures or control measures during slaughtering. A close to complete eradication of *Salmonella* on the farm may prevent much cross-contamination in the slaughterhouse, but if house flora remains an issue in slaughterhouses then carcasses will still be contaminated during slaughtering because such contamination may persist in the slaughterhouse over a much longer period.

The biotracing approach proposed in this thesis can be used to accomplish fast and accurate tracing of house flora in a slaughterhouse. This should lead to a thorough disinfection of the contaminated equipment and, hence, to a swift end of the contamination problems.

4. CONCLUSIONS AND FUTURE CHALLENGES

Currently used strategies in case of contamination events in pig slaughterhouses consist of thorough cleaning and disinfecting all potentially suspect equipment, checking all Critical Control Points (CCP) and (re-)instructing the plant workers (see Introduction). Prioritizing measures is based on the nature of the contamination problem, the contamination history of the specific slaughterhouse and on the experience and commitment of the quality manager in the slaughterhouse. In general, most attention is paid to preventing the contamination of carcasses instead of finding the source of contamination. In practice, solving the contamination problem is often an inefficient process, which may take up to several weeks in exceptional cases.

A more systematic approach to prevent and trace *Salmonella* in pig slaughterhouses is the use of

models in which historic and current data, expert knowledge of quality managers and other experts are combined. Most published quantitative microbial risk assessment (QMRA) models use Monte Carlo simulations to evaluate preventive measures, e.g. (EFSA 2010; Hill et al., 2003). Yet, in those models the link between the data and the parameters is not interactive which makes these models not suitable for reasoning backward in the chain. In this thesis, the development is discussed of models that can use downstream data to instantaneously update the parameter distributions in the model (e.g. Chapter 5, Fig.2). These models, implemented as BBNs, can be used to make predictions based on longitudinal data and on historical 'experience'. A specific model was developed that can be used for tracing and tracking of *Salmonella* in the pork slaughter chain The use of this model is a more systematic approach to find the source of contamination problems in a pig slaughterhouse.

The developed model is a proof-of-concept for a biotracing-system and its functionality using a targeted data collection was shown. Now that the use of this model has been clearly demonstrated, the time has come to initiate projects to study the practical and continuous application of such models in pig slaughterhouses. Possible future extensions and refinements of the model were discussed in section 2.2 of this chapter. Our results indicated that a continuous data input from a specific slaughterhouse was needed to learn the model parameters for that slaughterhouse. Data from process variables (time/temperature registrations in the slaughterhouse) may also be useful for an operational biotracing system, although the sensitivity analysis done in this thesis indicated that these were of minor importance in the current study. If the aim is to implement an operational biotracing system as a part of the quality control system of a slaughterhouse, then systematic microbiological surveillance data from specific steps in the chain should be collected. In this thesis we have shown which steps may be important, i.e. exsanguination, dehairing, polishing, belly opening, carcass splitting and meat inspection.

In general, minimal requirements for future biotracing systems include a more systematic collection of microbial data and process measurements and a more structural storage of those data in databases. This kind of biotracing is relatively easy to achieve as, already, many samples are collected and analyzed by pork producers because of EU requirements (EC/2073/2005) and ad hoc investigations to solve contamination problems. Simple Bayesian schemes such as Bayesian classifiers (Geenen et al., 2011) can then be used to link potential sources to measurable indications of contaminations problems, or (combinations of) corrective actions to possible effects of those actions. Maximal requirements for future biotracing systems include intense routine sampling efforts at multiple steps in the food production chain, analyses of those samples (quantification and subtyping of a pathogen) and monitoring of various process variables to generate a complete and actual understanding of the dynamics of pathogens through the chain and to relate these to environmental process variables. Both types of future biotracing systems may imply modifications to established work patterns, data collections and analyses in order to obtain data that is useful for biotracing.

Based on our experiences with the developed biotracing system for *Salmonella* in the pork chain, it is expected that the venue of operational biotracing systems in food chains will improve the quality and increase the speed of decision-making concerning the sources of food-borne hazards. This will lead to economic benefits of the pork industry and, eventually, to a reduction of human health risk.

REFERENCES

Anonymous. (2011). Scientific Opinion on the public health hazards to be covered by inspection of meat (swine). *The EFSA Journal, 9*(10).

Barker, G. C., Gomez, N., & Smid, J. (2009). An introduction to biotracing in food chain systems. *Tr Food Sci Tech, 20*(5), 220-226.

Berends, B. R., Van Knapen, F., Mossel, D. A. A., Burt, S. A., & Snijders, J. M. A. (1998a). Impact on human health of Salmonella spp. on pork in The Netherlands and the anticipated effects of some currently proposed control strategies. *Int J Food Microbiol, 44*(3), 219-229.

Berends, B. R., Van Knapen, F., Mossel, D. A. A., Burt, S. A., & Snijders, J. M. A. (1998b). Salmonella spp. on pork at cutting plants and at the retail level and the influence of particular risk factors. *Int J Food Microbiol, 44*(3), 207-217.

Berends, B. R., Van Knapen, F., Snijders, J. M. A., & Mossel, D. A. A. (1997). Identification and quantification of risk factors regarding Salmonella spp. on pork carcasses. *Int J Food Microbiol, 36*(2-3), 199-206.

Bollaerts, K., Messens, W., Aerts, M., Dewulf, J., Maes, D., Grijspeerdt, K., & Van der Stede, Y. (2010). Evaluation of scenarios for reducing human salmonellosis through household consumption of fresh minced pork meat. *Risk Anal, 30*(5), 853-865.

Borch, E., Nesbakken, T., & Christensen, H. (1996). Hazard identification in swine slaughter with respect to foodborne bacteria. *Int J Food Microbiol, 30*(1-2), 9-25.

Botteldoorn, N., Herman, L., Rijpens, N., & Heyndrickx, M. (2004). Phenotypic and molecular typing of Salmonella strains reveals different contamination sources in two commercial pig slaughterhouses. *Appl Environ Microbiol, 70*(9), 5305-5314.

Botteldoorn, N., Heyndrickx, M., Rijpens, N., Grijspeerdt, K., & Herman, L. (2003). Salmonella on pig carcasses: positive pigs and cross contamination in the slaughterhouse. *J Appl Microbiol, 95*(5), 891-903.

Delhalle, L., De Sadeleer, L., Bollaerts, K., Farnir, F., Saegerman, C., Korsak, N., Dewulf, J., De Zutter, L., & Daube, G. (2008). Risk factors for Salmonella and hygiene indicators in the 10 largest Belgian pig slaughterhouses. *J Food Prot, 71*(7), 1320-1329.

Duggan, S. J., Mannion, C., Prendergast, D. M., Leonard, N., Fanning, S., Gonzales-Barron, U., Egan, J., Butler, F., & Duffy, G. (2010). Tracking the Salmonella Status of Pigs and Pork from Lairage through the Slaughter Process in the Republic of Ireland. *J Food Prot, 73*(12), 2148-2160.

EFSA. (2006). Opinion of the Scientific Panel on Biological Hazards on "Risk assessment and mitigation options of Salmonella in pig production. *The EFSA Journal, 341, 1-131* (Vol. 2007).

EFSA. (2010). EFSA Quantitive Microbial Risk Assessment on *Salmonella* in Slaughter and Breeder pigs: Final Report.

Geenen, P. L., van der Gaag, L. C., Loeffen, W. L., & Elbers, A. R. (2011). Constructing naive Bayesian classifiers for veterinary medicine: a case study in the clinical diagnosis of classical swine fever. *Res Vet Sci, 91*(1), 64-70.

Grimont, P. A. D., & Weill, F.-X. (2007). Antigenic formulae of the Salmonella serovars (9th ed.). In. Paris: WHO Collaborating Center for Reference and Research on Salmonella, Institut Pasteur.

Hill, A., England, T., Snary, E., Cook, A., Kelly, L., Evans, S., & Wooldridge, M. (2003). A ‚Farm-To-Consumption' Risk Assessment for Salmonella Typhimurium in pigs. Weybridge: Department of Risk Research, Veterinary Laboratories Agency.

Hurd, H. S., Enoe, C., Sorensen, L., Wachmann, H., Corns, S. M., Bryden, K. M., & Greiner, M. (2008). Risk-based analysis of the Danish pork Salmonella program: past and future. *Risk Anal, 28*(2), 341-351.

Kjaerulff, U. B., & Madsen, A. L. (2008). *Bayesian Networks and Influence Diagrams*. New York: Springer.

Letellier, A., Beauchamp, G., Guevremont, E., D'Allaire, S., Hurnik, D., & Quessy, S. (2009). Risk factors at slaughter associated with presence of Salmonella on hog carcasses in Canada. *J Food Prot, 72*(11), 2326-2331.

Lo Fo Wong, D. M. A., & Hald, T. (2000). Salmonella in pork (SALINPORK): Pre-Harvest and Harvest control options based on epidemiologic, diagnostic and economic research - final report.

Malorny, B., Lofstrom, C., Wagner, M., Kramer, N., & Hoorfar, J. (2008). Enumeration of salmonella bacteria in food and feed samples by real-time PCR for quantitative microbial risk assessment. *Appl Environ Microbiol, 74*(5), 1299-1304.

de Man, J. C. (1983). MPN Tables, Corrected. *Eur J Appl Microbiol Biotech, 17*, 301-305.

Olsen, J. E., Brown, D. J., Madsen, M., & Bisgaard, M. (2003). Cross-contamination with Salmonella on a broiler slaughterhouse line demonstrated by use of epidemiological markers. *J Appl Microbiol, 94*(5), 826-835.

Richards, P. J., & Dodd, C. E. R. D. Contamination of pork with Salmonella during slaughter and possible intervention measures. *M01040 Pig Literature Review*. Nottingham: Division of Food Sciences, University of Nottingham.

Small, A., James, C., James, S., Davies, R., Liebana, E., Howell, M., Hutchison, M., & Buncic, S. (2006). Presence of Salmo-

nella in the red meat abattoir lairage after routine cleansing and disinfection and on carcasses. *J Food Prot, 69*(10), 2342-2351.

Spiegelhalter, D., & Lauritzen, S. L. (1990). Sequential updating of conditional probabilities on directed graphical structures. *Networks, 20*.

Swanenburg, M. (2000). *Salmonella in the pork production chain: sources of Salmonella on pork.* PhD thesis, University Utrecht, Utrecht, the Netherlands.

Swanenburg, M., Berends, B.R., Urlings, H.A.P., Snijders, J.M.A., Knapen, F., Van. (2001). Epidemiological investigations into the sources of Salmonella contamination of pork. *Berl Muench Tieraerztl Wochenschr, 114*, 356-359.

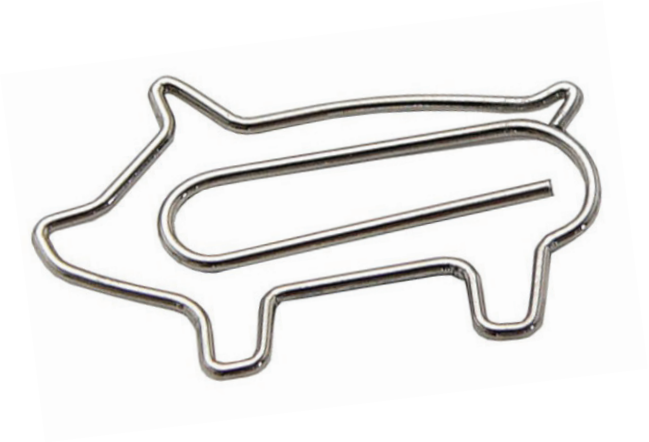
Vieira-Pinto, M., Tenreiro, R., & Martins, C. (2006). Unveiling contamination sources and dissemination routes of Salmonella sp. in pigs at a Portuguese slaughterhouse through macrorestriction profiling by pulsed-field gel electrophoresis. *Int J Food Microbiol, 110*(1), 77-84.

Warriner, K., Aldsworth, T. G., Kaur, S., & Dodd, C. E. (2002). Cross-contamination of carcasses and equipment during pork processing. *J Appl Microbiol, 93*(1), 169-177.

Wattiau, P., Weijers, T., Andreoli, P., Schliker, C., Veken, H. V., Maas, H. M., Verbruggen, A. J., Heck, M. E., Wannet, W. J., Imberechts, H., & Vos, P. (2008). Evaluation of the Premi Test Salmonella, a commercial low-density DNA microarray system intended for routine identification and typing of Salmonella enterica. *Int J Food Microbiol, 123*(3), 293-298.

Summary

Samenvatting



SUMMARY

Salmonella in pork causes annually between 5,000 and 8,000 human cases of salmonellosis in the Netherlands. The slaughter process plays an important role in the contamination of pork. *Salmonella* may be released in the slaughter chain by the incoming pigs carrying the pathogen on their skin or in feces in their intestinal tract. Alternatively, *Salmonella* may persist in the slaughterhouse environment as house flora, and be released to carcasses passing through the contaminated environment during processing.

Current strategies for solving microbial contamination problems in pig slaughterhouses typically consist of thoroughly cleaning and desinfecting all potentially suspect equipment, checking all Critical Control Points (CCP) and (re-)instructing the plant workers. Prioritizing measures is based on the nature of the contamination problem, the contamination history of the specific slaughterhouse and on the experience and commitment of the quality manager in that slaughterhouse. Typically, most attention is paid to preventing the contamination of carcasses instead of to finding the source of contamination. In practice, solving the contamination problem is often a time-consuming and inefficient process. The aim of this study was to develop a more systematic approach to preventing and tracing of Salmonella in pig slaughterhouses by the use of a model which combines historic and current data, expert knowledge of quality managers and other experts. This model was called a *biotracing* model.

Chapter 1 describes the development of the concept of *biotracing* and places it in the context of existing methods for microbial food safety, such as outbreak investigations, risk assessments (often using Monte Carlo methods) and source attribution. In particular, the importance of Bayes' theorem for inferential analysis was discussed and the use of Bayesian belief networks (BBNs) for constructing biotracing models was advocated.

Chapter 2 provides a more formal comparison between BBNs and Monte Carlo (MC) simulation models. Until now, the latter class of models is more currently used in microbial food safety to describe the dynamics of a pathogen in a food chain, in particular for microbial risk assessment. It was seen that BBNs have the advantage that downstream data can be incorporated in the model so that the uncertain values of the model parameters can be learned through Bayesian inference, and that the Bayesian mechanism of "reasoning backwards" allowed biotracing upstream sources of contamination. How this worked in practice was shown by a small example of a biotracing model.

In **Chapter 3**, a practical framework was proposed to construct a biotracing model. In seven successive steps, an MC model was converted to a BBN that can be used for biotracing. The proposed framework was illustrated with a BBN of *Salmonella* in the pork slaughter chain. The specific model equations, based on a previously published MC model, for describing this chain were given, but the microbial data that can be used for updating the model parameters were not yet shown in this chapter.

Chapter 4 describes the sampling experiment in a pork slaughterhouse in the Netherlands, from which data of *Salmonella* were obtained. Microbial samples were taken from the skin and from the rectum of pigs directly after exsanguinations, from the exterior and interior of the same carcasses after meat inspection at the re-work station and from the carcass splitter and the belly opener before and after each sampling day. These samples were analyzed to obtain prevalence, concentration and serotyping

data. The study showed a high Salmonella prevalence before slaughter, which had clearly declined at the re-work station. Seven different serovars were identified in the study, of which S. Derby and S. Typhimurium were the most prominent. Moreover, an endemic 'house flora' was present (S. Rissen).

Chapter 5 aggregates the results of the previous two Chapters. The BBN (developed in Chapter 3) was used as the *core model* in which the prevalence and concentration data of *Salmonella* were used as inputs for this model. The developed system was shown to allow an efficient tracing of the sources of contamination of individual pig carcasses that were found *Salmonella*-positive after slaughter and showed that resident *Salmonella* on the carcass splitter (house flora) was an important source of contamination for many *Salmonella*-positive carcasses, especially for those where *Salmonella* was found on the carcass interior (cutting section). The relevance of the different sources for carcass contamination varied over the (3 months) sampling period, with an increased relevance of house flora for exterior contamination of carcasses in June. Sensitivity analysis of the model showed that the biotracing results of the model were sensitive to the true values of the parameters describing the transfer of Salmonella from the equipment, which could only be estimated with limited accuracy from literature.

More realistic estimates of the cross-contamination rates were obtained by an additional laboratory experiment (**Chapter 6**). A novel Bayesian Network model was proposed which allows the combination of uncertainty within one experiment and variability over multiple experiments. The mean fraction of Salmonella that was transferred from contaminated pork to a knife during cutting was 0.19. The mean fraction of Salmonella that was transferred from a contaminated knife to pork during cutting was 0.58. Yet, the confidence intervals around these fractions, caused by uncertainty as a result of limited counting and by inherent variability of the transfer rates over different experiments, were large.

In **Chapter 7**, the results of the biotracing model (i.e. the BBN) were compared to the results of an epidemiological approach in which the *Salmonella* serotyping data as reported in Chapter 4 were used to find the most likely sources of *Salmonella*-positive carcasses after slaughter. The results of both approaches were similar for most Salmonella-positive carcasses at the re-work station. It was concluded by both approaches that house flora from one of the two carcass splitters was a relevant source of exterior contamination for carcasses slaughtered in June, and of interior contamination for carcasses slaughtered during the whole studied period. Nevertheless, some differences were observed, in particular the role of house flora was estimated to be more important when using the biotracing model instead of the serotyping approach.

This thesis describes the concept of biotracing and provides a theoretical and practical framework for the construction of a biotracing system. The model described in this thesis is a first attempt to describe a biotracing system of *Salmonella* in the pork slaughter chain and provides a tool to real-time monitor the process of contamination and trace sources of contamination. The results of the model will guide the discussion whether the development of such a system is -strategically- a good alternative to existing methods to trace back anomalies in hygienic procedures of the slaughterhouse. The model is a step in the process towards an operational biotracing system in which a stakeholder can initiate immediate responses to hazards imposed in the pork slaughterhouse, and can be seen as proof-of-principle of the functionality of a biotracing system in a food processing plant in general.

SAMENVATTING

Salmonella in varkensvlees veroorzaakt jaarlijks tussen de 5.000 en 8.000 gevallen van salmonellose in Nederland. Het slachtproces speelt een belangrijke rol in de besmetting van varkensvlees. *Salmonella* kan terechtkomen in de slachtlijn door de binnenkomende varkens die de bacterie op hun huid of in de ontlasting in hun darmkanaal hebben. Verschillende studies hebben ook uitgewezen dat *Salmonella* kan overleven in het slachthuis als huisflora, en varkens karkassen kan besmetten tijdens het slachtproces.

Huidige strategieën voor het oplossen van problemen van microbiële verontreiniging in varkensslachterijen bestaan meestal uit het grondig reinigen en het ontsmetten van alle potentieel verdachte apparatuur, het controleren van alle Critical Control Points (CCP) en (her-) instrueren van de werknemers in het slachthuis. Prioritering van maatregelen is in het algemeen gebaseerd op de aard van de besmetting, de besmettingsgeschiedenis van het specifieke slachthuis en op de ervaring en inzet van de kwaliteitsmanager in dat slachthuis. Gewoonlijk wordt de meeste aandacht besteed aan het voorkomen van de besmetting van karkassen in plaats van aan het vinden van de bron van besmetting. In de praktijk is het oplossen van microbiële besmettingsproblemen vaak een tijdrovend en inefficiënt proces.

Het doel van deze studie is om een meer systematische aanpak te ontwikkelen ter voorkoming van en voor de opsporing van *Salmonella* in varkensslachterijen door het gebruik van een model dat historische en actuele gegevens combineert met specialistische kennis van de kwaliteitsmanagers en andere deskundigen. Het ontwikkelde model wordt in dit proefschrift een “biotracing model” genoemd.

Hoofdstuk 1 beschrijft de ontwikkeling van het concept van “biotraceren” en plaatst het in de context van bestaande methoden voor microbiële voedselveiligheid, zoals het opsporen van uitbraken, microbiële risicoschattingen (die vaak gedaan worden met behulp van Monte Carlo methoden) en bronattributie. In het bijzonder wordt het belang van de stelling van Bayes besproken en het gebruik van Bayesiaanse netwerken (BBNs) bepleit voor het construeren van modellen voor biotraceren.

Hoofdstuk 2 geeft een meer formele vergelijking tussen BBNs en Monte Carlo (MC) simulatie modellen. Die laatste klasse van modellen is een tot nu toe vaak gebruikte methode in microbiële voedselveiligheid om de dynamiek van een ziekteverwekker te beschrijven in een voedselketen, met name voor microbiële risicoschattingen. BBNs hebben echter het voordeel dat gegevens, die stroomafwaarts in de voedselketen verzameld zijn, kunnen worden gebruikt in het model om de onzekere waarden van de modelparameters te leren door middel van het consequent toepassen van de stelling van Bayes in een ketenmodel. Bovendien kunnen door het Bayesiaanse mechanisme van “achteruit redeneren” microbiële besmettingsbronnen achterhaald worden. Hoe dit werkt in de praktijk is in dit hoofdstuk aangetoond door een klein praktisch voorbeeld van een biotracing model.

In hoofdstuk 3 wordt een praktisch kader geschetst om een biotracing model te construeren. In zeven opeenvolgende stappen wordt een MC model omgezet in BBN dat kan worden gebruikt voor biotracing. Dit kader wordt geïllustreerd met een BBN van *Salmonella* in de varkens slachtketen. De specifieke modelvergelijkingen, gebaseerd op een eerder gepubliceerd MC model dat deze keten beschrijft, worden gegeven maar in dit hoofdstuk is nog geen actuele Nederlandse microbiële data gebruikt om

het model te preciseren.

Hoofdstuk 4 beschrijft een bemonstering experiment in een varkensslachterij in Nederland, waardoor microbiële data van *Salmonella* werd verkregen. Microbiële monsters zijn genomen van de huid en van het rectum van varkens direct na verbloeding, van de huid en van het snijvlak van dezelfde karkassen na de vleeskeuring op het keurbordes. Ook zijn monsters genomen van de karkassplitter en de buikopener voor en na elke bemonsteringsdag. Deze monsters zijn geanalyseerd om prevalentie-, concentratie- en serotyperingsdata te verkrijgen. Uit het onderzoek blijkt een hoge prevalentie van *Salmonella* voorafgaand aan de slacht, die duidelijk is gedaald op het keurbordes. Zeven verschillende *Salmonella* serotypes zijn geïdentificeerd in de studie, waarvan *S.* Derby en *S.* Typhimurium het meest prominent. Bovendien zijn er sterke aanwijzingen voor het bestaan van een endemische 'huisflora' stam (*S.* Rissen).

Hoofdstuk 5 combineert de resultaten van de vorige twee hoofdstukken. De prevalentie- en de concentratiedata van *Salmonella* worden gebruikt als input voor het BBN (ontwikkeld in hoofdstuk 3). Het ontwikkelde systeem toont aan dat een efficiënte tracering mogelijk is van de besmettingsbronnen van individuele *Salmonella*-positieve varkens op het keurbordes. Persistente *Salmonella* op een van de twee karkassplitters (huisflora) is een belangrijke bron van besmetting voor veel *Salmonella*-positieve karkassen, in het bijzonder voor karkassen waarin *Salmonella* wordt gevonden op het snijvlak. De relevantie van de verschillende besmettingsbronnen voor het karkas is variabel gedurende de (3 maanden) bemonsteringsperiode, met een toegenomen relevantie van huisflora voor besmetting van de buitenkant van karkassen in juni. Uit sensitiviteitsanalyse van het model blijkt dat de resultaten van het model gevoelig zijn voor de waarden van de kruisbesmettings-parameters. Deze konden slechts met een beperkte nauwkeurigheid worden geschat uit de literatuur.

Meer realistische schattingen van de kruisbesmetting-parameters zijn verkregen door een extra laboratorium experiment (hoofdstuk 6). Dit hoofdstuk beschrijft een model die het mogelijk maakt onzekerheid over de waarde van een kruisbesmetting-parameter (als gevolg van een beperkt aantal metingen binnen een experiment) binnen een experiment te combineren met de inherente variabiliteit van die parameter over meerdere experimenten. De gemiddelde fractie van *Salmonella* die wordt overgedragen van besmet varkensvlees naar een mes tijdens het snijden is 0,19. De gemiddelde fractie van *Salmonella* die wordt overgedragen van een verontreinigd mes naar varkensvlees tijdens het snijden is 0,58. Maar de betrouwbaarheidsintervallen van deze fracties, veroorzaakt door onzekerheid en variabiliteit, zijn relatief groot.

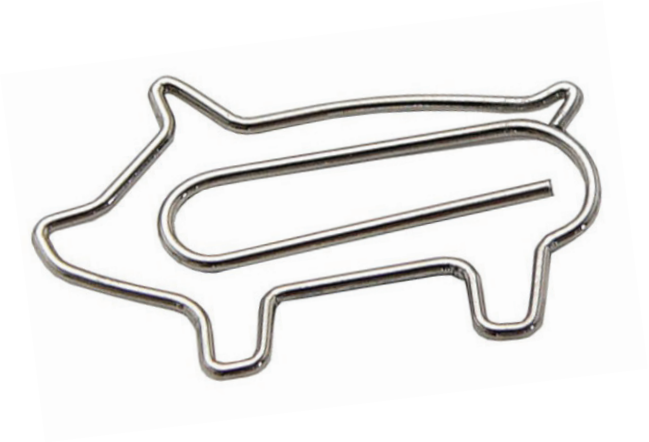
In hoofdstuk 7 worden de resultaten van het biotracing model vergeleken met die van een epidemiologische benadering waarin de *Salmonella* serotypering-data, zoals gerapporteerd in hoofdstuk 4, worden gebruik. De resultaten van beide modellen zijn vergelijkbaar voor de meeste *Salmonella*-positieve karkassen op het keurbordes. Beide modellen vinden dat huisflora afkomstig van één van de twee karkassplitters was een relevante besmettingsbron voor de buitenkant van karkassen geslacht in juni, en voor het snijvlak van karkassen in de gehele onderzochte periode. Toch zijn er een aantal discrepanties tussen de uitkomsten van beide modellen. Vooral de rol van huisflora wordt als belangrijker ingeschat door het biotracing model dan door het serotyperingsmodel.

Dit proefschrift beschrijft het concept van biotraceren en biedt een theoretisch en praktisch kader voor de bouw van een biotracing systeem. Het model beschreven in dit proefschrift is een eerste poging om een biotracing systeem voor *Salmonella* in de varkens slachtketen te beschrijven en biedt een instrument om het proces van besmetting real-time te monitoren en besmettingsbronnen op te sporen. De resultaten van het model kunnen gebruikt worden voor verder onderzoek om te bepalen of de ontwikkeling van een dergelijk systeem een goed alternatief is voor de bestaande methoden om afwijkingen in hygiënische procedures van het slachthuis te achterhalen. Het model is hierin een eerste stap in het proces naar een operationeel biotracing systeem in de varkensslachterij. De resultaten kunnen gezien worden als een proof-of-principle voor de functionaliteit van een biotracing systeem in een industrieel voedselproductieproces in het algemeen.

Curriculum Vitae

List of Publications

Dankwoord



CURRICULUM VITAE

Joost Smid was born December 8th 1979 in Noordscheschut and attended primary and secondary school in Zuidwolde and Hoogeveen, respectively. In the 5th class of secondary school Joost started a ten months home stay in Japan with a Japanese host family, where he attended classes in Japanese at Nara prefectural Ikoma high school. After having come back, Joost finished secondary school (VWO). Upon graduation, the taste of experiencing new countries seduced Joost to ten months of traveling in Asia together with a friend. In 2000, Joost began his study of Mathematics at the University of Groningen. Mathematics was chosen because of its broad applicability in many subjects. Six months of the Bachelor's phase were done at the Université Joseph Fourier in Grenoble. Joost specialized in Systems and Control theory (Applied Mathematics) for the Master's phase and performed a ten months research project for his Master's thesis at Philips Research in Eindhoven, where he studied processing theory of digital sound signals. He graduated in 2006 in the group of Prof. H. Trentelman.

In February 2007 Joost was employed at the National Institute for Public Health and the Environment (RIVM) in Bilthoven for a European Union funded project 'BIOTRACER'. He developed a Bayesian network model for 'biotracing' the source of *Salmonella* contamination in the pork slaughter chain, research which is described in this thesis. Joost was employed by the Institute for Risk Assessment Sciences (IRAS) in 2009, which is an interfacultary research institute within the faculties of Veterinary Medicine, Medicine, and Sciences of Utrecht University. Next to the BIOTRACER work, Joost worked on source attribution models for *Campylobacter* and spent a two months internship at the Hopkirk Research Institute of Massey University, New Zealand for this project. In 2012, Joost remains working on the practical application of biotracing models in pig slaughterhouses and on the development of a cost-benefit analysis to prioritize preventive measures in pig slaughterhouses, and he continues working on source attribution of *Campylobacter.*

LIST OF PUBLICATIONS

Published

An introduction to biotracing in food chain systems
Barker GC, Gomez-Tomé N, **Smid JH**
Trends in Food Science and Technology 2009; 20(5): 220-226

Strengths and weaknesses of Monte Carlo simulation models and Bayesian belief networks in microbial risk assessment
Smid JH, Verloo D, Barker GC, Havelaar AH
International Journal of Food Microbiology 2010; 139: 57–63

A Practical Framework for the Construction of a Biotracing Model: Application to *Salmonella* in the Pork Slaughter Chain
Smid JH, Swart AN, Havelaar AH, Pielaat A
Risk Analysis 2011; 31(9): 1434-1450

A quantitative approach towards a better understanding of the dynamics of *Salmonella* spp. in a pork slaughter-line
Van Hoek AHAM, De Jonge R, Van Overbeek WM, Bouw E, Pielaat A, **Smid JH**, Malorny B, Junker E, Löfström C, Pedersen K, Aarts HJM, Heres L
International Journal of Food Microbiology 2012; 153: 45–52

A Biotracing Model of *Salmonella* in the Pork Production Chain
Smid JH, Heres L, Havelaar AH, Pielaat A
Journal of Food Protection 2012; 75(2): 270–280

Submitted

Variability and uncertainty analysis of the cross-contamination rates of *Salmonella* during pork cutting.
Smid JH, De Jonge R, Swart A, Havelaar AH, Pielaat A

Quantifying the sources of *Salmonella* contamination in a Dutch pig slaughter plant.
Smid JH, Van Hoek AHAM, Aarts HJM, Havelaar AH, Heres L, De Jonge R, Pielaat A

DANKWOORD

Zonder de hulp van een groot aantal mensen was dit boekje er niet gekomen. In de eerste plaats wil ik hier Arie Havelaar en Annemarie Pielaat noemen. Arie, je deur was altijd open. Weliswaar zat je niet altijd achter die deur vanwege je vele buitenlandreizen, maar als je er wel zat dan kon je (bijna) altijd even tijd vrij maken voor het bespreken van mijn werk. Je meedenken en kritische opmerkingen hebben dit onderzoek enorm vooruit geholpen. Ik ben erg blij met je vertrouwen en daaropvolgende inzet voor contract-verlenging na mijn aanrijding vorig jaar, zodat dit proefschrift er toch is gekomen.

Annemarie, bedankt voor het bij elkaar brengen van de vele *Salmonella*- en varkens-mensen in dit project en het aansturen op een gemeenschappelijke onderzoeksvraag tijdens onze BIOTRACER bijeenkomsten. Dit waren mooie reisjes die ik altijd met veel plezier gemaakt heb (vaak met de wandelschoenen mee en een iets uitgestelde terugvlucht). Verder bedankt voor je uitgebreide commentaar op premature versies van mijn manuscripten en voor het uitlenen van jullie huis afgelopen zomer.

Iedereen met wie ik in BIOTRACER heb samengewerkt, bedankt. In het bijzonder Judith, Angela, Henk, Rob, Lourens en Gary. Judith, je antwoord op een kort technisch vraagje, over modelleren of over microbiologie, was bijna altijd lang en gedetailleerd. Ik heb hierdoor veel van je geleerd. Angela, bedankt voor de analyses van de monsters, de noodzakelijke inputs voor mijn modellen en de vele overzichten van de data. Rob heeft me voor het eerst (en totnogtoe laatst) een bacterie onder de microscoop laten zien, zodat ik weet waar ik eigenlijk mee bezig ben. Bedankt voor je enthousiaste suggesties voor verdere lab-experimenten. Alice, Ellen en Christiaan, bedankt voor het uitvoeren hiervan. Lourens heeft me zijn fascinerende werkomgeving laten zien, super dat we daar monsters konden nemen. Ik zal het rondhollende varken door het slachthuis niet snel vergeten.

Special thanks to Gary Barker, with whom I spent three weeks at IFR to learn tips and tricks in Bayesian Network modeling. Thanks for arranging a bicycle from one of your colleagues for this strange Dutch guy who wanted to cycle in your beautiful Norfolk countryside after one of these weeks. I enjoyed it. Another period during this project was spent abroad in New-Zealand, with the group of Nigel French. Although the research presented in this thesis does, eventually, not cover the work on *Campylobacter* studied with Nigel, I thank Nigel for his kind hospitality. You have chosen a splendid country.

Verder was er nog de begeleidingscommissie als (bij)sturing van dit project. Frans van Knapen, bedankt voor je vertrouwen en je begeleiding. Ook de resterende leden van de begeleidingscommissie (Wilfrid, Arno en Henk): bedankt. Arno, je initiële scepsis over Bayesiaanse technieken (prior *belief* in wiskundige vergelijkingen?) is goed geweest voor het gedachtenproces, bedankt voor het wiskundig sparren.

Dit promotie-onderzoek heb ik uitgevoerd op het Laboratorium voor Zoonosen en Omgevingsmicrobiologie (LZO) van het RIVM. Alle collega's van het LZO, bedankt voor de gezelligheid. In het bijzonder mijn kamergenoten (Arno, Arnout, Hetty), ex-kamergenoot Jasper en de "wandelclub". Even tussen de middag een rondje lopen in het Houdringe-bos (in weer en wind) was, als break van een dag computerwerk, erg prettig.

Afleiding en gezelligheid gaven me de afgelopen jaren nieuwe energie voor het promotiewerk. Die afleiding en gezelligheid was vaak afkomstig van vrienden en die wil ik daarvoor dan ook bedanken. Yvo, Annieka, Renske, Maarten, Leendert en Sjula: heel wat leuke klim-, fiets- en ski-vakanties hebben we gemaakt de afgelopen jaren. Op naar nieuwe avonturen! Ook oudere vrienden met wie ik contact heb gehouden, van de GSAC, de middelbare school tot en met de basisschool aan toe, bedankt. Ruud, Andalusië smaakte naar meer, wanneer trekken we de wandelschoenen weer aan?

Zonder mijn moeder zou dit proefschrift er in elk geval niet geweest zijn: zij zag de vacature toen ik nog op reis was. Lieve pap en mam en rest van de familie, bedankt voor jullie interesse en aanmoedigingen. Erik en Annemarie, ik ben erg vereerd dat jullie mijn paranimf willen zijn. Jorieke, dank voor je hulp bij het maken van de foto's. Nihon no okaasan to otousan, itsumo doumo arigatou gozaimashita.

Nog één dank voor mijn vriendinnetje. Lieve Heleen, het afgelopen jaar, na mijn aanrijding, was niet het gemakkelijkste jaar, zowel voor mij als voor jou. Je steun, warmte en liefde hebben me er doorheen geholpen met als resultaat dit proefschrift. Laat het fiets-avontuur beginnen dit najaar!