**The selected paper is about the application of Bayesian Network for Food Safety Risk in Cattle Slaughtering Industry**

1. Introduction

Nowadays, food safety has an important role for consumers. Fulfillment of the aspect of food safety is also encouraged to reduce consumer concerns about food safety, especially from the hazards of pesticide residues, preservatives, and hormones in livestock. The survey results in China show that consumer concern for food safety is indicated by buying food in supermarkets, expiration dates, food color and nutritional content.

The inconsistency of the food industry in implementing food safety has a negative impact on consumers. Foods that are processed without regard to food safety aspects has negative impact on human health in the short or long term. This negative impact arises because of the contamination of physics, biology (pathogens) or chemistry. Some phenomena due to food that do not meet food safety aspects have occurred in various countries. In the UK, an estimated 500,000 congenital diseases are caused by pathogens. In South Korea, it is estimated that every year 9.59 million people experience, 1.56 million are diagnosed and 140,000 people are hospitalized due to foodborne illness.

according to WHO, there are 600 million cases and 420000 deaths caused by foodborne diseases, so training activities, government, and private cooperation are needed to support the implementation of food safety.

This phenomenon shows the risk of changing safe status into unsafe. To keep the risk can be eliminated or eliminated, it is necessary to identify risky activities in the food manufacturing process.

This study aims to:

1. identify the risk event of food safety.
2. Measure the opportunities for these risks with Bayesian Network.

Therefore, this research contributes to the development of the use of the BN method to determine the risk of food contamination based on aspects of food safety.

1. Methodology

This research was conducted in three phase, namely observation, data collection, risk statistical analysis.

1. Observation

Observations were carried out at one of the Animal Slaughter Houses in Indonesia. This activity aims to observe the process of slaughtering animals (cattle) so that the risk event for food safety can be identified. In addition to observations, observations were also made through interviews with animal slaughterers.

2. Data collection

Data was collected using the checklist form which was arranged based on the observation stage. The code used on the form is Y (yes) and N (no). Code Y is be used if the observed activities identified are at risk of food safety. While the N code is used for activities that are not at risk of being contaminated with food safety.

3. Risk statistical analysis

At this stage, data is processed by one of the statistical methods, namely the Bayesian Network (BN). Bayesian network is done through steps:(a) compiling BN structure. (b) compiling BN node classification. (c) mapping BN causal relationship. (d) preparing conditional probability table (CPT)

1. Stages

Stages 1) Identification of Risk Event: This stage aims to identify risk events in cattle slaughtering industry which can lead to food safety contamination.

These risk events are nodes of our BN.

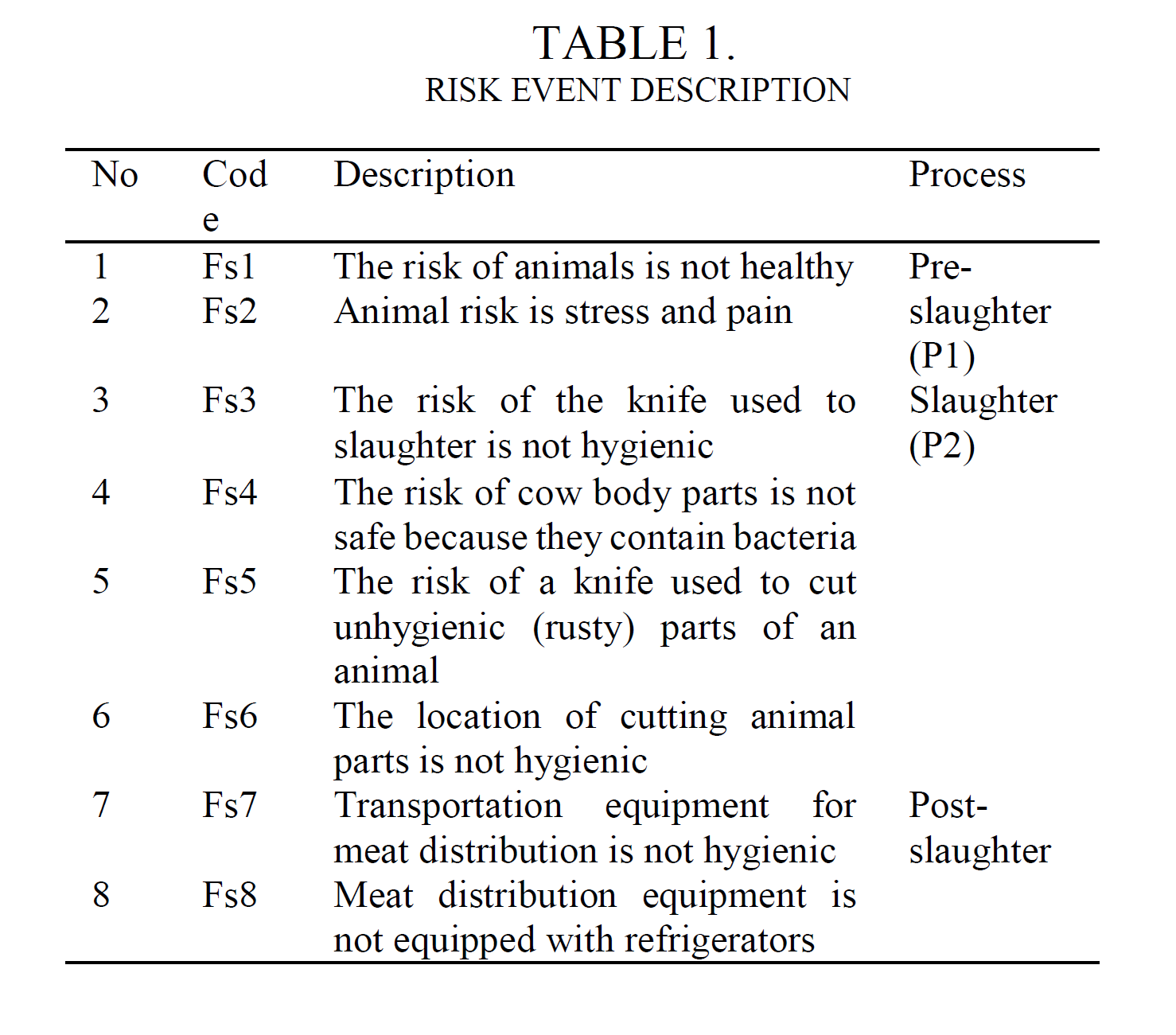


Table 1 shows that there are 8 risk events that are at risk for food safety. Furthermore, based on the results of risk identification, observations were made on the process of slaughtering animals. This observation aims to find out activities that are risky and not risky.

Stage 2) Determination of Probability of Risk Event: This stage aims to determine the opportunities for food safety contamination for each risk event that has been identified.

Calculation of opportunities is done by observing. Observations were carried out on 50 cows. Y indicates the risk event is potentially contaminated, while N is not contaminated.

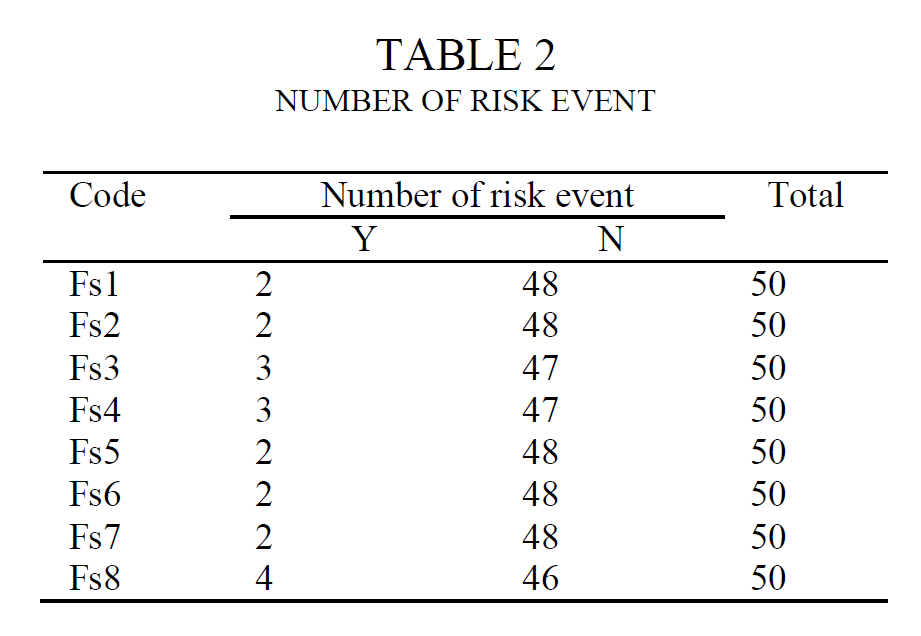
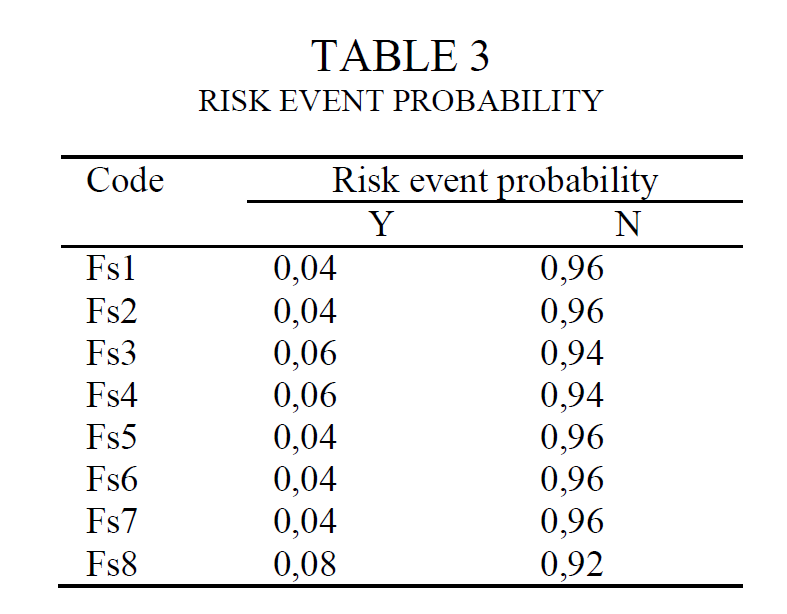


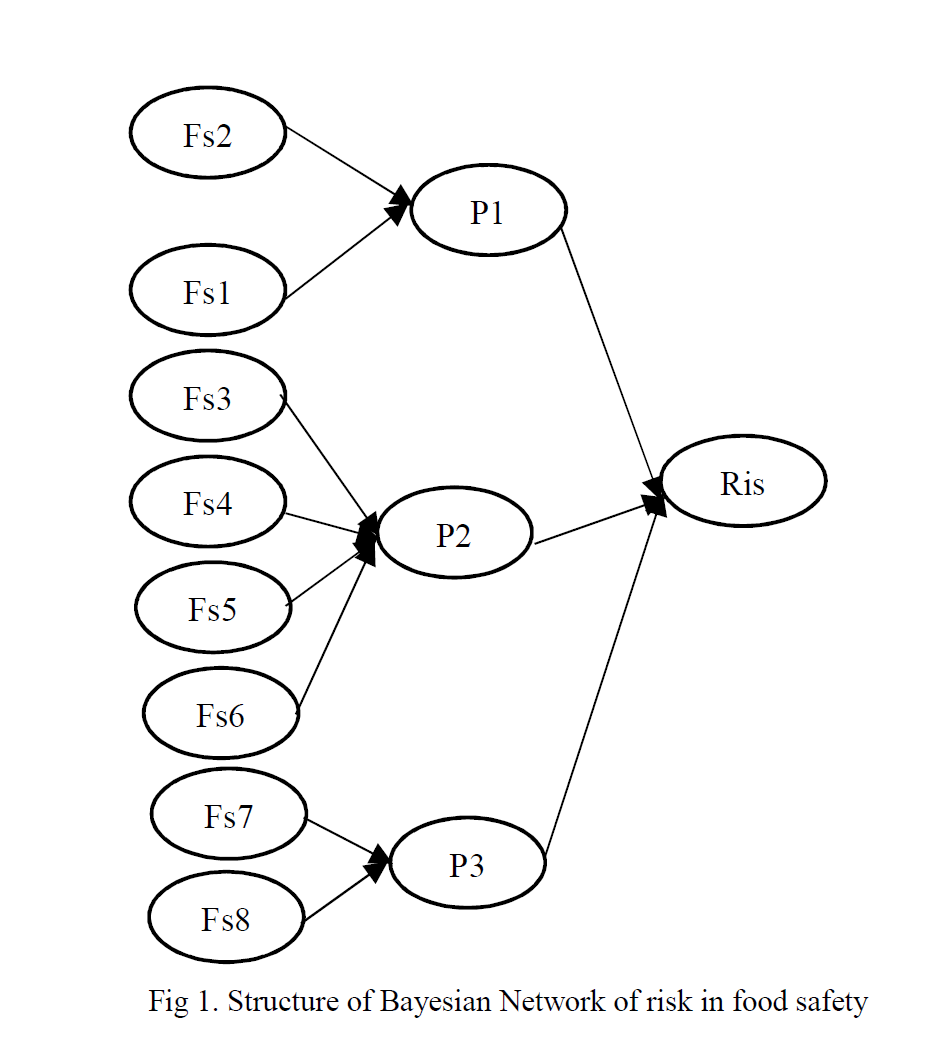
Table 2 is a risk recapitulation for each risk event that has been identified. Based on these results, the probability of each risk event is calculated.

Here we can see the prior probabilities of the root nodes:



Stage 3) Development of Bayesian network structure: This stage aims to develop the Bayesian Network structure. Furthermore, the structure will be used to show the relationship between risk events that occur in the cattle slaughtering industry.

This is the structure of our Bayesian network:

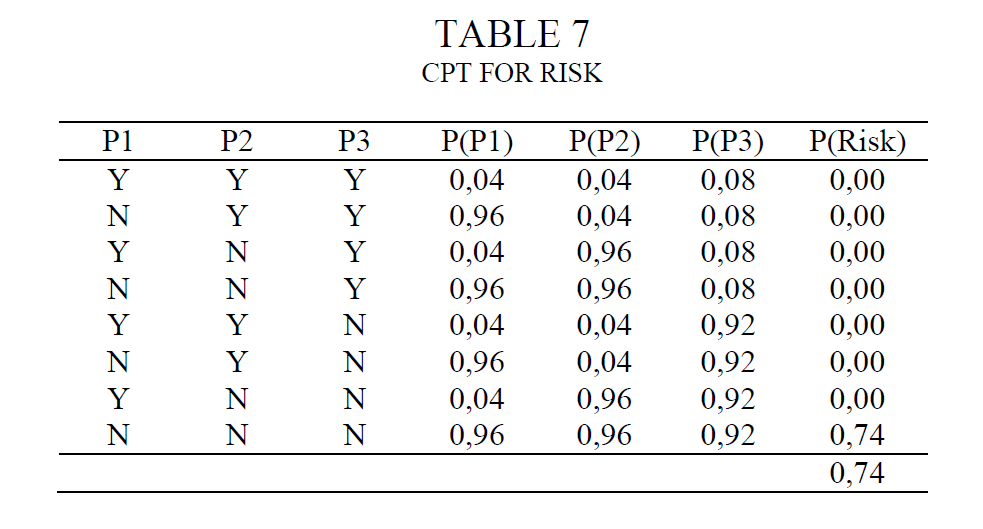
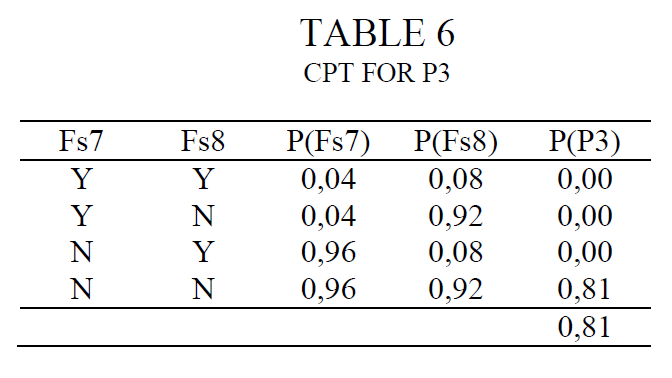
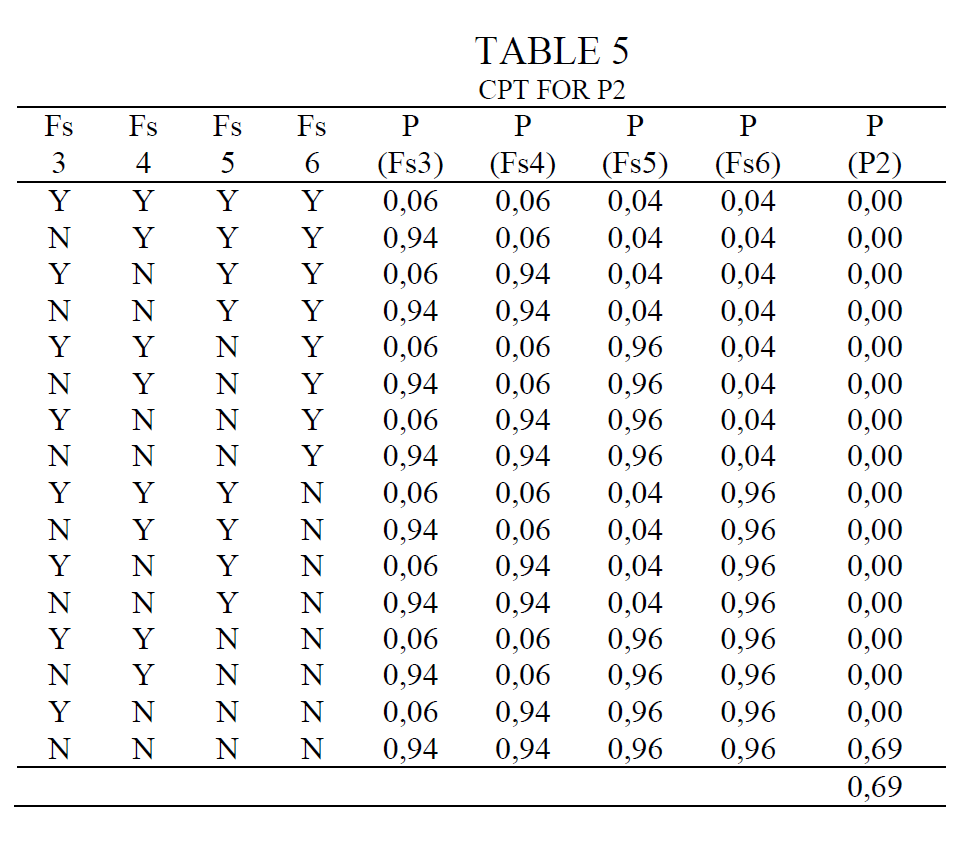
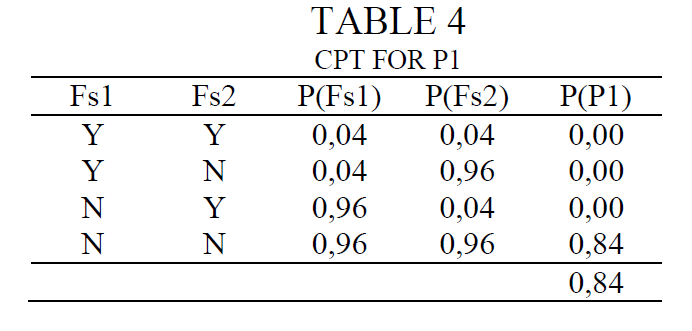


Stage 4) Calculation for Condition Probability Table (CPT):

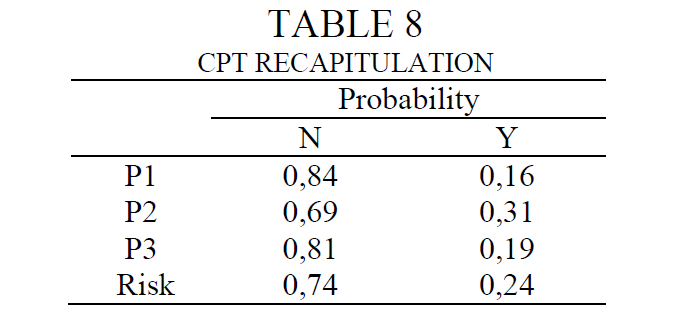
1. Risk event probability is used as the basis for calculations in the Bayesian Network. As a first step in the Bayesian Network, a Bayesian Network structure should be arranged. The structure of Bayesian network illustrates the relationship between food safety risks that have been identified. The Bayesian Network structure is used as the basis for preparing the CPT.

In this study, there are 4 CPTs, namely P1, P2, P3 and Risk node

Here is an example of CPT for node p1 which has 2 parent nodes:



1. Table 8 shows the results of CPT calculations for P1, P2, P3, and Risk. Based on the table, it is identified that the probability of food safety risks in the process of providing in cattle slaughtering industry is 24%. The probability shows that the opportunity for food safety contamination is 24%.



Stage 5)Analysis of Food Safety Risk in Cattle Slaughtering Industry: This stage is carried out to analyze the probability of the occurrence of food safety risks with the Bayesian Network. At this stage, a risk event ranking will be conducted as a basis for formulating preventive actions. For this reason, so that the probability of contamination can be eliminated, strategic action is needed. The identification of strategic actions that need to be done is based on the risk event that has the highest chance. From the process side, the priority for eliminating contamination is based on the risk priority on the CPT for the P value.

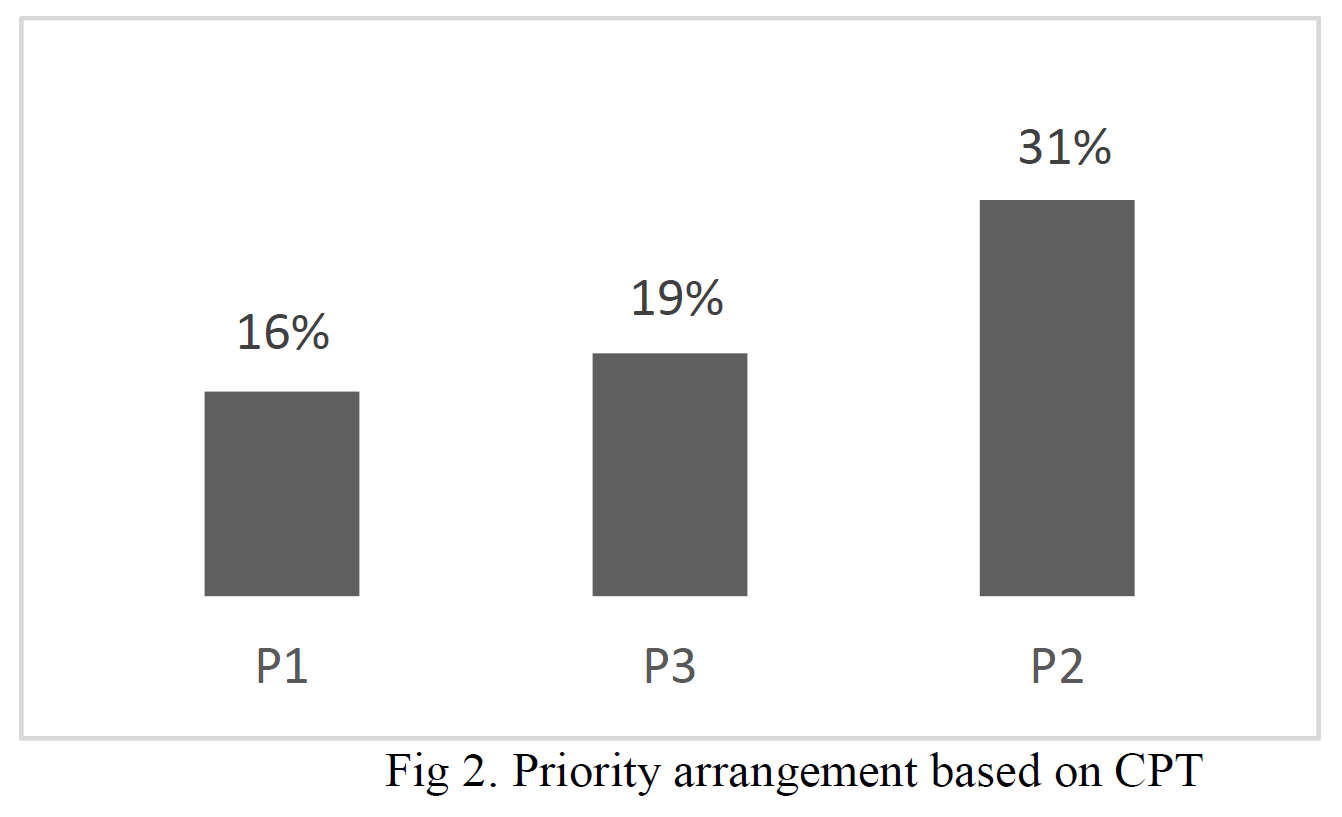
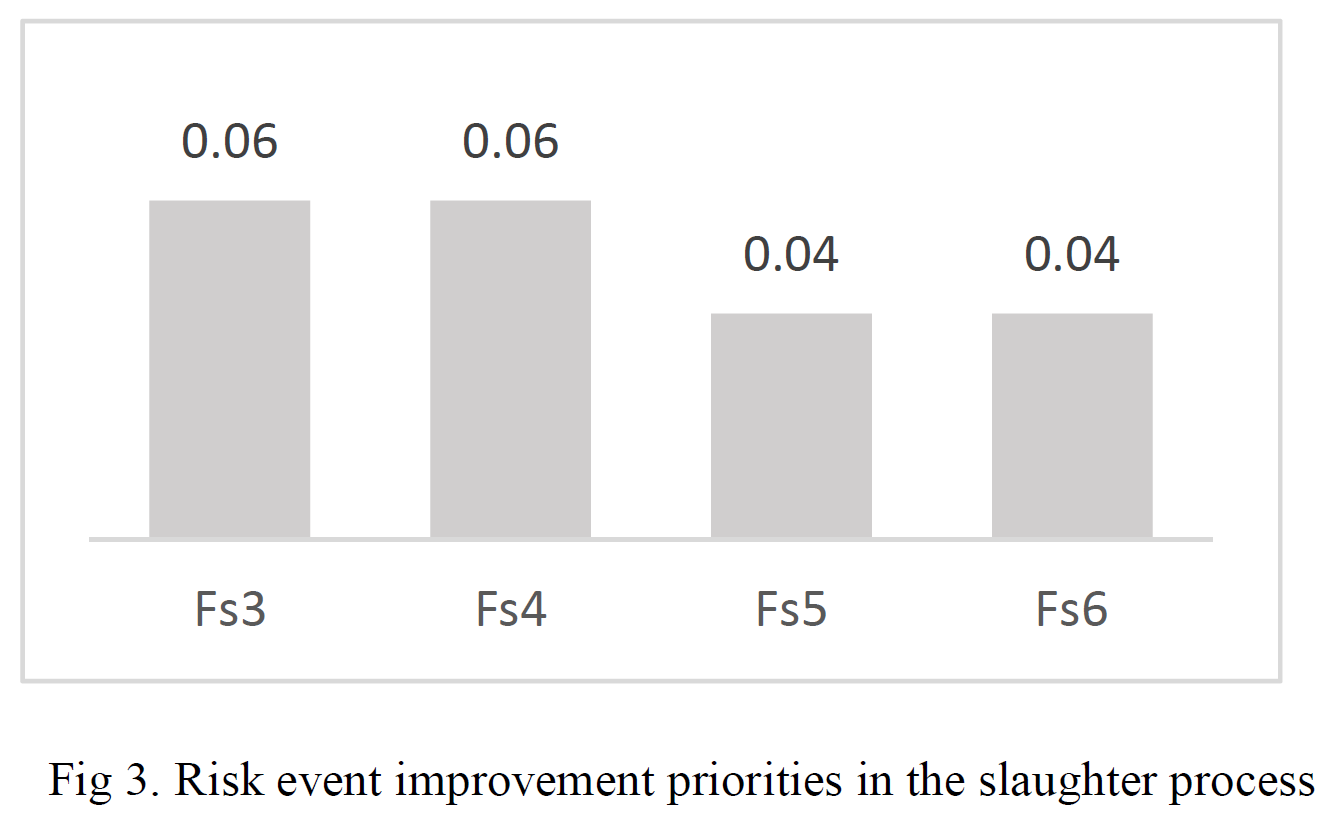


Figure 2 shows that the top priority of the process that needs to be corrected immediately to eliminate contamination is P2 (slaughtering process). This process has the highest risk of contamination with food safety. For this reason, the improvement process is focused on improving the slaughter process. In more detail, the slaughter process has 4 risk events that need attention (Fs3, Fs4, Fs5, Fs6).



Based on Figure 3, the priority of improvements to reduce the risk of contaminated food safety is at the risk event Fs3 (not sharp knife) and Fs4 (risk of cow body parts is not safe). For this reason, several corrective steps that need to be taken immediately are: 1. Providing equipment for sharpening knives so that the sharpness of the knife is maintained. 2. It provides a replacement knife (backup knife) that is ready for use when needed. 3. Optimize post mortem checks on animals that have been slaughtered.

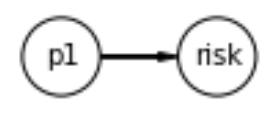
1. Flow of probabilistic influence

We investigate six types of probabilistic influence introduced in the class for our BN. Here is the representation of a sample of each type:

**4.1. Direct cause**

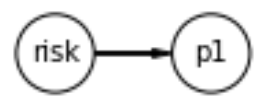
We can see from the network pre-slaughtering process (P1) directly affect on the risk in food safety (risk).

that means whenever some changes in pre-slaughtering process (P1) happens it will have effect on food safety (risk).



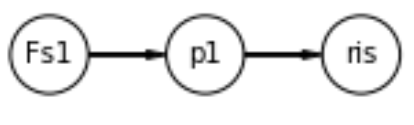
**4.2. Direct effect**

In the direct effect we can see that food safety (risk) can be an effect of pre-slaughtering process (P1).



**4.3. Causal trail**

any changes in the The risk of animals is not healthy (Fs1), will have effect on pre-slaughtering process (p1) and it will change the values in the Risk of food safety (risk). It means Fs1 indirectly affect on the risk. Now p1 is observed i.e., we know its value, any change in Fs1 won’t affect risk.



**4.4. Evidential trail**

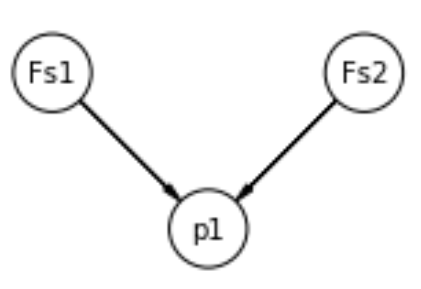
If we investigate the example of causal trail in reverse order, we get evidential trail which says evidence (risk) can influence (Fs1) via (P1) only if (P1) is unobserved. Observed (P1) blocks influence.

**4.5. Common cause**

For this network there is not common cause. Because there is not any node which have 2 children.

**4.6. Common effect**

This is a v-structure. When pre-slaughtering process (p1) is not observed, any change in (Fs1) reflects some changes in p1 but not in Fs2. But when p1 is observed, if Fs1 is also observed, this will change the probability of occurring Fs2.



1. Independency

**5.1. Conditional Independence**

A variable is conditionally independent of non-descendant variables given its parents.

As an example, we considered independencies of (p1). The result is as follows:

(p1 ⟂ p3, Fs8, Fs5, p2, Fs4, Fs7, Fs6, Fs3 | Fs2, Fs1)

It means : “pre-slaughter”(p1) ⟂ “post-slaughter”(p3), “meat distribution equipment is not equipped with refrigerators” (Fs8) , “the risk of a knife used to cut unhygienic parts of an animal” (Fs5), “slaughter”(p2) , “The risk of cow body parts is not safe because they contain bacteria”(Fs4), “Transportation equipment for meat distribution is not hygienic” (Fs7), “The location of cutting animal parts is not hygienic” (Fs6), “The risk of the knife used to slaughter is not hygienicko” (Fs3) given ( | ) “Animal risk is stress and pain” (Fs2) , “The risk of animals is not healthy” (Fs1).

Sometimes it does not have parents. --🡪 last exam of code

* 1. **Markov blanket**

A variable is conditionally independent of all other variables given its Markov blanket which is its parents, children, and children’s parents. Markov blanket nodes of “The risk of animals is not healthy” (Fs1) is as follows:

['Fs2', 'p1']

**5.3. Active trails**

When influence can flow from X to Y via Z then trail X—Z—Y is active. We investigated active trails for all the nodes in the network. Node “The risk of animals is not healthy” (Fs1) has the least active trail with 3 other nodes and the following nodes are those which have active trail to all the nodes in the network:

{'Fs1': {'p1', 'Fs1', 'ris'}}

* 1. **Direct separation**

Two nodes are d-separated if there is no active trail between them. According to this definition, nodes (Fs8) has active trail to (P1)

1. Inference

6.1. **Exact inference**

There are some algorithms which can be used for exact inference. These algorithms find the exact probability values for our queries. In this work we used variable elimination for inference.

**Inference by variable elimination:** Variable elimination is a dynamic programming algorithm that re-uses computation. For querying we select the risk factor with highest incidents frequency which is