ARTICLE IN PRESS

Journal of Safety Research xxx (xxxx) xxx

Contents lists available at ScienceDirect

Journal of Safety Research

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

A causal Bayesian network approach for consumer product safety and risk assessment

Joshua L. Hunte [a,*], Martin Neil [a,b], Norman E. Fenton [a,b]

[a] Risk and Information Management Research Group, School of Electronic Engineering and Computer Science, Queen Mary University of London, London E1 4NS, UK
[b] Agena Ltd, Cambridge, UK

ARTICLE INFO

Article history:
Received 29 January 2021
Received in revised form 19 July 2021
Accepted 5 December 2021
Available online xxxx

Keywords:
Bayesian network
Product risk
Risk assessment
RAPEX
Product safety

ABSTRACT

Introduction: Product risk assessment is the overall process of determining whether a product is judged safe for consumers to use. Among several methods for product risk assessment, RAPEX is the primary one used by regulators in the UK and EU. Despite its widespread use we identify several limitations of RAPEX, including a limited approach to handling uncertainty, especially in the absence of data, and the inability to incorporate causal explanations for using and interpreting the data. *Method:* We develop a Bayesian Network (BN) model to provide an improved systematic method for product risk assessment that resolves the identified limitations with RAPEX. BNs are a rigorous, normative method for modelling uncertainty and causality which are already used for risk assessment in domains such as medicine and finance, as well as critical systems generally. *Results:* We use the BN approach to demonstrate risk assessments for products where relevant test and product instance data are and are not available. Whereas RAPEX can only produce results given relevant data, the BN approach produce results for products with and with no relevant data – replicating RAPEX in the former and providing deeper insights in both cases. *Conclusion:* The BN approach is powerful and flexible for systematic product risk assessment. While it can complement more traditional methods like RAPEX, it is able to provide quantified, auditable assessments in situations where such methods cannot because of lack of data. *Practical Applications:* Safety regulators, manufacturers, and risk professionals can use the BN approach for all types of consumer product risk assessment, including for novel products or products with little or no historical data. They can also use it to validate the results of existing methods when data becomes available. It informs risk management decisions and helps understand the effect of those decisions on the consumer risk perception.

1. Introduction

It is essential that the products we use in our homes are acceptably safe. To ensure our safety, national regulators perform product risk assessments to limit consumer harm (RPA, 2006; PROSAFE, 2013; European Commission, 2015, 2018). There are several different methods used for product risk assessment, including Nomograph (RPA, 2006) and Matrix (RPA, 2006), but RAPEX (RPA, 2006; European Commission, 2015, 2018) is the primary one used by regulators in the UK and EU. In this paper, we identify a number of limitations of RAPEX and explain the need for a systematic method for product risk assessment that: improves the management of uncertainty; uses causal knowledge of both the testing and operational environment and the process by which data are generated; is able to produce auditable quantified risk assessments even where there is limited product testing and instance data; considers the user population at risk and the product risk tolerability.

We propose that Bayesian Networks (BNs) can provide such a systematic method as they are a rigorous, normative method for modelling uncertainty and causality (Fenton & Neil, 2018; Maglogiannis et al., 2006; Masmoudi et al., 2019; Torres-Toledano & Sucar, 1998; Welch & Thelen, 2000). We present a generic BN that significantly extends the previous work on BNs for product risk assessment. It incorporates hazard and injury data, manufacturer process information, product usage data and consumer utility and risk perception to estimate product risk. The proposed generic BN also complements traditional risk assessment methods like RAPEX.

The rest of this article is organized as follows: Section 2 provides an overview of product risk assessment, identifying the limitations of RAPEX. Section 3 describes our method for building the BN model, including selecting variables, determining the BN structure, and populating node probability tables (NPTs). Section 4 pre-

* Corresponding author.
E-mail address: j.l.hunte@qmul.ac.uk (J.L. Hunte).

https://doi.org/10.1016/j.jsr.2021.12.003
0022-4375/

Please cite this article as: J.L. Hunte, M. Neil and N.E. Fenton, A causal Bayesian network approach for consumer product safety and risk assessment, Journal of Safety Research, https://doi.org/10.1016/j.jsr.2021.12.003

sents the results of two case studies. The results are discussed in Section 5. Finally, conclusions and recommendations for further work are presented in Section 6.

2. Background: Product risk assessment overview

A 'product' is any physical non-food item offered in a market to meet consumer needs; it could be anything from a kitchen appliance to a toy. Product risk assessment is the overall process of determining whether a product is safe for consumers to use. Specifically, it is the process by which the level of risk associated with a particular (product) hazard is identified and categorized. The risk assessment process includes risk analysis and risk evaluation (European Commission, 2015; ISO/IEC, 2014):

1 *Risk Analysis:* This phase involves hazard identification and risk estimation (ISO/IEC, 2014): (a) *Hazard identification:* The process of finding, recognizing, and describing product hazards (European Commission 2015, 2018; ISO/IEC 2014). Hazards are potential sources of harm or injury and are intrinsic to the product (European Commission 2015, 2018; ISO/IEC 2014). (b) *Risk Estimation:* The process of determining the risk level of the product (European Commission 2015, 2018; ISO/IEC 2014). Risk is the combination of the likelihood of a hazard causing injury to a consumer and the severity of that injury (European Commission, 2015;; ISO/IEC, 2014). The risk level is the degree of the product risk on a scale from 'low' to 'serious' (European Commission 2015, 2018).
2 *Risk Evaluation:* The process by which the outcome of the risk analysis is combined with policy considerations to characterize the risk and inform decisions on risk management (European Commission, 2015). It includes determining if the risk is acceptable (European Commission, 2015; ISO/IEC, 2014).

As RAPEX (RPA 2006; European Commission 2015, 2018) is the most widely used method for product risk assessment, this article will review the RAPEX method (RPA 2006; European Commission 2015, 2018) and its limitations.

2.1. RAPEX overview and limitations

RAPEX was developed for the rapid exchange of information between the Member States of the European Union on the measures and actions concerning products that pose a serious risk to the safety and health of consumers (European Commission 2015, 2018). An essential component of RAPEX is product risk assessment, which determines product risk and informs risk management response (European Commission 2015, 2018). The following steps or guidelines describe the method used by RAPEX for product risk assessment:

1. *Describe the product*: Product details such as name, brand and model are documented during this stage.
2. *Describe product hazards*: Hazards are identified by tests and standards or by the manufacturers' product labelling and instructions. Identified hazards are classified using RAPEX's hazard taxonomy (e.g., electrical energy, extreme temperatures and toxicity).
3. *Identify consumers at risk*: Consumer types include intended users, non-intended users and vulnerable users.
4. *Describe the injury scenario*: Injury scenarios that causally describes how the product hazard may harm the consumer via a series of steps are developed. For example, an axe breaks and the ejected part strikes the user's head.
5. *Determine the probability of injury*: Probabilities are assigned to each step of the injury scenario to determine the probability of injury. For example, to determine the probability of injury while using an axe, we combine the following probabilities:
   a. Probability of axe breaking = 1/100
   b. Probability of broken part hitting body = 1/10
   c. Probability of broken part hitting head = 1/10
   Total probability of injury = $0.01 \times 0.1 \times 0.1 = 0.0001$

Probabilities used in this step are assumed to be independent and are obtained from what are assumed to be reliable sources such as the European Injury Database and hospital injury databases.

6. *Determine the severity of the injury*: The severity of the injury is determined by the type of medical intervention required for the injury scenario. The injury severity level and associated medical intervention are shown in Table 1.

For example, we assign a severity level of 2 for the injury scenario "an axe breaks and the ejected part strikes the user's head," since it may require a visit to A&E.

7. *Determine the risk level*: The risk level is determined by using a risk level matrix that combines the severity of the injury and the probability of the injury occurring as described in the injury scenario. The estimated risk level of the product will contain some level of uncertainty since the probability of injury and severity of injury are estimated parameters. RAPEX handles uncertainty in the estimated risk level using a sensitivity analysis that determines how variations in the estimated parameters (i.e., probability of injury and the severity of injury) affect the overall risk result. It entails repeating the risk assessment process using different probabilities for the steps in the injury scenario and different injury severity levels. If the sensitivity analysis shows that there is no significant change in the risk level, then there is increased confidence in the initial estimated risk level. In contrast, a significant change will reduce confidence and require a review of the estimated parameters. For instance, if the initial risk assessment for the axe scenario determines that the risk level is 'low' and the sensitivity analysis also shows that there is no significant change in the risk level, then the risk level of the axe is confidently considered as 'low.' However, a product can have many risk levels due to many hazards, many injury scenarios, or varying probabilities or severities of injuries. In these situations, the risk level of the product is the highest risk level identified for that product.

Despite the widespread use of the RAPEX method, it has the following limitations:

1. *Limited approach to handling uncertainty:* In RAPEX, probabilities are assigned using point values instead of distributions (i.e., the assignment of probability values to each of the possible states of a random variable). RAPEX attempts to handle second-

Table 1
Injury severity level and associated medical intervention.

Injury Severity Level	Medical Intervention
1	First Aid
2	Visit Accident and Emergency Department (A&E)
3	Hospitalization
4	Fatal or loss of a limb(s)

order uncertainty (i.e., the uncertainty in the estimation of the parameters of interest (Briggs et al., 2012))using a sensitivity analysis that entails repeating the risk assessment process using different probabilities for the steps in the injury scenario and different injury severity levels. This method of handling uncertainty is not practical for probabilities that are not directly observable, nor where there is uncertainty about the data.

2. *Cannot be applied where there is little or no product data*: RAPEX cannot produce risk assessments for genuinely novel products (i.e., those for which little or no relevant historical data exist) or products for which limited testing data are available.
3. *Does not incorporate causal explanations for using and interpreting the data*: RAPEX provides no systematic or rigorous method for taking account of causal knowledge and explanations of the statistical data it uses, which may lead to inaccurate results. Also, RAPEX does not consider the causal factors that generate the data it uses since it assumes that the data are reliable because it is obtained from credible sources. The most general example is that lack of incident data for a product may be due to lack of reporting on the product rather than lack of incidents while, at the other extreme, multiple incidents associated with a product may be the result of testing the product beyond its intended scope.
4. *Does not differentiate between different types of users – i.e. their usage profile and risk tolerability*: In the RAPEX method, product risk is based on the likelihood of a product causing injury to a 'generic' user and the severity of that injury without any consideration of the context of use (RPA, 2006; European Commission, 2015, 2018). Hence, a product formally classified as 'high risk' may actually be 'low risk' or 'tolerable' for different classes of users taking into account the way they use the product, the benefits they receive from it and risk controls and mitigants. Risk controls and mitigants vary for different types of users due to their knowledge of the hazard and the environment in which they use the product. For instance, users that are aware of a fire hazard from a device are likely to have a smoke alarm installed nearby thus reducing the likelihood of injury (e.g., burn) even if the hazard occurs.
5. *Does not consider different product combinations and interactions with different classes of users when estimating product risk*: RAPEX's injury scenario assumes that the events leading to an injury are independent and that the product is used by a user independent of other classes of products and users. Hence RAPEX cannot assess the injury scenarios with different product combination interactions with different classes of users. For example, the risk of an axe used by a student supervised by a trainer.
6. *Does not consider the user exposure to the risk*: RAPEX does not include the usage frequency when determining the probability of a product causing injury to a user. Usage frequency is essential to determining the probability of injury since injury can only occur during product use. For instance, a consumer that uses a product often will have a higher probability of being injured due to repeated exposure to the hazard when compared to a consumer that rarely uses the product.
7. *Does not include information on risk tolerability*: Risk tolerability is the trade-off between risk and utility. For instance, a 'high risk' product may be considered 'tolerable' for some users since they value the utility of the product sufficiently high and are willing to tolerate the 'high risk' as a trade-off for the utility. Hence, risk tolerability is an essential component of product risk assessment since it informs risk management response to a non-compliant product.
8. *Does not consider increased risk of hazards over the lifetime of a product*: Due to wear and tear the 'hazard rate' of a product will generally increase over time, with different classes of products having very different increasing hazard rates. An estimated hazard rate of a product – based only on testing instances of the product when new – will underestimate the true hazard rate of the product in operation.
9. *Cannot assess the risk of products with unknown hazards or unknown product usage information*: RAPEX cannot assess the risk of products, especially novel products, with unknown hazards or unknown product usage information since it requires an injury scenario to estimate product risk. Nor does it provide a method for recognizing when novelty in hazard or usage arises.

3. Method

3.1. Bayesian networks (BNs)

Bayesian networks (BNs) are a type of probabilistic graphical model that explicitly describe dependencies between a set of variables using a directed acyclic graph (DAG) and a set of node probability tables (NPTs) (Fenton & Neil, 2018; Pearl, 2009; Spohn, 2008). The directed acyclic graph consists of nodes and directed arcs; nodes represent variables and arcs are used to link nodes (Fenton & Neil, 2018; Spohn, 2008). The arcs assume that there is a causal influence or statistical relation between nodes (although in our work we generally assume the relationship is causal, and emphasize this by referring to them as causal BNs). For instance, given two nodes X and Y, an arc from X to Y assumes that X directly influences Y; as a result, X is called the parent of Y, and we also say Y is dependent on X (although note that this means statistical dependence – it does not mean it is the only thing that can 'cause' Y). Each node in a DAG has a node probability table (NPT) also called a conditional probability table (CPT) which describes the probability distribution of the node given its parents. Any node without a parent is called a root node, and the NPT for that given node is its probability distribution (Fenton & Neil, 2018).

Once the DAG and NPTs of the BN are defined, inferences are done via Bayes Theorem, which updates our prior belief of a hypothesis given new evidence. Our prior belief is called the prior probability (or the prior), and our revised belief is called the posterior probability (or the posterior; Pearl, 2009). Given evidence, we can do three types of reasoning in a BN to answer questions: observational reasoning (which includes diagnostic and predictive), interventional reasoning, and counterfactual reasoning.

1 *Observational Reasoning*: This entails entering some observed evidence in a BN and making inferences. The two main types of observational inference are diagnostic (backward) and predictive (forward). Diagnostic inference entails discovering the cause of an observation, whereas predictive inference entails discovering the effect of an observation (Pearl, 2009; Wiegerinck, Burgers, & Kappen, 2013).
2 *Interventional Reasoning:* This entails determining the effect of a variable on an outcome by modelling the variable as an intervention on the outcome. Modelling a variable as an intervention involves fixing its value by making the variable independent of its causes. This is done in a BN by performing "graph surgery," where all the arcs entering the intervened variable are removed (Pearl, 2009).
3 *Counterfactual Reasoning:* This entails predicting what would have happened if events were different from what is actually observed. This type of reasoning combines observational and interventional reasoning in a BN using the twin network method proposed by Balke and Pearl (1994). In this method, two identical BNs representing real and counterfactual worlds are connected via shared background or exogenous variables where an observation in the real world is modelled as an intervention in the counterfactual world (Pearl, 2009).

There has also been previous work on using BNs for product risk assessments. For instance, Suh (2017) developed a product risk assessment system using a BN to assess product risk based on injury information from the Korea Consumer Agency. They evaluated 33 children's products and compared the results with RAPEX. Berchialla et al. (2010) used a BN to estimate the risk of ingestion, inhalation, and insertion of consumer products in children aged 0–14. They also compared the BN approach to other quantitative risk assessment methods such as neural networks, classification trees, and logistic models (Berchialla et al., 2016). Their results indicate that BNs are the best method for assessing this safety risk because they are easier to interpret and provide accurate predictions (Berchialla et al., 2016). However, these previous works have not explicitly described or proposed a generic BN for product risk assessment. For instance, the Berchialla et al. (2010) BN cannot be used for product risk assessment since its structure and parameters are not applicable. Though Suh (2017) used a BN in their product risk assessment system, it is unclear if this BN is generalizable since they did not provide a causal diagram (a diagram that describes the causal relationship between the variables in the BN).

3.2. The causal Bayesian network for product risk assessment

A schematic for our proposed generic BN for product risk assessment is shown in Fig. 1 (see Fig. 3 for complete BN model).

The proposed generic BN model provides an improved systematic method for product risk assessment that resolves the limitations identified with RAPEX discussed in Section 2.1. It can complement traditional methods like RAPEX and provide quantified, auditable assessments in situations where such methods cannot because of lack of data. Though its variables and structure are generic and can be used for any type of consumer product risk assessment, its probability tables (see Appendix) are not generic as these would be revised for different classes of products. The following steps describe the risk assessment process using the BN model:

i. *Determine the probability of hazard per demand:* A demand is a single use of a particular product. Hazard testing information (e.g., testing strategy) are combined with the manufacturer process information (e.g., country of origin) to estimate the probability of hazard per demand.
ii. *Determine the probability of hazard occurrence:* The estimated probability of hazard per demand is combined with product usage information (e.g., number of demands) to determine the probability of hazard occurrence.
iii. *Determine the probability of injury (major and minor) during the product lifetime:* The estimated probability of hazard occurrence is combined with the probability of the uncontrolled hazard causing an injury to determine the probability of injury during the product lifetime.
iv. *Determine the probability of product instances causing major and minor injuries:* The estimated probability of injury is combined with product instances information to determine the probability of product instances causing major and minor injuries.
v. *Determine the risk level:* The estimated probability of product instances causing major and minor injuries are combined to determine the risk level of the product.
vi. *Determine the risk tolerability:* The product utility information and estimated product risk level are combined to determine risk tolerability.
vii. *Determine consumer risk perception given government intervention, negative media stories and warnings:* The initial consumer risk perception for a particular product is revised given government intervention information (e.g., recall, negative media stories and warnings).

In Sections 3.3 and 3.4, the methods used for selecting variables, determining the BN structure, and populating the node probability tables are discussed.

3.3. Variables selection and BN structure

A knowledge-based approach was used to identify relevant variables for building the BN model. It entails using domain expert knowledge to build the BN structure and to define model parameters where data are limited or missing (Constantinou & Fenton, 2018). In this study, product risk assessment literature (RPA, 2006; PROSAFE, 2013; ISO/IEC, 2014; European Commission 2015, 2018) was reviewed, and a core team of three domain experts identified relevant variables and an initial model structure. This was presented for discussion in a workshop with six senior government safety and risk experts and was revised accordingly

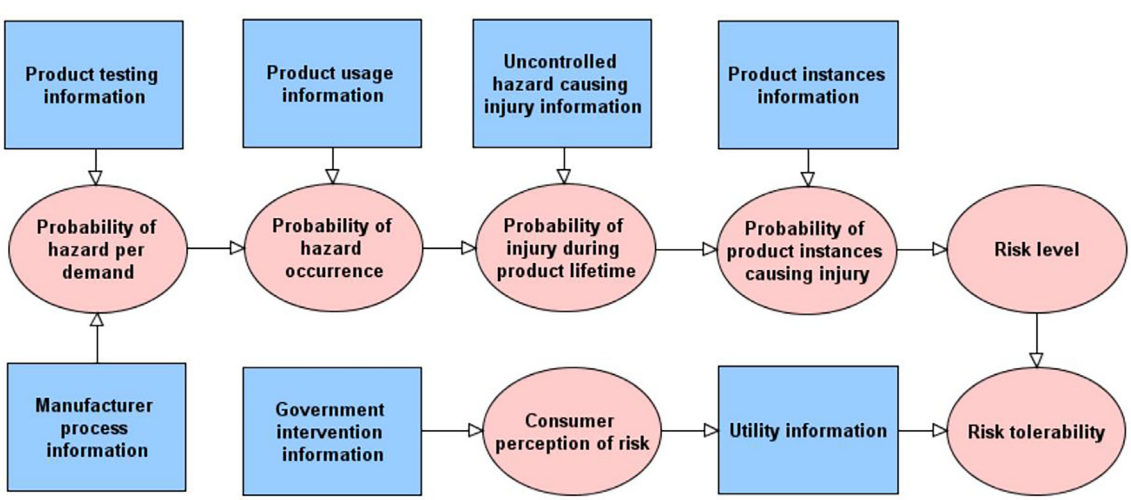

Fig. 1. Schematic view of the BN model for Product risk assessment.

both at the workshop and in subsequent iterations (mainly through email as, due to the Covid-19 crisis, no further in-person workshops were conducted). A consensus on the proposed BN model was reached when the model included all relevant variables (connected causally) required for doing a product risk assessment. The identified variables are shown in the Appendix, and the BN structure shown in Fig. 3.

The initial core BN structure (which was retained throughout) was built using idioms. Idioms are fragments of a BN that represent generic types of uncertain reasoning (Neil, Fenton, & Nielsen, 2000). The identified variables were organized into idioms representing different components of product risk assessment. For instance, the measurement idiom shown in Fig. 2a is used to model the uncertainty concerning the measurement of a variable. Fig. 2b shows an instance of this idiom in the model. The product risk assessment idioms were then linked to build the proposed BN model.

3.4. Node probability tables (NPTs)

The node probability tables (NPTs) shown in the Appendix were defined using: (1) expert knowledge, (2) ranked nodes (Fenton et al., 2007), (3) standard mathematical or statistical assumptions and distributions, and (4) comparative expressions.

1. *Ranked nodes:* These nodes are used in the BN model to represent discrete variables with states expressed on an ordinal scale (e.g., *Customer satisfaction* node with states; *very low, low, medium, high, very high*). A ranked node maps the variable states to an underlying numerical scale ranging from 0 to 1 in equal intervals (Fenton et al., 2007). Given the underlying numerical scale, the NPT for a ranked node can be defined as a statistical distribution such as a truncated normal distribution (TNormal) with mean μ and variance $\sigma^{2,}$ i.e., *TNormal* (μ, σ^2). In the BN model, the NPT for a ranked node without parents is a uniform distribution (i.e., the probability of each state is the same). The NPT for a ranked node with parents is a TNormal distribution with mean μ defined as the weighted average of its parents and variance σ^2.
2. *Standard mathematical or statistical assumptions and distributions*: The NPT for numeric variables (nodes) in the BN model is defined using standard mathematical or statistical assumptions and distributions. For example, the NPT for *Number of times hazard observed in tests* node is a Binomial(n,p) distribution where n is the number of demands made during testing and p is the probability of observing a hazard per demand. The NPT for some numeric nodes are deterministic and self-explanatory; for instance, the NPT for *Probability the hazard causes a major injury* node is an arithmetic expression (i.e., *probability of uncontrolled hazard causing a major injury* $\times$ (*1 − probability of control stops injury*). The NPT for numeric nodes without parents is a uniform distribution, and those with parents is a TNormal distribution. The mathematical expressions and statistical distributions used to define the NPT for each numeric node is dependent on the function of the node, its input, and its output.
3. *Comparative expressions:* These expressions (e.g., IF statements) are used to define the NPT for discrete variables with two states and parents. For instance, the NPT for *Government intervention required given risk level* node with states (True, False) and parent *Risk level* is an IF statement (i.e., *if(risk_level > 0.5,"True", "False")*).

Although the proposed BN structure and variables are relevant for assessing the risk of any consumer product, it is important to note that the NPT for some of these variables will be revised given specific data about a particular product or class of product. However, NPTs with prior statistical distributions are defined with a sufficiently large variance to enable them to be applied to the very different product examples used in Section 4. For example, the NPT for numeric (probability) nodes without parents are defined as uniform distributions and those with parents are defined as TNormal distributions. All model development was done using AgenaRisk software (Agena Ltd, 2021).

3.5. Bayesian network propagation and inference

All calculations (inferences) in the proposed BN model are done using the junction tree algorithm (Kim & Pearl, 1983; Lauritzen & Spiegelhalter, 1988). The junction tree algorithm transforms the Bayesian network into a tree structure with clusters (i.e., sets comprising one or more variables). Computations are done locally within the clusters and then propagated to other clusters in the tree structure as 'messages' (this method is called message passing). The junction tree algorithm allows the BN model to provide global answers using local calculations. In this study, it is implemented by AgenaRisk software (Agena Ltd, 2021).

4. Results

In this section, we compare the process and results of the BN model and RAPEX in terms of their ability to assess the risk of products with relevant data (e.g., teddy Bear; Section 4.1) and products with limited or no relevant data (e.g., a new uncertified kettle; Section 4.2).

4.1. Teddy bear risk assessment using the BN model

4.1.1. Background

In the UK and the EU, the RAPEX method is used to assess the risk associated with different toys to prevent harm to children. We will use the BN model to assess the risk associated with a teddy

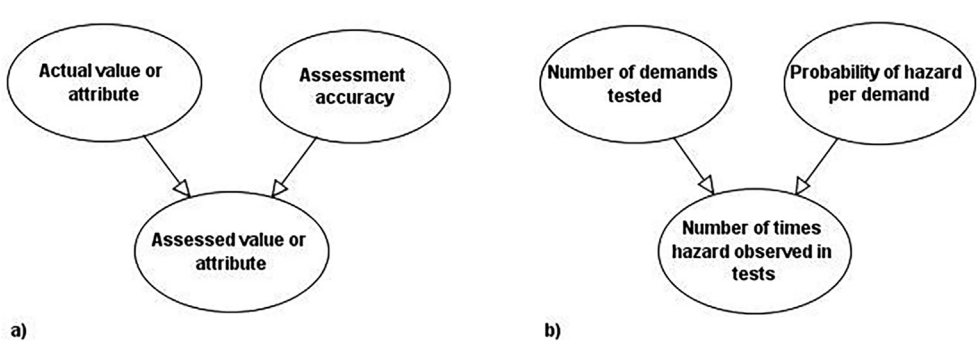

Fig. 2. Measurement Idiom (a) Generic (b) Instance in BN model.

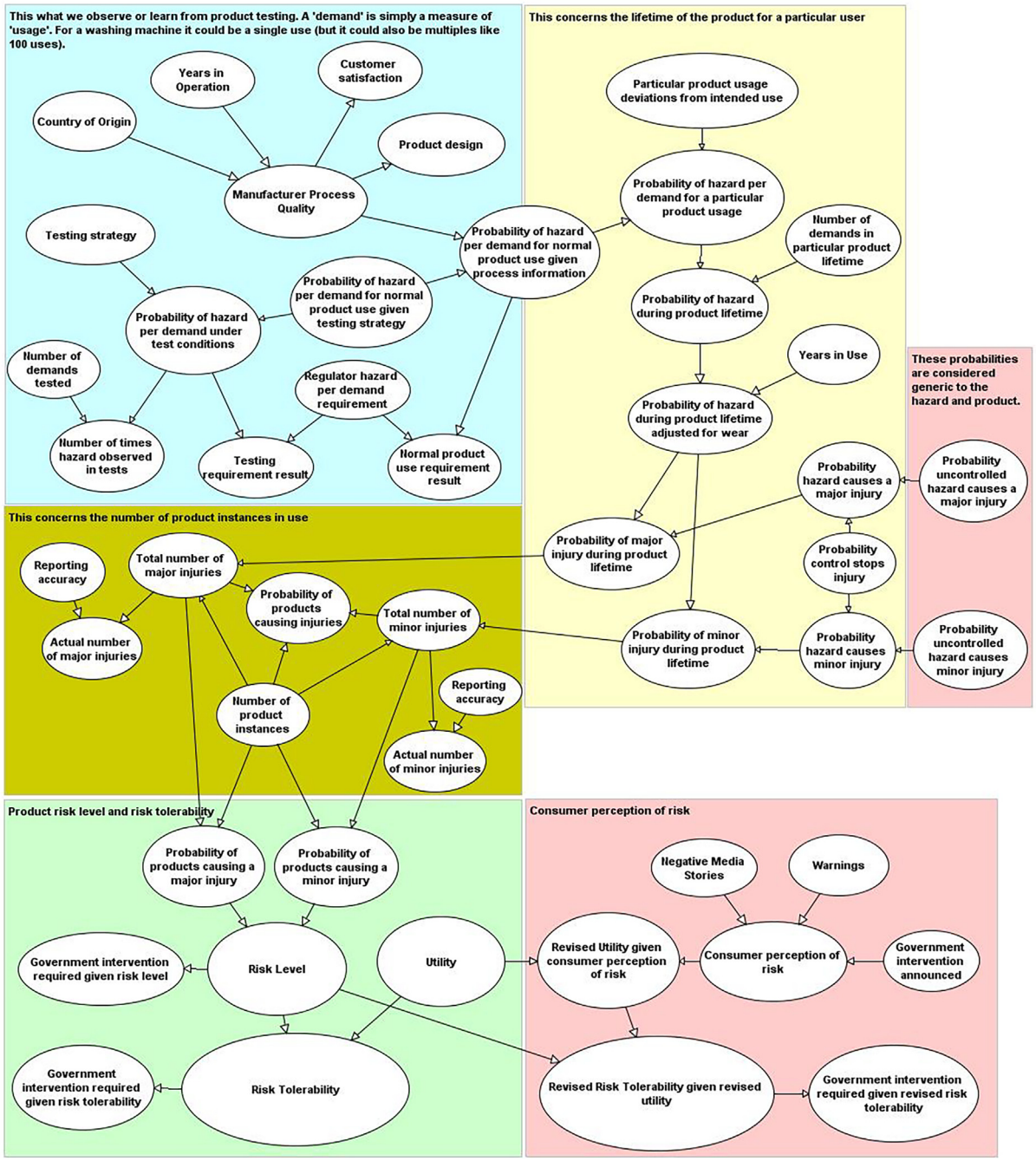

Fig. 3. BN model for Product risk assessment.

bear using two scenarios and compare the results to the RAPEX method.

4.1.2. Scenarios

We assume that the hazard is a small part (e.g., a teddy bear eye), which is swallowed by a child resulting in an injury.

1 *Scenario 1:* We assume that the teddy bear is used by a child aged 0–36 months as intended for one year with a high number of demands (i.e., 4000) and no caregiver intervention (i.e., the child is not sufficiently supervised so the caregiver cannot take away the small detached part (e.g. teddy bear eye) before it is ingested by the child).

2 *Scenario 2:* We assume that the teddy bear is used by a child aged 0–36 months as intended for one year with a low number of demands (i.e., 200) and an 80% chance of caregiver intervention (i.e., the child is sufficiently supervised so that the caregiver can take away the small detached part (e.g., teddy bear eye) before it is ingested by the child).

We also assume the probability of the uncontrolled hazard causing a major injury is 0.1 and the probability of causing a minor injury is 0.2. We assume that the product was tested 'typical of normal use' and for 5000 demands we observe one hazard (e.g., the teddy bear eye is dislodged). We assume that there are 20,000 product instances and the utility of the teddy bear is 'med-

ium.' Finally, we assume that the manufacturer has been in operation for 5–10 years and is from a country with a good safety record for toys. The manufacturer also has a 'high' customer satisfaction rating, and there are no changes in product design (i.e., product appearance is the same as previous similar products).

4.1.3. Results

1. *Scenario 1*: The BN model shown in Fig. 4 learns using the junction tree algorithm (see Section 3.5) that the risk level for the teddy bear is 'very high' with little uncertainty for the child. The BN model calculates that the mean probability of a major injury for this scenario is 0.07 and for a minor injury it is 0.14. It also calculates that the mean number of major and minor injuries for 20,000 product instances is 1,387 and 2,773, respectively. Finally, the BN model shows that the risk tolerability for the teddy bear is 'low' or 'very low' given a 'medium' utility and that a government intervention such as recall is required. One of the limitations of the RAPEX method is that it does not consider the number of demands for a particular product when determining risk. So, although we are unable to make a direct comparison to the BN model, we can compare the product risk result of the BN model to the RAPEX method by using the mean probability of a major injury learnt by the BN model as the probability of injury for the RAPEX system. We set the injury severity level to '3′ as this corresponds to a major injury such as internal airway obstruction. The RAPEX method assesses the risk level of the teddy bear as 'serious' as

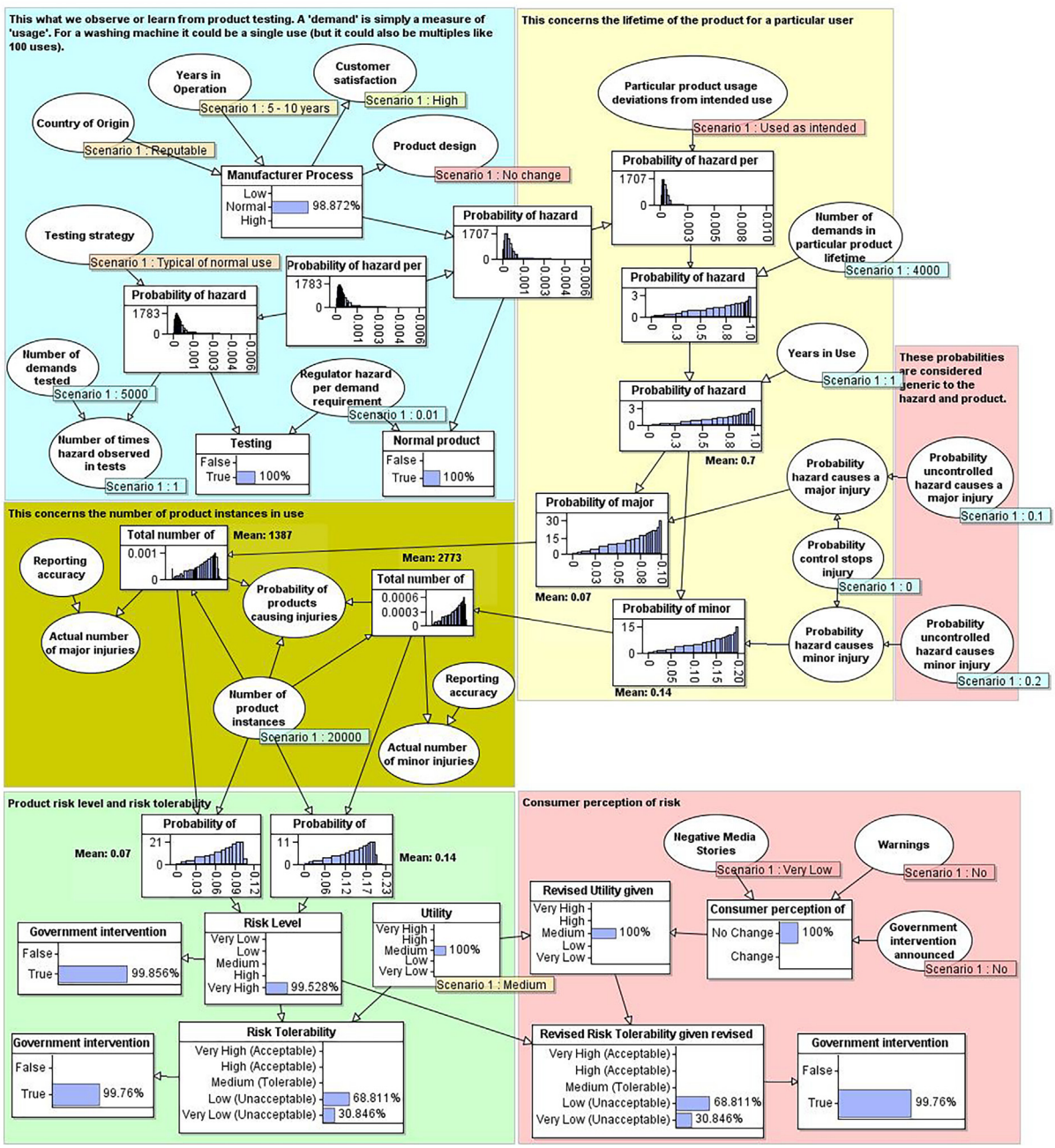

Fig. 4. BN Model for teddy bear Scenario 1.

shown in Fig. 6. This result is the same as the BN model even though the BN model also uses the probability of a minor injury to compute the product risk.

2. *Scenario 2*: The BN model shown in Fig. 5 learns that there is 67% chance that the risk level for the teddy bear is 'low' or 'very low' for the child. The BN model also calculates the mean probability of a major injury, which for this scenario is 0.001 and for a minor injury it is 0.003. The BN model calculates that the mean number of major and minor injuries for 20,000 product instances is 35 and 68, respectively. Finally, the BN model shows that there is a 58% chance that the risk tolerability will be 'high' or 'very high' for the teddy bear given a moderate utility and recommends no government intervention such as a recall with some uncertainty. We compare the product risk result of the BN model to the RAPEX method by using the mean probability of a major injury calculated by the BN model as the probability of injury for the RAPEX system. We set the injury severity level to '3′ as this corresponds to a major injury such as internal airway obstruction. The RAPEX method assesses the risk level of the teddy bear as 'serious' as shown in Fig. 7. This result is not the same as the BN model for the given probability of a major injury since the BN model also uses the probability of a minor injury to compute the product risk.

4.2. A new uncertified kettle risk assessment using the BN model

4.2.1. Background

Every year there are new uncertified products available on the market that pose a serious risk to the health and safety of consumers. However, regulators may find it difficult to assess the risk for these products using RAPEX, since they may not have access to the manufacturer testing data generated during product develop-

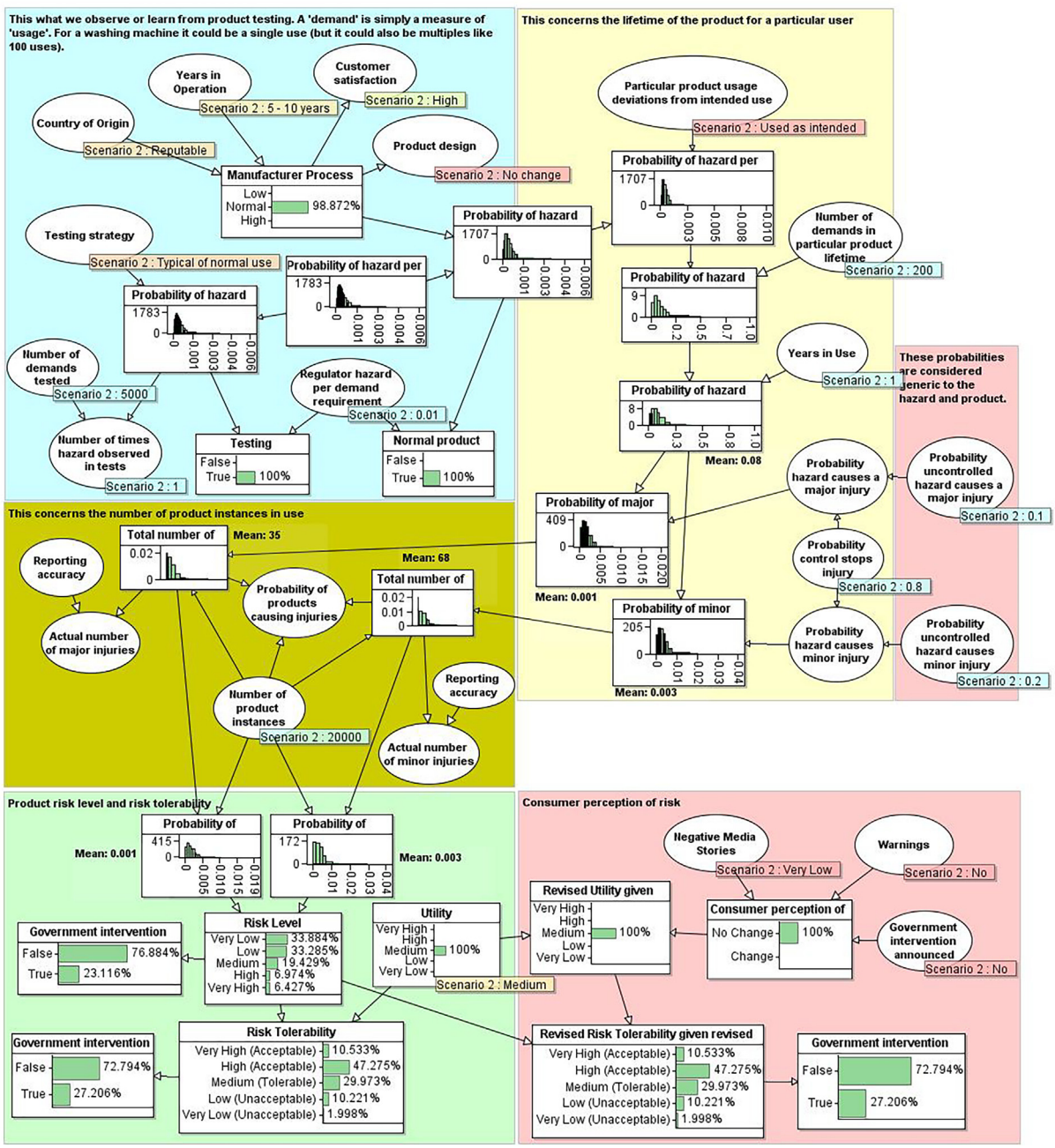

Fig. 5. BN Model for teddy bear Scenario 2.

European Commission | RAG - Risk Assessment

Scenario 1 : Very young children - Product is or contains small part

1 | Product hazard

Hazard Group: **Size, shape and surface**
Hazard Type: **Product is or contains small part**

2 | Consumer

Consumer type: **Very young children - 0 to 36 months (Very vulnerable consumers)**

3 | How the hazard causes an injury to the consumer

Injury scenario: **Person (child) swallows small part; the part gets stuck in larynx and blocks airways**

4 | Severity of Injury

Injury: **Ingestion**
Level: **3 Internal organ injury (Refer also to internal airway obstruction in case the ingested object gets stuck high in the oesophagus.)**

5 | Probability of the steps to injury

Step	Step(s) to Injury	Probability
1	Teddy bear scenario 1: mean probability of a major injury	0.07

Calculated probability	Overall probability	Risk of this scenario
0.07	> 1/100	Serious risk

Fig. 6. RAPEX results for teddy bear Scenario 1 (European Union, 2019).

European Commission | RAG - Risk Assessment

Scenario 1 : Very young children - Product is or contains small part

1 | Product hazard

Hazard Group: **Size, shape and surface**
Hazard Type: **Product is or contains small part**

2 | Consumer

Consumer type: **Very young children - 0 to 36 months (Very vulnerable consumers)**

3 | How the hazard causes an injury to the consumer

Injury scenario: **Person (child) swallows small part; the part gets stuck in larynx and blocks airways**

4 | Severity of Injury

Injury: **Ingestion**
Level: **3 Internal organ injury (Refer also to internal airway obstruction in case the ingested object gets stuck high in the oesophagus.)**

5 | Probability of the steps to injury

Step	Step(s) to Injury	Probability
1	Teddy Bear Scenario 2: mean probability of a major injury	0.001

Calculated probability	Overall probability	Risk of this scenario
0.001	= 1/1000	Serious risk

Fig. 7. RAPEX results for teddy bear Scenario 2 (European Union, 2019).

ment (even if such data were collected) and the number of product instances is unknown. We will use the BN model to assess the risk for two different scenarios, both involving a new uncertified kettle on the market for which there are no testing data, and the number of product instances is unknown. The hazard is an ignition source that causes a fire resulting in a burn injury.

4.2.2. Scenarios

1. *Scenario 1*: There are no reported injuries. The manufacturer has been in operation for 4 years and is from a country with a poor safety record for consumer electrical appliances. The manufacturer also has a 'low' customer satisfaction rating, and there are no changes in product design (i.e., product appearance is the same as previous similar products).
2. *Scenario 2*: There has been one major injury reported. The manufacturer has been in operation for 20+ years and is from a country with a very good safety record for consumer electrical appliances. The manufacturer also has a 'very high' customer satisfaction rating, and there are major improvements in product design.

4.2.3. Method

Using the BN model, we estimate the probability distribution of hazard per demand for the new uncertified kettle by assigning priors (i.e., initial probability values) to the following nodes: (1) number of demands tested, (2) number of hazards observed during testing, and (3) testing strategy. These priors are based on testing data for three similar kettles. We assume that the data indicate that testing was 'typical of normal use' and the number of demands for the kettles are in the range of 7500–10,000 with one hazard observed. Since we are uncertain about the 'true' number of demands at which the hazard will appear for this particular kettle, we set the lower and upper bounds for the distribution for the number of demands tested as 7,500 and 10,000, respectively, and we set the number of hazards observed to one. We then combine the estimated probability distribution of hazard per demand with the manufacturer process information such as country of origin, to estimate the 'true' probability distribution of the hazard per demand.

We calculate the probability distribution of hazard occurrence by combining the user behavior with the estimated probability distribution of hazard per demand. We do this by assigning priors to the following nodes: (1) product usage, (2) number of demands, and (3) years in use. Since we assume that this kettle will be used on average 3,000 times, we use this value as the mean of the distribution for the number of demands. Also, since we are uncertain about the consumer behavior during use, we assume that the kettle is used as intended 90% of the times with major and minor deviations of 7% and 3%, respectively, based on the data for similar kettles. We then calculate the probability distribution of the kettle causing a major and minor injury given the estimated probability distribution of hazard occurrence and the estimated probability distributions of the uncontrolled hazard causing a major and a minor injury, respectively. We assume that the probability of the uncontrolled hazard causing a major and a minor injury are 0.1 and 0.2, respectively, and the probability of the control to stop the hazard causing an injury is 0.5. Finally, we estimate the number of product instances causing major and minor injuries and the overall product risk for the kettle. We assume that there are approximately 50,000–100,000 product instances based on data for similar kettles. Since we are uncertain about the 'true' number of product instances, we set the lower and upper bounds for distribution of the number of product instances as 50,000 and 100,000, respectively. The BN model then predicts the number of product instances causing major and minor injuries and the risk of the new uncertified kettle.

4.2.4. Results

1. *Scenario 1*: The BN model shown in Fig. 8 learns that the risk level for the new uncertified kettle is 'very high' with some uncertainty. The BN model also calculates that the mean probability of a major injury is 0.02 and a minor injury is 0.05. It calculates the mean probability of hazard per demand and hazard occurrence as 0.0003 and 0.5, respectively. It also estimates that the mean number of major and minor injuries for product instances in the range of 50000–100000 is 1,828 and 3,657, respectively. Finally, the BN model shows that there is a 98% chance that the risk tolerability for the kettle is 'low' or 'very low' given a 'medium' utility and recommends a government intervention such as recall with some uncertainty.
2. *Scenario 2:* The BN model shown in Fig. 9 learns that the risk level for the new uncertified kettle is 'very low' when the number of product instances causing a major injury is 1. The BN model also calculates that the mean probability of a major injury is 0.00005 and a minor injury is 0.0001. It also estimates that the mean number of minor injuries is 7, and the mean probability of the hazard per demand and hazard occurrence as 0.0000003 and 0.001, respectively. Finally, the BN model shows that the risk tolerability for the kettle is 'high' or 'very high' given a 'medium' utility and recommends no government intervention with little uncertainty.

Note that, although in Scenario 1 there were no reported injuries from using the product (whereas there was such a report in Scenario 2) the method uses prior and process information to produce recommendations to intervene in Scenario 1 but not Scenario 2.

5. Discussion

5.1. Product risk assessment methods' results comparison

The teddy bear case study results (See Section 4.1.3) show that the BN model and RAPEX method may estimate different product risk levels for a particular product. In the teddy bear scenario 1, the BN model and the RAPEX method estimated the risk level as 'very high' or 'serious' (see Fig. 4 and Fig. 6). However, in teddy bear scenario 2, the BN model shows that there is a 67% chance that the risk level is 'low' or 'very low' (see Fig. 5), whereas the RAPEX method predicted the risk level as 'serious' (see Fig. 7). This difference in risk level estimates is due to the method used by the BN model to estimate product risk. As shown in Fig. 1 and Fig. 3, the BN model includes additional information such as manufacturer process information and product usage information (such as number of demands and product wear) when estimating product risk. This allows the BN model risk estimates to be comprehensive as it incorporates all relevant factors that affect product risk. Also, since these factors are linked causally, they support ease of interpretability and explanation of risk level estimates.

The case study results for the new uncertified kettle in Section 4.2.4 show that, while RAPEX is unable to assess the risk level of novel products for which there is little or no available data, the BN model can provide auditable, quantified risk assessments. For these scenarios, the BN model estimates product risk by combining manufacturer process information with estimates from previous similar products. This estimated product risk is also revised given new data (e.g., reported injuries; see Fig. 9). This ability of the BN model to revise the risk level given new data is essential for regu-

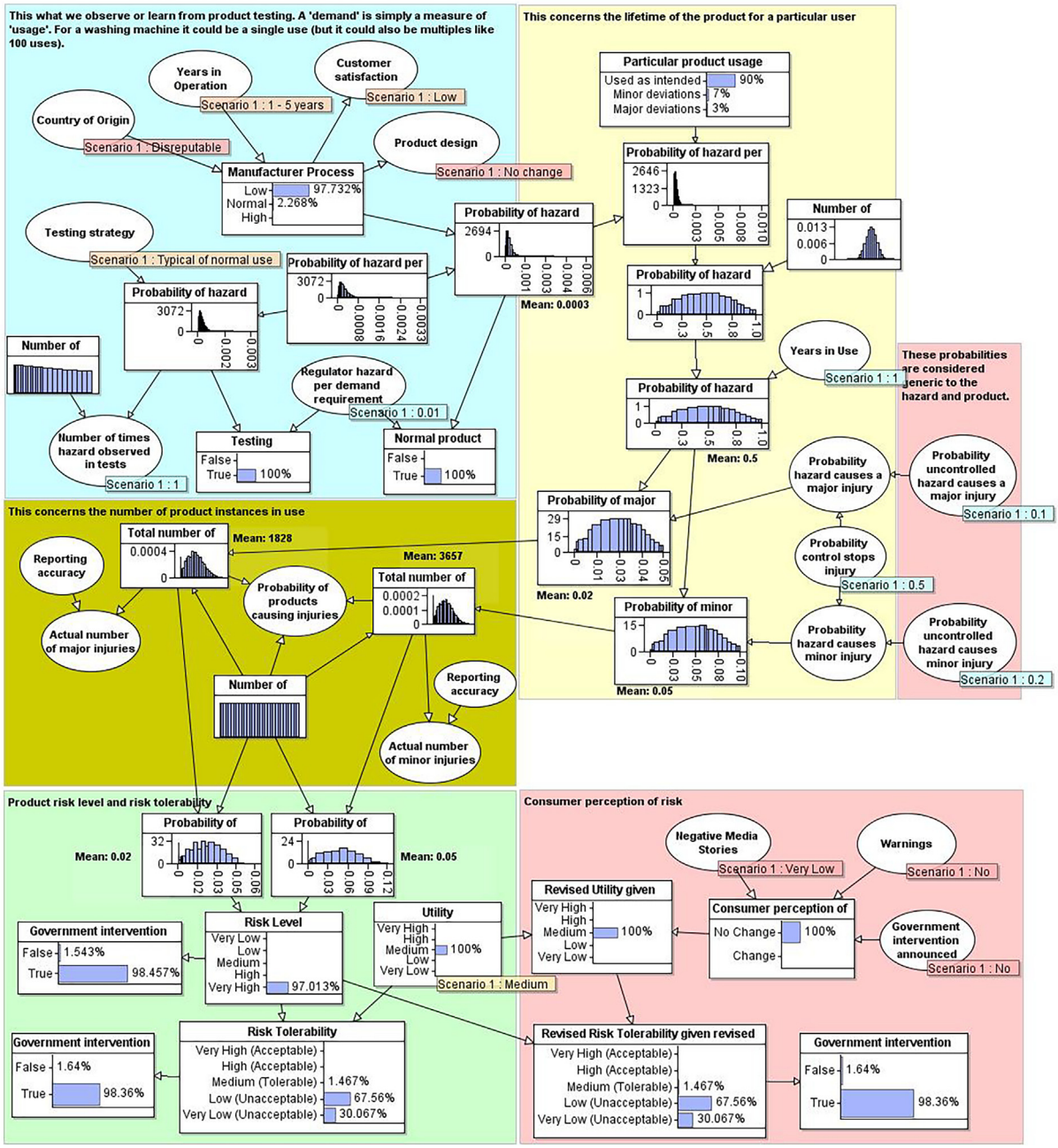

Fig. 8. BN model for a new uncertified Kettle Scenario 1.

lators to adequately assess and monitor the risk of novel products over time. In fact, the BN model will also perform better than RAPEX for novel products given new data since it incorporates all the factors that affect product risk in a causal manner, and it takes full account of uncertainty when estimating product risk.

5.2. Benefits of using the BN model for product risk assessment

The case study results also show that the BN model resolves the issues with RAPEX discussed in Section 2.1 and provides the following improvements to the product risk assessment:

1. *Properly handles uncertainty about probabilities assigned during risk assessment*: The BN model handles second-order uncertainty by incorporating distributions, rather than point values for probabilities that are not directly observable.
2. *Can be applied where there is little or no product data:* In situations where it will neither be feasible nor possible to get any extensive data from testing or details on product instances, the BN model can incorporate expert judgement and/or data from previous similar products together with process information about the manufacturer and their reputation. This results in product risk assessments that are still quantified and auditable.
3. *Incorporates causal explanations for using and interpreting the data*: The BN model explicitly describes the risk assessment process and the causal relationship between the data used.
4. *Considers the usage behavior for different types of users and the number of product instances when determining product risk*: The BN model can take full account of the distributions of different types of users when estimating product risk by simply assigning priors to the 'particular product usage' node that capture the

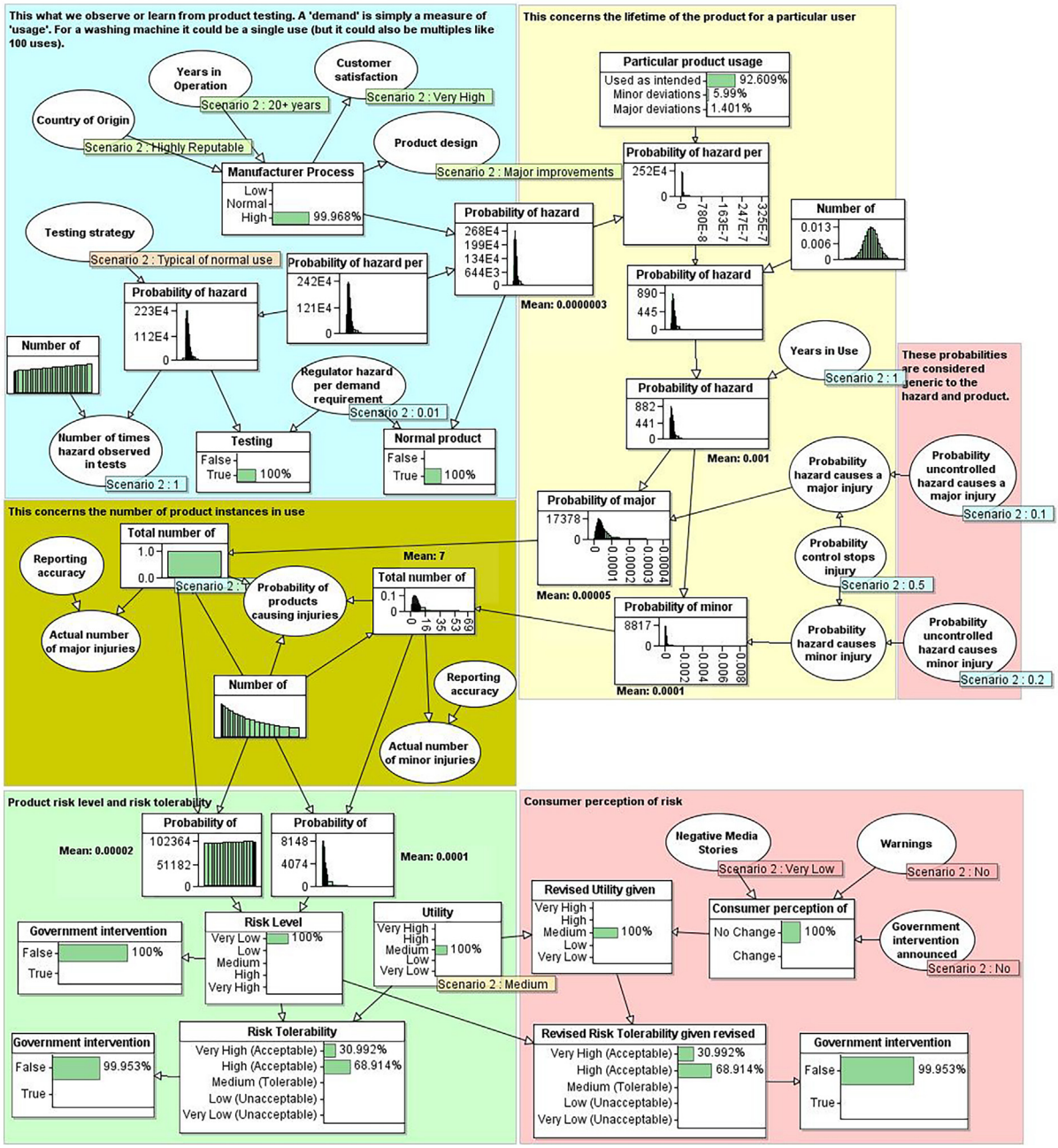

Fig. 9. BN model for a new uncertified Kettle Scenario 2.

population distribution. For instance, if for a particular product we estimate that only 30% of the population will 'use it as intended' then we set the prior probability of that node state to 30%. In addition, the BN model explicitly includes 'controls' that can prevent a hazard from causing an injury. For example, in households where there is a smoke alarm and fire extinguisher, the probability that a fire from a washing machine leads to injury is greatly reduced. In households where young children are under close supervision, there is a much lower probability that a hazard from a toy (such as an eye pulled off a teddy bear) will lead to injury compared to households where children are left unsupervised. Lastly, the BN model can provide individualized risk assessments. For instance, for a particular user, the model can estimate the probability that this user will suffer an injury during the product lifetime.

5. *Considers the user exposure to risk:* The BN model uses the usage frequency of the product (i.e., the number of demands) to determine the probability of injury for a particular user or class of user.
6. *Includes information on risk tolerability*: The BN model combines utility and the product risk to determine risk tolerability for a particular user or class of user.
7. *Considers increased risk of hazards over the lifetime of a product:* The BN model considers the effect of wear and tear on the 'hazard rate' of the product.

The BN model also improves product risk assessment by modelling:

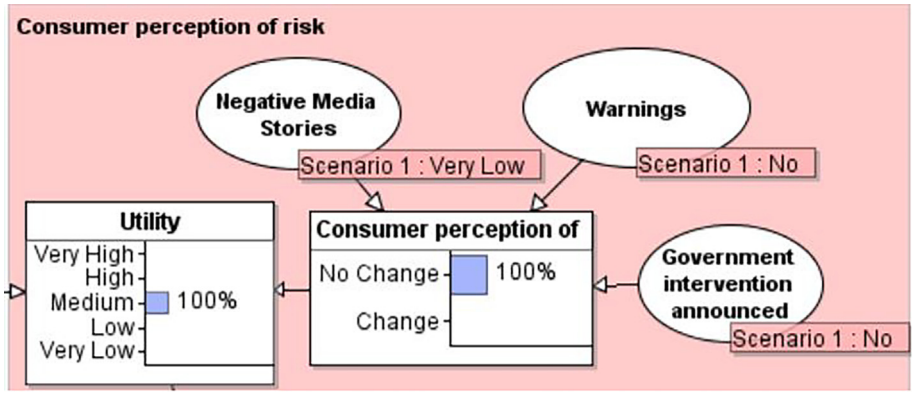

Fig. 10. BN fragment showing no change in consumer perception of risk.

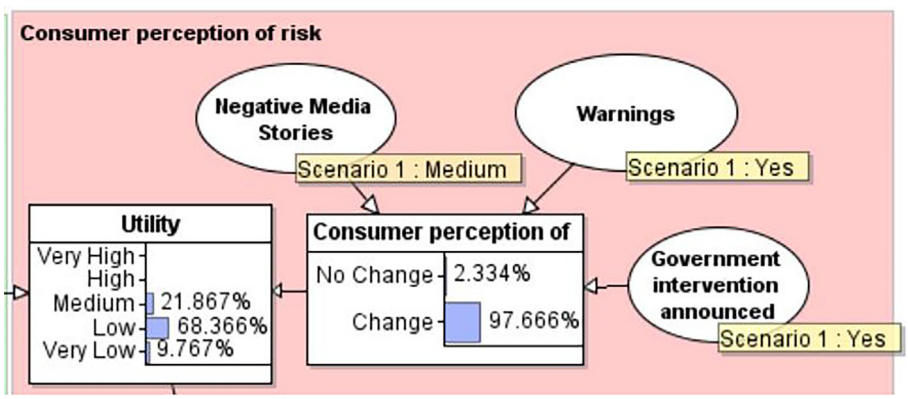

Fig. 11. BN fragment showing the change in consumer perception of risk.

a. *The effect of a government intervention, negative media stories, and warnings on the consumer perception of the risk*: For instance, as shown in Fig. 10, if there are little or no negative media stories, no warnings, and no government intervention announced for a particular product, the BN model predicts no change in consumer perception of the risk. However, as shown in Fig. 11, if there are negative media stories, warnings and a government intervention is announced, the BN model predicts the consumer perception of the risk will most likely change.
b. *The mean number of product instances causing major and minor injuries, respectively*: For instance, as shown in Fig. 12, the BN model can estimate the mean number of product instances causing major and minor injuries for a particular product based on the total number of product instances and the probability distribution of major and minor injuries, respectively. In this example, for a particular product with 519,000 instances and mean probabilities of causing major and minor injuries as 0.018 and 0.036, respectively, for a particular user or class of user, the BN model estimates that the mean number of product instances causing major and minor injuries are 9,335 and 18,668, respectively.
c. *The probability of hazard per demand using testing data and process information about the manufacturer and its reputation*: For instance, as shown in Fig. 13, the BN model can estimate the probability distribution of the hazard per demand for normal product use based on the probability distribution of hazard per demand observed during testing. In this example, for a particular product with 2,000 demands, 1 observed hazard and a testing strategy 'typical of normal use,' the BN model estimates that the probability distribution of hazard per demand for normal product use (mean 0.001) is the same as the probability distribution of hazard per demand observed during testing (mean 0.001). If product testing was 'poor,' as shown in Fig. 14, the probability distribution of hazard per demand for normal product use (mean 0.02) varies significantly from the distribution observed during testing (mean 0.001). This variation is due to Bayes' inference, which revises the probability distribution of the hazard per demand for normal product use based on the testing strategy.

6. Conclusion and recommendations

The proposed BN approach is a more powerful and flexible approach for systematic product risk assessment than traditional methods like RAPEX. In particular, it can: produce quantified, auditable assessment with limited or no data; properly handle second-order uncertainty; incorporate causal explanations for using and interpreting data; allow for different types of users including different exposure to risk and risk tolerability; and incorporate increased risk of hazards over the lifetime of a product. However, it is important to note that it can also complement traditional methods like RAPEX. For instance, since the BN approach estimates product risk using additional parameters such as product usage data and manufacturer process information, it can be used in the interim to validate RAPEX risk assessments. In situations where it will neither be feasible nor possible to get any extensive data from testing or injury databases, the BN approach can be used to estimate product risk due to its ability to handle incomplete data,

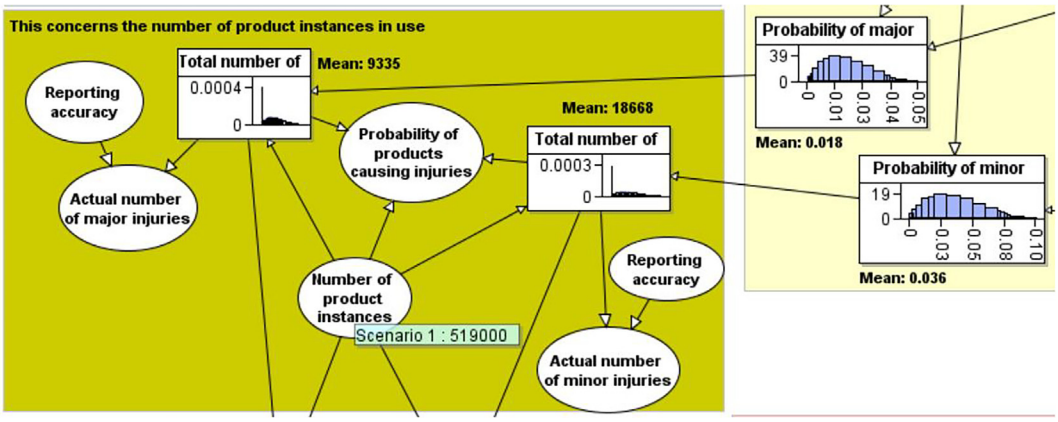

Fig. 12. BN fragment showing the mean number of product instances causing major and minor injuries.

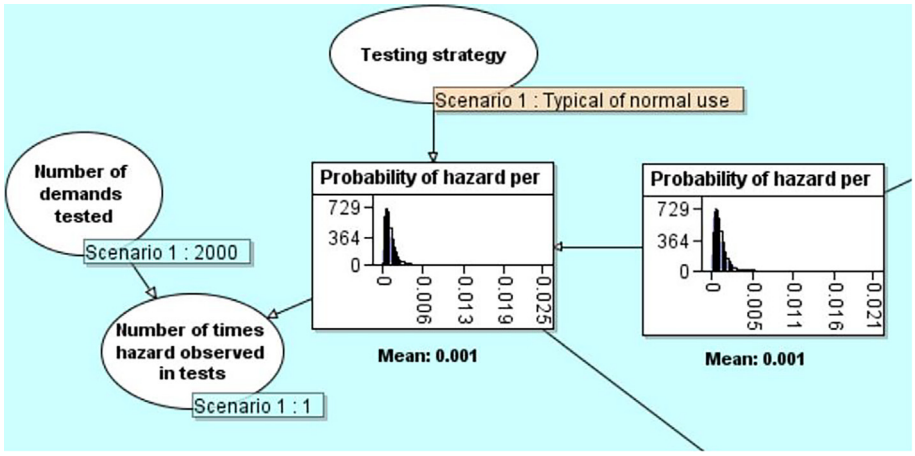

Fig. 13. BN fragment showing the probability distribution of hazard per demand for a particular product with a testing strategy typical of normal use.

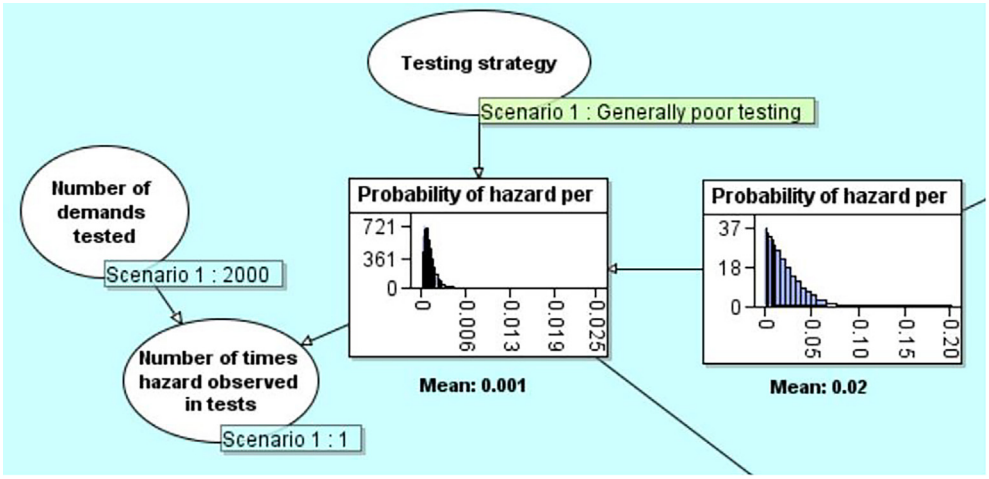

Fig. 14. BN fragment showing the probability distribution of hazard per demand for a particular product with a poor testing strategy.

combine objective and subjective evidence and revise risk estimates given new data. Also, it informs risk management decisions and helps understand the effect of those decisions on consumer risk perception. For these reasons, safety regulators, product manufacturers, and risk professionals will find the BN approach suitable for all types of consumer product risk assessment, including for novel products or products with little or no historical data.

Table 2
Variables, node types and NPTs.

Variable name	Abbrev	Node type	NPT
Number of demands tested	*ndt*	Simulation (integer interval)	*Uniform(0, 1000000)*
Number of times hazard observed in tests	*nho*	Simulation (integer interval)	*Binomial(ndt, p_h_testcond)*
Testing strategy	*ts*	Labelled (Less strenuous than normal use, Typical of normal use, More strenuous than normal use, Generally poor testing)	*Less strenuous than normal use: 0.1, Typical of normal use 0.6, More strenuous than normal use 0.2, Generally poor testing 0.1*
Probability of hazard per demand under testing conditions	*p_h_testcond*	Simulation (continuous interval)	*Partitioned Expression (Less strenuous than normal use: p_h_strat-0.5*p_h_strat, Typical of normal use: p_h_strat, More strenuous than normal use: p_h_strat + 0.5*p_h_strat, Generally poor testing: TNormal(p_h_strat,0.001,0,1))*
Probability of hazard per demand given testing strategy	*p_h_strat*	Simulation (continuous interval)	*TNormal(0.001, 0.01, 0,1)*
Testing requirement result	*test_req*	Boolean (True, False)	*if(reg_hpd>=p_h_testcond,"True","False")*
Regulator hazard per demand requirement	*reg_hpd_req*	Simulation (continuous interval)	*TNormal(0,0.001,0,1)*
Normal product use requirement result	*norm_req*	Boolean (True, False)	*if(reg_hpd>=p_h_normal_use,"True","False")*
Country of origin	*country_of_origin*	Ranked (Disreputable, Reputable, Highly Reputable)	*Disreputable: 0.33333334, Reputable: 0.33333334, Highly Reputable: 0.33333334*
Years in operation	*years_operating*	Ranked (<1 year, 1–5 years, 5–10 years, 10–20 years, 20 + years)	*< 1 year: 0.2, 1–5 years: 0.2, 5–10 years: 0.2, 10–20 years: 0.2, 20 + years: 0.2*
Customer satisfaction	*cust_sat*	Ranked (Very Low, Low, Medium, High, Very High)	*TNormal(m_quality,0.05, 0, 1)*
Product design	*prod_design*	Ranked (No change, Minor improvements, Major improvements)	*TNormal(m_quality,0.05, 0, 1)*
Manufacturer process quality	*m_quality*	Ranked (Low, Normal, High)	*TNormal(wmean(1.0, years_operating,2.0,country_of_origin),0.001,0,1)*
Probability of hazard per demand for normal product use given process information	*p_h_normal_use*	Simulation (continuous interval)	*Partitioned Expression (Low:p_h_strat*1.1, Normal: p_h_strat, High: 0.9*p_h_strat)*
Particular product usage deviations from intended use	*prod_usage*	Labelled (Used as intended, Minor deviations, Major deviations)	*Used as intended: 0.9, Minor deviations: 0.07, Major deviations: 0.03*
Probability of hazard per demand for a particular product usage	*p_h_usage*	Simulation (continuous interval)	*Partitioned Expression (Used as intended: p_h_normal_use, Minor deviations: p_h_normal_use + 0.1*p_h_normal_use, Major deviations: p_h_normal_use + 0.5*p_h_normal_use)*
Number of demands in particular product lifetime	*demands*	Simulation (integer interval)	*TNormal(100, 1000, 0, 1E8)*
Probability of hazard during product lifetime	*p_h_demands*	Simulation (continuous interval)	*1.0-(1.0-p_h_usage)^demands*
Years in use	*years*	Simulation (continuous interval)	*TNormal(0, 10, 0, 30)*
Probability of hazard during product lifetime adjusted for wear	*p_h_wear*	Simulation (continuous interval)	min*(1.0,p_h_demands + p_h_demands*years^2.0/1000.0)*
Probability uncontrolled hazard causes a major injury	*p_uh_major*	Simulation (continuous interval)	*Uniform(0,1)*
Probability uncontrolled hazard causes a minor injury	*p_uh_minor*	Simulation (continuous interval)	*Uniform(0,1)*
Probability control stops injury	*p_control*	Simulation (continuous interval)	*Uniform(0,1)*
Probability hazard causes a major injury	*p_h_major*	Simulation (continuous interval)	*p_uh_major*(1.0-p_control)*
Probability hazard causes a minor injury	*p_h_minor*	Simulation (continuous interval)	*p_uh_minor*(1.0-p_control)*
Probability of major injury during product lifetime	*p_major_L*	Simulation (continuous interval)	*p_h_wear*p_h_major*
Probability of minor injury during product lifetime	*p_minor_L*	Simulation (continuous interval)	*p_h_wear*p_h_minor*
Number of product instances	*t_prod*	Simulation (integer interval)	*Uniform(0,1000000000)*
Total number of major injuries	*t_major*	Simulation (integer interval)	*Binomial(t_prod, p_major_L)*
Actual number of major injuries	*at_major*	Simulation (continuous interval)	*Partitioned Expression (Underreporting: t_major −0.25*t_major, Accurate: t_major, Overreporting: t_major + 0.25*t_major)*
Total number of minor injuries	*t_minor*	Simulation (integer interval)	*Binomial(t_prod, p_minor_L)*
Actual number of minor injuries	*at_minor*	Simulation (continuous interval)	*Partitioned Expression (Underreporting: t_minor −0.25*t_minor, Accurate: t_minor, Overreporting: t_minor + 0.25*t_minor)*
Probability of products causing injuries	*prod_injury*	Simulation (continuous interval)	*(t_major + t_minor)/t_prod*
Reporting accuracy	*rep_acc*	Labelled (Underreporting, Accurate, Overreporting)	*Underreporting: 0.04* *Accurate: 0.95* *Overreporting: 0.01*

(*continued on next page*)

Table 2 (*continued*)

Variable name	Abbrev	Node type	NPT
Probability of products causing a major injury	*prod_major_injury*	Simulation (continuous interval)	*t_major / t_prod*
Probability of product causing a minor injury	*prod_minor_injury*	Simulation (continuous interval)	*t_minor / t_prod*
Risk level	*risk_level*	Ranked (Very Low, Low, Medium, High, Very High)	*TNormal((*min*(1.0,100.0*(prod_major_injury + 0.5*prod_minor_injury)),0.001, 0,1)*
Utility	*util*	Ranked (Very High, High, Medium, Low, Very Low)	*Very High: 0.2, High: 0.2, Medium: 0.2, Low: 0.2, Very Low: 0.2*
Government intervention required given risk level	*gov_int_req*	Boolean(True, False)	*if(risk_level > 0.5,"True","False")*
Risk tolerability	*risk_toler*	Ranked (Very High (Acceptable), High (Acceptable), Medium (Tolerable), Low (Unacceptable), Very Low (Unacceptable))	*TNormal((wmean(2.0,risk_level,1.0,util)), 0.001, 0, 1)*
Government intervention required given risk tolerability	*gov_int_req2*	Boolean(True, False)	*if(risk_toler > 0.5,"True","False")*
Consumer perception of risk	*c_risk_per*	Ranked (No change, Change)	*TNormal((wmean(1.0,govt_int_ann,1.0,media_stories,1.0, warnings)), 0.001,0,1)*
Warnings	*warnings*	Ranked (No, Yes)	*No: 0.5, Yes: 0.5*
Negative media stories	*media_stories*	Ranked (Very Low, Low, Medium, High, Very High)	*Very Low: 0.2, Low: 0.2, Medium: 0.2, High: 0.2, Very High: 0.2*
Government intervention announced	*govt_int_ann*	Ranked (No, Yes)	*No: 0.5, Yes: 0.5*
Revised Risk tolerability given revised utility	*risk_toler2*	Ranked (Very High (Acceptable), High (Acceptable), Medium (Tolerable), Low (Unacceptable), Very Low (Unacceptable))	*TNormal((wmean(2.0,risk_level,1.0,util2)),0.001,0,1)*
Government intervention required given revised risk tolerability	*gov_int_req3*	Boolean(True, False)	*if(risk_toler2 > 0.5,"True","False")*
Revised Utility given consumer perception of risk	*util2*	Ranked (Very High, High, Medium, Low, Very Low)	*See* Table 3

Table 3
Revised Utility given consumer perception of risk NPT.

Consumer perception of risk	No change					Change				
Utility	Very High	High	Medium	Low	Very Low	Very High	High	Medium	Low	Very Low
Very High	*1*	*0*	*0*	*0*	*0*	*0.2*	*0*	*0*	*0*	*0*
High	*0*	*1*	*0*	*0*	*0*	*0.4*	*0.2*	*0*	*0*	*0*
Medium	*0*	*0*	*1*	*0*	*0*	*0.3*	*0.4*	*0.2*	*0*	*0*
Low	*0*	*0*	*0*	*1*	*0*	*0.1*	*0.3*	*0.7*	*0.5*	*0*
Very Low	*0*	*0*	*0*	*0*	*1*	*0*	*0.1*	*0.1*	*0.5*	*1*

Future work should use behavioral insights to determine consumer behavior during a particular product lifetime and their perception of risk given a government intervention (e.g., recall). This will inform the BN model and improve product risk estimates.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of Interest/Conflict of Interest

Norman Fenton and Martin Neil are Directors of Agena Ltd.

Acknowledgements

This work was supported by the UK Government Department for Business, Energy and Industrial Strategy, Office for Product Safety and Standards (OPSS) and Agena Ltd. The views expressed in this article are those of the authors exclusively, and not necessarily those of the UK Government Department for Business, Energy and Industrial Strategy, Office for Product Safety and Standards (OPSS).

Appendix

References

Agena Ltd. (2021). AgenaRisk: Bayesian Network Software. Retrieved from www.agenarisk.com

Balke, A., & Pearl, J. (1994). Probabilistic evaluation of counterfactual queries. *Proceedings of the National Conference on Artificial Intelligence, 1*, 230–237.

Berchialla, P., Scarinzi, C., Snidero, S., & Gregori, D. (2010). Adaptive Bayesian Networks for quantitative risk assessment of foreign body injuries in children. *Journal of Risk Research, 13*(3), 367–377. https://doi.org/10.1080/13658810903233419.

Berchialla, P., Scarinzi, C., Snidero, S., Gregori, D., Lawson, A. B., Lee, D., & MacNab, Y. (2016). Comparing models for quantitative risk assessment: An application to the European Registry of foreign body injuries in children. *Statistical Methods in Medical Research, 25*(4), 1244–1259. https://doi.org/10.1177/0962280213476167.

Constantinou, A. C., & Fenton, N. (2018). Things to know about Bayesian networks: Decisions under uncertainty, part 2. *Significance, 15*(2), 19–23. https://doi.org/10.1111/j.1740-9713.2018.01126.x.

European Commission. (2015). EU general risk assessment methodology (Action 5 of Multi-Annual Action Plan for the surveillance of products in the EU (COM (2013)76) (Vol. 2015-IMP-M). Retrieved from http://ec.europa.eu/DocsRoom/documents/17107/attachments/1/translations/

European Commission. (2018). Commission Implementing Decision (EU) 2019/417 of 8 November 2018 laying down guidelines for the management of the European Union Rapid Information System 'RAPEX' established under Article 12

of Directive 2001/95/EC on general product safety. *Official Journal of the European Union*, *2018*(November 2018). Retrieved from http://data.europa.eu/eli/dec/2019/417/oj

European Union. (2019). Risk Assessment for RAPEX. Retrieved from version 3.3.1 website: https://ec.europa.eu/consumers/consumer-safety/rag/#/screen/home

Fenton, N., & Neil, M. (2018). Risk assessment and decision analysis with bayesian networks. *In Risk Assessment and Decision Analysis with Bayesian Networks*. https://doi.org/10.1201/b21982.

Fenton, N., Neil, M., & Caballero, J. G. (2007). Using ranked nodes to model qualitative judgments in bayesian networks. *IEEE Transactions on Knowledge and Data Engineering, 19*(10), 1420–1432. https://doi.org/10.1109/TKDE.2007.1073.

ISO/IEC. Safety aspects — Guidelines for their inclusion in standards. , Iso/Iec Guide 51 § (2014).

Kim, J. H., & Pearl, J. (1983). Computational model for causal and diagnostic reasoning in inference systems Retrieved from. *Dl.Acm.Org, 1*, 190–193 https://dl.acm.org/doi/abs/10.5555/1623373.1623417.

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: Series B (Methodological), 50*(2), 157–194. https://doi.org/10.1111/j.2517-6161.1988.tb01721.x.

Maglogiannis, I., Zafiropoulos, E., Platis, A., & Lambrinoudakis, C. (2006). Risk analysis of a patient monitoring system using Bayesian Network modeling. *Journal of Biomedical Informatics, 39*(6), 637–647. https://doi.org/10.1016/j.jbi.2005.10.003.

Masmoudi, K., Abid, L., & Masmoudi, A. (2019). Credit risk modeling using Bayesian network with a latent variable. *Expert Systems with Applications*. https://doi.org/10.1016/j.eswa.2019.03.014.

Neil, M., Fenton, N., & Nielsen, L. (2000). Building large-scale Bayesian networks. *Knowledge Engineering Review, 15*(3), 257–284. https://doi.org/10.1017/S0269888900003039.

Pearl, J. (2009). Causality: Models, Reasoning, and Inference. 2nd edition. In *Cambridge University Press*. https://doi.org/10.1017/S0266466603004109.

PROSAFE. (2013). *Best Practice Techniques in Market Surveillance*. Retrieved from https://www.prosafe.org/index.php/best-practice/item/best-practices-techniques-in-market-surveillance

RPA. (2006). Establishing a Comparative Inventory of Approaches and Methods Used by Enforcement Authorities for the Assessment of the Safety of Consumer Products Covered by Directive 2001/95/EC on General Product Safety and Identification of Best Practices. Retrieved from https://rpaltd.co.uk/uploads/report_files/j497-consumer-products.pdf

Spohn, W. (2008). Bayesian Nets Are All There Is to Causal Dependence. *Causation, Coherence, and Concepts*, (June 2001), 99–111. https://doi.org/10.1007/978-1-4020-5474-7_4

Suh, J. (2017). Development of a Product Risk Assessment System using Injury Information in Korea Consumer Agency. *Journal of Digital Convergence, 15*(4), 181–190 https://doi.org/10.14400/jdc.2017.15.4.181.

Torres-Toledano, J. É. G., & Sucar, L. E. (1998). Bayesian networks for reliability analysis of complex systems. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 1484*, 195–206. https://doi.org/10.1007/3-540-49795-1_17.

Welch, R., & Thelen, T. (2000). Dynamic reliability analysis in an operational context: the Bayesian network perspective. *Dynamic Reliability: Future Directions*.

Wiegerinck, W., Burgers, W., & Kappen, B. (2013). Bayesian Networks, Introduction and Practical Applications. *Intelligent Systems Reference Library, 49*, 401–431. https://doi.org/10.1007/978-3-642-36657-4_12.

Joshua L. Hunte is a PhD student in the School of Electronic Engineering and Computer Science, Queen Mary University of London, London, UK. He is a member of the Risk and Information Management Research Group at the same university. His research interest includes Bayesian modelling, decision analysis, risk and safety assessment.

Professor Martin Neil is a Professor of Computer Science and Statistics in the School of Electronic Engineering and Computer Science, Queen Mary University of London, London, UK. He is a member of the Risk and Information Management Research Group at the same university. He is also a Director of Agena Ltd, a company that develop and distribute AgenaRisk, a software product for modelling risk and uncertainty. He is also a fellow of the Alan Turing Institute. His research interest covers Bayesian modelling and risk quantification in diverse areas.

Professor Norman Fenton is a Professor of Risk and Information Management in the School of Electronic Engineering and Computer Science, Queen Mary University of London, London, UK. He is also the Director of the Risk and Information Management Research Group at the same university. He is also a Director of Agena Ltd, a company that develop and distribute AgenaRisk, a software product for modelling risk and uncertainty. He is also a fellow of the Alan Turing Institute. His research interest covers Bayesian modelling and risk quantification in diverse areas.