Development of food fraud media monitoring system based on text mining

Y. Bouzembrak[a,*], B. Steen[a], R. Neslo[b], J. Linge[c], V. Mojtahed[d], H.J.P. Marvin[a]

[a] RIKILT Wageningen University and Research, Akkermaalsbos 2, 6708 PD, Wageningen, The Netherlands
[b] Julius Centre for Health Sciences and Primary Care, University Medical Centre Utrecht, Utrecht, The Netherlands
[c] European Commission – Joint Research Centre (JRC), Ispra, VA, Italy
[d] Fera Science Ltd., York, YO41 1LZ, UK

ABSTRACT

Food fraud is receiving considerable attention with the growing body of literature that recognises its importance. No system exists that collects media reports on food fraud. In this study, we used the infrastructure provided by the European Media Monitor (EMM), in particular it's MedISys portal for this purpose. We developed a food fraud tool (MedISys-FF) that collects, processes and presents food fraud reports published world-wide in the media. MedISys-FF is updated every 10 min 24/7. Food fraud reports were collected with MedISys-FF for 16 months (September 2014 to December 2015) and benchmarked against food fraud reports published in Rapid Alert for Food and feed (RASFF), Economically Motivated Adulteration Database (EMA) and HorizonScan. The results showed that MedISys-FF collects food fraud publications with high relevance > 75% and the top 4 most reported fraudulent commodities in the media were i) meat, ii) seafood, iii) milk and iv) alcohol. These top stories align with those found in RASFF and EMA but differences in frequency are apparent. Analysis of the collected articles can help understanding food fraud issues in the origin countries and can facilitate the development of control measures and to detect food fraud in the food supply chain.

1. Introduction

It is accepted within the EU that food fraud covers cases where there is a violation of EU food law, which is committed intentionally to pursue an economic or financial gain through consumer deception. Food fraud in the food supply chain can arise as a result of misrepresentation associated with: product integrity (e.g. counterfeit product, expiration date), process integrity (e.g. diversion of products outside of intended markets), people integrity (e.g. characterizations such as the cyber criminals and hacktivist) and data integrity (e.g. improper, expired, fraudulent or missing common entry documents or health certificates) of information accompanying the food item throughout the supply chain (Manning & Soon, 2018; Manning, 2016).

Food fraud incidents have been reported in many countries with a direct link to public health problems and even deaths. In Norway in 2002–2004, a liquor adulteration with methanol killed 9 people and 51 people were admitted to hospital with methanol poisoning. The liquor responsible for these cases contained 20% methanol and 80% ethanol and probably came from the same producer in southern Europe (Hovda et al., 2005). In China in 2008, melamine was used to adulterate protein levels in infant powder milk produced locally in China. The

adulteration resulted in illness in 294 000 individuals, in hospitalisation of 50 000 infants and in six deaths (Domingo, Tirelli, Nunes, Guerreiro, & Pinto, 2014; Liu, Liu, Zhang, & Gao, 2015; Qiao, Guo, & Klein, 2010; Xiu & Klein, 2010; Zhang & Xue, 2016). In Europe in 2013, several EU countries found traces of horsemeat in products fraudulently labelled as beef (O'Mahony, 2013; Stanciu, 2015). The Rapid Alert System for Food and Feed (RASFF) (RASFF, 2017) was used to exchange relevant information during this incident and to react quickly to protect consumers. Traceability checks began on the same day and the European Commission (EC) started an implement action plan to fight against food fraud to strengthen the European Union (EU) system and to restore consumer confidence (EU, 2013).

1 Recently, various developments have taken place in the field of food fraud systems to understand and document food fraud incidents. In the UK, HorizonScan database was created by Fera.[1] HorizonScan focuses on global food and feed integrity issues such as incidents related to adulteration, substitution and fraud, as well as microbial contaminants, allergens, pesticides and drug residues. In EU, the Food Fraud Network (FFN) was established to handle requests for cross-border cooperation and to ensure the rapid exchange of information between national authorities and the Commission in cases of suspected

---

fraudulent practices. In 2017, 597 cases have been exchanged by the FFN (EC, 2018) using a dedicated IT system, the Administrative Assistance and Cooperation (AAC) system. In the USA, other databases have been created, such as the Economically Motivated Adulteration database (EMA) (EMA, 2017) and the USP food fraud database (USP, 2018). EMA database contains food fraud incidents since 1980 and is housed at the National Center for Food Protection and Defence. The database provides information about the food product, fraud incident year, adulterant, type of fraud, health consequences, country of origin and how the incident was discovered (Bouzembrak & Marvin, 2016; Marvin et al., 2016). Food fraud systems have a very different fraud type classification when compared to each other. RASFF includes both intentional and unintentional food fraud notifications, which are classified in six different fraud types (Table 1): improper, fraudulent, missing or absent health certificates; illegal importation; tampering; improper, expired, fraudulent or missing common entry documents or import declarations; expiration date; and mislabelling.

EMA database proposes nine different types of food fraud (Table 1), of which the most important are: substitution, artificial enhancement, dilution, transhipment, counterfeit, and misbranding. HorizonScan database includes mainly six types of food fraud (Table 1): adulteration/substitution, fraudulent health certificate/documentation, produced without inspection, unapproved premises, expiry date changes, unauthorised/unsuitable transport.

None of the above mentioned systems provide cases of food fraud reported in media globally and it is advocated that this information may be a useful additional intelligence source that will help the authorities and industries to design their control strategies to increase the chance of detecting fraudulent activities. Therefore, the aim of this study was to develop this extra information source by exploiting a powerful infrastructure that is already in place (i.e. the MedISys portal of the European Media Monitor (EMM)). MediSys[2] is a media monitoring system providing event-based surveillance to rapidly identify potential public health threats using information from media reports. MedISys is a text mining system that continuously monitors about 900 specialist medical sites plus all generic EMM news, i.e. over 20 000 RSS feeds and HTML pages sites from 7000 generic news portals and 20 commercial news wires in altogether 70 languages. This system is used by a number of Health Agencies, including ECDC,[3] EFSA[4] and WHO[5] (Mantero, Cox, Linge, van der Goot, & Coulombier, 2010; Rortais, Belyaeva, Gemo, van der Goot, & Linge, 2010). However, MedISys does not collect media reports on food fraud. Therefore, we developed a filter in MedISys dedicated for food fraud (MedISys-FF). The developed MedISys-FF collects food fraud reports in the media worldwide in eight different languages and was tested for 16 months (September 2014 to December 2015). In this period all reports collected by MedISys-FF were retrieved and compared to food fraud reports published in three food fraud systems, which are: RASFF, EMA, and HorizonScan. It was concluded that the newly developed system collects complementary information and therefore may be a useful additional intelligence source to combat food fraud.

2. Material and methods

In this section, we present the development steps of the MedISys-FF and introduce three food fraud systems (RASFF, EMA, and HorizonScan).

Table 1
Example food fraud categorizations in RASFF, EMA and HorizonScan.

RASFF	EMA	HorizonScan
• Improper, fraudulent, missing or absent health certificates	• Mislabelling	• Adulteration/ substitution
• Illegal importation	• Artificial enhancement	• Fraudulent health certificate/ documentation
• Tampering	• Substitution	• Produced without inspection
• Improper, expired, fraudulent or missing common entry documents or import declarations	• Dilution	• Unapproved premises
• Expiration date	• Transhipment/origin masking	• Expiry date changes
• Mislabelling	• Counterfeit	• Unauthorised/ unsuitable transport
	• Intentional distribution of contaminated product	
	• Substitution with a non-food-grade substance	
	• Theft and resale	

2.1. MedISys-FF tool

The construction of the food fraud filter consists of the following steps: (i) identification of food fraud keywords, (ii) validation of keywords by experts in food fraud, (iii) development of the MedISys food fraud filter, (iv) and evaluation and improvement of the filter (Fig. 1).

2.1.1. Step 1: Identification of food fraud keywords

A literature research was carried out to obtain an overview of the most recent trends in food fraud and to identify the list of food fraud keywords. Several digital libraries were used to search for scientific articles in food fraud, such as Scopus, Science Direct, Springer Link, and Google Scholar. In addition, food fraud databases were analysed to determine the keywords, such as the Rapid Alert System for Food and Feed (RASFF)[6] and the Economically Motivated Adulteration database (EMA)[7] and HorizonScan.

2.1.2. Step 2: Validation of food fraud key words by experts

Three senior experts with long term experience in this topic from the Netherlands and USA validated the identified keywords. The validation resulted in 59 English food fraud keywords, which are listed in the following table (Table 2).

All the keywords cited in Table 2 were translated into 8 different languages to collect articles in various languages. The following languages were used: Arabic, Chinese, Dutch, French, German, Italian, Portuguese, and Spanish. This translation resulted in 531 food fraud keywords.

2.1.3. Step 3: The design of MedISys food fraud filter

The keywords method was used to create the food fraud filter by using the defined list of keywords (Table 2) and the keywords that should be excluded, such as artificial food, how to make fake milk, how to make fake alcohol, and counterfeit medicine. This keywords method was preferred because of its effectiveness and precision. If a keyword consisted of two or more words, the symbol "+" was used. For example, the keyword "food fraud" was converted to "food + fraud", which means that the words "food" and "fraud" should appear together and in the same order in the article.

2.1.4. Step 4: Evaluation and improvement

The newly built filter with keywords related to food fraud was set operational for 2 weeks. All articles collected by the filter in this period

---

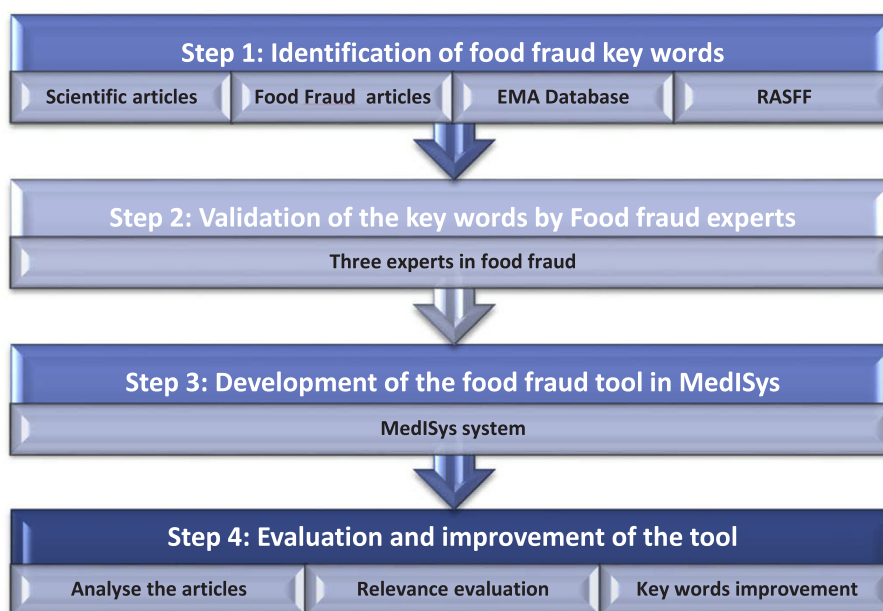**Fig. 1.** The design steps of the food fraud filter.

Table 2
List of English food fraud keywords.

English Keywords		
caviar fraud	false food	mislabelled milk
coffee adulteration	fish fraud	mislabelled olive oil
coffee fraud	food adulteration	mislabelled saffron
counterfeit coffee	food fraud	mislabelled wine
counterfeit fish	formaldehyde fish fraud	misleading of food
counterfeit food	honey fraud	misleading of food
counterfeit honey	honey mislabelling	mozzarella fraud
counterfeit meat	intentional substitution of food	olive oil dilution
counterfeit milk	juice adulteration	olive oil fraud
counterfeit olive oil	juice fraud	saffron adulteration
counterfeit olive oil	meat adulteration	saffron fraud
counterfeit orange juice	meat fraud	salmon fraud
counterfeit saffron	melamine in milk	seafood fraud
counterfeit wine	milk adulteration	smuggling food
economic adulteration food	milk fraud	spices fraud
fake alcohol	mislabelled coffee	tampering of food
fake food	mislabelled fish	tampering of food
fake juice	mislabelled food	wine adulteration
fake milk	mislabelled juice	wine fraud
fake saffron	mislabelled meat	

were retrieved from the hyperlink provided by the system and assessed for relevance to 'food fraud'. The relevance was about 30%. An article was considered relevant if it describes an event about food *and* fraud. To improve the relevance of the filter, the set of keywords were adapted and tested for relevance after another week. The relevance had increased to 67%. This process was repeated twice until a stable level of 80% relevance was reached. Once this optimal performance was achieved, the keywords combinations were kept constant and the food fraud articles were collected from 15/09/2014 to 31/12/2015. The collected articles are presented on the MedISys filter as a hyperlink to the original publication including a short summary (in the original languages and translated into English) generated by the system itself. This link and summary stays on the filter as long the original article is online. Once it is removed by its publisher, the link and summary disappears. In order to be able to evaluate the articles found by the

MedISys filter for a long time period it is preferable to extract these from the filter and store them in an own (not publically available) database. To this end the RSS feed functionality of MedISys was used to automatically extracted the articles (hyperlink and summary) and store them in a local database using MongoDB and Elasticsearch (see Fig. 2).

2.2. Food fraud reporting systems

2.2.1. Rapid Alert for Food and Feed (RASFF)

RASFF portal is a dynamic online alert tool that aims to rapidly share any information concerning health risks derived from food or feed (RASFF, 2017) between the European Commission (EC) and the control authorities of Member States (MS).

RASFF ensures that urgent notifications are sent, received and responded collectively and efficiently. RASFF gives public access to summary information about the most recently transmitted RASFF notifications and also allows users to search for any notification issued in the past. This portal is compiled from many sources: i) official control on the market, ii) border control, iii) and consumer complaint. Notification details can include the notification type, the basis of notification, the product in question, the countries involved, the action taken, the distribution status, and the type of hazard identified. All fraud/adulteration notifications under the product category food reported in RASFF in the period 01/01/2000 to 31/12/2015 were extracted and analysed.

2.2.2. Economical Motivated Adulteration database (EMA)

The EMA database focuses on the intentional adulteration of food for economic gain or for food fraud. This database is compiled from literature and media searches of EMA incidents in food products since 1980. Sources include LexisNexis, PubMed, Google, FDA Consumer and FDA recall records, state reports, and RASFF portal. Food fraud reports collected by EMA in the period 01/01/2000 to 12/2015 were analysed in this study.

2.2.3. HorizonScan

HorizonScan focuses on global food and feed integrity issues and food fraud is one of the food safety hazards. Food fraud incidents
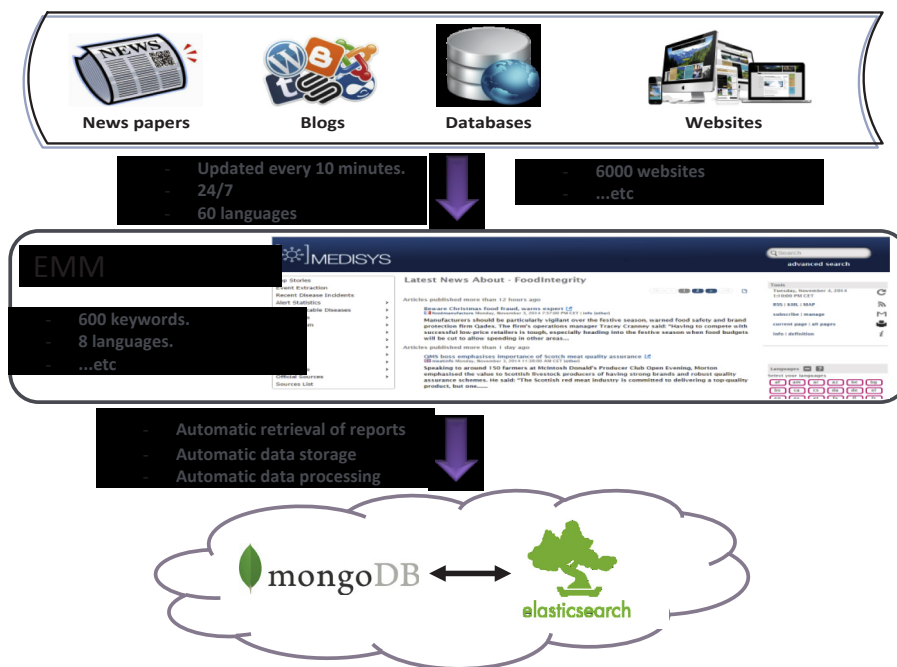
Fig. 2. The structure of MedISys-FF.

recorded in the database are related to adulteration, substitution and fraud. HorizonScan tracks more than 536 commodities in 202 countries from official websites of more than 65 countries with more than 100 data sources that are scanned daily. Every day on average more than 30 new records are added to the database. All food fraud reports collected by HorizonScan in the period 01/01/2000 to 31/12/2015 were extracted and analysed.

3. Results

3.1. MedISys-FF

3.1.1. Food fraud publications (2014–2015)

In the period from 15/09/2014 to 31/12/2015, 1144 newspapers articles on food fraud were automatically collected by MedISys-FF. The relevant articles (851 articles) were assessed manually and the irrelevant ones (293 articles) were removed. The percentage of articles reporting a food fraud incident was 58% of the relevant articles (Fig. 3).

3.1.2. Product category and origin country

The relevant articles were further analysed for the following characteristics: (i) the food products reported as fraudulent, (ii) the type of fraud detected, and (iii) the country in which the media report appeared. The main topics in these media articles were food fraud in general followed by fraud with meat and fraud with seafood (Table 3). Food fraud media articles were collected from 78 countries. The number articles per country is shown in Fig. 4. The countries with the highest number of food fraud articles in the media in the period
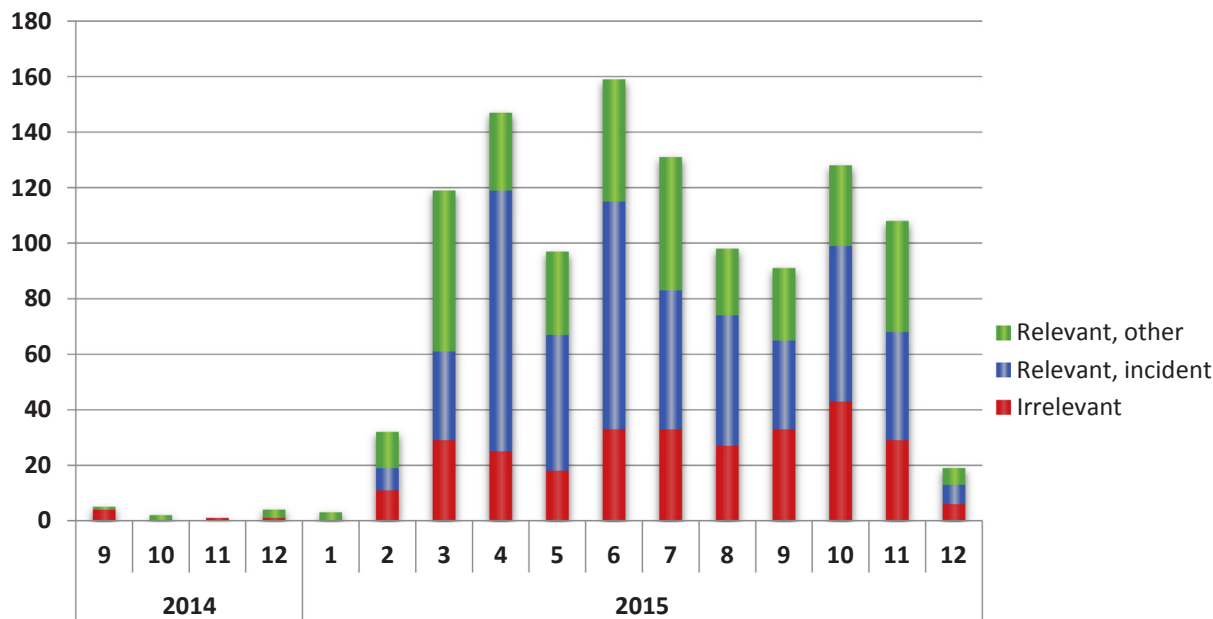
Fig. 3. Number of articles collected with MedISys-FF (2014–2015).

Table 3
Number of times food or specific types food products were reported as fraudulent in media articles collected by MedISys-FF (2014–2015).

Products	Number of articles	Percentage (%)
Food	352	41
Meat	185	22
Seafood	91	11
Milk	71	8
Alcohol	39	5
Poultry	33	4
Other products	30	4
Cheese	16	2
Olive Oil	20	2
Honey	10	1
Coffee	4	< 1
Total	851	100

Legend
- no articles
- 1-5 articles
- 6-10 articles
- 10-25 articles
- over 25 articles with number of articles

Fig. 4. An overview of published articles on food fraud collected by MedISys-FF.

investigated are Egypt (94 articles), United States (85 articles), United Kingdom (84 articles) and Saudi Arabia (73 articles). However, we did not succeed in collecting food fraud publications from several countries in Latin America (e.g., Argentina, Chile, Uruguay, Peru, and Bolivia) or in Africa (e.g., Angola, Zim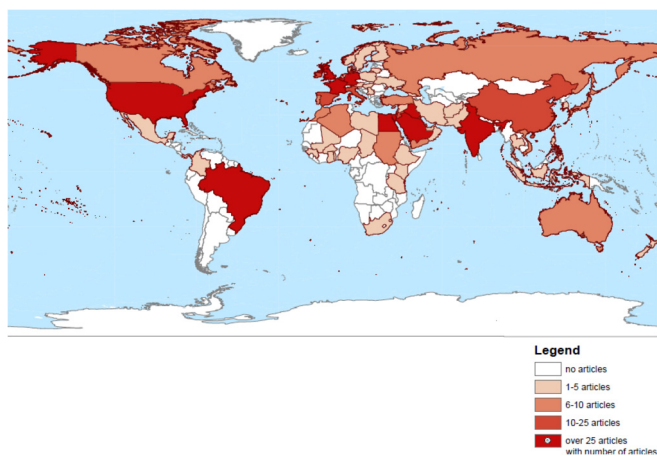babwe, Mozambique, and Congo-Kinshasa). This may be due to the lack of geographical locations in the articles, language barriers or no media coverage on food fraud in the analysed period. The vast majority of the media reports (> 95%) dealt with food fraud conducted in the country of publication. Interesting difference between countries was observed. For example, in US the most publications deals with fraud related to seafood whereas in United Kingdom reported fraud is mainly related to alcohol and meat (Appendix Table A1).

3.1.3. Types of food fraud

The type of food fraud used in the various system reporting food fraud varies (see Table 1) and clearly there is a need for harmonisation. For the purpose of this study, we used the classification of food fraud based on RASFF and defined by Bouzembrak and Marvin (Bouzembrak & Marvin, 2016) and are shown in Table 4. However, not all food fraud reports collected by MedISys-FF could be classified in this way and therefore we added three other categories being: food fraud (in general,

Table 4
Types of food fraud and the topic in articles collected with MedISys-FF (9/2014 to 12/2015).

Type of food fraud and article topic	Number of articles	%
Food fraud	235	28%
Expired date	194	23%
Mislabelling	106	13%
Regulation	106	13%
Illegal importation	81	10%
Tampering	76	9%
Disrespect of regulation	43	5%
HC	7	1%
Total	848	100%

hence not specified), regulation and disrespect of regulation. The distribution of the food fraud reports collected by MedISys-FF over these categories is shown in (Table 4). Most of the reports were on food fraud in general (28%), followed by expired food (23%), mislabelling (13%) and regulation (13%).

3.2. RASFF

3.2.1. Food fraud notifications (2000–2015)

All RASFF notification under the hazard category fraud/adulteration from 01/01/2000 to 31/12/2015 were collected and analysed. In total, 918 food fraud notifications were registered (Fig. 5).

Fig. 5 shows that the number of food fraud notifications recorded in RASFF increased from 3 notifications in 2000 to 91 notifications in 2015, with a peak of 158 notifications in 2013. The food fraud notifications comprised only a small proportion (2.2%) of the total notifications in the RASFF reported between 2000 and 2015 (i.e. 41 909 notifications). A detailed description of these reports is shown in the Appendix (Table A4).

3.2.2. Product category and origin country

The product categories of the food fraud notification and their frequencies are shown in Table 5. In the analysed period (2000–2015), more than 50% of food fraud notifications were reported in four product categories: nuts products and seeds (17.4%), meat and meat products (17.1%), fish and fish products (9.80%), and fruits and vegetables (8.3%) (Table 5).

The number of fraud notifications related to meat and meat products increased from 1 notification in 2000 to 59 notifications in 2013. This is partly explained by the horse meat scandal in Europe in 2013. Another example is the increase of nuts products and seeds fraud notifications from 1 notification in 2001 to 40 notifications in 2015 (Appendix Table A2).

3.2.3. Types of food fraud

The next analysis was devoted to the type of food fraud recorded in RASFF. Six food fraud categories are recorded (Bouzembrak & Marvin, 2016): missing of health certificate (HC), illegal importation, tampering, missing of common entry document (CED), expired date, and mislabelling, as described in Table 6. The HC category represented 40% of the notifications in the database, followed by illegal importation (28%) and tampering (24%).
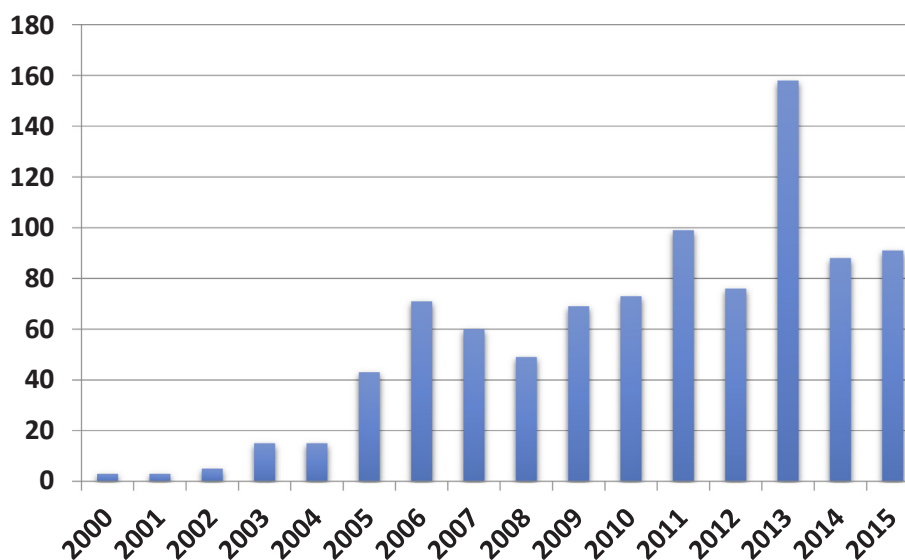
Fig. 5. Food fraud notifications in RASFF (2000–2015).

Table 5
The main product categories reported in RASFF related to food fraud (2000–2015).

Product categories	Number of notifications	Percentage (%)
Nuts, nut products and seeds	160	17.4
Meat and meat products	157	17.1
Fish and fish products	90	9.8
Fruits and vegetables	76	8.3
Food products/mixed	55	6.0
Prepared dishes and snacks	50	5.4
Cereals and bakery products	35	3.8
Bivalve molluscs and products thereof	34	3.7
Poultry meat and poultry meat products	34	3.7

Table 6
Number of notifications per food fraud type in RASFF (2000–2015).

Fraud type	Number of notifications	Percentage (%)
HC	436	40
Illegal importation	304	28
Tampering	260	24
CED	53	5
Expired Date	28	3
Mislabelling	19	2
Total number	1100	100

3.3. EMA database

3.3.1. Food fraud incidents (2000–2015)

In line with the RASFF results, the number of food fraud incidents recorded in EMA database increased from 2 cases in 2000 to 38 cases in 2015, with peaks in 2011 (147), 2012 (117) and 2013 (136) (Fig. 6).

3.3.2. Product categories and origin country

The main food categories and their frequencies are recapitulated in Table 7. In the period analysed (2000–2015), 642 distinct food fraud incidents were recorded dealing with the following products: fish and seafood (26.8%), fruits and vegetables (14.8%), dairy (10.0%), and meat (9.8%). These food categories represented more than 60 percent of the food fraud incidents records in the database. The complete table can be seen in appendix (Table A3).

3.3.3. Types of food fraud

Table 8 provides a classification of EMA incidents by type of fraud: mislabelling (28%), artificial enhancement (16%), substitution (15%) and dilution (12%) (Appendix Table A5).

3.4. HorizonScan

3.4.1. Food fraud incidents (2000–2015)

There are no records of food fraud prior to 2006 in HorizonScan. In the period of 2006–2015, the number of recorded food fraud incidences increased from 7 to 377 with the peak being in 2015 (377) (see Fig. 7).

3.4.2. Product categories and origin country

In the period 2000–2015, 1515 incidents were reported for 149 distinct food commodities. The main food categories and their frequencies are recapitulated in Table 9. The 10 reported commodities in Table 9 account for 55% of food fraud incidences. Each of the remaining commodities constitutes less than 2% of total reported commodities subject to fraud.

3.4.3. Types of food fraud

Table 10 provides the fraud categories of HorizonScan database which are adulteration (60%), fraudulent health certificates (20%), produced without inspection (11%), unapproved premises (4%) and changing expiry date (4%).

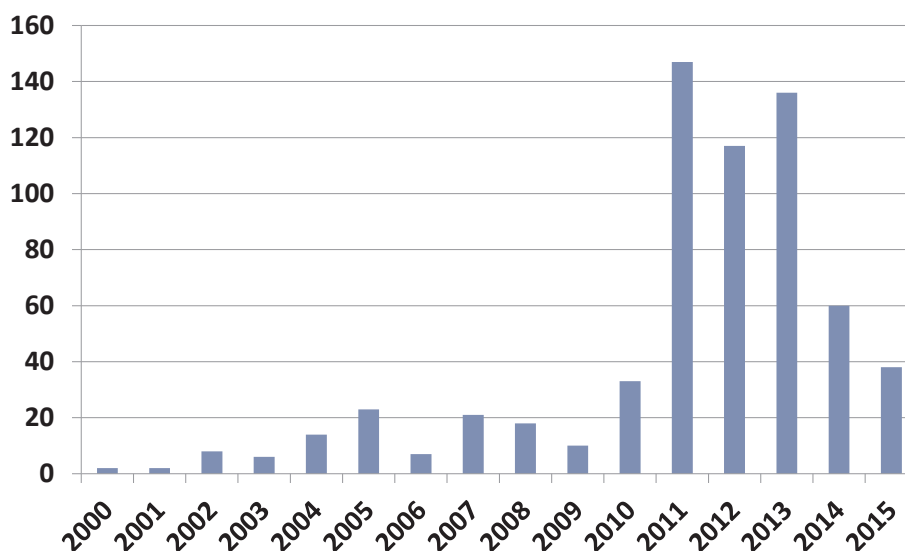**Fig. 6.** EMA Food fraud incidents (2000–2015).

Table 7
The main product categories reported in EMA (2000–2015).

Product categories	Number of notifications	Percentage (%)
Fish and seafood	172	26.8
Fruits and vegetables	95	14.8
Dairy	64	10.0
Meat	63	9.8
Oils and fats	44	6.9
Alcoholic	42	6.5
Honey	30	4.7
Herbs and spices	26	4.0
Grain products	25	3.9

Table 8
Types of fraud in EMA database (2000–2015).

Fraud type	Number of incidents	Percentage (%)
Mislabelling	177	28
Artificial enhancement	104	16
Substitution	97	15
Dilution	76	12
Dilution with a non-food-grade substance	57	9
Transhipment/origin masking	45	7
Counterfeit	41	6
Intentional distribution of contaminated product	19	3
Substitution with a non-food-grade substance	15	2
Theft and resale	11	2
Total number	642	100

3.5. Comparison of food fraud cases in MedISys-FF, RASFF, EMA and HorizonScan in the period September 2014 to December 2015

In this section, the result of MedISys-FF is compared with three existing food fraud systems that collect and report food fraud issues being: RASFF, EMA and HorizonScan (Table 11).

The origins of the reports presented in the four different systems range from media publications (MedISys-FF) to official notifications from authorities (RASFF) and this is also reflected in the commodity most frequently found, the type of fraud reported and the country of origin of this fraudulent product. Table 11 shows for the same time span (September 2014–December 2015), the four most frequent reported commodities, type of fraud and the origin country of the fraudulent product. Note that in the table these three characteristics not necessarily are linked. Interestingly, the two systems that collect food fraud reports from authorities (RASFF and HorizonScan) have different commodities in the four most frequently reported ones. In addition, the top four countries of origin of the fraudulent product are different between the two systems. Note that for this study the reports in HorizonScan that originate from RASFF could be identified and were removed. This was not possible for the EMA data and therefore overlap due to duplication will exist between RASFF and EMA. MedISys-FF and EMA both have fish products in common with RASFF. Overlap between EMA and RASFF is expected since EMA also collect food fraud reports from the RASFF. In the period analysed RASFF reported 90 fraud cases with seafood which accounts for almost 50% of the fraud cases in seafood in the EMA database. Such duplication does not exist between the MedISys-FF and RASFF.

The main countries from which food fraud articles were collected from the media by MedISys-FF are Egypt (11%), United States (10%), United Kingdom (10%) and Saudi Arabia (9%). For RASFF the main countries from which products were imported into EU and that were found to be fraudulent were China (28%), India (12%), Turkey (11%), and Nigeria (11%). EMA reported food fraud in this period mainly from India (22%), United States (18%) and China (11%). HorizonScan reported food fraud from the Czech Republic (27%), the United States (8%), Moldova (5%), and the Slovak Republic (4%).

There were other differences between the systems as well. There seems to be differences in the type of fraud reported of the fraudulent products (Table 11). In RASSF, the most common food fraud types were: HC (63%), illegal importation (25%), CED (10%), and expired date (1%). In EMA, the most important food fraud types recorded were: tempering (73%), origin labelling (22%), and theft and resale (4%). The food fraud types in MedISys-FF were: food fraud not specified (28%), expired food (23%), mislabelling (13%) and regulation (13%). In HorizonScan, the most common food fraud types were adulteration/
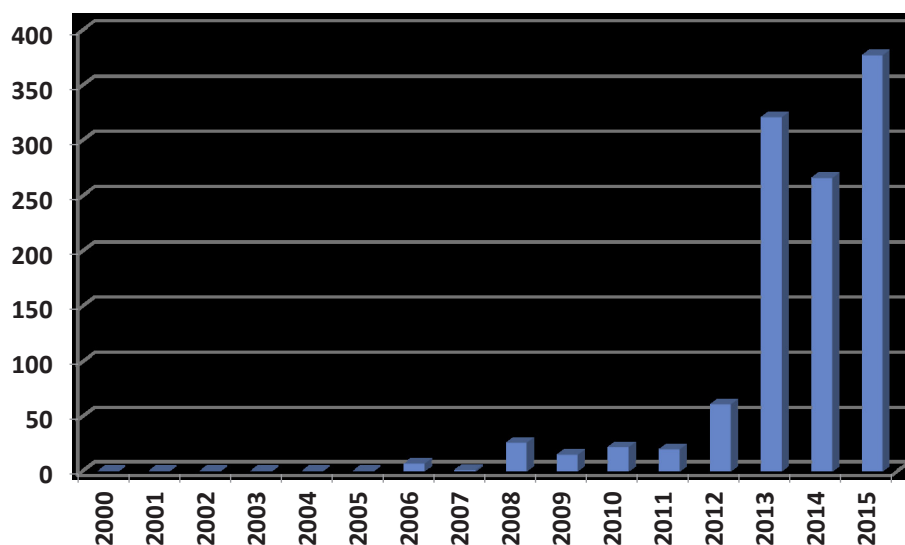
Fig. 7. HorizonScan Food Fraud incidents (2000–2015).

Table 9
The main product categories reported in HorizonScan (2000–2015).

Product categories	Number of records	Percentage (%)
Wines (red, white, rose, sparkling, fortified)	261	17
Meat products and meat preparations - other	130	9
Honey	108	7
Other prepared foods	52	3
Spirits and liqueurs	51	3
Ware potatoes	49	3
Milk products - cheese	49	3
Meat products - sausages (fuet, chorizo, salame etc.)	49	3
Fish - fresh/frozen	47	3
Offal - bovine (beef, veal, cow)	36	2

Table 10
Types of fraud in HorizonScan database (2000–2015).

Description	Number of notifications	Percentage (%)
Adulteration/substitution	914	60
Fraudulent health certificate/documentation	302	20
Produced without inspection	162	11
Unapproved premises	72	4
Expiry date changes	63	4
Unauthorised/unsuitable transport	2	1
Total number	1515	100

substitution (72%), fraudulent health certificates (11%), produced without inspection (10%), and expiry date change (4%).

4. Discussions

In this study, a tool has been developed that specifically collects articles on food fraud from the media world-wide and the collected articles were compared with three existing systems that report on food fraud. The four systems analysed in this study showed little overlap regarding the fraud incidents they report.

Such differences may be due the non-consistency in the systems (i.e. products categories, type of fraud, origin country), trade between countries (e.g. EU countries are not importing milk from India or Brazil), relatively short period used for this comparison, lack of information (e.g., EU cases of food fraud are missing in RASFF). Furthermore, it may be partly due to the origin of the data and the purpose of the data bases. RASFF is a European portal that is maintained by the European Commission. The objective of RASFF is to provide food and feed control authorities within EU, including Norway, Liechtenstein, Iceland and Switzerland with an effective tool to exchange information about measures taken responding to serious risks detected in relation to food or feed. HorizonScan, on the other hand, collects reports from the websites of food safety authorities (worldwide) and obviously authorities not necessarily report on their website cases that are also disclosed in RASFF since these often deals with incompliance in documentations. MedISys-FF collects specifically reports on food fraud that are discussed in the media. The majority of these reports deals with food fraud in the country where the media report appears and generally does not involve discussions on fraudulent documentations, such as mainly reported in RASFF. Based on the results, RASFF seems to tackle most of the cases originating from outside the borders of the EU by mainly preventing document fraud, such as HC, illegal importation or improper or missing documentation (Tähkäpää, Maijala, Korkeala, & Nevas, 2015). Furthermore, a bias is expected in the RASFF data since often the testing is risk-based and is performed by the border inspections on food products imported into the EU. The focus is on fraud that may impact human health and less on fraud that has only economic effects (cheating). However, the latter may receive much media attention which is picked up by MedISys-FF.

As determined by the newly developed food fraud tool, seafood in the US is the most often reported fraudulent product (Appendix Table A1). Fraud in seafood in US has been reported in the literature as well. Manning and Soon (2014) published a report on the types of food fraud incidents identified in US restaurants and retail outlets in 2012 and concluded that 58% of the sampled retail outlets (81 retail outlets) sold mislabelled fish with small markets having a higher incidence of fraud (40%) than national chain grocery stores (12%). Furthermore, all sushi bars sampled (16 bars) sold mislabelled fish and 94% of the tuna tested

Table 11
The most common food fraud cases reported in MedISys-FF, RASFF, EMA, and HorizonScan (2014–2015).

Systems	Product category		Country of origin		Type of fraud	
	Product	%	Country	%	Fraud type	%
MedISys-FF	Food	41%	Egypt	11%	Food fraud	28%
	Meat	22%	United States	10%	Expired food	23%
	Seafood	10%	United Kingdom	10%	Mislabelling	13%
	Milk	8%	Saudi Arabia	9%	Regulation	13%
RASFF	Nuts and seeds	38%	China	28%	HC	63%
	Fruits and vegetables	16%	India	12%	Illegal importation	25%
	Fish and fish products	8%	Turkey	11%	CED	10%
	Mixed food	6%	Nigeria	11%	Expired Date	1%
EMA	Meat	27%	India	22%	Tampering	73%
	Dairy	20%	United States	18%	Origin labelling	22%
	Alcoholic	9%	Unknown	13%	Theft and resale	4%
	Oils fats	8%	China	11%	–	–
HorizonScan (Non RASFF)	Wines	20%	Czech republic	27%	Adulteration/substitution	72%
	Meat	8%	United States	8%	HC	11%
	Honey	7%	Moldova	5%	Produced without inspection	10%
	Ware potatoes	5%	Slovak Republic	4%	Expired date	4%

was not tuna at all (Manning & Soon, 2014). In India in 2011, a survey of adulteration in liquid milk (1791 samples) found that 68% of the randomly collected samples tested were fraudulent milk. In urban areas, 69% of milk sold was tampered milk (detergent addition to milk (8%); skimmed milk powder (45%) and glucose addition to milk (27%)). In seven Indian states all samples taken were found to be impure. Interestingly, fraud in milk apparently continued to be an issue in India since the MedISys-FF detected high number of media articles on milk fraud in India in the period analysed (2014–2015) (see Appendix Table A1). These two cases demonstrate that MedISys-FF can contribute to the identification of food fraud issues occurring in a specific country in the world.

Food fraud is reported in various systems but have, as discussed in this paper, different objectives and uses food fraud classifications and product categories which are not harmonised. Therefore, a full picture of the extent of fraud (e.g. actual level of food fraud) is difficult to obtain. This is further hampered by the often risk-based approach used to by the authorities to select the samples to be tested. Nevertheless, trends in fraudulent products observed in these systems may be useful for the development of predictions systems especially when fraudulent findings are linked to economic and/or human behaviour factors. Linking food fraud reports of RASFF and EMA to economic and other data using Bayesian Networks, Marvin et al., 2016 (Marvin et al., 2016) could predict the type of fraud with more than 80% accuracy demonstrating the practical use of the current data.

The newly developed MedISys-FF system collects information on food fraud cases mentioned in the media and therefore potentially provides additional information since the other systems do not cover the media. Some media coverage of food fraud may be driven by publications of the authorities and therefore picked up by systems that scan the authorities' websites such as the HorizonScan. However, media reports on food fraud due to research of newspapers themselves will not be found by HorizonScan. Differences in technology used in HorizonScan and MedISys-FF could also cause differences in detection

of reports between both systems. A systematic comparison over a long period (2–3 years) could help to clarify this point.

5. Conclusions

Within MedISys a food fraud tool has been constructed that collects food fraud media reports world-wide every 10 min 24/7 and therefore provides an actual overview of media articles on food fraud on a global level.

The accuracy of the newly developed tool MedISys-FF was high (75%). The most reported fraudulent commodities in the media were i) meat, ii) seafood, iii) milk and iv) alcohol. The analysis of these articles can facilitate the development of control measures to protect the food supply chain. Furthermore, the comparison performed in this study revealed a lack of consistency in the terminology between the data sources studied (MedISys-FF, RASFF, EMA and HorizonScan). For instance, there is no clear definition or presentation of food fraud types or product categories in EMA and RASFF. This type of information is essential to detect fraud, to improve control policies and programmes, and to evaluate any actions taken.

The results shows that the media based food fraud filter adds to the current systems in place and that it may be a useful input source for quality managers and food safety authorities to inform their control programmes.

Appendix

Table A1
Frequency of media reported incidents (MedISys-FF), by food per publishing country (2014–2015).

Labels	Alcohol	Cheese	Chili sauce	Coffee	Eggs	Food fraud	Honey	Juice	Meat	Milk	Olive Oil	Oregano	Oregano Oil	Peanuts	Poultry	Rice	Seafood	Spices	Tea	(blank)	Grand Total
Egypt		2				23			37	1	1				27	1	1				93
United States	6					26	1		5	2	2			1			39			1	83
United Kingdom	10	1		1		32			16	2	2		1	2	2	2	6	1		3	81
Saudi Arabia		1				61	3		5						1		1				72
Belgium		3				8			27								4				42
Germany	1			1	1	4			29								1				37
France	2	1				4			26								2				35
Brazil	1					2	3		1	24							1				32
Iraq						24			2	1	1										28
India	2	1				3			1	18											25
(blank)	2	1				7			5	3	1									1	23
Ireland	1	1		1		1			4	2	2	1	1				2			1	17
Spain		2				5			3		3		2	1							16
China	4					4			2	2						1					13
Jordan						8		2	2								1				13
Kuwait	1					9			1											1	12
Italy		2				7			1								1			1	12
Qatar						12															12
Israel			1			8			1	1							1				12
Morocco						9															9
Palestine						6	1		1		1										9
Yemen						8															8
Canada	1					2			2						1		2				8
Japan				1							2						4			1	8
- Unknown -						3			2								3				8
Russia	3	1				2			1												7
Sudan		1				4				2											7
Turkey						7															7
Taiwan						2											2		1	1	6
Vatican City						5					1										6
Portugal	1										5										6
Vietnam	1					4						1									6
Nigeria						3			1								1				5
New Zealand	1								2	2											5
Australia						2	1		1								1				5
Algeria						5															5
United Arab Emirates						2	1		1												4
Norway						2			1											1	4
Netherlands					1												3				4
Lebanon						4															4
Oman						2		1							1						4
Pakistan						2				2											4
Syria						1			1						1						3
Thailand						1											1	1			3
Philippines						1											2				3
Indonesia						1										1					2
American Samoa																	2				2
Isle Of Man										1							1				2
Libya						2															2
Colombia						1											1				2
Hungary	1					1															2

(continued on next page)

Table A1 (*continued*)

Labels	Alcohol	Cheese	Chili sauce	Coffee	Eggs	Food fraud	Honey	Juice	Meat	Milk	Olive Oil	Oregano	Oregano Oil	Peanuts	Poultry	Rice	Seafood	Spices	Tea	(blank)	Grand Total
Mexico																	2				2
Mali						2															2
Kenya	1									1											2
South Africa									1	1											2
Afghanistan									1												1
Yugoslavia										1											1
Panama						1															1
Denmark																	1				1
Finland						1															1
Georgia														1							1
Ethiopia									1												1
Micronesia																	1				1
Mauritius									1												1
Korea, Republic of										1											1
Tanzania						1															1
Iran						1															1
Togo										1											1
Ghana						1															1
Ukraine				1																	1
Austria						1															1
Guinea									1												1
Swaziland										1											1
Chad									1												1
Sweden						1															1
Burkina Faso						1															1
Montserrat						1															1
Poland									1												1
Cyprus						1															1
Lithuania																				1	1
Total	39	16	1	4	2	344	10	3	185	71	20	2	4	5	33	5	91	3	1	12	851

Table A2
Frequency of food fraud notifications reported in RASFF, by food per country of origin (2014–2015).

Country of origin	alcoholic beverages	Bivalve molluscs	Cereals and bakery products	Confectionery	Crustaceans	Dietetic, supplements foods	Eggs and egg products	Fats and oils	Fish and fish products	Fruits and vegetables	Herbs and spices	Honey and royal jelly	Meat and products	Milk and milk products	Nuts products and seeds	Mixed food	Poultry meat and products	Prepared dishes and snacks	Soups, sauces and condiments	Total
Bangladesh										14										14
Belgium																	2			2
Belize									1											1
Brazil													1							1
Canada															1					1
China	1		8				3		3	1		1			21	4		2	4	48
Côte d'Ivoire					1															1
Egypt															4					4
Ghana									1						7	1				9
Hong Kong			4																	4
India										8	4				10	2		1		25
Iran									1					1	1					3
Italy				1																1
Japan								1								1				2
Morocco										1										1
Myanmar					1															1
Nigeria										10					9	1				20
Pakistan																1				1
Poland									1											1
Russia													1							1
Senegal		1							3											4
Serbia															1					1
South Korea		2																		2
Thailand										3						1			1	5
Turkey										5					9	1			1	16
Ukraine								1												1
United States						3									1					4
Vietnam		2								1	2									5
Total	1	5	12	1	2	3	3	2	10	43	6	1	2	1	64	12	2	3	6	179

Table A3

Frequency of fraud incidents reported in EMA database, by food per publishing country (2014–2015).

Origin country	Alcoholic	Beverages	Dairy	Dietetic foods	Feed	Fish Seafood	Food additives	Fruits vegetables	Grain products	Herbs spices	Honey	Meat	Nuts	Oils Fats	Poultry	Grand Total
India	2		11						3	1				5		**22**
United States	2		3	1	1		3		1			6			1	**18**
Unknown	1				1	1					2	7		1		**13**
China		1	1				1		2			4			2	**11**
United Kingdom	1		1									3	1			**6**
Taiwan				1			1					1		1		**4**
Ukraine	2									1	1					**4**
Thailand		1					2									**3**
Mexico						2						1				**3**
Brazil			2									1				**3**
Bulgaria												2				**2**
Turkey											1					**1**
Russian Federation			1													**1**
Uganda			1													**1**
Italy														1		**1**
Saudi Arabia		1														**1**
Ireland											1					**1**
Netherlands												1				**1**
Bolivia				1												**1**
Nigeria	1															**1**
Grand Total	**9**	**3**	**20**	**3**	**2**	**3**	**3**	**4**	**6**	**2**	**5**	**26**	**1**	**8**	**3**	**98**

Table A4

Frequency of fraud notifications in RASFF, by food per year (2000–2015).

Product categories	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	Notifications	%
Nuts, nut products and seeds				3			4	12	4	8	7	20	16	22	24	40	160	17.4
Meat and meat products	1	2		1	3	21	18	8	8	15	11	2	6	59	1	1	157	17.1
Fish and fish products		1	2	1	3		3	5	9	8	18	10	8	12	2	8	90	9.8
Fruits and vegetables				1			2	2	2	5	1	11	1	8	29	14	76	8.3
Other food product/mixed				1	1	7	7	7	5	1	6	3	2	3	8	4	55	6.0
Prepared dishes and snacks				1			12		2			4	1	27	2	1	50	5.4
Cereals and bakery products				1			4			2	1	4		8	10	2	35	3.8
Bivalve molluscs and products thereof							6	7	2	9	2		2	1		5	34	3.7
Poultry meat and poultry meat products				1		13	4	2	1	3	2	3	3		2		34	3.7
Confectionery							1	4	1			5	18	1			30	3.3
Milk and milk products	1		1			1	2	1	7	5	3	4	2	2		1	30	3.3
Crustaceans and products thereof			1	1	1		3	2		2	9	2		2	1	1	25	2.7
Fats and oils	1						1	2	4	4	1	3	3	1	2		22	2.4
Herbs and spices					3	1		1			2	1	1	4		6	19	2.1
Eggs and egg products								2	1	2	5	2	2			3	17	1.9
Dietetic, supplements, fortified foods							1		2	2			4	2	2	1	14	1.5
Honey and royal jelly			1	1	1		1		1	1	1	3	1	1		1	13	1.4
Soups, broths, sauces and condiments							1	1			1	2	1	1	4	2	13	1.4
Cocoa, coffee and tea												7	2	1			10	1.1
Cephalopods and products thereof								3		1	1	1		3			9	1.0
Non-alcoholic beverages											1	4					5	0.5
Alcoholic beverages												3				1	4	0.4
Wild caught crustaceans and products thereof					4												4	0.4
Food additives and flavourings										1		2					3	0.3
Gastropods								1			1			1			3	0.3
Natural mineral water												3					3	0.3
Molluscs and products thereof					2												2	0.2
Ices and desserts						1											1	0.1
Total	3	3	5	15	15	43	71	60	49	69	73	99	76	158	88	91	918	100.0

Table A5

Frequency of fraud incidents reported in EMA database, by food per year (2000–2015).

Product categories	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	Total notifications	(%)
Fish and Seafood			2		3	8	1	8	2		10	6	68	61		3	172	26.8
Fruits and vegetables				1				3	1	1	4	71	2	8	3	1	95	14.8
Dairy					2	1		3	2		2	13	10	11	15	5	64	10.0
Meat						1	1	1		1	6	7	1	19	16	10	63	9.8
Oils and fats	2	2		2		1	2	1	2	1	1	9	4	9	2	6	44	6.9
Alcoholic			2			1		1	3	1	3	6	11	5	4	5	42	6.5
Honey			4	1	2	1			2	2		1	10	2	3	2	30	4.7
Herbs and spices					2	8			1		4	5	1	3	1	1	26	4.0
Grain products								1	3		1	7	1	6	3	3	25	3.9
Beverages					1	1		1				4	3	7	3		20	3.1
Dietetic foods				2	3						1	10			3		19	3.0
Poultry					1					3		3		2	2	1	12	1.9
Eggs							3				1	1	1	2			8	1.2
Coffee and tea						1				1		1	3				6	0.9
Feed								3							1	1	5	0.8
Confectionary										1		2		1			4	0.6
Food additives															3		3	0.5
Nuts												1	1		1		3	0.5
Soups and sauces													1				1	0.2
Total notifications	2	2	8	6	14	23	7	21	18	10	33	147	117	136	60	38	642	100.0

References

Bouzembrak, Y., & Marvin, H. J. P. (2016). Prediction of food fraud type using data from Rapid Alert System for Food and Feed (RASFF) and Bayesian network modelling. *Food Control, 61*, 180–187.

Domingo, E., Tirelli, A. A., Nunes, C. A., Guerreiro, M. C., & Pinto, S. M. (2014). Melamine detection in milk using vibrational spectroscopy and chemometrics analysis: A review. *Food Research International, 60*, 131–139.

EC (2018). *The EU food fraud network and the system for administrative assistance & food Fraud: Annual report 2017.* European Commission.

EMA (2017). *The economically motivated adulteration (EMA) databases.* http://ema. foodshield.org/.

EU (2013). *Commission Recommendation of 19 February 2013 on a coordinated control plan with a view to establish the prevalence of fraudulent practices in the marketing of certain foods.* 2013/99/EU.

Hovda, K. E., Hunderi, O. H., Tafjord, A. B., Dunlop, O., Rudberg, N., & Jacobsen, D. (2005). Methanol outbreak in Norway 2002–2004: Epidemiology, clinical features and prognostic signs. *Journal of Internal Medicine, 258*(2), 181–190.

Liu, Y., Liu, F., Zhang, J., & Gao, J. (2015). Insights into the nature of food safety issues in Beijing through content analysis of an internet database of food safety incidents in China. *Food Control, 51*, 206–211.

Manning, L. (2016). Food fraud: Policy and food chain. *Current Opinion in Food Science, 10*, 16–21.

Manning, L., & Soon, J. M. (2014). Developing systems to control food adulteration. *Food Policy, 49*(Part 1), 23–32.

Manning, L., & Soon, J. M. (2018). The value of product sampling in mitigating food adulteration. *A handbook of food crime* (pp. 127–140). Policy Press at the University of Bristol.

Mantero, J., Cox, A., Linge, J., van der Goot, E., & Coulombier, D. (2010). The European Centre for Disease Prevention and Control strengthening MedISys as a tool to accelerate detection of threats to human health from communicable diseases in the EU. *International Journal of Infectious Diseases, 14*(Supplement 1), e274–e275.

Marvin, H. J. P., Bouzembrak, Y., Janssen, E. M., van der Fels- Klerx, H. J., van Asselt, E. D., & Kleter, G. A. (2016). A holistic approach to food safety risks: Food fraud as an example. *Food Research International, 89*, 463–470.

O'Mahony, P. J. (2013). Finding horse meat in beef products–a global problem. *QJM, 106*(6), 595–597.

Qiao, G., Guo, T., & Klein, K. K. (2010). Melamine in Chinese milk products and consumer confidence. *Appetite, 55*(2), 190–195.

RASFF (2017). *Rapid alert system for food and feed (RASFF). Directorate general for health and consumer protection.* Brussels: European Commission.

Rortais, A., Belyaeva, J., Gemo, M., van der Goot, E., & Linge, J. P. (2010). MedISys: An early-warning system for the detection of (re-)emerging food- and feed-borne hazards. *Food Research International, 43*(5), 1553–1556.

Stanciu, S. (2015). Horse meat consumption − between scandal and reality. *Procedia Economics and Finance, 23*, 697–703.

Tähkäpää, S., Maijala, R., Korkeala, H., & Nevas, M. (2015). Patterns of food frauds and adulterations reported in the EU rapid alert system for food and feed and in Finland. *Food Control, 47*(0), 175–184.

USP (2018). *Food fraud database.* http://www.foodfraud.org/.

Xiu, C., & Klein, K. K. (2010). Melamine in milk products in China: Examining the factors that led to deliberate use of the contaminant. *Food Policy, 35*(5), 463–470.

Zhang, W., & Xue, J. (2016). Economically motivated food fraud and adulteration in China: An analysis based on 1553 media reports. *Food Control, 67*, 192–198.