

3 Surveillance data in Denmark

This chapter delves into the data collection methods and quality assurance procedures within the Danish surveillance system. Moreover, it introduces the case studies selected for this master's thesis, which include *Listeriosis*, *Shigellosis*, Shiga toxin (verotoxin)-producing *Escherichia coli*, and *Salmonellosis*. These diseases will be referred to using either their full name or their related case definition:

- LIST: *Listeriosis*
- SHIG: *Shigellosis*
- STEC: Shiga toxin (verotoxin)-producing *Escherichia coli*
- SALM: *Salmonellosis*

3.1 Data collection and data quality

In Denmark, the surveillance of infectious diseases is conducted by Statens Serum Institut (SSI). This surveillance system plays a pivotal role in national and international disease preparedness. It encompasses more than just the collection and registration of disease data; it also involves the prompt and ongoing dissemination of knowledge to the relevant authorities responsible for treatment, prevention, and control. This comprehensive approach ensures efficient communication and facilitates appropriate measures to address infectious diseases.

The quality of the Danish surveillance registers is maintained at a high standard, thanks to The National Board of Health Statutory Order on Physicians' Notification of Infectious Diseases (<https://www.retsinformation.dk/eli/ita/2000/277>). This order specifies that several diseases¹ are individually notifiable by physicians and general practitioners. Notifications consist of essential patient information and are submitted in paper form to both the Ministry of Health and to SSI. This rigorous notification process ensures accurate and comprehensive data collection for disease surveillance purposes in Denmark.

In addition to the individually notifiable diseases, SSI has implemented a laboratory notification system for numerous microorganisms. Clinical-microbiological laboratories are obligated to report the identification of specific microorganisms, along with relevant patient information. These data are then stored in the Danish Microbiology Database (MiBa), which was established by SSI in 2010. MiBa is a nationwide and automatically updated database specifically designed to collect and store microbiological test results. In order to utilize the data from MiBa, the information in the test results needs to have a standardized structure with common codes and terminology. MiBa employs national standards to harmonize the data, which initially may be structured in diverse formats. The standards currently used are XRPT05, which is widely employed for the exchange of microbiological test results in the healthcare system, and a specific standard called XRPT06. These standards are regularly revised and exist in various versions². The national surveillance system focuses on diseases of a severe nature, those that are highly contagious, and the majority of vaccine-preventable diseases.

¹For a full list of diseases see https://www.ssi.dk/sygdomme-beredskab-og-forskning/anmeldelse-af-sygdomme/lovpligtige-meldesystemer/individ_anmeldelses_sygdomme

²Information on national standards and codes within the healthcare domain can be found on MedCom's website (<https://medcom.dk/>).

3.2 Introducing the case studies

For the scope of this master's thesis, only a specific subset of diseases from the mandatory notification system will be considered. This subset consists of *Listeriosis*, *Shigellosis*, Shiga toxin (verotoxin)-producing *Escherichia coli*, and *Salmonellosis*. These diseases have been chosen for analysis and investigation based on various factors, such as seasonality, incidence, and severity. Additionally, these diseases have been associated with documented outbreaks investigated by SSI in the past decade, which adds to their relevance for the study.

The count observations are observed in the period from January 2008 to December 2022. An epidemic curve graph for an excerpt of the data for each of the diseases considered in this master's thesis is shown in Figure 3.1.

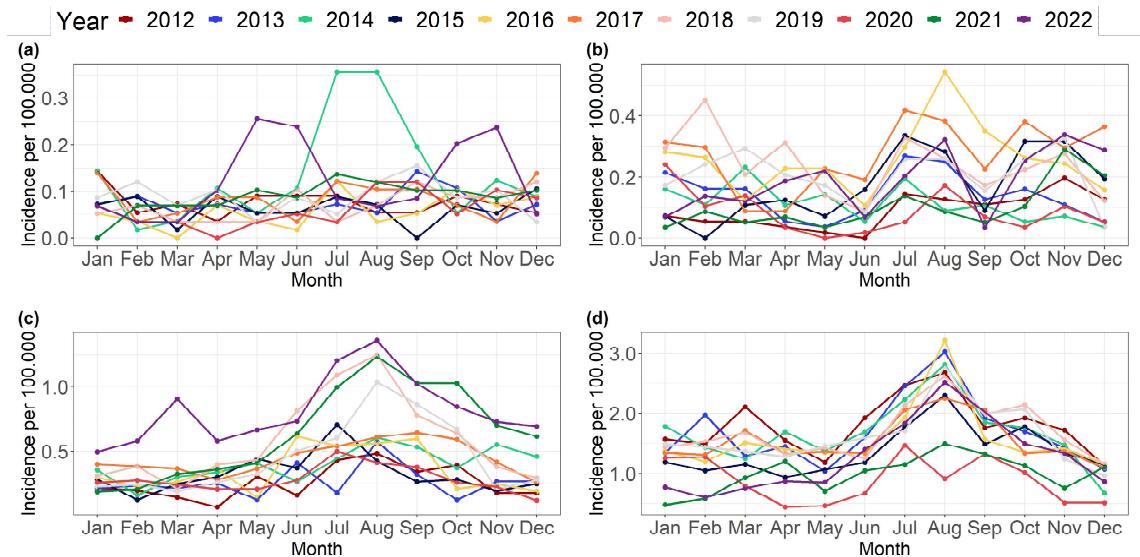


Figure 3.1: Epidemic curve showing the incidence per 100,000 in Denmark, 2012-2022, for the subset of diseases considered in this master thesis. **(a)** LIST, **(b)** SHIG, **(c)** STEC, and **(d)** SALM.

Evidently, the subset of diseases in Figure 3.1 exhibits varying incidences and different levels of seasonal patterns on an annual basis. It is interesting to note that the incidences peak in August, which can be attributed to several factors, including:

- Increased travel activity: Especially when individuals travel to countries with unsafe drinking water, poor sanitation, and insufficient hygiene practices.
- Large social gatherings: Events such as weddings, large festivals, or other gatherings where a significant number of people consume potentially contaminated food or drinking water.

Moreover, one could hypothesize that the warmer climate during the summer could potentially directly influence the proliferation of bacteria, which in turn may have an impact on the transmission and spread of these diseases.

In general, there is a noticeable decrease in the number of observed cases starting from March 2020 and continuing until January 2021. This decline can be attributed to the strict lockdown measures implemented in Denmark in response to the Covid-19 pandemic. These measures, which involved restrictions on movement and social interactions, likely

played a significant role in reducing the transmission of infectious diseases, including the ones being investigated.

From Figure 3.1a, it is evident that the incidence of LIST remains relatively low but stable over time. However, there are six specific months that stand out due to higher incidence rates. These months include July and August 2014, May and June 2022, and October and November 2022. These periods show a notable increase in the number of reported cases compared to the rest of the time series.

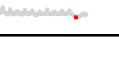
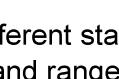
In Figure 3.1b, a clear annual seasonality of SHIG is observed. The incidence of the disease shows a consistent pattern over the years, with peaks occurring in the summer months of 2016 and 2017. These years exhibit the highest incidences of the disease. On the other hand, the lowest incidences are observed in 2012 and 2020.

Furthermore, in Figure 3.1c, a significant increase in the amplitude of the seasonal variation in STEC can be observed starting from 2018, with incidences doubling compared to the preceding years. At a first glance, this increase in the incidences might be recognized as a serious, reoccurring outbreak of the disease, but a more reasonable explanation can be found. Up to 2018, most departments of clinical microbiology used culture-based methods as a diagnostic test for bacterial pathogens and the process of changing the test method to polymerase chain reaction (PCR) methods was ongoing (Svendsen et al. 2023). In general, PCR resulted in higher incidences compared to other test methods, which is to no surprise as higher sensitivity is well documented for PCR (Buss et al. 2015; Knabl, Grutsch, and Orth-Höller 2016).

Overall, the highest incidences among the diseases considered in this master's thesis are observed for SALM in Figure 3.1d. SALM exhibits a notable pattern of high incidence throughout the observed period. However, there is a significant drop in incidences observed in 2020, which is consistent with the impact of the COVID-19 pandemic on disease surveillance and reporting. Additionally, 2021 shows generally lower incidences compared to previous years.

Some summary statistics for each of the diseases considered in this master's thesis are gathered in Table 3.1.

Table 3.1: Summary statistics of the monthly count observations for the subset of diseases considered in this master's thesis. Boxplot: median (red line), IQR (grey box), whiskers (1.5 IQR), outliers (points). Time series: normalized observations (0-1), first time points minimum and maximum count (red)

Case definition	Min	Max	Mean	Median	Std. Deviation	Boxplot	Time series
LIST	0	20	4.446	4.0	2.934	•	
SHIG	0	28	9.078	8.0	5.652	•	
STEC	2	80	21.933	16.5	14.902	•+	
SALM	0	583	108.012	85.5	82.725	— — · · · · —	

In Table @ref{tab:summaryStat}, it is readily seen that the diseases exhibit different statistical properties, including variations in mean incidence, standard deviation, and range. These variations indicate that each disease has its own unique epidemiological characteristics.

From an epidemiological perspective, all the diseases within the selected subset pose a

significant risk of infection and can vary in terms of severity for affected individuals. Therefore, early identification of disease outbreaks is of utmost importance in order to promptly implement necessary interventions. Timely detection allows for swift and targeted actions to control the spread of these diseases and mitigate their impact on public health.

3.2.1 Listeriosis

LIST is a foodborne illness that is caused by consuming food contaminated with *Listeria monocytogenes*. This disease primarily affects pregnant women, unborn or newborn babies, the elderly, and individuals with weakened immune systems. The disease is associated with high mortality (Goulet et al. 2012) and manifests in three ways: sepsis, meningitis, and mother-to-child transmission. Pregnancy-associated LIST can have severe consequences for the fetus or newborn, including miscarriage, stillbirth, neonatal sepsis, and meningitis (Awofisayo et al. 2015). LIST is uncommon among individuals in other demographic groups. The bacteria is ubiquitous in the environment, found in moist environments, soil, water, decaying vegetation, and animals. Furthermore, it can survive and even grow under refrigeration and other food preservation measures.

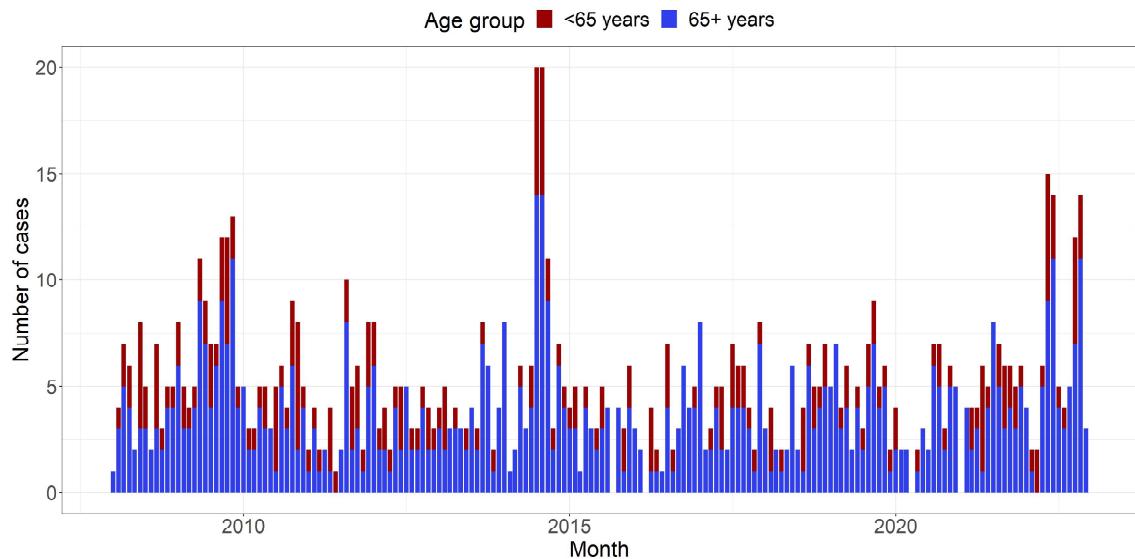


Figure 3.2: A stacked bar graph illustrating the number of LIST cases observed in the period from 2008 to 2022 for the age groups below and above 65 years.

In general, SSI employs whole-genome sequencing (WGS) as the state-of-the-art method to detect disease outbreaks caused by *Listeria monocytogenes*. This method involves mapping the entire DNA of the bacteria and enables SSI to identify cases where patients are infected with identical Listeria bacteria. However, it is important to note that for this master's thesis, the DNA typing data is unavailable for use.

The first notified outbreak in the 20th century occurred in 2009 (Smith et al. 2011). This outbreak affected 8 people, of which two died. The patient samples were isolated between the 6th and 11th of May and it was believed that the cause was infected beef meat from a meals-on-wheels delivery.

Another notable outbreak investigated by SSI occurred between September 2013 and October 2014 (Kvistholm Jensen et al. 2016). This LIST outbreak involved a total of 41 cases, resulting in 17 deaths. Deli meat products from a specific company were identified as the source of the outbreak. The high mortality rate may be attributed to the

consumption of these products in nursing homes and hospitals, where patients are more vulnerable. Following the discovery of Listeria at the facility, the Danish Veterinary and Food Administration recalled all products from the company.

In another LIST outbreak investigated by SSI, the source was traced back to cold-smoked and cured salmon products (Schjorring et al. 2017). A total of 5 related cases were identified, with 4 of them occurring in August 2017, and the fifth case in May 2017.

In some cases, despite extensive investigations, the source of contamination in an outbreak cannot always be identified. Such was the case in an unresolved outbreak that took place between the 13th of May and the 6th of June, 2022. During this period, a total of nine cases were infected with the same type of Listeria, with the majority of affected patients located in the Capital Region of Denmark. Despite thorough efforts, the specific source of contamination remained unknown.

Early identification of outbreaks caused by *Listeria monocytogenes* is crucial to implement timely interventions and mitigate the impact of the disease. Otherwise, these outbreaks can persist over an extended period. SSI has successfully resolved several long-spanned outbreaks in the last decade. For example, one investigation revealed that a single outbreak was actually two simultaneous outbreaks caused by the consumption of smoked fish. Each outbreak consisted of ten cases and spanned from May 2013 to July 2015 (Gillesberg Lassen et al. 2016).

Other documented long-spanned outbreaks investigated by SSI include:

- A cold-smoked fish outbreak with 9 cases spanning from December 2016 to February 2019.
- A prolonged outbreak with 6 cases from 2016 to 2019, traced back to a local green-grocer.
- An outbreak with 8 cases from October 2021 to June 2022, caused by a deli meat product.
- Two unresolved outbreaks with 9 cases and 12 cases from the end of 2018 to November 2021 and October 2020 to May 2022, respectively.

The use of WGS in these outbreaks provided the ability to link cases that occurred over a period of years and revealed that they were, in fact, continuous-source outbreaks.

In a recent outbreak investigated by SSI, the Danish Veterinary and Food Administration, and the National Food Institute at the Technical University of Denmark, fish patties were identified as the source of contamination. This outbreak occurred from August 2022 to December 2022 and affected a total of 11 cases.

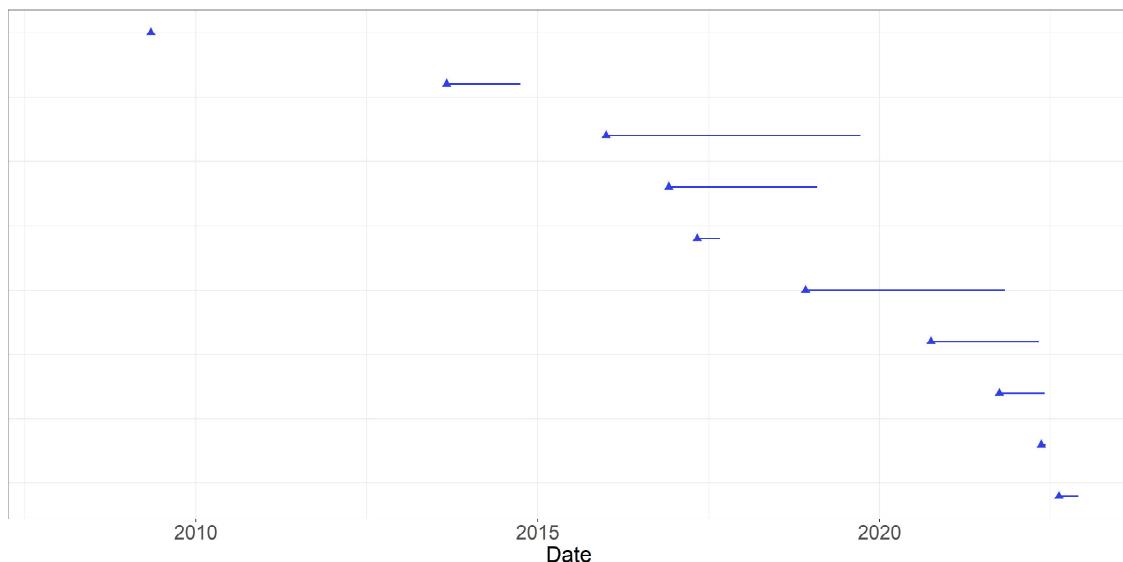


Figure 3.3: Timeline plot indicating the start (triangle) and duration (line) of documented outbreaks of LIST by SSI and possible collaborators.

3.2.2 *Shiggellosis*

SHIG is a diarrheal illness that is caused by a group of bacteria called *Shigella*. There are four types of *Shigella* bacteria, namely: *Shigella dysenteriae*, *Shigella boydii*, *Shigella flexneri*, and *Shigella sonnei*. The latter is the most common species in Denmark. The bacteria are highly contagious and can be transmitted through direct person-to-person contact, consumption of contaminated food, or ingestion of water contaminated with human feces. SHIG infections are most commonly observed in children under the age of 5, individuals traveling to regions with poor sanitation and unsafe water and food practices, as well as gay, bisexual, and other men who have sex with men.

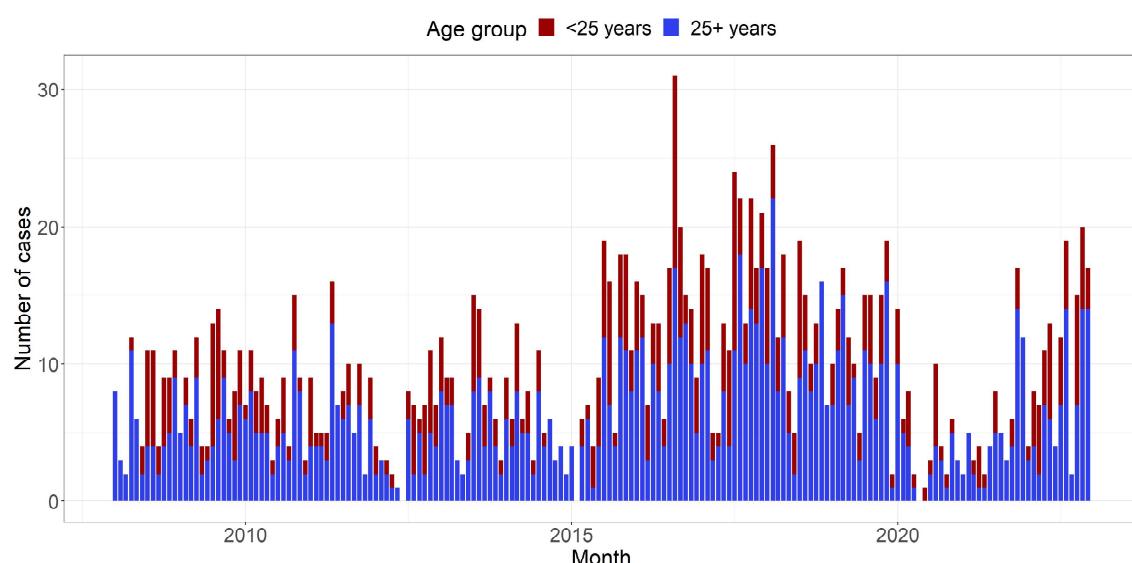


Figure 3.4: A stacked bar graph illustrating the number of SHIG cases observed in the period from 2008 to 2022 for the age groups below and above 25 years.

In Denmark, another significant cause of SHIG outbreaks is the importation of contami-

nated vegetables. This was evident in several incidents, including a 2007 outbreak where 215 individuals fell ill after consuming imported contaminated baby corn (Lewis et al. 2009), a smaller outbreak in 2009 linked to sugar snap peas from Kenya (Muller et al. 2009), and a 2020 outbreak associated with fresh mint as the source of infection.

The 2020 outbreak is indeed a significant focus of this study, as it serves as a benchmark for evaluating the effectiveness of outbreak detection algorithms. It took place from the 22th of August to the 9th of September and was investigated by SSI in collaboration with the Danish Veterinary and Food Administration and the National Food Institute at the Technical University of Denmark. The outbreak affected 44 patients, mainly concentrated in the Capital Region of Denmark. During the investigation, at least five events were identified where individuals subsequently developed *Shigellosis*.

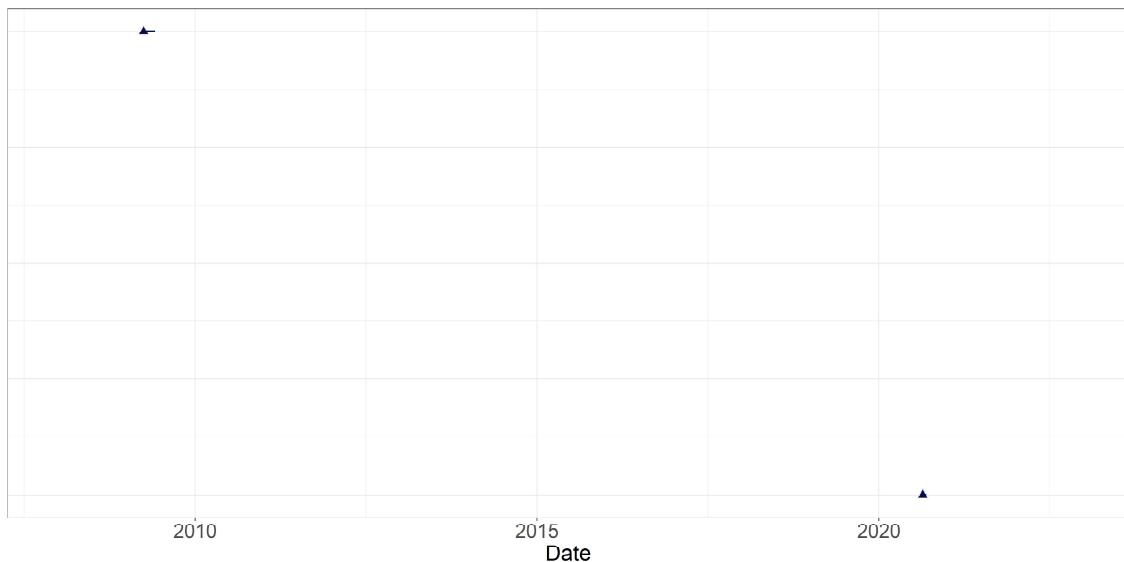


Figure 3.5: Timeline plot indicating the start (triangle) and duration (line) of documented outbreaks of SHIG by SSI and possible collaborators.

3.2.3 Shiga toxin (verotoxin)-producing *Escherichia coli*

STEC primarily spreads through contaminated food. Less common sources of infection include contaminated drinking and bathing water, as well as direct or indirect contact with infected animals. Cattle and other ruminants are primary reservoirs for STEC serotypes that are frequently associated with human disease (Menge 2020). Therefore, in Denmark, the source of infection is often products derived from beef, non-heat-treated dairy products, or other foods such as ready-to-eat vegetables, leafy greens, vegetable sprouts, and berries contaminated with feces from cows. *Hemolytic uremic syndrome* (HUS) is a severe complication that, in some cases, particularly in children, can develop following an infection with STEC.

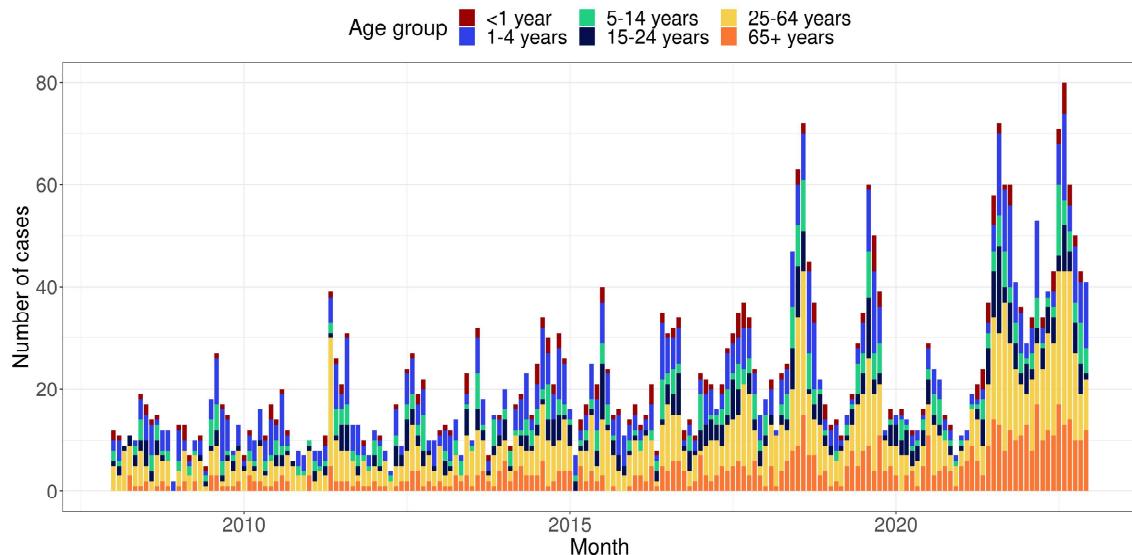


Figure 3.6: A stacked bar graph illustrating the number of STEC cases observed in the period from 2008 to 2022 for the six age groups.

In general, stool samples are commonly used for diagnostic purposes in cases of STEC infections. Until 2018, most clinical microbiology departments relied on culture-based methods to detect and identify STEC bacteria in stool samples. However, in recent years, PCR methods have been increasingly adopted as a replacement for culture-based methods in the diagnosis of STEC infections (Svendsen et al. 2023). PCR methods offer advantages such as increased sensitivity and faster turnaround time, contributing to their growing popularity in clinical laboratories.

It is important to note that not all patients are routinely tested for STEC, and therefore, physicians need to specifically request STEC testing when submitting stool samples.

One of the earliest documented STEC outbreaks occurred in 2007, involving 18 laboratory-confirmed cases over a six-week period. The outbreak primarily affected children in day-care settings, and most patients experienced mild symptoms without bloody diarrhea. Investigations indicated a specific brand of organic beef sausage as the likely source of infection.

In September to October 2012, a STEC outbreak with a high risk of HUS was observed. Thirteen cases were diagnosed, with eight individuals developing HUS. Epidemiological investigations suggested that ground beef was the vehicle of the outbreak (Soborg et al. 2013).

More recent outbreaks include a 38-case outbreak from September to November 2018, with a suspected association with beef sausage as the source of infection. Additionally, there were two unresolved outbreaks with 11 and 14 cases occurring from May to July 2019 and from December 2021 to January 2022, respectively. The latter outbreak included three cases of HUS.

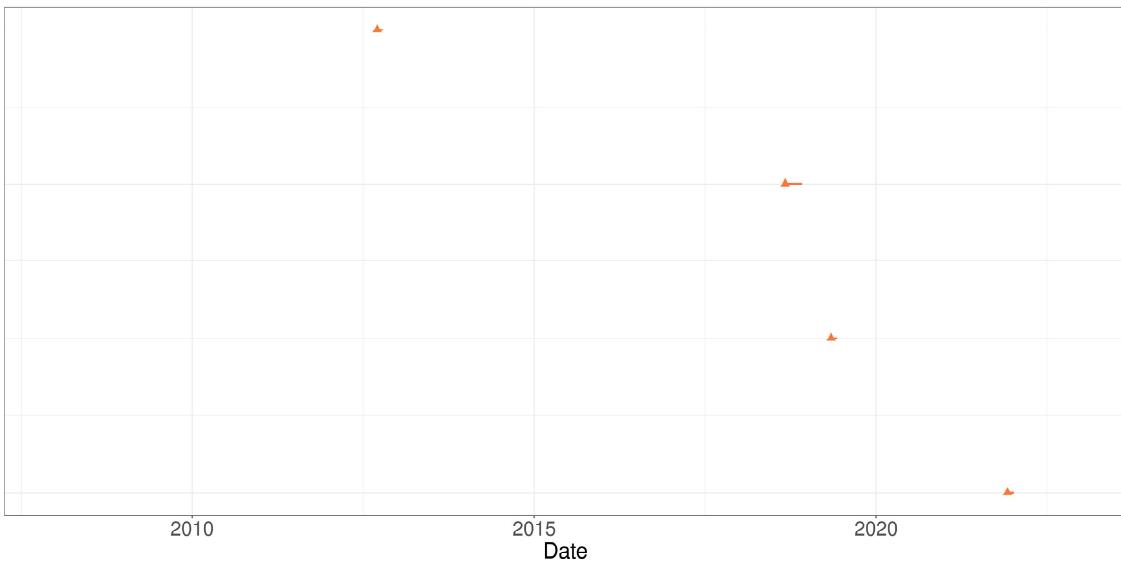


Figure 3.7: Timeline plot indicating the start (triangle) and duration (line) of documented outbreaks of STEC by SSI and possible collaborators.

3.2.4 *Salmonellosis*

SALM is a bacterial disease that primarily affects the intestinal tracts of humans. The *Salmonella* bacteria are commonly found in the intestines of animals and humans and are excreted in feces. Human infection typically occurs through the consumption of contaminated food or water. *Salmonella* infections are often associated with the consumption of raw or undercooked meat, poultry, eggs or egg products, as well as unpasteurized milk.

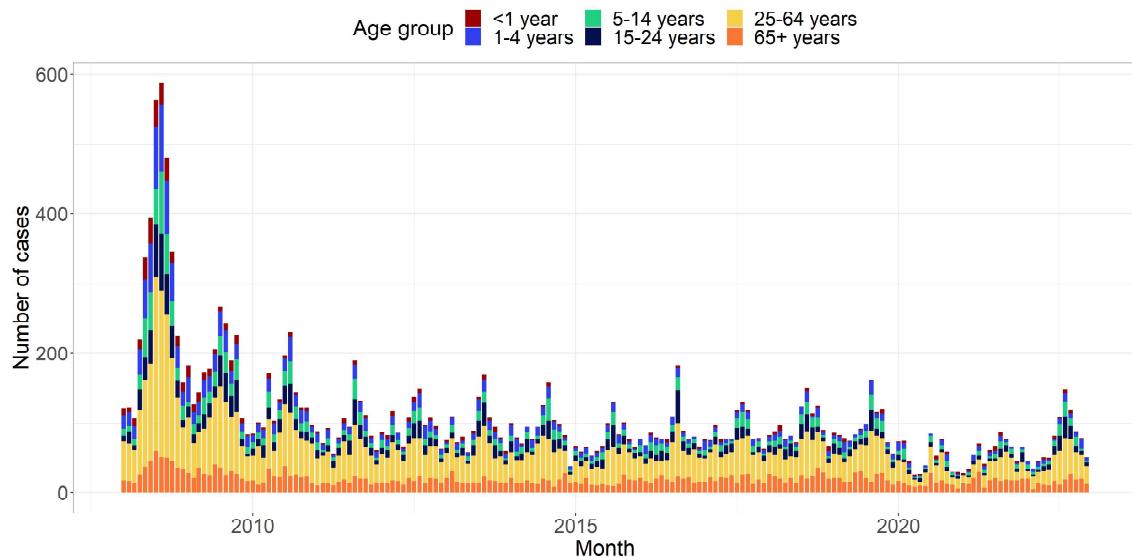


Figure 3.8: A stacked bar graph illustrating the number of SALM cases observed in the period from 2008 to 2022 for the six age groups.

It is worth noting that an increasing proportion of infections in Denmark are now observed in connection with international travel, particularly since *Salmonella* has been eliminated from commercial chicken flocks in Denmark, making Danish eggs and poultry meat free from the bacteria. However, imported meat products can still pose a risk of contamination.

In 2015, three outbreaks of SALM with patients in two or more regions were investigated (Helwigh, Christensen, and Müller 2016):

- An outbreak caused by *S. Newport* affecting 6 people in the period from March to April 2015.
- A long-lasting outbreak caused by *S. Oranienburg* with 14 genetically linked cases from July 2015 to January 2016.
- An outbreak with 6 patients from November 2015 to January 2016.

Other resolved and well-documented outbreaks include:

- An outbreak with 49 cases from October 2018 to January 2019, where Mediterranean sausage was identified as a possible source of contamination.
- An outbreak with 45 cases from November 2020 to April 2021, linked to the consumption of the natural remedy HUSK Psyllium.
- An international outbreak from March to July 2021, with more than 300 cases in Europe, including 39 cases in Denmark. Imported melons were suspected as the source of infection.
- An outbreak caused by eggs from a Danish producer, resulting in 24 cases registered from September to November 2021.
- An international outbreak with a total of 392 cases across 12 countries in the EU-/EEA and the UK, including 4 cases in Denmark. Kinder chocolate products were identified as the source of infection.

There were also outbreaks where it was not possible to identify the source of contamination, including:

- An outbreak with 26 cases in the period from May to August 2019.
- An outbreak with 11 cases in the period from June to July 2020.
- An outbreak with 24 cases in the period from March to September 2022.
- An outbreak with 15 cases in the period from August to September 2022.

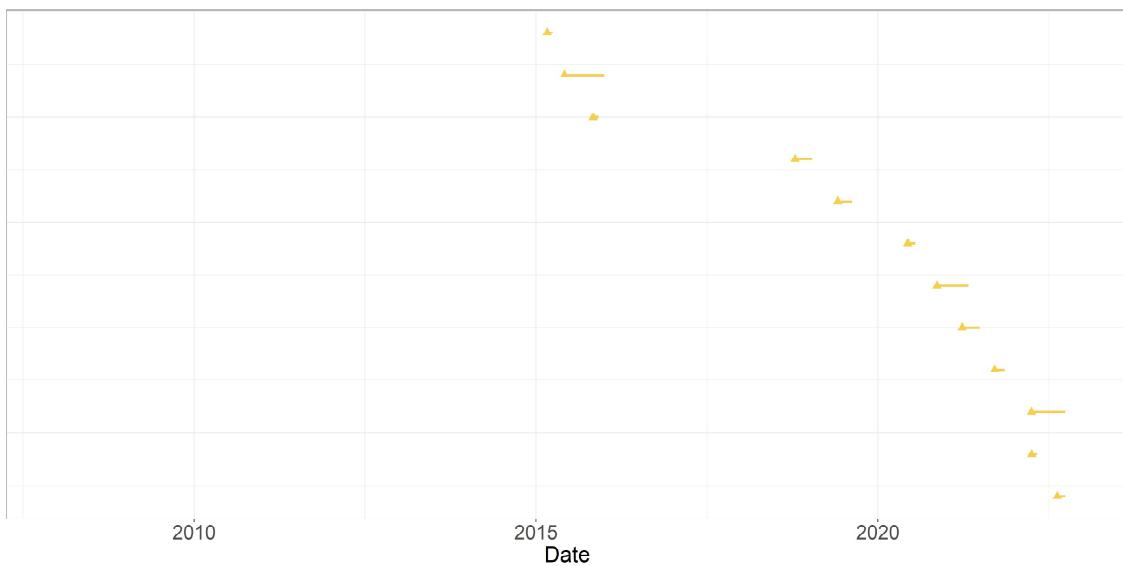


Figure 3.9: Timeline plot indicating the start (triangle) and duration (line) of documented outbreaks of SALM by SSI and possible collaborators.

