# Analysis, interpretation, reporting and use of data

Module 1.4

# Overview

- Analysis and interpretation of data
- Surveillance bulletins
- Using surveillance data for advocacy

# General approach to data analysis

- Analyse the surveillance data on a continuous basis – plan to analyse on at least a weekly basis.
- Typically report:
  - Total number of cases
  - Incidence or notification rates – adjust for size of underlying population
  - Proportions
- Core descriptive analyses:
  - Time (day, week, month, year)
  - Place (district, region, country)
  - Person (age, sex, occupation, race, ethnicity
- Present results in tables, graphs and maps

# Outputs of data analysis

## Target analyses to address surveillance objectives  and questions

| Surveillance objectives | Analytical outputs that can address these objectives | Frequency of analysis |
|---|---|---|
| Identify temporal trends and detect possible outbreaks | Line graph of incidence over time | Weekly |
| Identify groups who are at higher risk of WRID | Table of total number of cases and incidence or prevalence rate by age, sex and geographic area | Weekly |
| Detect possible outbreaks or clusters of cases;<br>identify areas associated with higher rates of disease | Table or map of the number of cases or the incidence rate by geographical area | Weekly |
| Estimate disease burden | Table of frequency of cases | Quarterly or annually |
| Evaluate the impact of control measures, such as implementing a new water-treatment step | Incidence of disease before and after changes in the water treatment | Based on needs |

# Calculating an incidence or notification rate

$$\text{Notification rate per 100,000 persons} = \frac{\text{Number of cases (notifications)}}{\text{Total population}} \times 100{,}000$$

| Surveillance week | Number of notifications | Population estimate | Notification rate / 100,000 persons |
|---|---|---|---|
| 12 | 525 | 1,291,850 | 40.6 |
| 13 | 489 | 1,291,850 | 37.9 |
| 14 | 501 | 1,291,850 | ? |
| 15 | 579 | 1,291,850 | ? |

***Exercise: Using the formula, calculate the notification rates for weeks 14 and 15***

# Calculating an incidence or notification rate

$$\text{Notification rate per 100,000 persons} = \frac{\text{Number of cases (notifications)}}{\text{Total population}} \times 100{,}000$$

| Surveillance week | Number of notifications | Population estimate | Notification rate / 100,000 persons |
|---|---|---|---|
| 12 | 525 | 1,291,850 | 40.6 |
| 13 | 489 | 1,291,850 | 37.9 |
| 14 | 501 | 1,291,850 | **38.8** |
| 15 | 579 | 1,291,850 | **44.8** |

# Analysis by time – monitoring trends

# Analysis by time – monitoring trends

- Different ways of presenting the data over time will illustrate different information and will convey different messages:
  - Is the rate or burden of disease increasing or decreasing?
  - How does this year compare to previous years?
  - Is there any seasonality in the incidence of disease?
- Can apply alert thresholds to detect outbreaks or identify the start of seasonal epidemics

# Alert thresholds and moving averages

- Alert thresholds provide a signal that the number of cases exceeds a defined level
  - Possible outbreak or start of seasonal epidemic → sign that action may be needed
- Threshold depends on severity and epidemic potential of a pathogen and the local epidemiology
- Defined based on number of cases or by comparing number of cases in current surveillance period to historical data over previously defined time periods:
  - 5 year moving average

# Example Alert Thresholds

| Surveillance outcome | Alert threshold |
| --- | --- |
| Bloody diarrhoea | 5 or more cases in one place in one day<br><br>Double the 5-year weekly average of cases |
| Acute gastroenteritis | Increase above the five-year average for that reporting period<br><br>Two standard deviations above the five-year average for that reporting period |

# Calculating a – 5 year weekly moving average

5-year moving average of weekly cases =

$$\frac{\text{Total Yr 1 + total Yr 2 + total Yr 3 + total Yr 4 + total Yr 5}}{5}$$

| Surveillance week | Weekly notifications per year | | | | | 5-year total | 5-year average |
|---|---|---|---|---|---|---|---|
| | 2015 | 2016 | 2017 | 2018 | 2019 | | |
| **12** | 10 | 10 | 10 | 10 | 10 | 50 | 10 |
| **13** | 53 | 49 | 61 | 43 | 57 | 263 | 53 |
| **14** | 48 | 37 | 45 | 54 | 51 | ? | ? |

**Exercise: Using the formula on this slide, calculate the 5 year average for week 14**

# Calculating a – 5 year weekly moving average

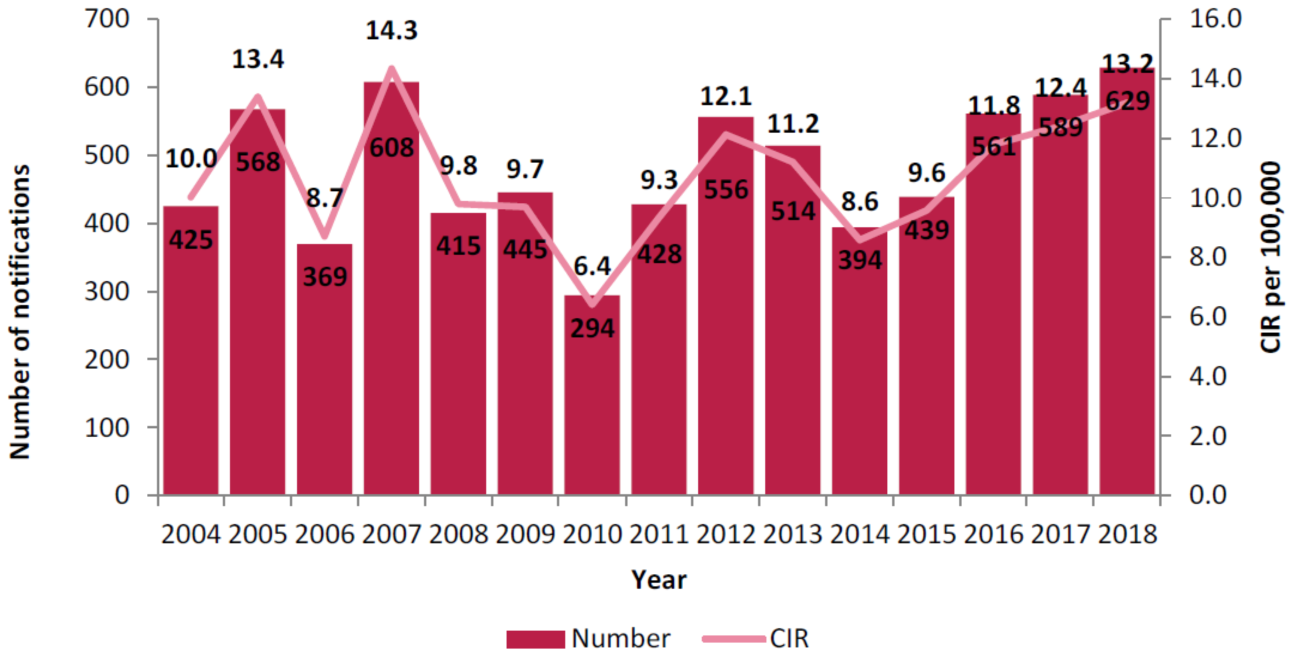
5-year moving average of weekly cases =

$$\frac{\text{Total Yr 1} + \text{total Yr 2} + \text{total Yr 3} + \text{total Yr 4} + \text{total Yr 5}}{5}$$

| Surveillance week | Weekly notifications per year | | | | | 5-year total | 5-year average |
|---|---|---|---|---|---|---|---|
| | 2015 | 2016 | 2017 | 2018 | 2019 | | |
| **12** | 10 | 10 | 10 | 10 | 10 | 50 | 10 |
| **13** | 53 | 49 | 61 | 43 | 57 | 263 | 53 |
| **14** | 48 | 37 | 45 | 54 | 51 | **235** | **47** |

# Number of cases and crude incidence rate (CIR) over time

**Figure 1.** Number and CIR cryptosporidiosis per 100,000 population, Ireland, 2004-2018

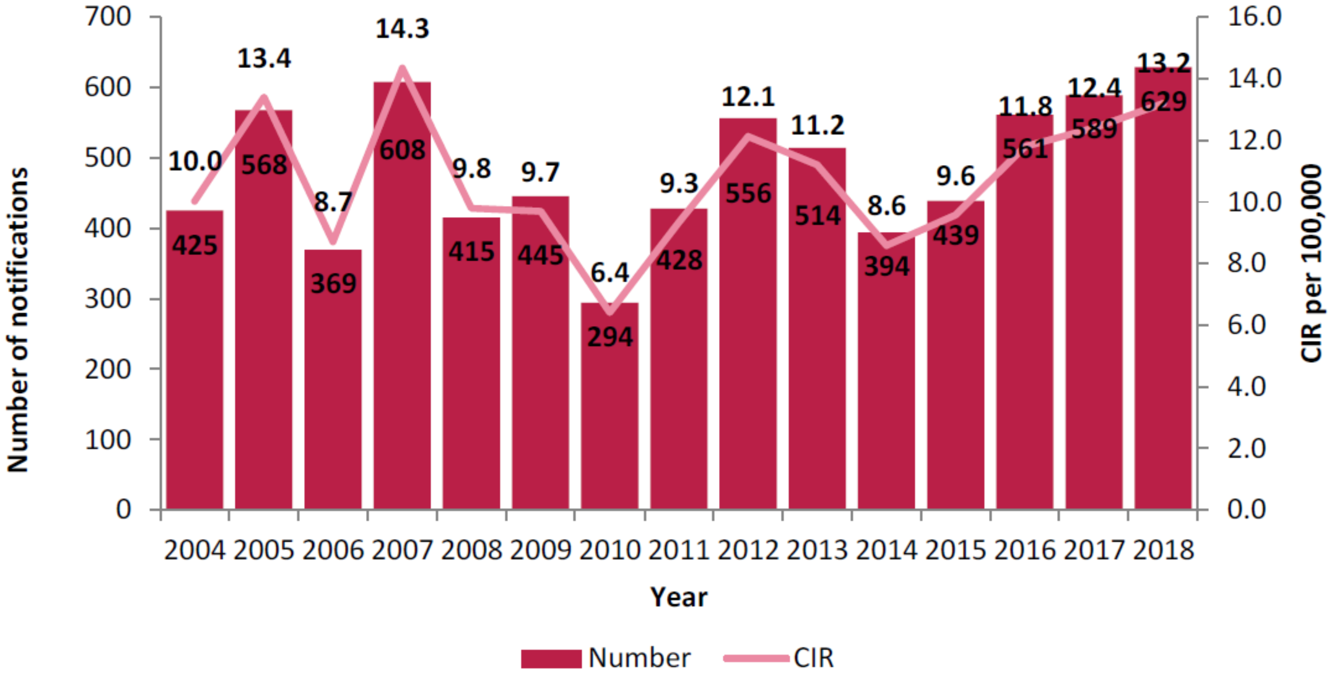



**Discussion:** *What is your interpretation of this graph?*


**HSE Health Protection Surveillance Centre (2019):** Cryptosporidiosis in Ireland, 2018.
https://www.hpsc.ie/a-z/gastroenteric/cryptosporidiosis/publications/epidemiologyofcryptosporidiosisinirelandannualreports/Crypto%20Annual%20Report%202018.pdf

# Number of cases and crude incidence rate (CIR) over time

**Figure 1.** **Number and CIR cryptosporidiosis per 100,000 population, Ireland, 2004-2018**

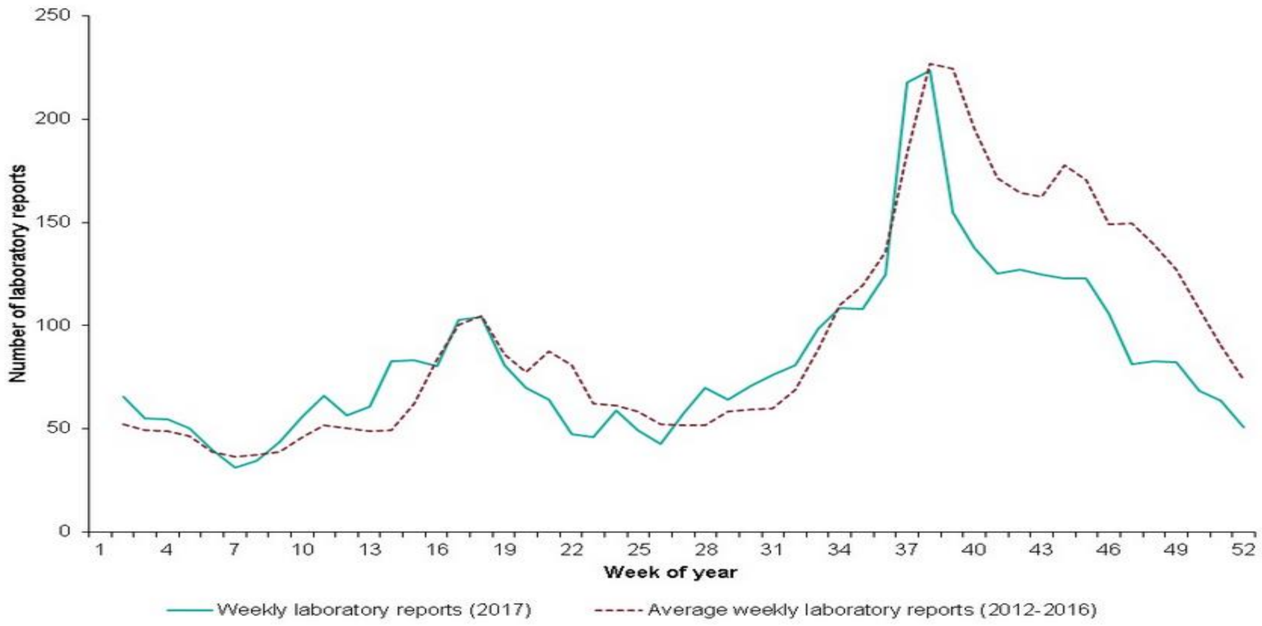



*Interpretation: The number of notifications and the population based incidence rate have increased over the past five years; the burden of disease is increasing over time*


**HSE Health Protection Surveillance Centre (2019):** Cryptosporidiosis in Ireland, 2018. https://www.hpsc.ie/a-z/gastroenteric/cryptosporidiosis/publications/epidemiologyofcryptosporidiosisinirelandannualreports/Crypto%20Annual%20Report%202018.pdf

# Weekly notifications (or incidence) compared to average notifications (or incidence) for the previous 5 years

**Figure 3: Number of laboratory reports of Cryptosporidium spp in England and Wales by week in 2017, and average number of reports by week in the period 2012 to 2016.**

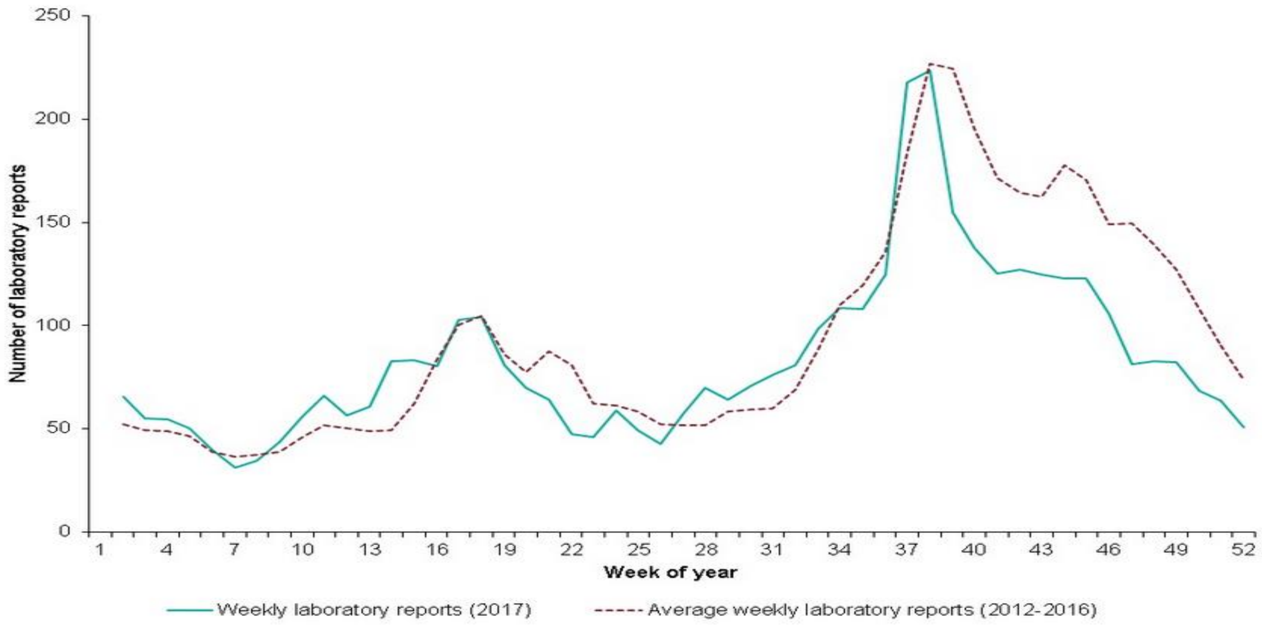



**Discussion:** *What is your interpretation of this graph?*

**Public Health England (2019):** Cryptosporidium data 2008 to 2017. https://www.gov.uk/government/publications/cryptosporidium-national-laboratory-data/cryptosporidium-data-2008-to-2017

# Weekly notifications (or incidence) compared to average notifications (or incidence) for the previous 5 years

**Figure 3: Number of laboratory reports of Cryptosporidium spp in England and Wales by week in 2017, and average number of reports by week in the period 2012 to 2016.**

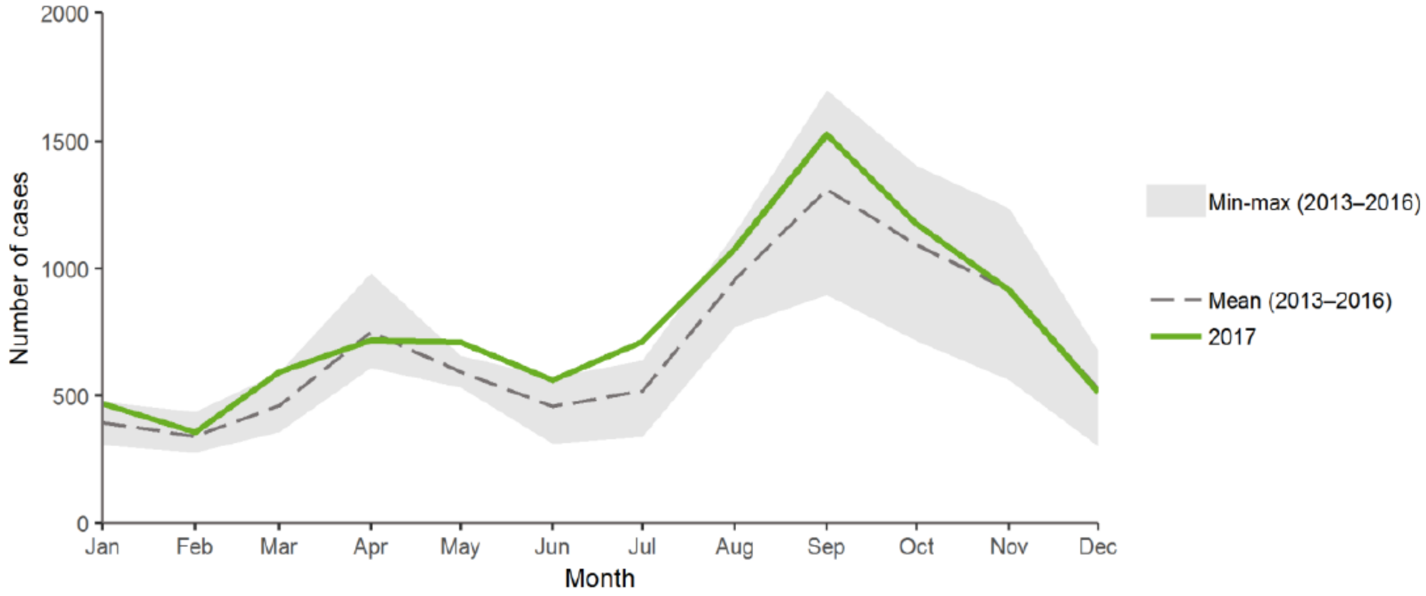



*Interpretation: The temporal distribution of Cryptosporidium in 2017 is similar to the previous 5 years, with cases peaking at a similar time. The overall number of cases for 2017 appears to be lower than the average for the previous 5 years.*


**Public Health England (2019):** Cryptosporidium data 2008 to 2017. https://www.gov.uk/government/publications/cryptosporidium-national-laboratory-data/cryptosporidium-data-2008-to-2017

# Monthly notifications compared to mean, minimum and maximum notifications for the previous 5 years

**Figure 3.** **Distribution of confirmed cryptosporidiosis cases by month, EU/EEA, 2017 and 2013–2016**



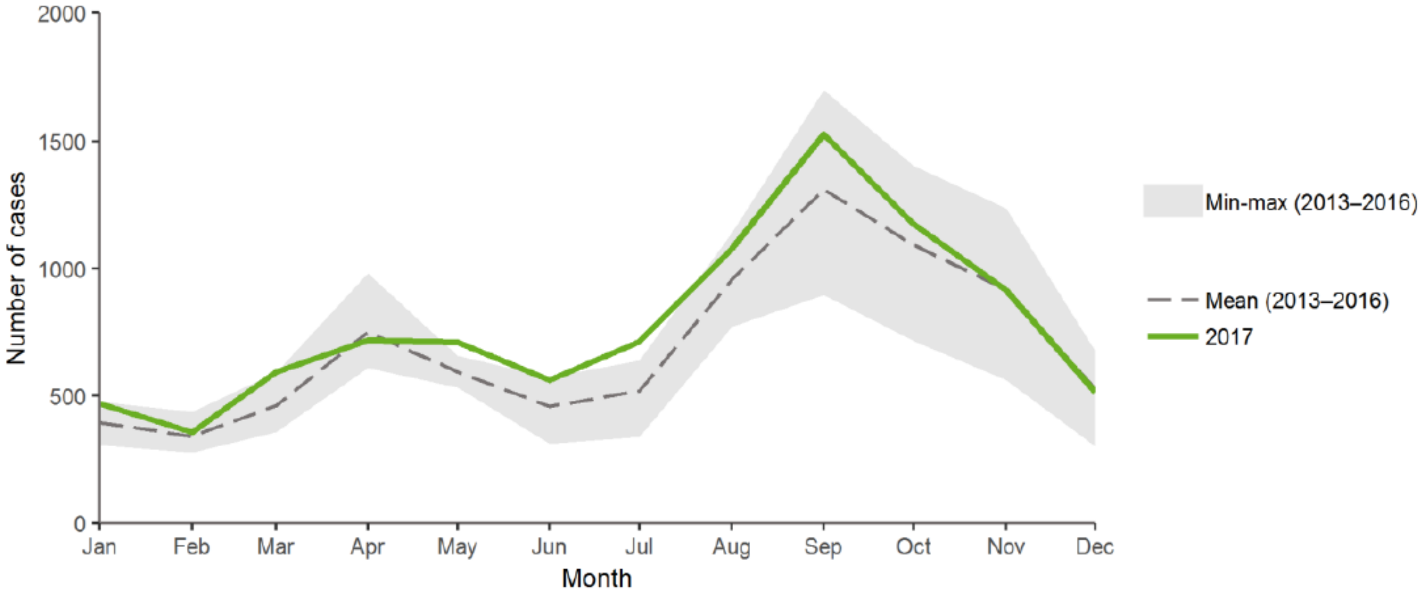

*Source: Country reports from Cyprus, the Czech Republic, Estonia, Finland, Germany, Hungary, Iceland, Ireland, Latvia, Lithuania, Malta, Norway, Poland, Romania, Slovakia, Slovenia, Spain, Sweden and the United Kingdom.*

**Discussion:** *What is your interpretation of this graph?*


**ECDC (2019):** Cryptosporidiosis. Annual Epidemiological Report for 2017. https://www.ecdc.europa.eu/sites/default/files/documents/AER_for_2017-cryptosporidiosis.pdf

# Monthly notifications compared to mean, minimum and maximum notifications for the previous 5 years

**Figure 3.** **Distribution of confirmed cryptosporidiosis cases by month, EU/EEA, 2017 and 2013–2016**



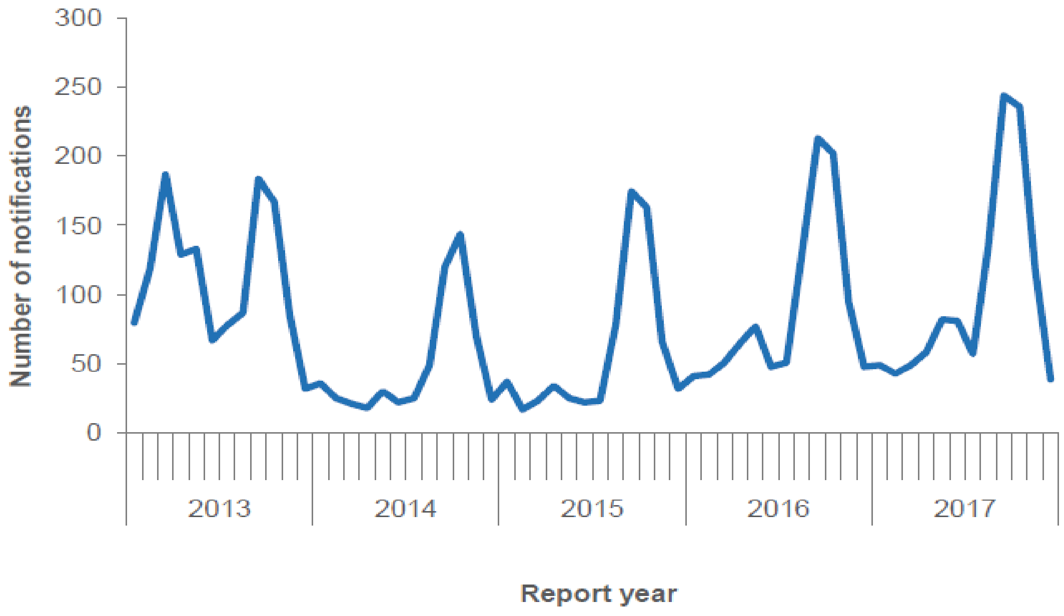

*Source: Country reports from Cyprus, the Czech Republic, Estonia, Finland, Germany, Hungary, Iceland, Ireland, Latvia, Lithuania, Malta, Norway, Poland, Romania, Slovakia, Slovenia, Spain, Sweden and the United Kingdom.*

*Interpretation: The number of monthly notifications of cryptosporidiosis are higher than the 5-year average and are at the higher limit of notifications observed over the past 5 years. There is a higher burden of cryptosporidiosis this year compared to previous years.*


**ECDC (2019):** Cryptosporidiosis. Annual Epidemiological Report for 2017. https://www.ecdc.europa.eu/sites/default/files/documents/AER_for_2017-cryptosporidiosis.pdf

# Monthly notifications over time

**Figure 6. Cryptosporidiosis notifications by month, January 2013–December 2017**

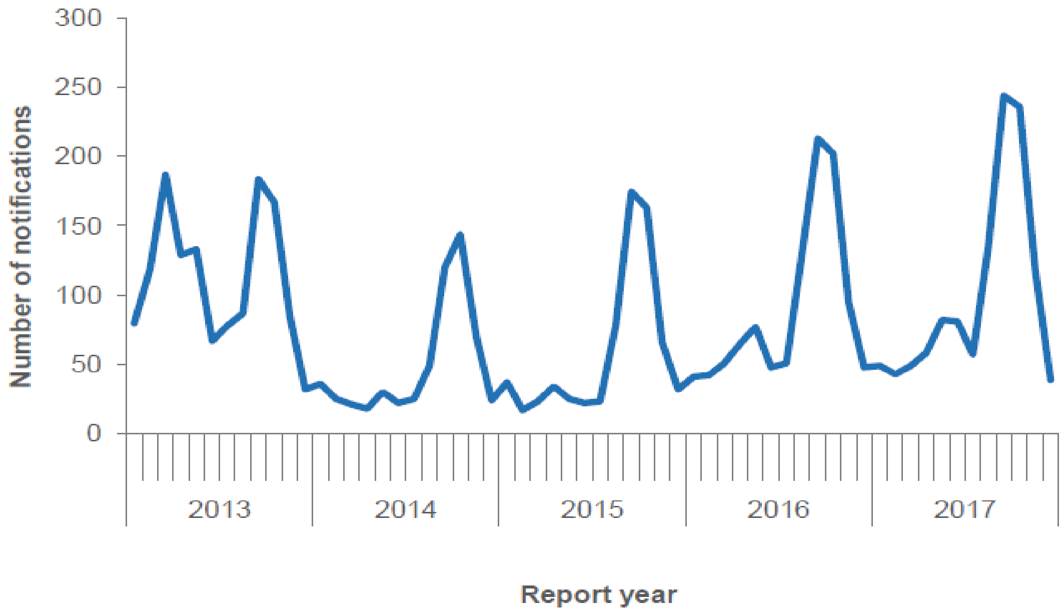



**Discussion:** *What is your interpretation of this graph?*

**The Institute of Environmental Science and Research Ltd. (2019):** Notifiable Diseases in New Zealand: Annual Report 2017. https://surv.esr.cri.nz/PDF_surveillance/AnnualRpt/AnnualSurv/2017/2017AnnualNDReport_FINAL.pdf

# Monthly notifications over time

**Figure 6. Cryptosporidiosis notifications by month, January 2013–December 2017**




*Interpretation: Cryptosporidium follows a seasonal pattern, with most notifications occurring between October and November. There has been an upward trend in notifications over the past 4 years.*


**The Institute of Environmental Science and Research Ltd. (2019):** Notifiable Diseases in New Zealand: Annual Report 2017. https://surv.esr.cri.nz/PDF_surveillance/AnnualRpt/AnnualSurv/2017/2017AnnualNDReport_FINAL.pdf

# Analysis by place

# Spatial analyses

- Identify high-risk areas for WRID
- Simple analyses using tables and graphs.
- Use geographic information systems to map the distribution of surveillance indicators by geographical area or water supply zone
  - Number of cases
  - Incidence rates
  - Complaints to water companies
- Need a geographical marker
  - Postcode
  - Place of residence
  - Location of medical facility

# Simple tables of cases and rates

**Table 2: Regional distribution[2] of laboratory reports of Cryptosporidium in England and Wales: 2017**

| Country | Region | Number of laboratory reports | per 100,000 population |
|---|---|---|---|
| England | East Midlands | 378 | 7.9 |
| England | East of England | 539 | 8.7 |
| England | London | 250 | 2.8 |
| England | North East | 275 | 10.4 |
| England | North West | 554 | 7.6 |
| England | South East | 582 | 6.4 |
| England | South West | 590 | 10.6 |
| England | Yorkshire and The Humber | 450 | 8.3 |
| England | West Midlands | 414 | 7.1 |
| Wales | Wales | 260 | 8.3 |

- Compare number of cases and notification rates by region
- **Discussion**: *What is your interpretation of this table?*

**Public Health England (2019):** Cryptosporidium data 2008 to 2017. https://www.gov.uk/government/publications/cryptosporidium-national-laboratory-data/cryptosporidium-data-2008-to-2017

# Simple tables of cases and rates

**Table 2: Regional distribution[2] of laboratory reports of Cryptosporidium in England and Wales: 2017**

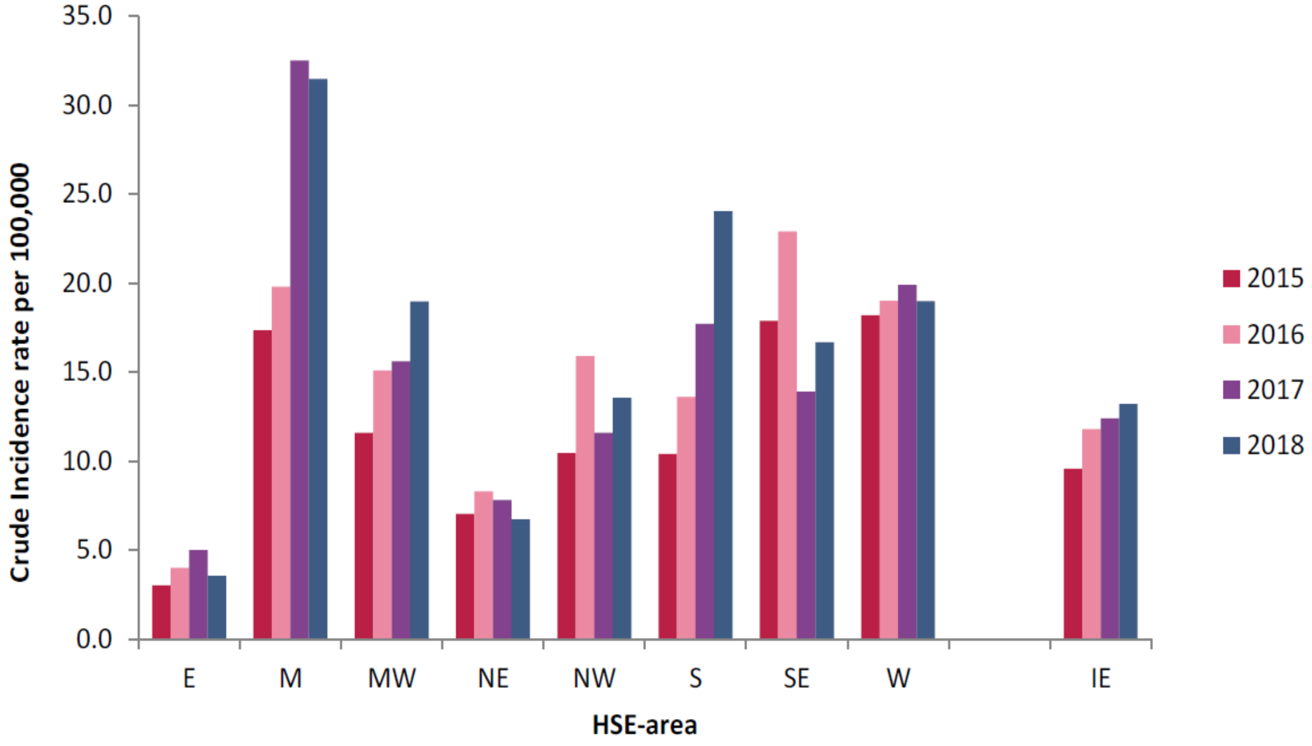
| Country | Region | Number of laboratory reports | per 100,000 population |
|---|---|---|---|
| England | East Midlands | 378 | 7.9 |
| England | East of England | 539 | 8.7 |
| England | London | 250 | 2.8 |
| England | North East | 275 | 10.4 |
| England | North West | 554 | 7.6 |
| England | South East | 582 | 6.4 |
| England | South West | 590 | 10.6 |
| England | Yorkshire and The Humber | 450 | 8.3 |
| England | West Midlands | 414 | 7.1 |
| Wales | Wales | 260 | 8.3 |

*Interpretation: The highest number of cases and the highest notification rate was reported from the South West region. The burden of Cryptosporidium is highest in the South West*

**Public Health England (2019):** Cryptosporidium data 2008 to 2017.
https://www.gov.uk/government/publications/cryptosporidium-national-laboratory-data/cryptosporidium-data-2008-to-2017

# Graph of rates by place over time

Figure 3. Regional crude incidence rates (CIR) cryptosporidiosis, Ireland, 2015-2018




## Discussion:

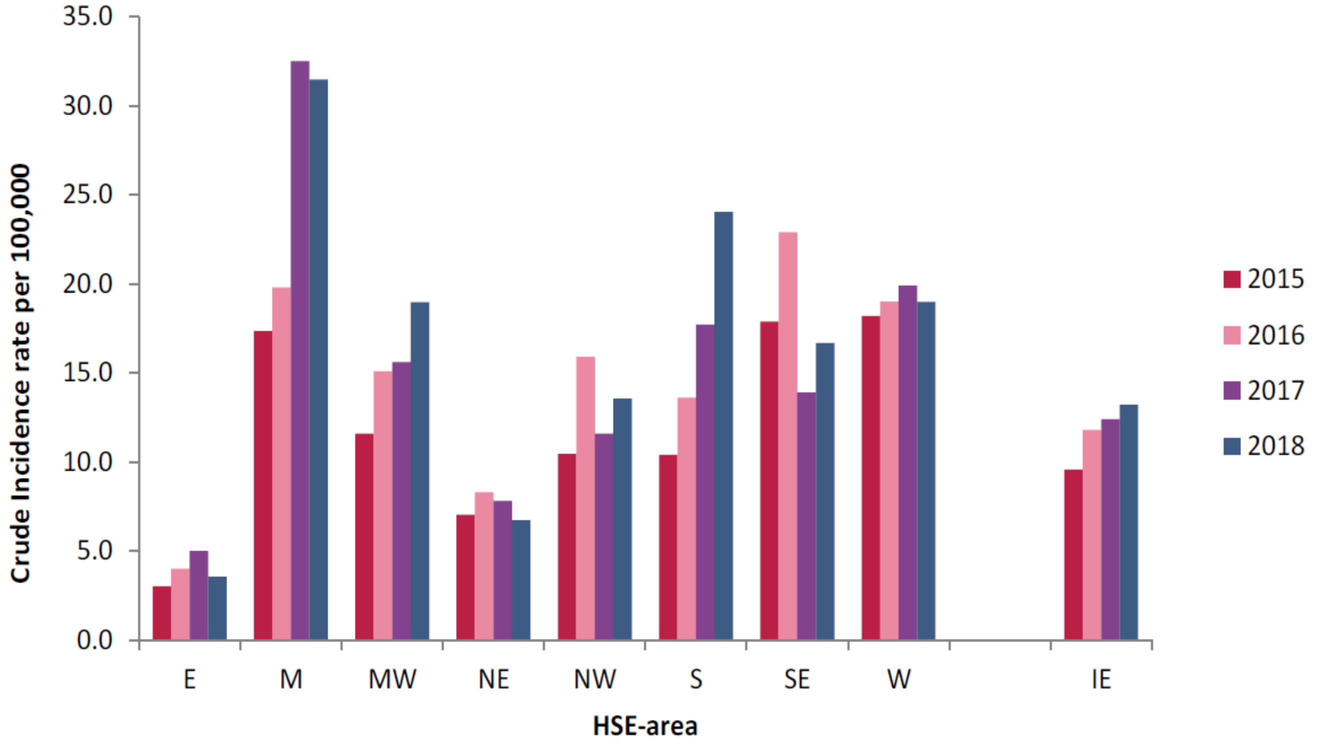
*What is your interpretation of this graph?*

*What are the possible explanations for the different distribution of cryptosporidiosis by region and over time?*


**HSE Health Protection Surveillance Centre (2019):** Cryptosporidiosis in Ireland, 2018. https://www.hpsc.ie/a-z/gastroenteric/cryptosporidiosis/publications/epidemiologyofcryptosporidiosisinirelandannualreports/Crypto%20Annual%20Report%202018.pdf

# Graph of rates by place over time

Figure 3. Regional crude incidence rates (CIR) cryptosporidiosis, Ireland, 2015-2018




*Interpretation: The highest notification rate was reported from the midlands. Consistently over the past 5 years, and particularly in the last two years, the burden of cryptosporidiosis has been highest in the midlands. The eastern region has the lowest burden of disease.*

*Over the past 5 years the incidence of cryptosporidium in Ireland has been increasing*


**HSE Health Protection Surveillance Centre (2019):** Cryptosporidiosis in Ireland, 2018. https://www.hpsc.ie/a-z/gastroenteric/cryptosporidiosis/publications/epidemiologyofcryptosporidiosisinirelandannualreports/Crypto%20Annual%20Report%202018.pdf

# Maps of cases and rates

Figure 7. Cryptosporidiosis notifications by DHB, 2017

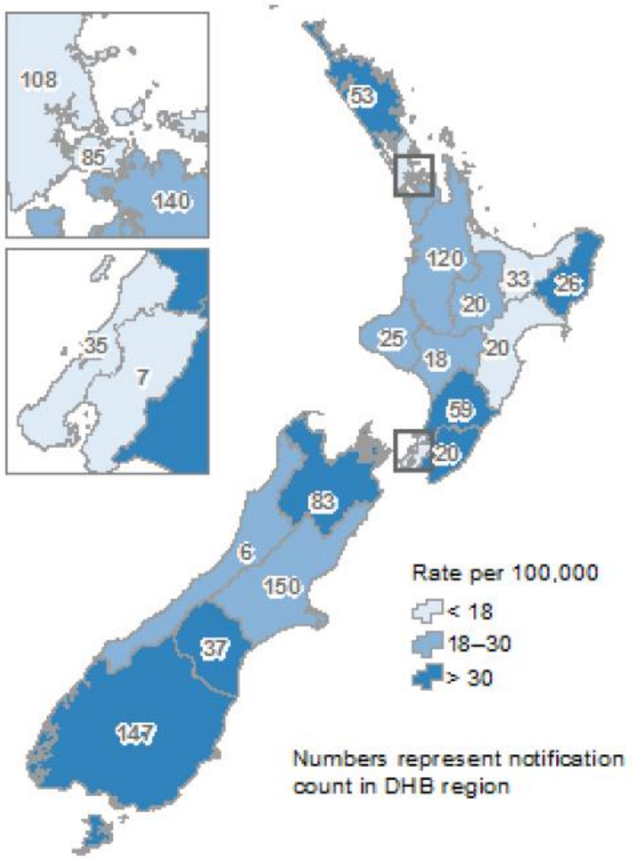
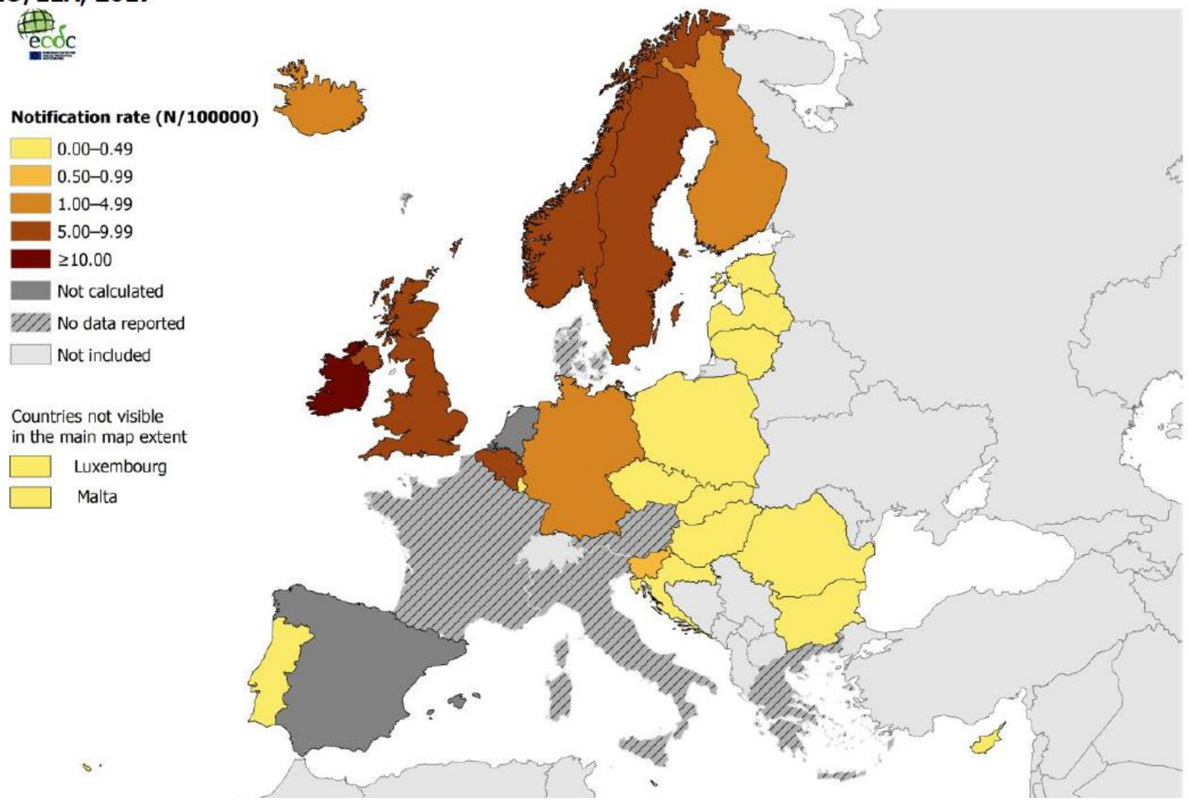



Figure 1. Distribution of confirmed cryptosporidiosis cases per 100 000 population by country, EU/EEA, 2017



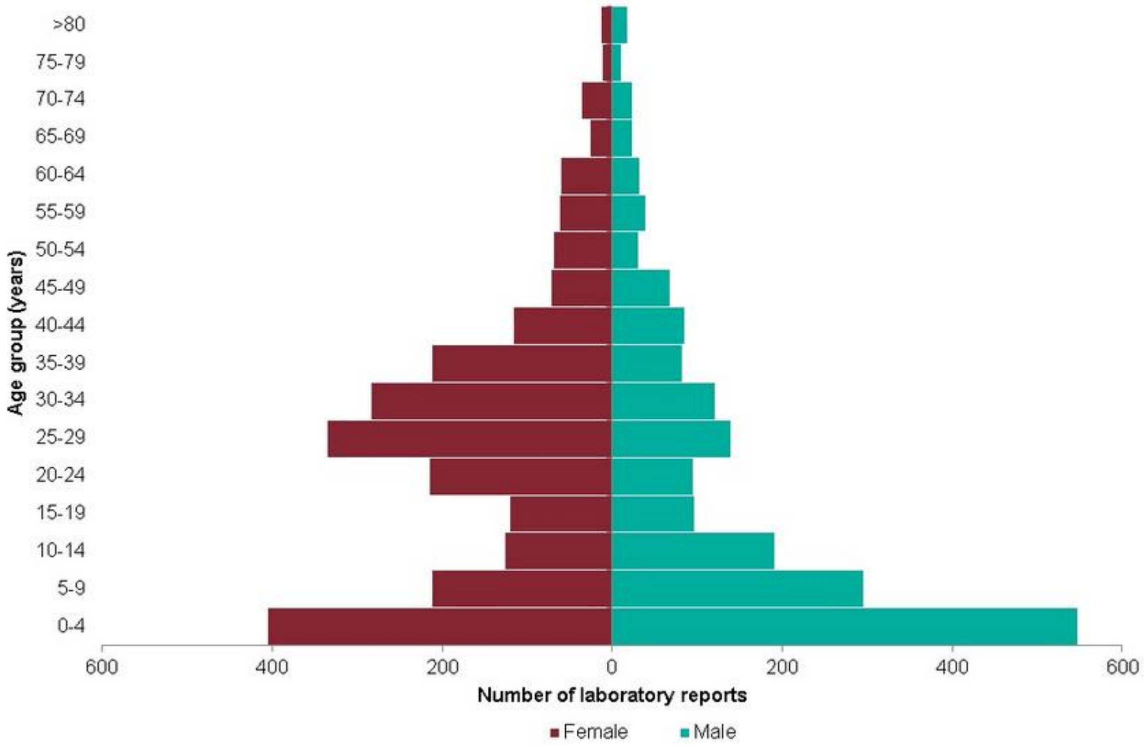

**The Institute of Environmental Science and Research Ltd. (2019):** Notifiable Diseases in New Zealand: Annual Report 2017. https://surv.esr.cri.nz/PDF_surveillance/AnnualRpt/AnnualSurv/2017/2017AnnualNDReport_FINAL.pdf

**ECDC (2019):** Cryptosporidiosis. Annual Epidemiological Report for 2017. https://www.ecdc.europa.eu/sites/default/files/documents/AER_for_2017-cryptosporidiosis.pdf

# Analysis by person

# Analysis by age and sex – number of notifications

**Figure 2: Age and sex distribution of laboratory reports of Cryptosporidium spp reported in England and Wales: 2017[3].**



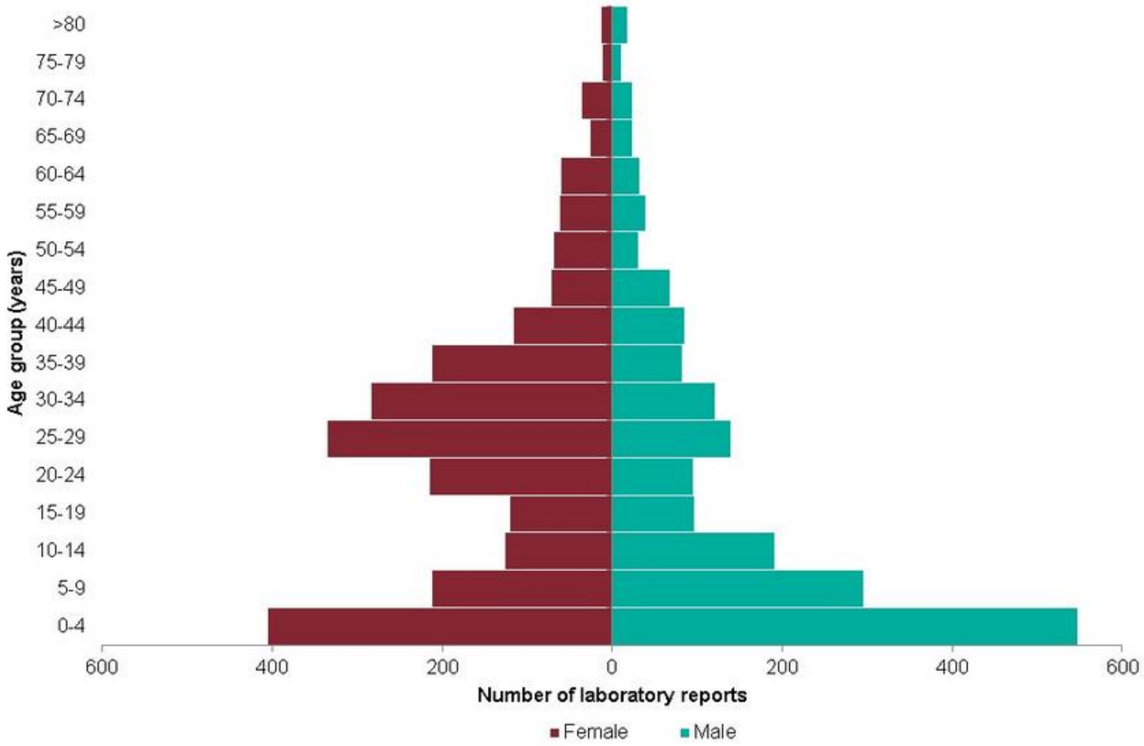

## Discussion:

*What is your interpretation of this graph?*

**Public Health England (2019):** Cryptosporidium data 2008 to 2017. https://www.gov.uk/government/publications/cryptosporidium-national-laboratory-data/cryptosporidium-data-2008-to-2017

# Analysis by age and sex – number of notifications

**Figure 2: Age and sex distribution of laboratory reports of Cryptosporidium spp reported in England and Wales: 2017[3].**



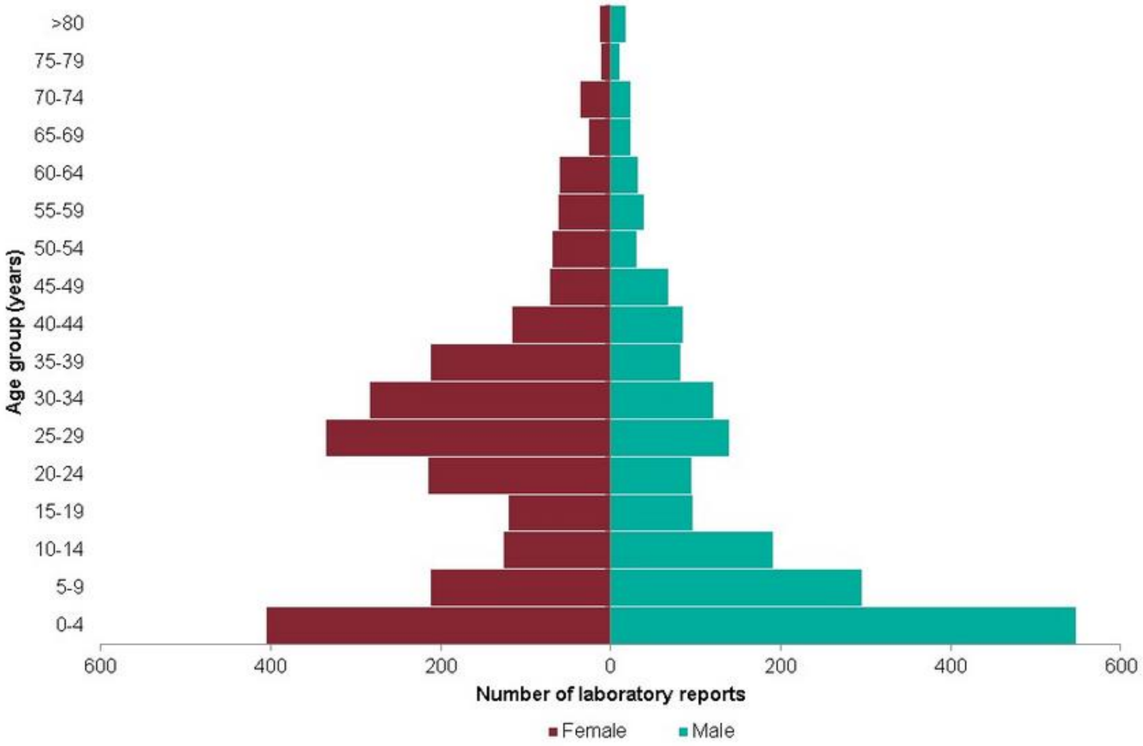

*Interpretation: The highest number of laboratory reports of cryptosporidium occurs in children aged 0-4 years old. In this age-group, the burden is highest in males. The burden of cryptosporidium is also high among women aged between 20 and 39.*

***What are the possible explanations for the different distribution of cryptosporidium by age and sex?***


**Public Health England (2019):** Cryptosporidium data 2008 to 2017. https://www.gov.uk/government/publications/cryptosporidium-national-laboratory-data/cryptosporidium-data-2008-to-2017

# Analysis by age and sex – number of notifications

**Figure 2: Age and sex distribution of laboratory reports of Cryptosporidium spp reported in England and Wales: 2017[3].**




***Young children:***

- Environmental exposure including exposure to animals
- Greater susceptibility
- More severe disease and greater care seeking

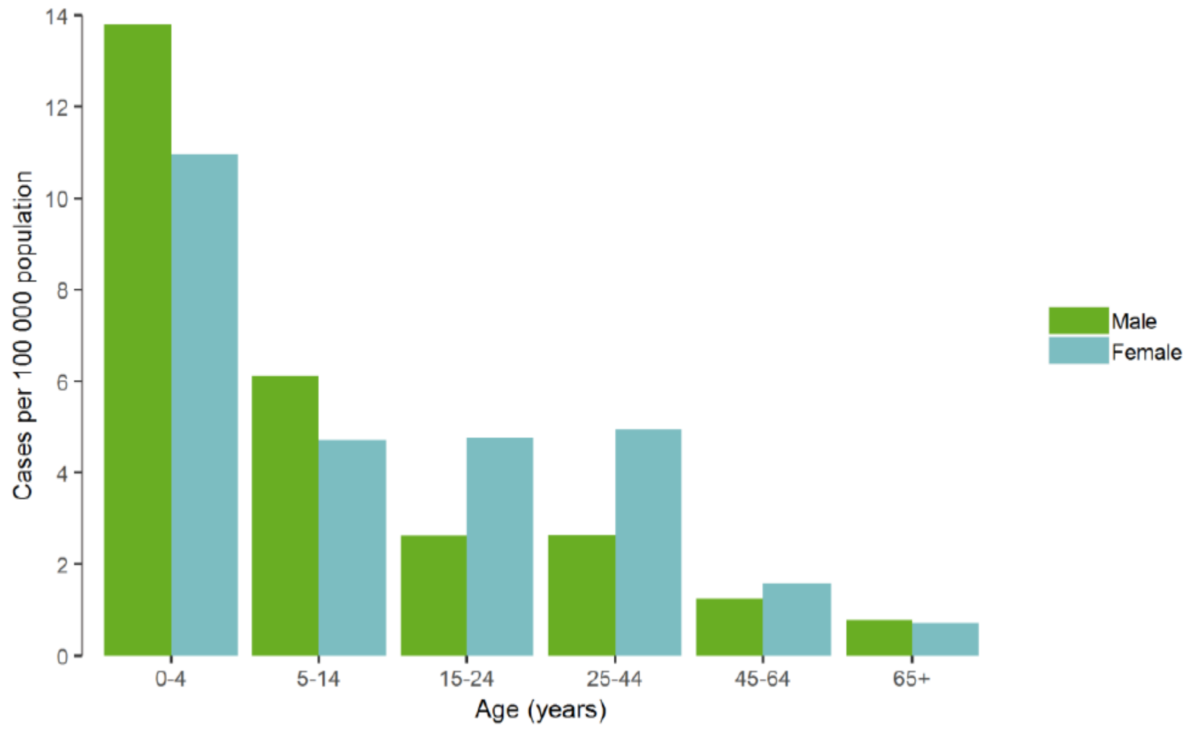
***Women aged 20-40:***

- Drink more water?
- Eat more salad and raw vegetables?
- More likely to seek care?


**Public Health England (2019):** Cryptosporidium data 2008 to 2017. https://www.gov.uk/government/publications/cryptosporidium-national-laboratory-data/cryptosporidium-data-2008-to-2017

# Analysis by age and sex

**Figure 4.** **Distribution of confirmed cryptosporidiosis cases per 100 000 population, by age and gender, EU/EEA, 2017**




## Discussion:

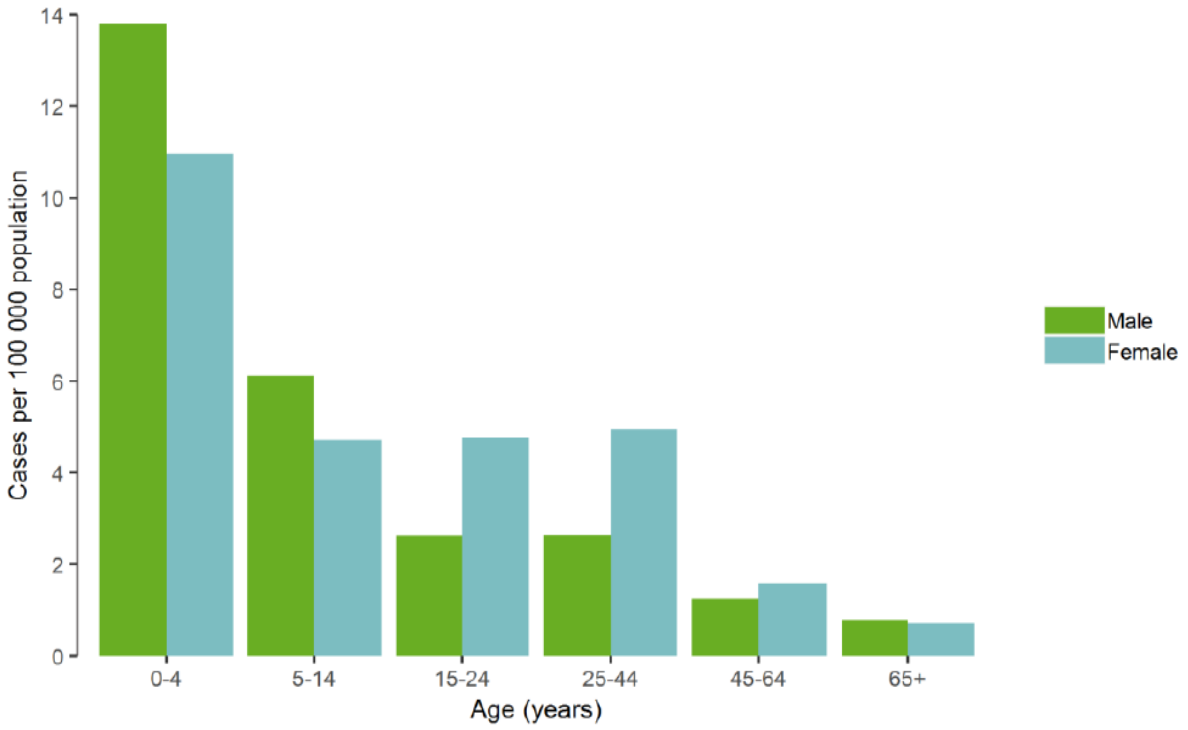
*What is your interpretation of this graph?*

*What is the advantage of this type of graph compared to the previous graph?*

**ECDC (2019):** Cryptosporidiosis. Annual Epidemiological Report for 2017. https://www.ecdc.europa.eu/sites/default/files/documents/AER_for_2017-cryptosporidiosis.pdf

# Analysis by age and sex

**Figure 4.** **Distribution of confirmed cryptosporidiosis cases per 100 000 population, by age and gender, EU/EEA, 2017**




*Interpretation: The notification rate is highest in those aged 0-4 years old, and particularly in males aged 0 to 4 years. Higher notification rates are also observed in women aged 15 to 24 and 25 to 44.*

**Advantages**:

Adjust for size of underlying population in each age and sex group


**ECDC (2019):** Cryptosporidiosis. Annual Epidemiological Report for 2017. https://www.ecdc.europa.eu/sites/default/files/documents/AER_for_2017-cryptosporidiosis.pdf

# Analysis by person – risk factors for infection

**Table 1. Number of cases (and percentage of cases where information available) where selected risk factors were reported for cryptosporidiosis cases (n=629), Ireland, 2018**

| Risk factor | Yes | No | UNK/NS | % of known |
|---|---|---|---|---|
| Travel outside of Ireland[a] | 43 | 435 | 151 | 9.0% |
| Lives/cared for on farm | 167 | 386 | 76 | 30.2% |
| Visited farm | 164 | 329 | 136 | 33.9% |
| *Lives/works on or visited farm*[b] | 296 | 213 | 120 | 58.1% |
| Swimming pool visit | 166 | 385 | 78 | 30.1% |
| Other water based activities | 45 | 397 | 187 | 10.2% |
| Contact with domestic pets | 365 | 164 | 100 | 69.0% |

Data source: CIDR

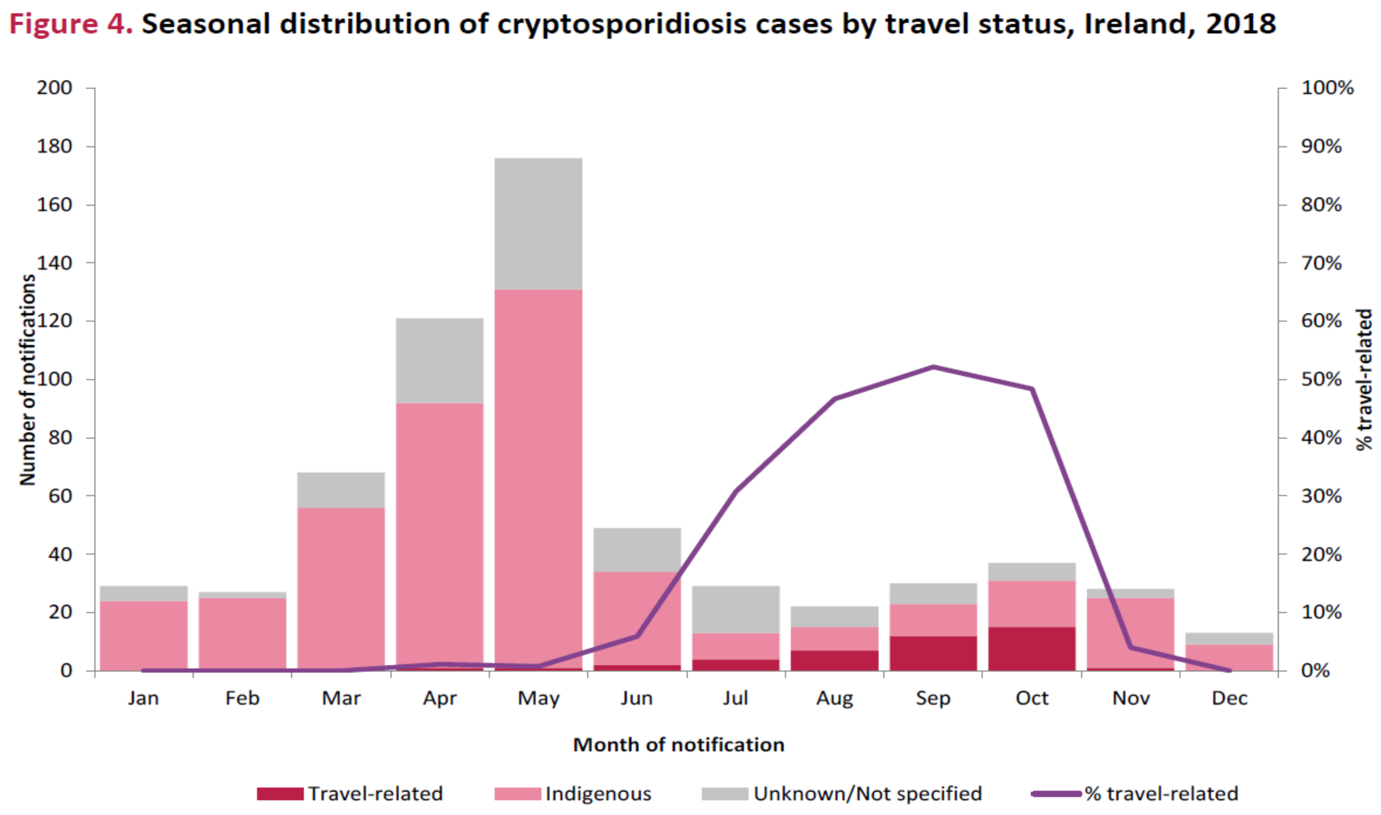
[a] *Based on country of infection variable*
[b] *Composite of the two previous variables*

***Discussion**: What is your interpretation of this table?*

**HSE Health Protection Surveillance Centre (2019):** Cryptosporidiosis in Ireland, 2018. https://www.hpsc.ie/a-z/gastroenteric/cryptosporidiosis/publications/epidemiologyofcryptosporidiosisinirelandannualreports/Crypto%20Annual%20Report%202018.pdf

# Other types of analyses

# Analysis by person and time –risk factors for infection




## Discussion:

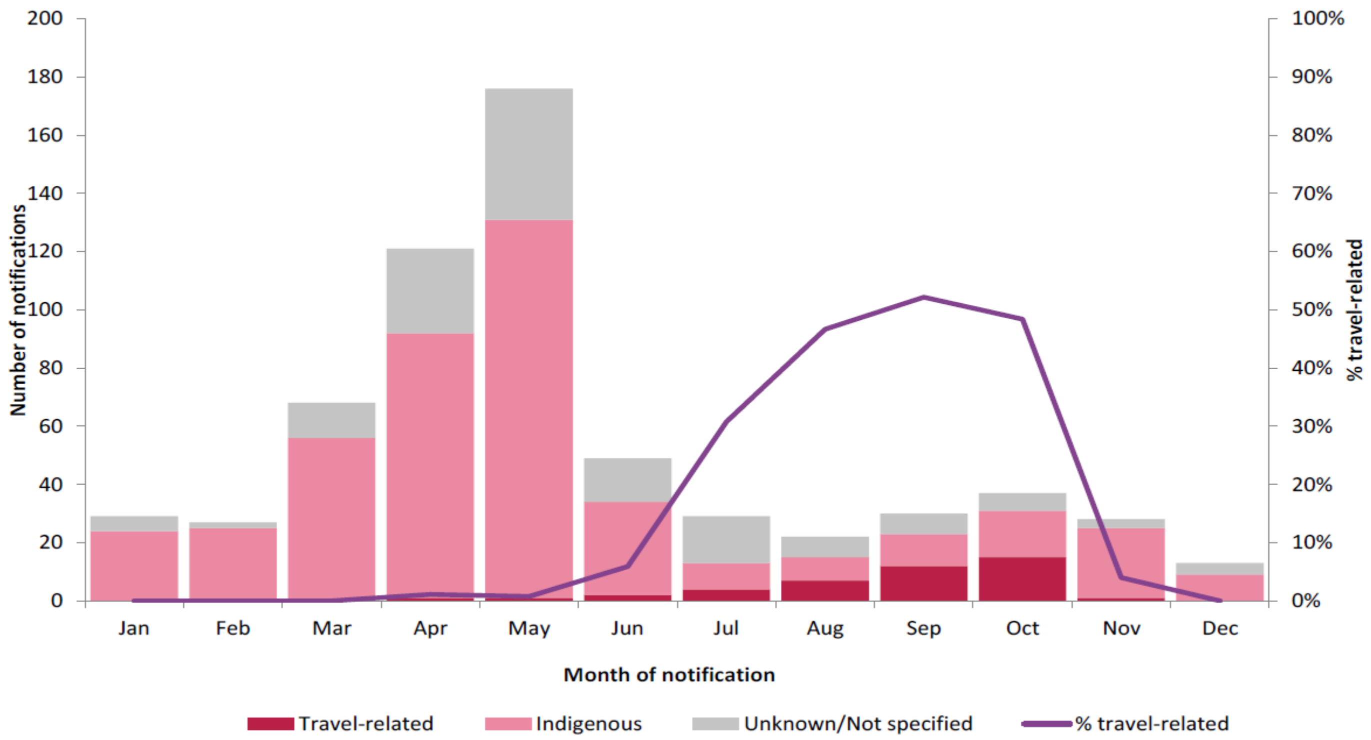
*What is your interpretation of this graph?*

*What are the possible explanations for the varying distribution of cryptosporidiosis in this graph?*

**HSE Health Protection Surveillance Centre (2019):** Cryptosporidiosis in Ireland, 2018. https://www.hpsc.ie/a-z/gastroenteric/cryptosporidiosis/publications/epidemiologyofcryptosporidiosisinirelandannualreports/Crypto%20Annual%20Report%202018.pdf

# Analysis by person and time –risk factors for infection

**Figure 4.** **Seasonal distribution of cryptosporidiosis cases by travel status, Ireland, 2018**




*Interpretation:*

*Cryptosporidiosis notifications peak in the spring. Travel associated cases are most frequently reported between July and October. The percentage of travel associated cases peaks in October.*

***Explanation:***

*Most rain falls in spring → increase in environmental exposure (water and farm exposures) and domestic notifications*

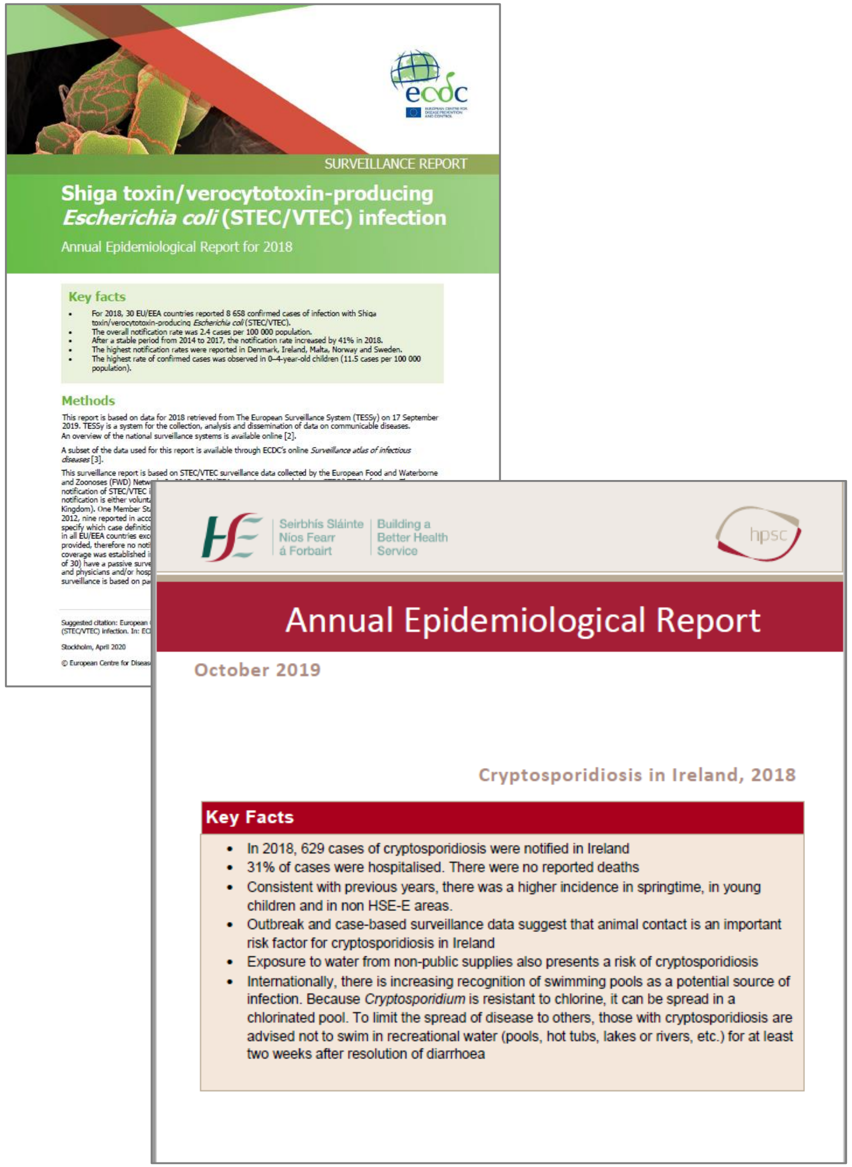
*Lambing and calving in spring*

*Most people travel overseas in the summer – higher numbers of travel associated cases*

**HSE Health Protection Surveillance Centre (2019):** Cryptosporidiosis in Ireland, 2018. https://www.hpsc.ie/a-z/gastroenteric/cryptosporidiosis/publications/epidemiologyofcryptosporidiosisinirelandannualreports/Crypto%20Annual%20Report%202018.pdf

# Surveillance Bulletins

- Regularly communicate results of surveillance to stakeholders (weekly, monthly, quarterly)
  - Inform decision making for public health action
  - Demonstrate the purpose and usefulness of surveillance to those working on surveillance
- Incorporate into existing surveillance bulletins (enteric pathogens, food and waterborne illness bulletin, or communicable diseases bulletin)
- Disseminate to stakeholders (water providers, regulators etc)
- Make publicly available (public health agency website)



# Outline for a surveillance bulletin

- Key messages / summary
- Introduction (brief)
- Methods (brief)
- Epidemiology
  - Time (trends in notifications or rates)
  - Person (age, sex, other risk factors (travel)
  - Place
- Outbreaks
- Discussion / conclusions

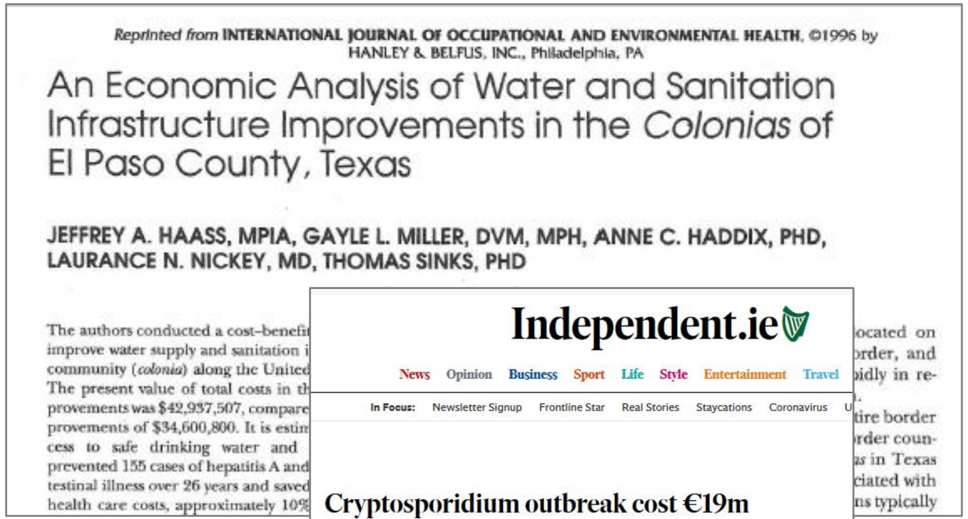
# Key messages

- Summarise the main findings and take home messages of the report
  - ***What is the ONE message you want the audience to take away from this report?***
  - ***What is the ONE message the reader needs to understand?***

- Focus the key messages on:
  - The most important conclusions arising from the analyses
  - The most important facts you want to communicate to the reader (3 or 4 facts)

# Using surveillance data for advocacy

- Inform development of policy, regulations and guidelines
- Identify priorities and where to target resources for improving the water system
- Estimate impact of WRID –disability adjusted life years, quality adjusted life years, direct costs (healthcare utilisation) and indirect costs (work absenteeism and productivity losses)
- Evaluate impact of control measures
  - impact on incidence after the introduction of the control measure
  - cost benefit analyses



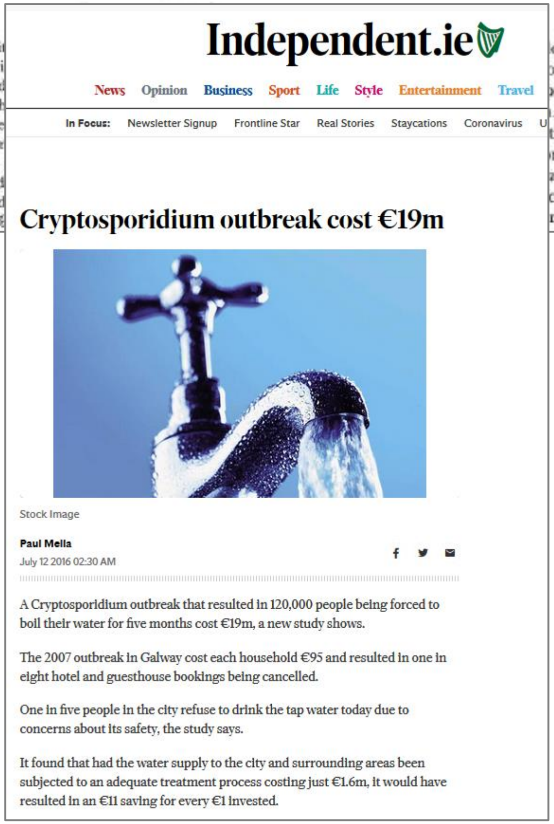
Reprinted from INTERNATIONAL JOURNAL OF OCCUPATIONAL AND ENVIRONMENTAL HEALTH, ©1996 by HANLEY & BELFUS, INC., Philadelphia, PA

An Economic Analysis of Water and Sanitation Infrastructure Improvements in the *Colonias* of El Paso County, Texas

JEFFREY A. HAASS, MPIA, GAYLE L. MILLER, DVM, MPH, ANNE C. HADDIX, PHD, LAURANCE N. NICKEY, MD, THOMAS SINKS, PHD

The authors conducted a cost–benefi[...] improve water supply and sanitation i[...] community (*colonia*) along the United[...] The present value of total costs in th[...] provements was $42,937,507, compare[...] provements of $34,600,800. It is estim[...] cess to safe drinking water and [...] prevented 155 cases of hepatitis A and [...] testinal illness over 26 years and saved [...] health care costs, approximately 10%[...]

[...]ocated on [...]order, and [...]pidly in re-[...] [...]tire border [...]order coun-[...]s in Texas [...]ciated with [...]ns typically



Independent.ie

News Opinion Business Sport Life Style Entertainment Travel

In Focus: Newsletter Signup Frontline Star Real Stories Staycations Coronavirus

Cryptosporidium outbreak cost €19m

Stock Image

Paul Melia

July 12 2016 02:30 AM

A Cryptosporidium outbreak that resulted in 120,000 people being forced to boil their water for five months cost €19m, a new study shows.

The 2007 outbreak in Galway cost each household €95 and resulted in one in eight hotel and guesthouse bookings being cancelled.

One in five people in the city refuse to drink the tap water today due to concerns about its safety, the study says.

It found that had the water supply to the city and surrounding areas been subjected to an adequate treatment process costing just €1.6m, it would have resulted in an €11 saving for every €1 invested.

**Questions?**