

Identifying Signature Features of Epidemic Diseases in 19th Century All-cause Mortality Data

Rasmus Kristoffer Pedersen

Postdoc, PandemiX Center
Dept. Science and Environment,
Roskilde University, Denmark
Email: rakrpe@ruc.dk

Joint work with

*Mathias Mølbak Ingholt, Maarten van Wijhe,
Viggo Andreasen & Lone Simonsen*

Epidemics 9, November 30th, 2023



Danmarks
Grundforskningsfond
Danish National
Research Foundation

Introduction

- ▶ Historical data provides us with more examples of epidemics than modern data alone.

Introduction

Background and data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

Results and discussion

Representative signature features

Grouping crises

General discussion



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

General discussion

Introduction

- ▶ Historical data provides us with more examples of epidemics than modern data alone.
- ▶ Our response to emerging diseases come from historical experience.





Image from The New York Times article "*The Mask Slackers of 1918*", Aug. 3, 2020
<https://www.nytimes.com/2020/08/03/us/mask-protests-1918.html>



Introduction

Background and data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

Results and discussion

Representative signature features

Grouping crises

General discussion

Introduction

Background and data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

Results and discussion

Representative signature features

Grouping crises

General discussion

- ▶ Historical data provides us with more examples of epidemics than modern data alone.
- ▶ Our response to emerging diseases come from historical experience.
 - ▶ Quarantine - e.g. plague
 - ▶ Restriction of movement (*cordon sanitaire*) - e.g. cholera
 - ▶ Social distancing - e.g. 1918 influenza
 - ▶ Masks - e.g. 1918 influenza



Introduction

Background and data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

Results and discussion

Representative signature
features

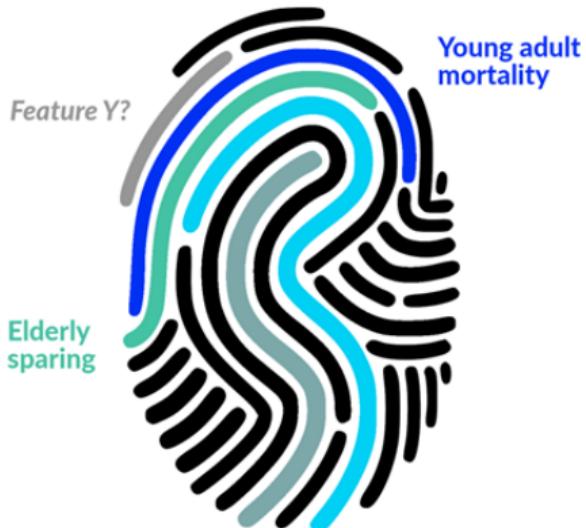
Grouping crises

General discussion

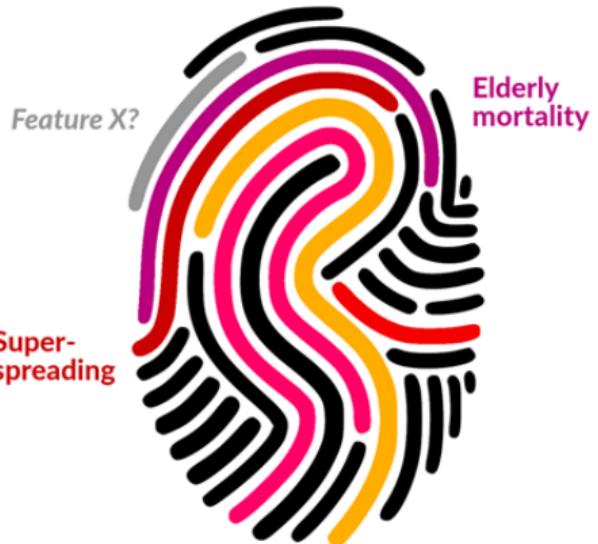
- ▶ Historical data provides us with more examples of epidemics than modern data alone.
- ▶ Our response to emerging diseases come from historical experience.
 - ▶ Quarantine - e.g. plague
 - ▶ Restriction of movement (*cordon sanitaire*) - e.g. cholera
 - ▶ Social distancing - e.g. 1918 influenza
 - ▶ Masks - e.g. 1918 influenza
- ▶ The pandemics of recent years may only be a subset of potential threats to consider for surveillance.



1918 Influenza



SARS-CoV-2



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

General discussion

Overview of talk

In this talk, I will talk about:

- Our recent study of epidemics in 19th century Denmark.



Introduction**Background and
data handling**

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

**Results and
discussion**Representative signature
features

Grouping crises

General discussion

Overview of talk

In this talk, I will talk about:

- ▶ Our recent study of epidemics in 19th century Denmark.
- ▶ Mortality baseline calculation.



Introduction**Background and
data handling**

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

**Results and
discussion**Representative signature
features

Grouping crises

General discussion

Overview of talk

In this talk, I will talk about:

- ▶ Our recent study of epidemics in 19th century Denmark.
- ▶ Mortality baseline calculation.
- ▶ Age pattern analysis.



Introduction**Background and
data handling**

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

**Results and
discussion**Representative signature
features

Grouping crises

General discussion

Overview of talk

In this talk, I will talk about:

- ▶ Our recent study of epidemics in 19th century Denmark.
- ▶ Mortality baseline calculation.
- ▶ Age pattern analysis.
- ▶ Reflect on what we learned.



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

Results and
discussionRepresentative signature
features

Grouping crises

General discussion

Data source

No.	Dødsdagen.	Begravelsedagen.	Den Dødes Navn og tilnavn.	Stand, Haandtering og Befolkningsd.	Ålder.	Hvem an- først i det almindeligt gæ. Den foretak Registret.	Emner
35.	29. August	4. Sept.	Jønders Jønson	Gjenn i Kjønsgå	53 Års	692. 138	
36.	31. August	4. Sept.	Hans Carlsen	Først i Skælværet	63 Års	692. 139	
37.	30. August	3. Sept.	Ole Jensen	Møgdomsmann i Glæslefjord	39 Års	692. 140	
38.	3. Sept.	7. Sept.	Olger Larsen	Gjenn i Kjønsgå	70 Års	692. 141	
39.	31. August	4. Sept.	Hans Olen	Gjenn og Senge i Kjønsgå	42 Års	692. 142	
40.	4. Sept.	6. Sept.	Niels Pedersen	Først i den store	61 Års	692. 143	
41.	5. Sept.	9. Sept.	Ole Hansen	Møgdomsmann i Glæslefjord	63 Års	692. 144	
42.	4. Sept.	9. Sept.	Niels Christensen	Gjenn i Kysten	57 Års	692. 145	
43.	7. Sept.	12. Sept.	Niels Løyen	Gjenn i Lægen	80 Års	692. 146	
44.	6. Sept.	12. Sept.	Jens Andersen	Giftmord og døden i Glæslefjord	72 Års	692. 147	
45.	8. Sept.	13. Sept.	Hans Lachsen	Gjenn i Kysten	42 Års	692. 148	
46.	5. Sept.	9. Sept.	Lars Christophersen	Møgdomsmann i Glæslefjord	9.3. Års	692. 149	
47.	12. Sept.	16. Sept.	Kristian Hærumesen	Gjenn i Glæslefjord	78 Års	692. 150	
48.	11. Sept.	14. Sept.	Tharins Larsen	Gjenn i Glæslefjord	49 Års	692. 151	



Data source

- Parish registers for Danish church parishes between 1815-1915

Sn.	Dato	Baptismus.	En. Dato. En. og Eller.	Mord. Kværling og Dødsdag.	Sted. af begravelse.	Tændes.
25.	17. Februar.	Jørgen. Linders. Sørensen.	Januar i 1815.	-	Blaa. 1815.	
26.	31. Januar.	August. Hans. Carlsen.	Januar i 1815.	-	Blaa. 1815.	
27.	30. Februar.	3. marts. Ole. Jensen.	1815.	-	Blaa. 1815.	
28.	3. Mars.	7. marts. Hager. Larsen.	Januar i 1815.	-	Blaa. 1815.	
29.	31. Januar.	4. marts. Hans. Olof.	Januar i 1815.	-	Blaa. 1815.	
30.	1. Februar.	6. marts. Ole. Andersen.	Januar i 1815.	-	Blaa. 1815.	
31.	3. Februar.	9. marts. Ole. Hansen.	Januar i 1815.	-	Blaa. 1815.	
32.	4. Februar.	Gratius. Ole. Nielsen. Carstensen.	Januar i 1815.	-	Blaa. 1815.	
33.	7. Februar.	Bjørn. Ole. Larsen.	Januar i 1815.	-	Blaa. 1815.	
34.	10. Februar.	Bjørn. Jens. Andersen.	Januar i 1815.	-	Blaa. 1815.	
35.	8. Februar.	3. marts. Hans. Leth.	Januar i 1815.	-	Blaa. 1815.	
36.	3. Mars.	4. marts. Lars. Christensen.	Januar i 1815.	-	Blaa. 1815.	
37.	11. Februar.	12. marts. Peter. Hansen.	Januar i 1815.	-	Blaa. 1815.	
38.	11. Februar.	14. marts. Hans. Larsen.	Januar i 1815.	-	Blaa. 1815.	

Scan of parish register for "Fakse" parish.

Data source

- Parish registers for Danish church parishes between 1815-1915
- Approximately 4 million burials

MS 1831

Døde.

Ramfjord.

Sn.	Eddate.	Begravelsessted.	Døde navn	Åldre ved døden	Begravelsesdato.	Tilhørige.
25.	1. Februar	Jørgen Brøndum	Jørgen i Jydske	58	1831 feb 1	Brøndum
26.	3. Februar	Hans Carlsten	Jørgen i Hals	59	1831 feb 3	Carlsten
27.	30. Januar	3. marts	Ol. Jensen	58	1831 jan 30	Jensen
28.	3. Februar	7. marts	Hans Larsen	58	1831 feb 3	Larsen
29.	31. Januar	4. marts	Hans Olof	58	1831 jan 31	Olof
30.	1. Februar	6. marts	Niels Andersen	58	1831 feb 1	Andersen
31.	2. Februar	8. marts	Ol. Hansen	58	1831 feb 2	Hansen
32.	3. Februar	9. marts	Andreas Christensen	58	1831 feb 3	Christensen
33.	4. Februar	9. marts	Jørgen i Jydske	58	1831 feb 4	
34.	5. Februar	10. marts	Ol. Vold Larsen	58	1831 feb 5	Larsen
35.	7. Februar	10. marts	Jens Andersen	58	1831 feb 7	Andersen
36.	8. Februar	3. april	Hans Løgten	58	1831 feb 8	Løgten
37.	9. Februar	4. april	Andreas Christensen	58	1831 feb 9	Christensen
38.	10. Februar	11. april	Andreas Christensen	58	1831 feb 10	Christensen
39.	11. Februar	14. april	Hans Larsen	58	1831 feb 11	Larsen

Scan of parish register for "Fakse" parish.

Data source

- Parish registers for Danish church parishes between 1815-1915
- Approximately 4 million burials
- Individual level information
 - Includes date, age, gender and parish

Sn.	Født/død	Begravelsesdag	Em. dødsdag og tilhører	Døds. omstændighed og dødsårsak	Åldre	Tilh. dødsårsak	Ålder ved døden	Tilh. dødsårsak
25.	27. Februar	10. Marts	Lindore. Sørensen	Avd. i gennemgående	58	Avd. i gennemgående	58	Avd. i gennemgående
26.	31. Des.	8. Januar	Hans Carlsen	Avd. i gennemgående	58	Avd. i gennemgående	58	Avd. i gennemgående
27.	30. Des.	3. Januar	Ol. Jensen	Udledes med dødehånd	94	Udledes med dødehånd	94	Udledes med dødehånd
28.	3. Januar	7. Januar	Hager Larsen	gærd i gennemgående	70	gærd i gennemgående	70	gærd i gennemgående
29.	31. Des.	4. Januar	Hans Olof	Udledes med dødehånd	70	Udledes med dødehånd	70	Udledes med dødehånd
30.	1. Januar	8. Januar	Vilh. Andersen	Surf. i dødehånd	68	Surf. i dødehånd	68	Surf. i dødehånd
31.	3. Januar	9. Januar	Ol. Hansen	Udledes med dødehånd	68	Udledes med dødehånd	68	Udledes med dødehånd
32.	4. Januar	9. Januar	Niels Christensen	Surf. i gennemgående	57	Surf. i gennemgående	57	Surf. i gennemgående
33.	7. Januar	12. Januar	Ol. Vald. Larsen	gærd i sengen	80	gærd i sengen	80	gærd i sengen
34.	8. Januar	13. Januar	Jens. Sørensen	Udledes med dødehånd	68	Udledes med dødehånd	68	Udledes med dødehånd
35.	8. Januar	13. Januar	Hans. Lethsen	Surf. i gennemgående	58	Surf. i gennemgående	58	Surf. i gennemgående
36.	9. Januar	14. Januar	Niels. Christensen	Udledes med dødehånd	68	Udledes med dødehånd	68	Udledes med dødehånd
37.	11. Januar	16. Januar	M. Hansen	Surf. i gennemgående	70	Surf. i gennemgående	70	Surf. i gennemgående
38.	11. Januar	16. Januar	Hans. Larsen	Surf. i gennemgående	70	Surf. i gennemgående	70	Surf. i gennemgående

Scan of parish register for "Fakse" parish.

Data source

- Parish registers for Danish church parishes between 1815-1915
- Approximately 4 million burials
- Individual level information
 - Includes date, age, gender and parish
- Property of the Danish National Archives, but digitized and transcribed by *Ancestry*

Nr.	Fødselsdag.	Begravelsesdag.	En. Dato. En. og Eller.	Overs. Begravelse og Begravelses-	Klare. Begravelse og Begravelses-	Tidspunkt.	Tidspunkt.
				størrelse	størrelse	størrelse	størrelse
25.	1. Februar	1. Februar	Lindore. Sørensen	Jens i 1815	1815	Februar 1815	Februar 1815
26.	3. Februar	3. Februar	Hans. Carlsen	Jens i 1815	1815	Februar 1815	Februar 1815
27.	30. Januar	3. Februar	Ø. Jensen	Jens i 1815	1815	Februar 1815	Februar 1815
28.	3. Februar	7. Februar	Hans. Larsen	Jens i 1815	1815	Februar 1815	Februar 1815
29.	31. Desember	4. Januar	Hans. Olesen	Jens i 1815	1815	Februar 1815	Februar 1815
30.	1. Januar	8. Januar	Niels. Andersen	Jens i 1815	1815	Februar 1815	Februar 1815
31.	3. Januar	9. Januar	Ø. Hansen	Jens i 1815	1815	Februar 1815	Februar 1815
32.	4. Januar	9. Januar	Niels. Christensen	Jens i 1815	1815	Februar 1815	Februar 1815
33.	7. Januar	13. Januar	Ø. Vald. Larsen	Jens i 1815	1815	Februar 1815	Februar 1815
34.	10. Januar	17. Januar	Jens. Andersen	Jens i 1815	1815	Februar 1815	Februar 1815
35.	8. Januar	13. Januar	Hans. Larsen	Jens i 1815	1815	Februar 1815	Februar 1815
36.	3. Januar	9. Januar	Niels. Christensen	Jens i 1815	1815	Februar 1815	Februar 1815
37.	11. Januar	17. Januar	Hans. Hansen	Jens i 1815	1815	Februar 1815	Februar 1815
38.	11. Januar	17. Januar	Hans. Hansen	Jens i 1815	1815	Februar 1815	Februar 1815

Scan of parish register for "Fakse" parish.

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

Results and
discussionRepresentative signature
features

Grouping crises

General discussion

Data cleaning and managing

Date of burial	Name	Age	Gender	Amt	Sogn
1857-01-02	Ane Kirstine Christensen	2	Female	Thisted Amt	Jannerup Sogn (Thisted Amt)
1857-01-02	Birthe Marie Christensen.	0	Female	Thisted Amt	Hundborg Sogn
1857-01-02	Ane Marie Mortensdatter	81	Female	Thisted Amt	Vejerslev Sogn (Mors)
1857-01-02	Gjertrud Jensdatter	82	Female	Thisted Amt	Thisted Sogn
1857-01-02	Karen Christensdatter Wiilsbøll	52	Female	Thisted Amt	Vester Vandet Sogn
1857-01-02	Karen Marie Jensen	21	Female	Thisted Amt	Sennels Sogn
1857-01-02	Ane Christensdatter Krogh	76	Female	Thisted Amt	Vester Vandet Sogn
1857-01-03	Anders Hansen Tülfang	79	Male	Thisted Amt	Gøttrup Sogn
1857-01-03	Marcus Christensen	6	Male	Thisted Amt	Hunstrup Sogn
1857-01-04	Maren Jensen	0	Female	Thisted Amt	Hillerslev Sogn (Thisted Amt)
1857-01-04	Thomas Jensen	0	Male	Thisted Amt	Flade Sogn (Thisted Amt)
1857-01-04	Niels Madsen Thÿstrup	74	Male	Thisted Amt	Skjoldborg Sogn
1857-01-04	Poul Pedersen	72	Male	Thisted Amt	Villerslev Sogn
1857-01-04	Oline Christine Christensen	1	Female	Thisted Amt	Kollerup Sogn (Thisted Amt)
1857-01-04	Maren Cathrine Nielsen	2	Female	Thisted Amt	Kollerup Sogn (Thisted Amt)



Data cleaning and managing

Date of burial	Name	Age	Gender	Amt	Sogn
1857-01-02	Anne Kirstine Christensen	2	Female	Thisted Amt	Jannerup Sogn (Thisted Amt)
1857-01-02	Birthe Marie Christensen.	0	Female	Thisted Amt	Hundborg Sogn
1857-01-02	Ane Marie Mortensdatter	81	Female	Thisted Amt	Vejerslev Sogn (Mors)
1857-01-02	Gjertrud Jensdatter	82	Female	Thisted Amt	Thisted Sogn
1857-01-02	Karen Christensdatter Wilsbøll	52	Female	Thisted Amt	Vester Vandet Sogn
1857-01-02	Karen Marie Jensen	21	Female	Thisted Amt	Sennels Sogn
1857-01-02	Ane Christensdatter Krogh	76	Female	Thisted Amt	Vester Vandet Sogn
1857-01-03	Anders Hansen Tøylfang	79	Male	Thisted Amt	Gettrup Sogn
1857-01-03	Marcus Christensen	6	Male	Thisted Amt	Hunstrup Sogn
1857-01-04	Maren Jensen	0	Female	Thisted Amt	Hillerslev Sogn (Thisted Amt)
1857-01-04	Thomas Jensen	0	Male	Thisted Amt	Flade Sogn (Thisted Amt)
1857-01-04	Niels Madsen Thystrup	74	Male	Thisted Amt	Skjoldborg Sogn
1857-01-04	Poul Pedersen	72	Male	Thisted Amt	Villerslev Sogn
1857-01-04	Oline Christine Christensen	1	Female	Thisted Amt	Kollerup Sogn (Thisted Amt)
1857-01-04	Maren Cathrine Nielsen	2	Female	Thisted Amt	Kollerup Sogn (Thisted Amt)

- ▶ Temporal resolution:
- ▶ Daily

- ▶ Geographical resolution:
- ▶ Individual parishes



Data cleaning and managing

Date of burial	Name	Age	Gender	Amt	Sogn
1857-01-02	Anne Kirstine Christensen	2	Female	Thisted Amt	Jannerup Sogn (Thisted Amt)
1857-01-02	Birthe Marie Christensen.	0	Female	Thisted Amt	Hundborg Sogn
1857-01-02	Ane Marie Mortensdatter	81	Female	Thisted Amt	Vejerslev Sogn (Mors)
1857-01-02	Gjertrud Jensdatter	82	Female	Thisted Amt	Thisted Sogn
1857-01-02	Karen Christensdatter Wilbøll	52	Female	Thisted Amt	Vester Vandet Sogn
1857-01-02	Karen Marie Jensen	21	Female	Thisted Amt	Sennels Sogn
1857-01-02	Ane Christensdatter Krogh	76	Female	Thisted Amt	Vester Vandet Sogn
1857-01-03	Anders Hansen Tøylfang	79	Male	Thisted Amt	Gettrup Sogn
1857-01-03	Marcus Christensen	6	Male	Thisted Amt	Hunstrup Sogn
1857-01-04	Maren Jensen	0	Female	Thisted Amt	Hillerslev Sogn (Thisted Amt)
1857-01-04	Thomas Jensen	0	Male	Thisted Amt	Flade Sogn (Thisted Amt)
1857-01-04	Niels Madsen Thystrup	74	Male	Thisted Amt	Skjoldborg Sogn
1857-01-04	Poul Pedersen	72	Male	Thisted Amt	Villerslev Sogn
1857-01-04	Oline Christine Christensen	1	Female	Thisted Amt	Kollerup Sogn (Thisted Amt)
1857-01-04	Maren Cathrine Nielsen	2	Female	Thisted Amt	Kollerup Sogn (Thisted Amt)

► Temporal resolution:

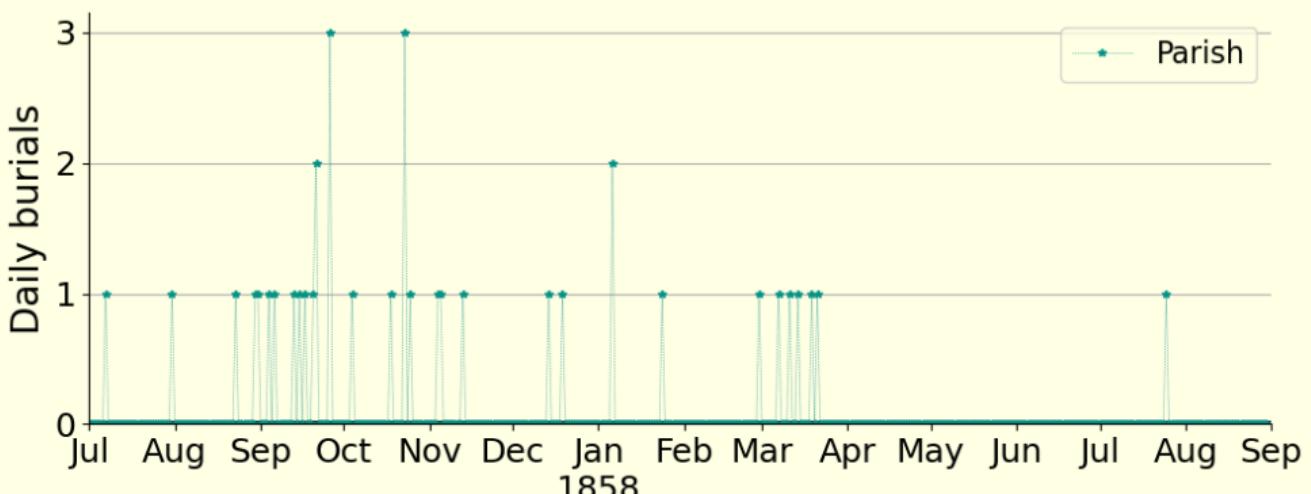
- Daily
- Weekly
- Monthly
- Yearly

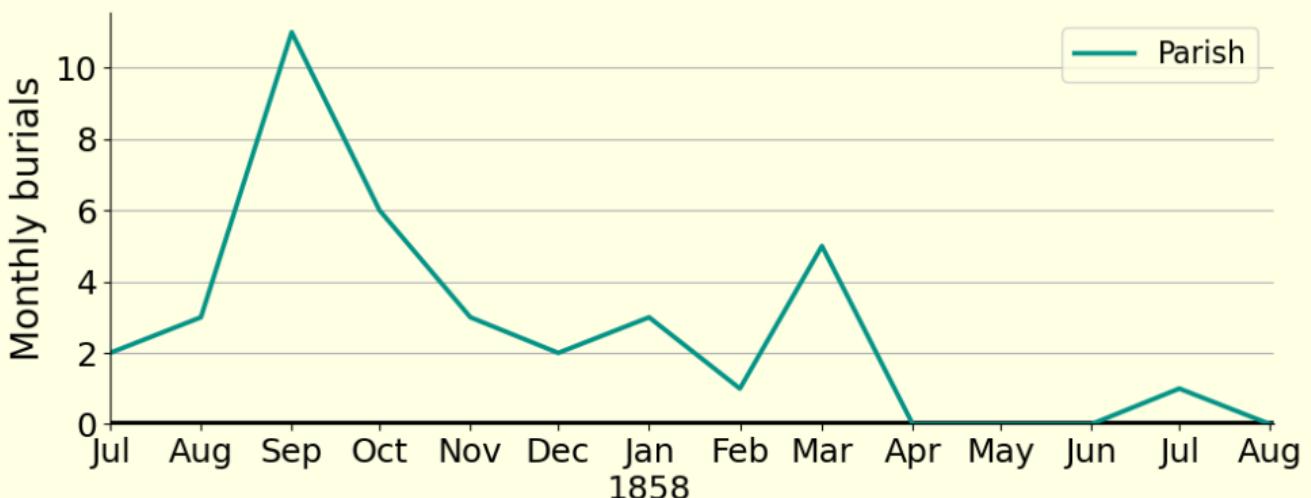
► Geographical resolution:

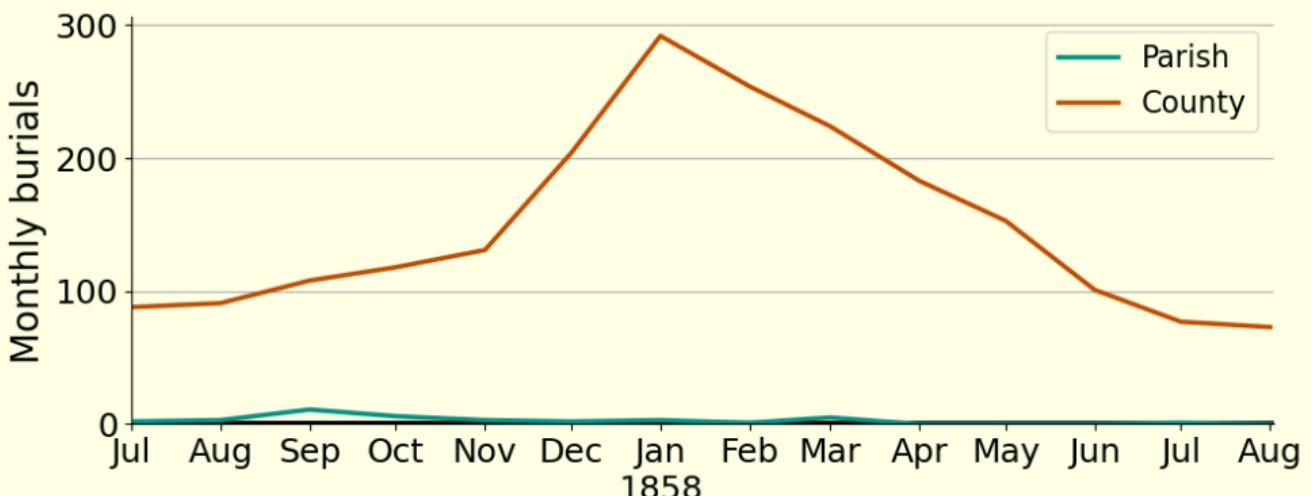
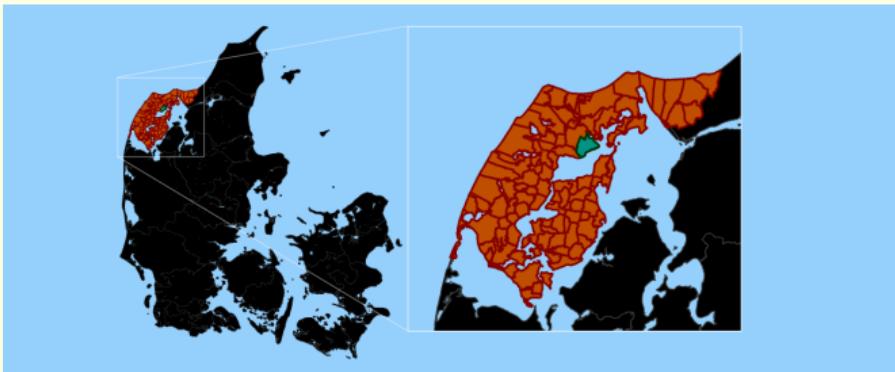
- Individual parishes
- Shire
(groups of 5 to 10 parishes)
- Counties
(groups of 5 to 10 shires)



Data cleaning and managing







Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

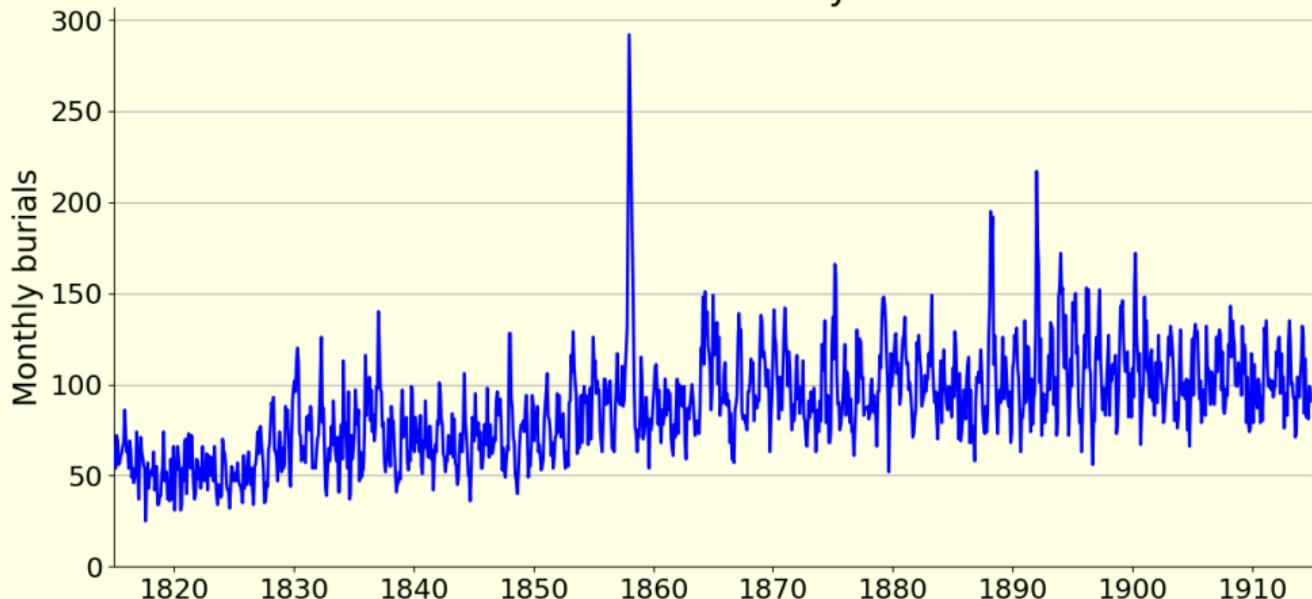
Results and
discussionRepresentative signature
features

Grouping crises

General discussion

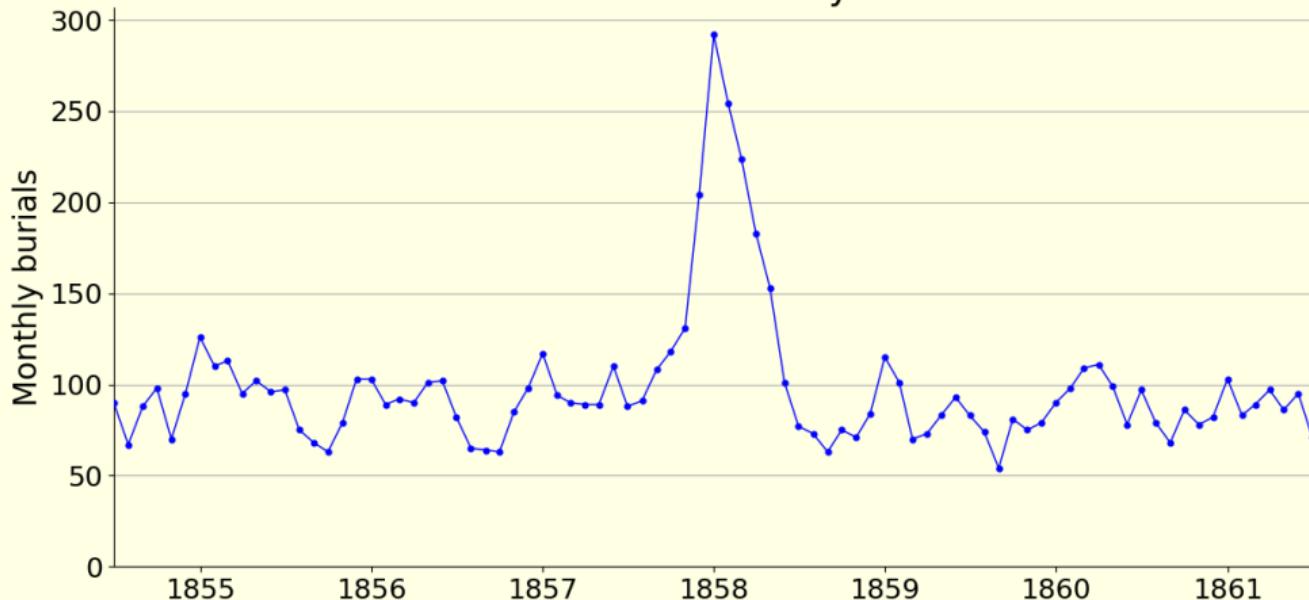
Calculating the mortality baseline

Thisted county



Calculating the mortality baseline

Thisted county



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

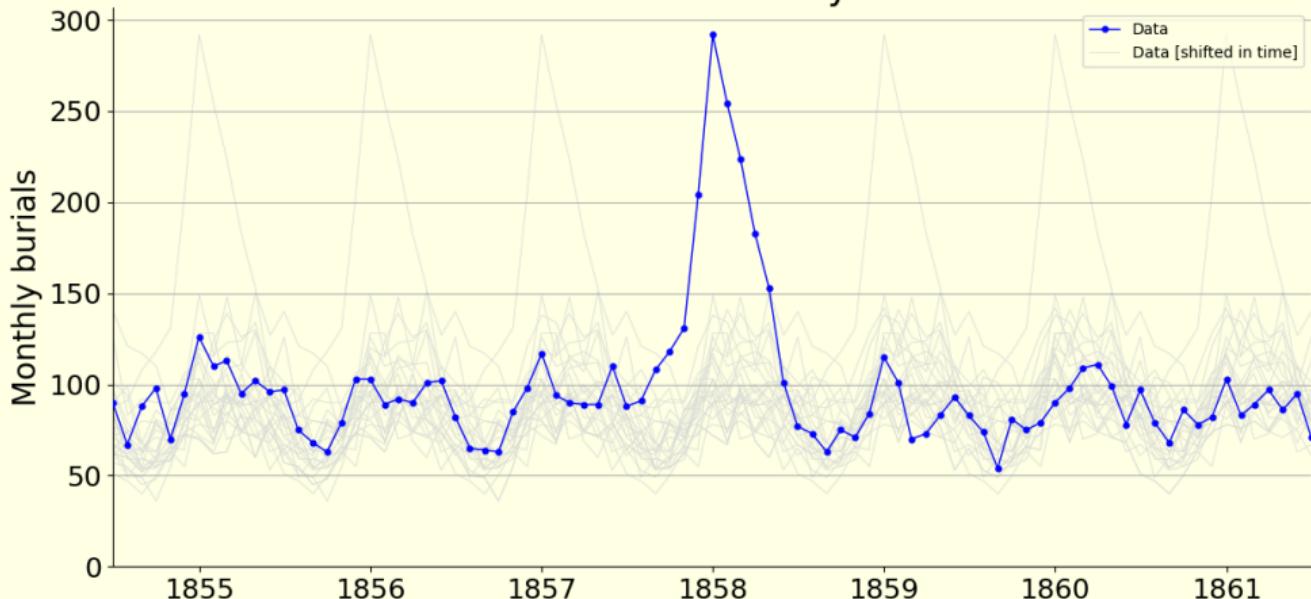
Results and
discussionRepresentative signature
features

Grouping crises

General discussion

Calculating the mortality baseline

Thisted county



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

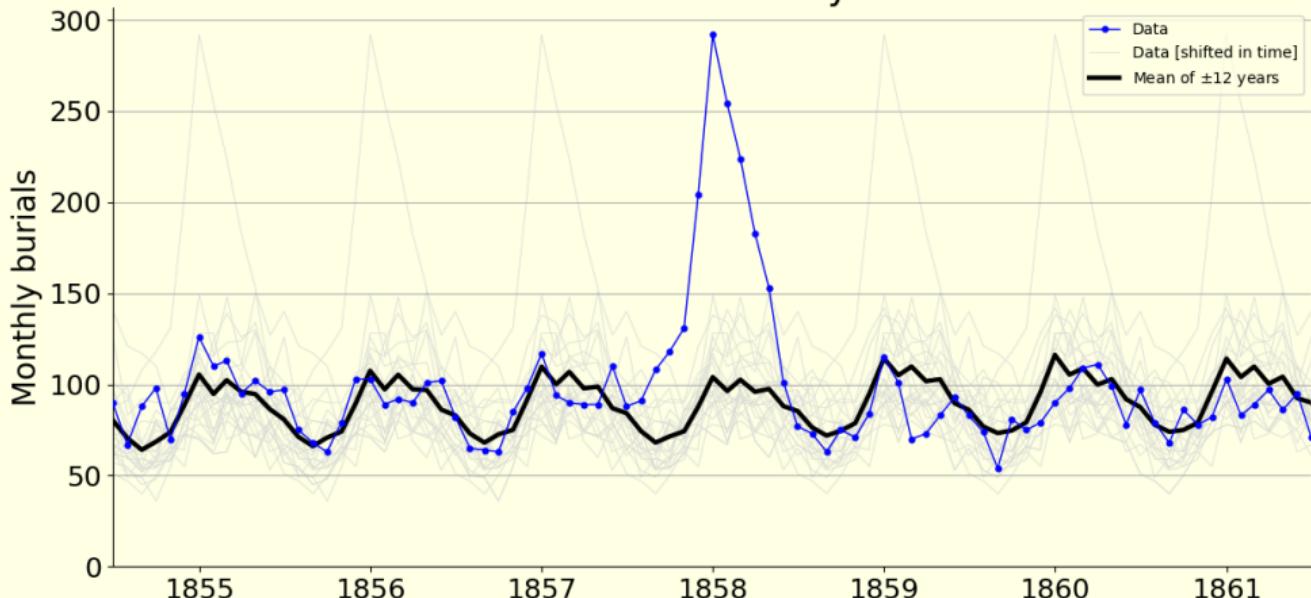
Results and
discussionRepresentative signature
features

Grouping crises

General discussion

Calculating the mortality baseline

Thisted county



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

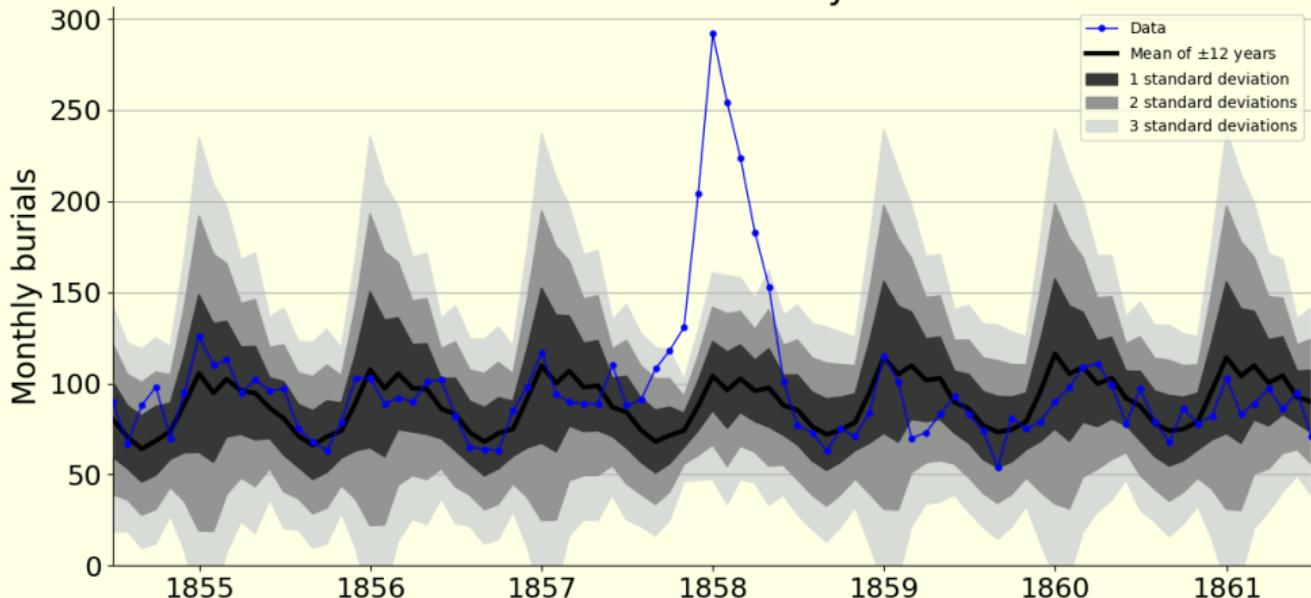
Results and
discussionRepresentative signature
features

Grouping crises

General discussion

Calculating the mortality baseline

Thisted county



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

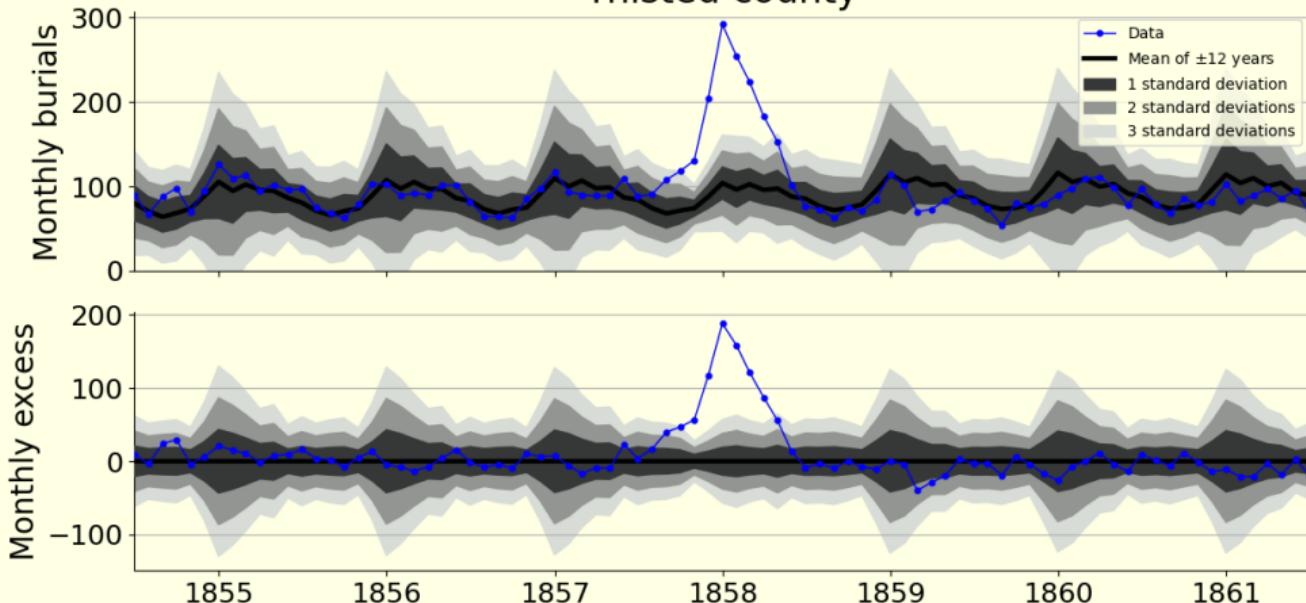
Results and
discussionRepresentative signature
features

Grouping crises

General discussion

Calculating the mortality baseline

Thisted county



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

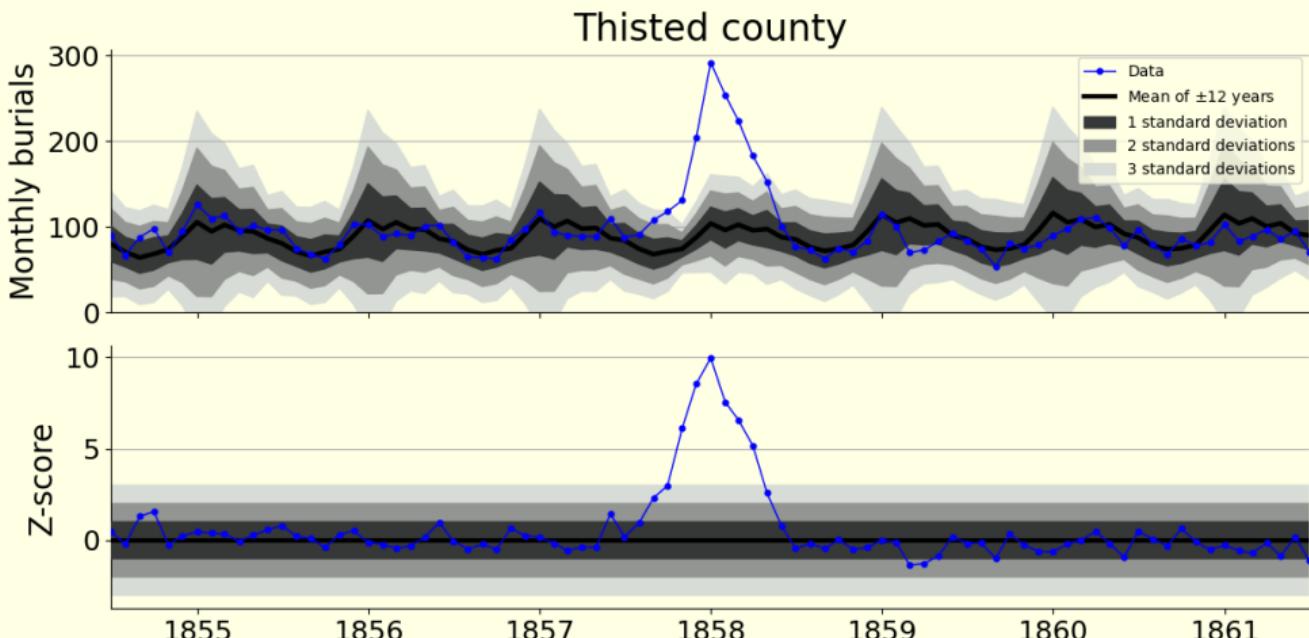
Comparing age patterns

Results and
discussionRepresentative signature
features

Grouping crises

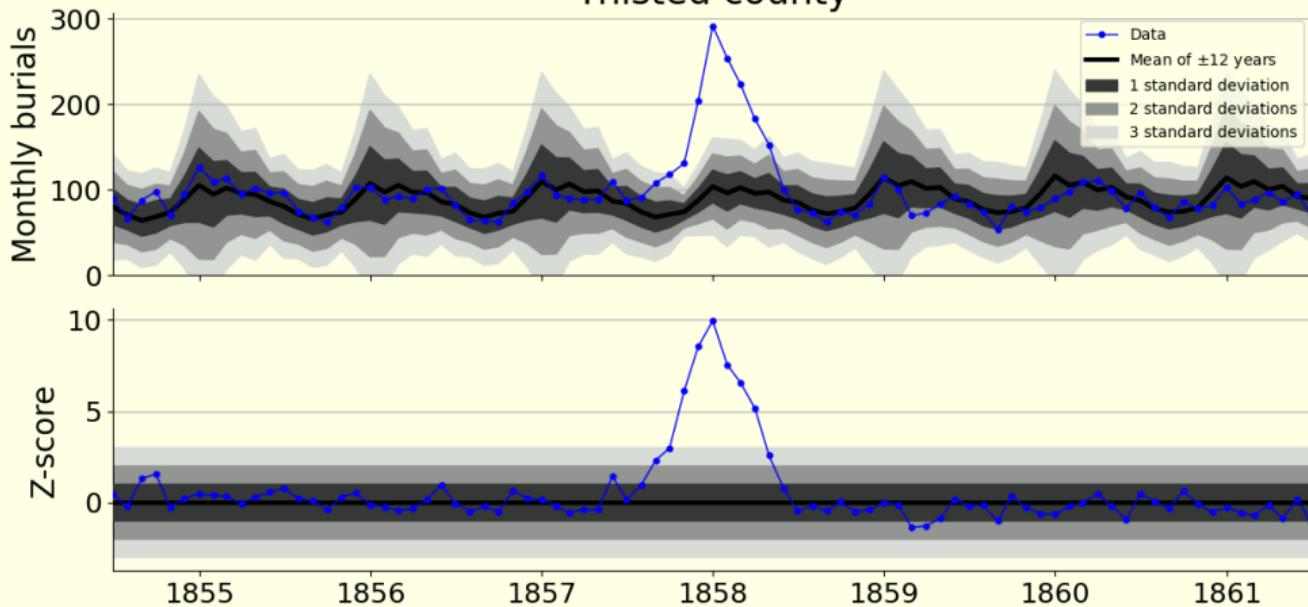
General discussion

Calculating the mortality baseline



Calculating the mortality baseline

Thisted county



From historical demography¹: "Mortality crisis" when Z-score above two.

¹ A. Hinde (2010) "A review of methods for identifying mortality 'crises' using parish record data" - Local Population Studies



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

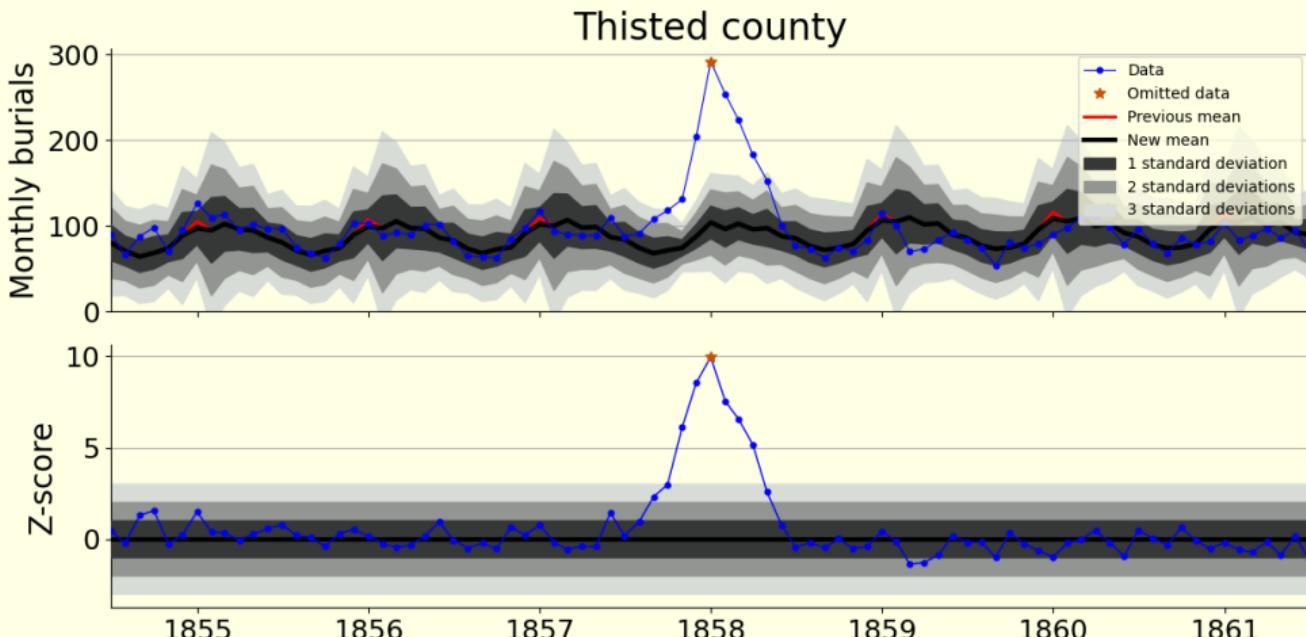
Comparing age patterns

Results and
discussionRepresentative signature
features

Grouping crises

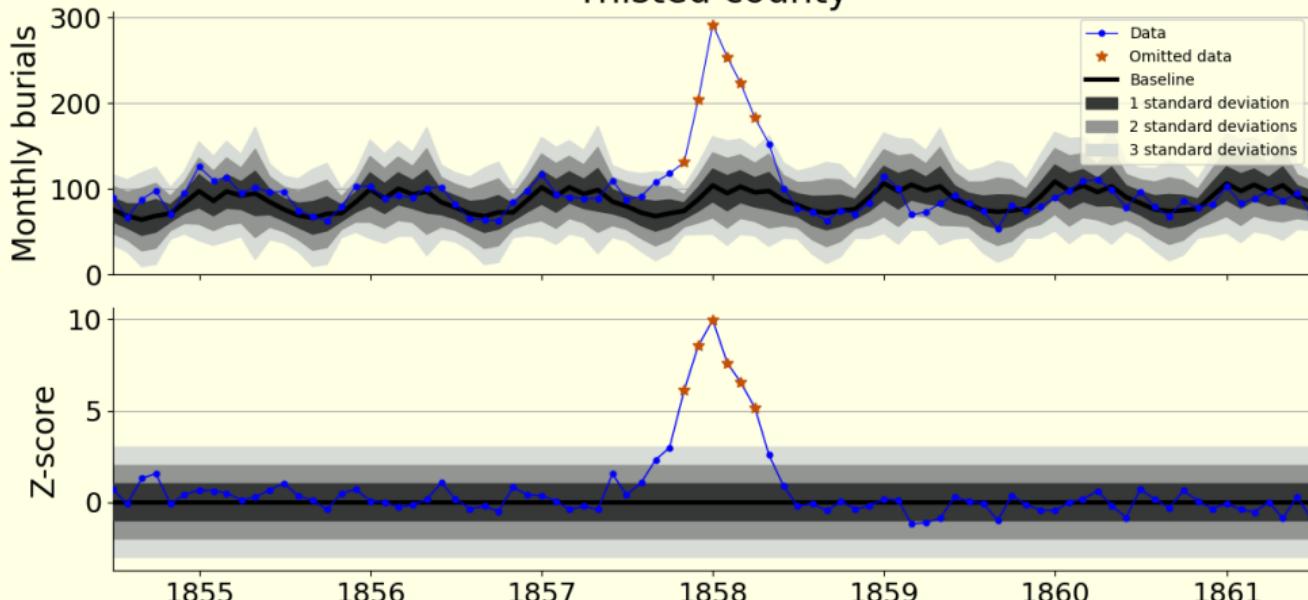
General discussion

Calculating the mortality baseline



Calculating the mortality baseline

Thisted county



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

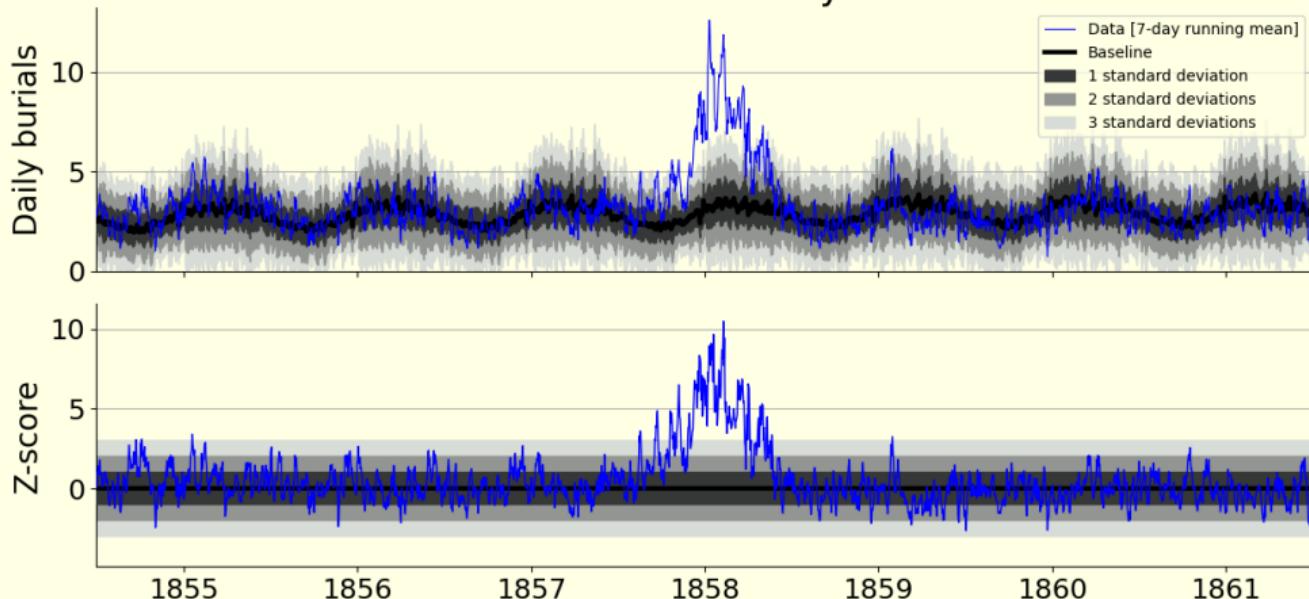
Results and
discussionRepresentative signature
features

Grouping crises

General discussion

Calculating the mortality baseline

Thisted county



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

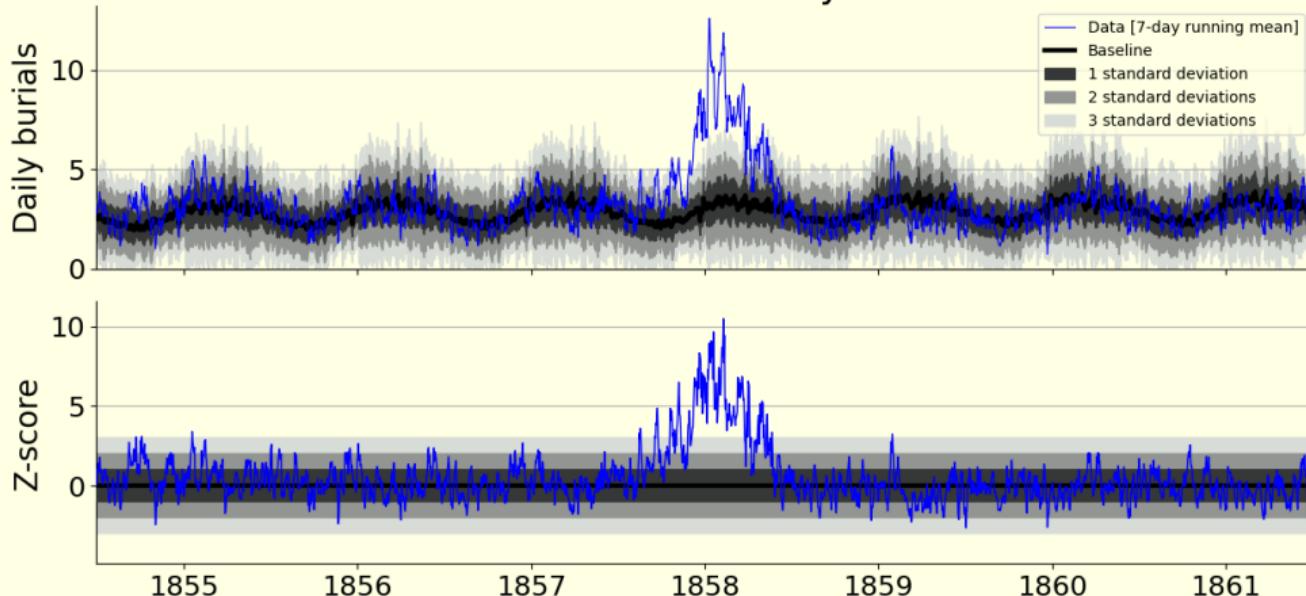
Results and
discussionRepresentative signature
features

Grouping crises

General discussion

Identification of "mortality crises"

Thisted county



We wish to identify continuous periods of excess mortality.



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

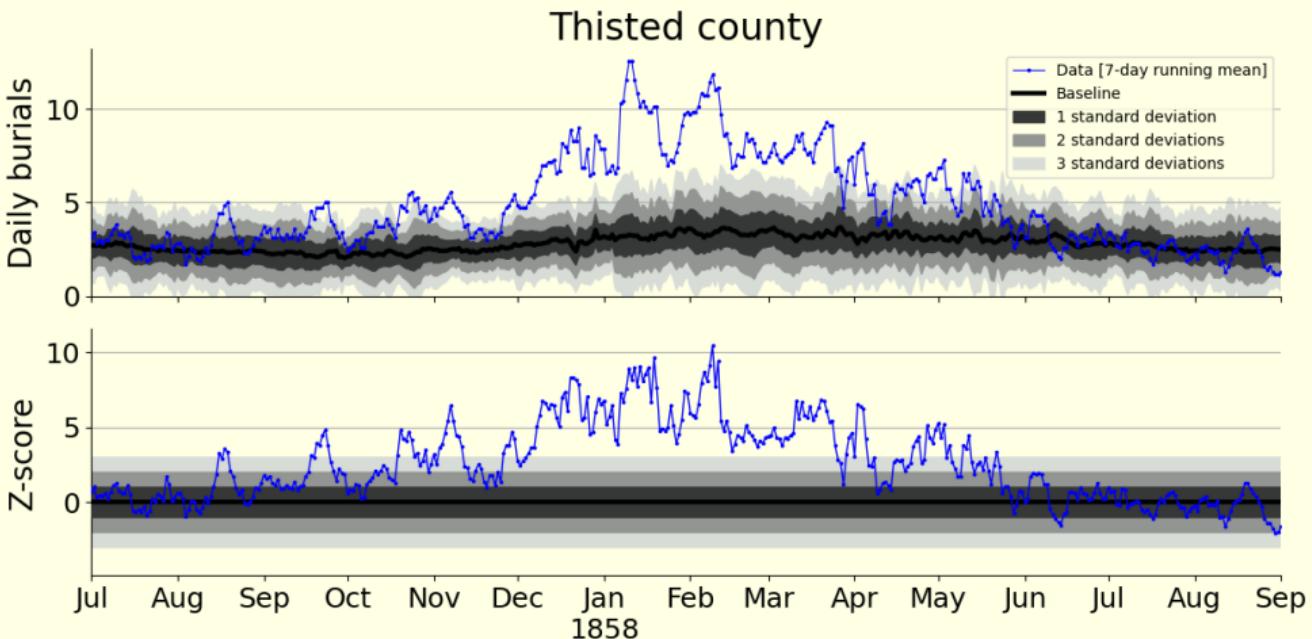
Comparing age patterns

Results and
discussionRepresentative signature
features

Grouping crises

General discussion

Identification of "mortality crises"



Zooming in...



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

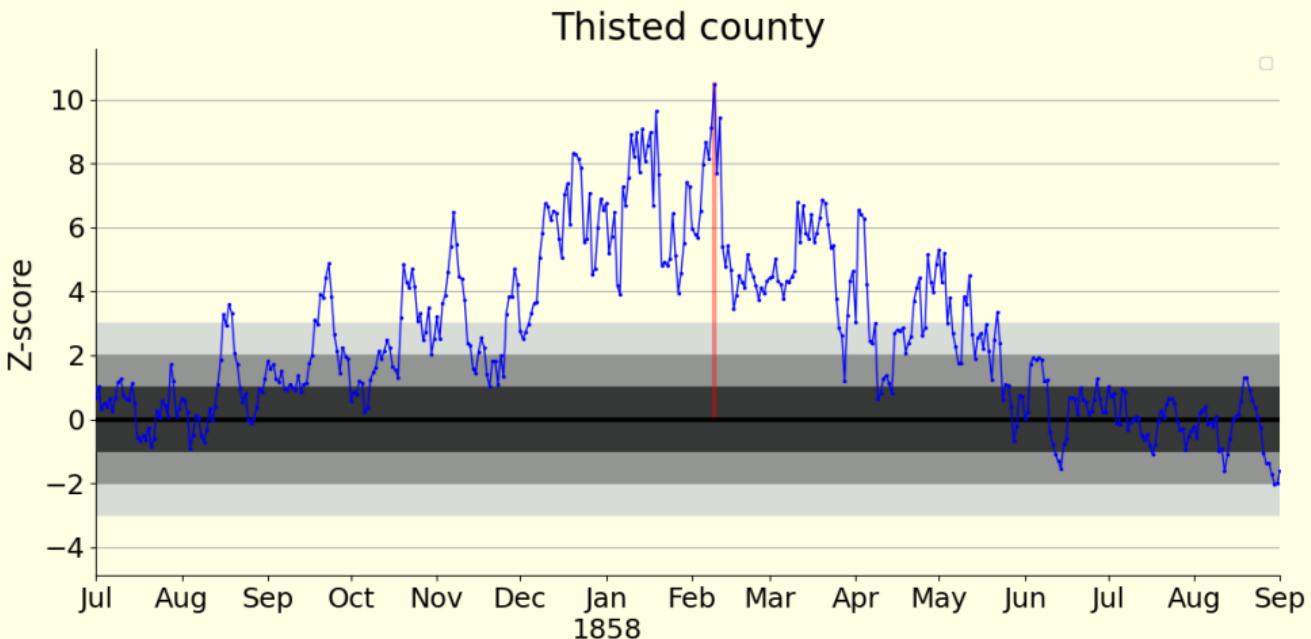
Comparing age patterns

Results and
discussionRepresentative signature
features

Grouping crises

General discussion

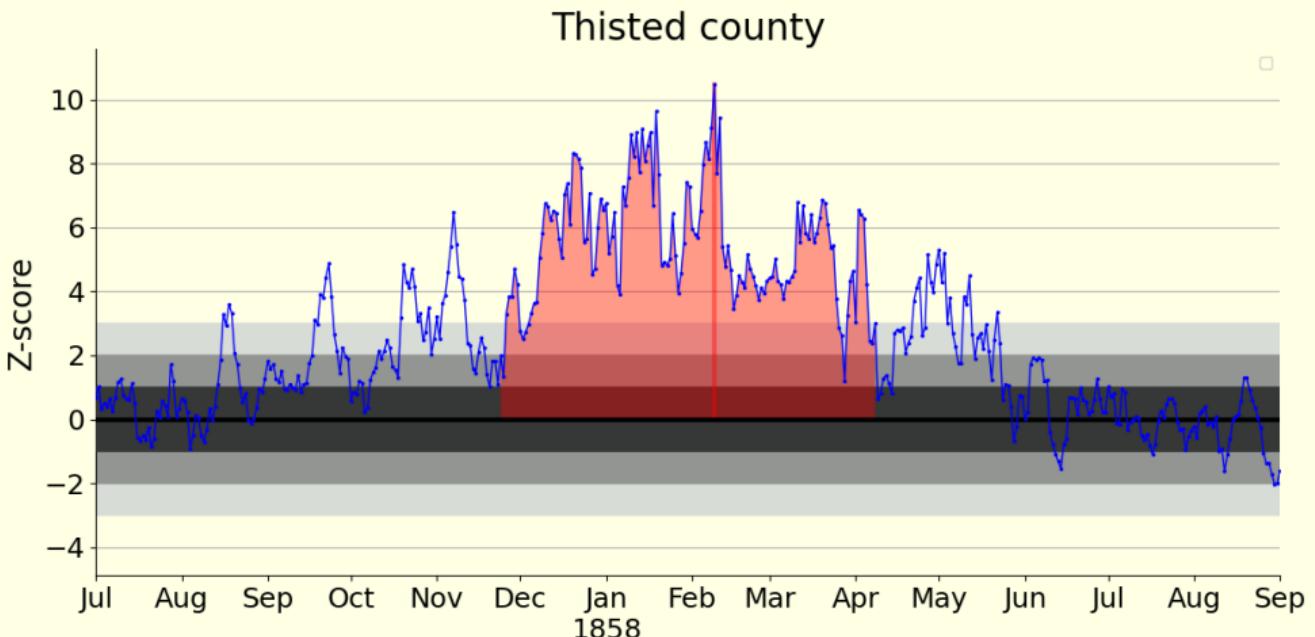
Identification of “mortality crises”



Starting from the date with the highest Z-score...



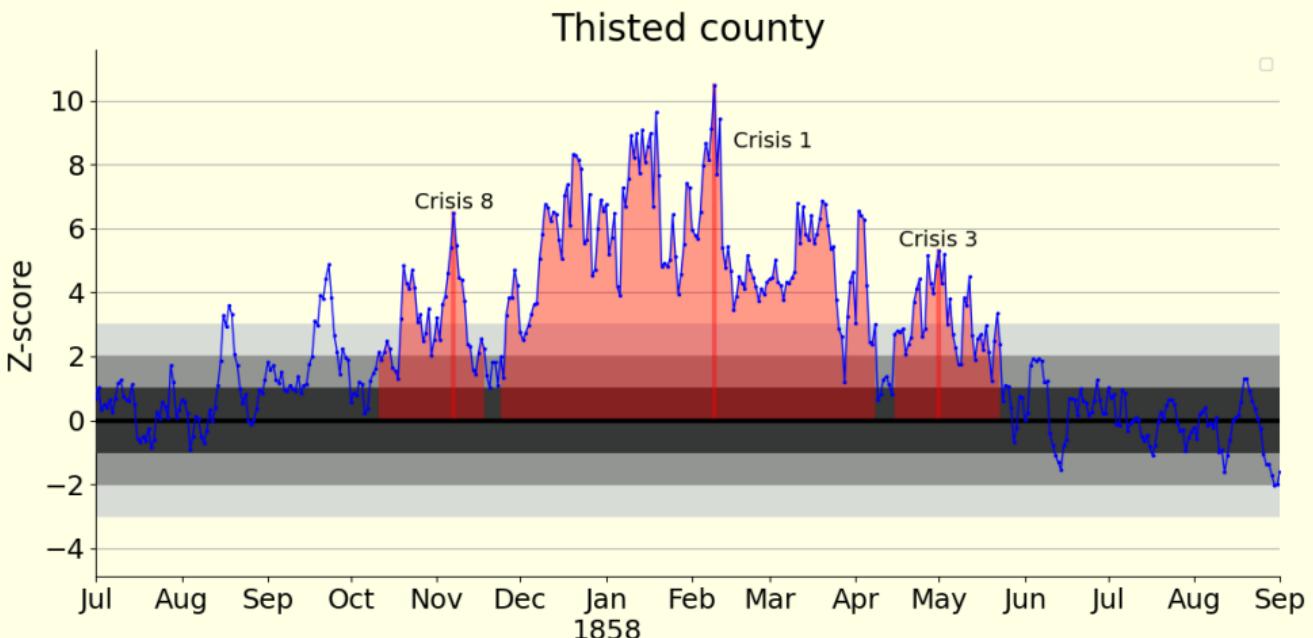
Identification of "mortality crises"



... we group all days with Z-score above three, until the Z-score drops below two for *four* days or more.



Identification of “mortality crises”



All crises with at least *seven* days above the Z-score threshold of three are considered “mortality crises”.



Identifying main "signature features"

Using this methodology, we identify 320 mortality crises.

County	#			
Thisted	1			
Thisted	2			
Thisted	3			
:	:			
Copenhagen	1			
Copenhagen	2			
Copenhagen	3			
:	:			
Aarhus	1			
Aarhus	2			
:	:			



Identifying main "signature features"

Using this methodology, we identify 320 mortality crises.

For each crisis,

County	#			
Thisted	1			
Thisted	2			
Thisted	3			
:	:			
Copenhagen	1			
Copenhagen	2			
Copenhagen	3			
:	:			
Aarhus	1			
Aarhus	2			
:	:			



Identifying main "signature features"

Using this methodology, we identify 320 mortality crises.

For each crisis,

- ▶ Date with most burial.

County	#	Peak-date		
Thisted	1	1858-01-10		
Thisted	2	1892-01-10		
Thisted	3	1858-04-25		
⋮	⋮	⋮		
Copenhagen	1	1892-01-10		
Copenhagen	2	1831-09-04		
Copenhagen	3	1853-08-02		
⋮	⋮	⋮		
Aarhus	1	1892-01-15		
Aarhus	2	1853-08-21		
⋮	⋮	⋮		



Identifying main "signature features"

Using this methodology, we identify 320 mortality crises.

For each crisis,

- ▶ Date with most burial.
- ▶ Excess burials.

Significant on a county level
(Population-sizes $\approx 100,000$).

County	#	Peak-date	Total excess	
Thisted	1	1858-01-10	640	
Thisted	2	1892-01-10	137	
Thisted	3	1858-04-25	100	
⋮	⋮	⋮	⋮	
Copenhagen	1	1892-01-10	716	
Copenhagen	2	1831-09-04	578	
Copenhagen	3	1853-08-02	575	
⋮	⋮	⋮	⋮	
Aarhus	1	1892-01-15	478	
Aarhus	2	1853-08-21	200	
⋮	⋮	⋮	⋮	



Identifying main "signature features"

Using this methodology, we identify 320 mortality crises.

For each crisis,

- ▶ Date with most burial.
- ▶ Excess burials.
Significant on a county level
(Population-sizes $\approx 100,000$).
- ▶ Duration.
e.g. "lasting two months"

<i>County</i>	<i>#</i>	<i>Peak-date</i>	<i>Total excess</i>	<i>Duration</i>
Thisted	1	1858-01-10	640	135 days
Thisted	2	1892-01-10	137	36 days
Thisted	3	1858-04-25	100	38 days
⋮	⋮	⋮	⋮	⋮
Copenhagen	1	1892-01-10	716	71 days
Copenhagen	2	1831-09-04	578	60 days
Copenhagen	3	1853-08-02	575	60 days
⋮	⋮	⋮	⋮	⋮
Aarhus	1	1892-01-15	478	63 days
Aarhus	2	1853-08-21	200	49 days
⋮	⋮	⋮	⋮	⋮



Identifying main "signature features"

Using this methodology, we identify 320 mortality crises.

For each crisis,

- ▶ Date with most burial.
- ▶ Excess burials.
Significant on a county level
(Population-sizes $\approx 100,000$).
- ▶ Duration.
e.g. "lasting two months"
- ▶ Timing and seasonality.
e.g. "peaking in winter" or
"late summer"

<i>County</i>	<i>#</i>	<i>Peak-date</i>	<i>Total excess</i>	<i>Duration</i>
Thisted	1	1858-01-10	640	135 days
Thisted	2	1892-01-10	137	36 days
Thisted	3	1858-04-25	100	38 days
⋮	⋮	⋮	⋮	⋮
Copenhagen	1	1892-01-10	716	71 days
Copenhagen	2	1831-09-04	578	60 days
Copenhagen	3	1853-08-02	575	60 days
⋮	⋮	⋮	⋮	⋮
Aarhus	1	1892-01-15	478	63 days
Aarhus	2	1853-08-21	200	49 days
⋮	⋮	⋮	⋮	⋮



Identifying main "signature features"

Using this methodology, we identify 320 mortality crises.

For each crisis,

- ▶ Date with most burial.
- ▶ Excess burials.
Significant on a county level
(Population-sizes \approx 100,000).
- ▶ Duration.
e.g. "lasting two months"
- ▶ Timing and seasonality.
e.g. "peaking in winter" or
"late summer"

<i>County</i>	<i>#</i>	<i>Peak-date</i>	<i>Total excess</i>	<i>Duration</i>
Thisted	1	1858-01-10	640	135 days
Thisted	2	1892-01-10	137	36 days
Thisted	3	1858-04-25	100	38 days
:	:	:	:	:
Copenhagen	1	1892-01-10	716	71 days
Copenhagen	2	1831-09-04	578	60 days
Copenhagen	3	1853-08-02	575	60 days
:	:	:	:	:
Aarhus	1	1892-01-15	478	63 days
Aarhus	2	1853-08-21	200	49 days
:	:	:	:	:
:	:	:	:	:

But we also have data on age.



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

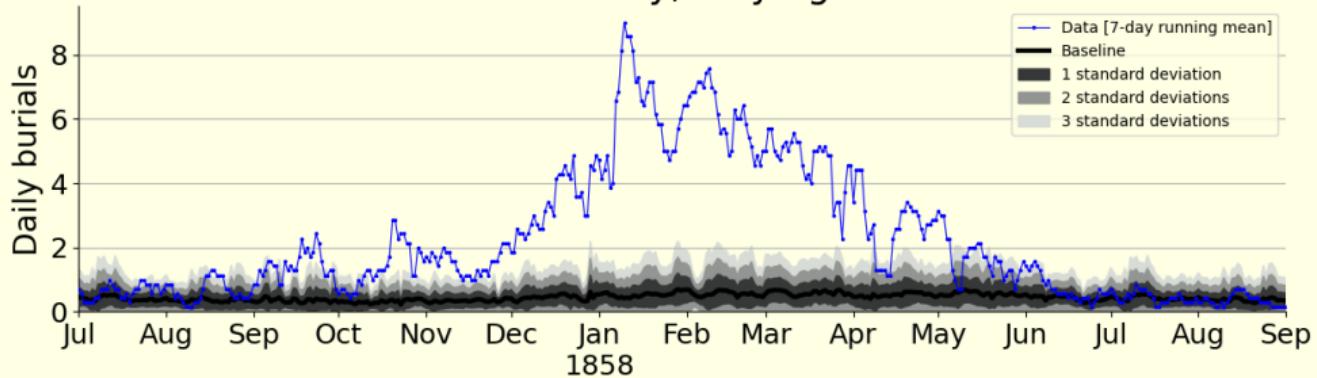
Results and
discussionRepresentative signature
features

Grouping crises

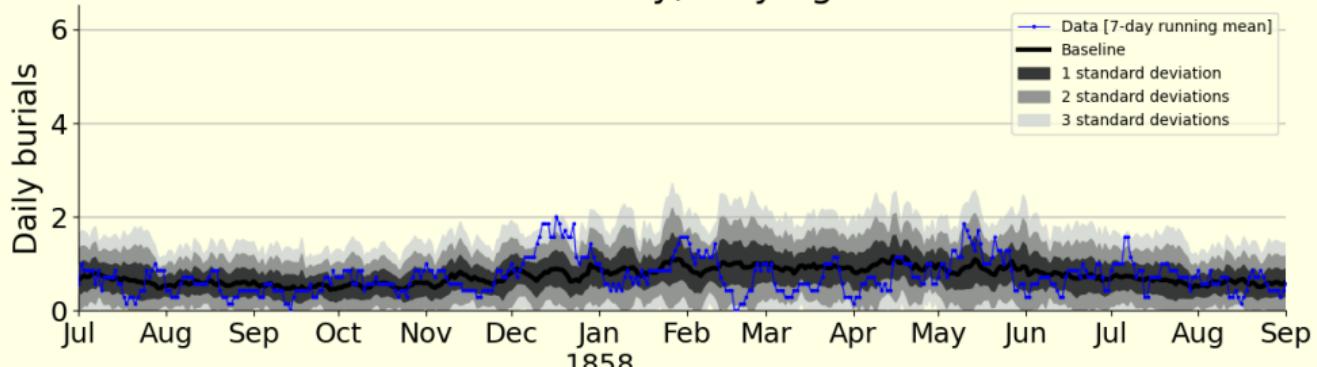
General discussion

Analyzing age-patterns

Thisted county, only ages 1-14



Thisted county, only ages 60+



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

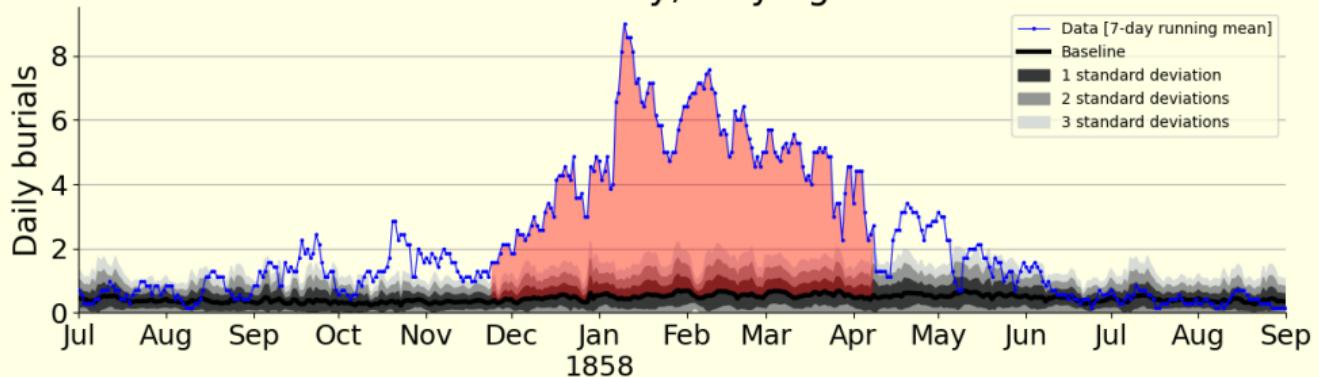
Results and
discussionRepresentative signature
features

Grouping crises

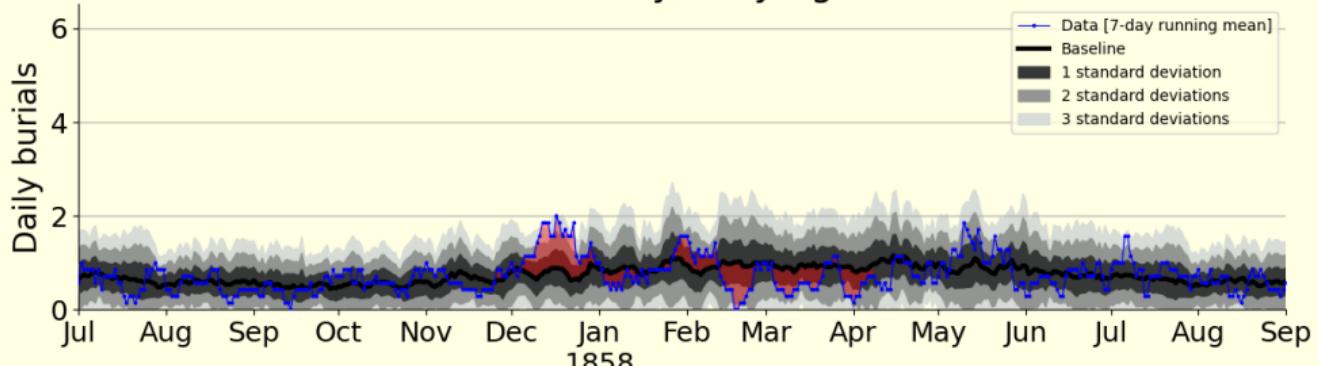
General discussion

Analyzing age-patterns

Thisted county, only ages 1-14



Thisted county, only ages 60+



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

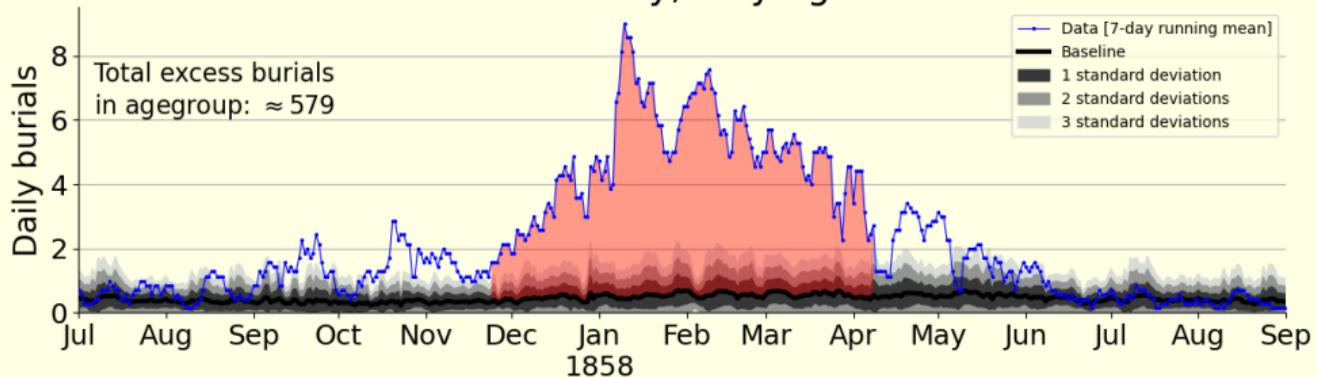
Results and
discussionRepresentative signature
features

Grouping crises

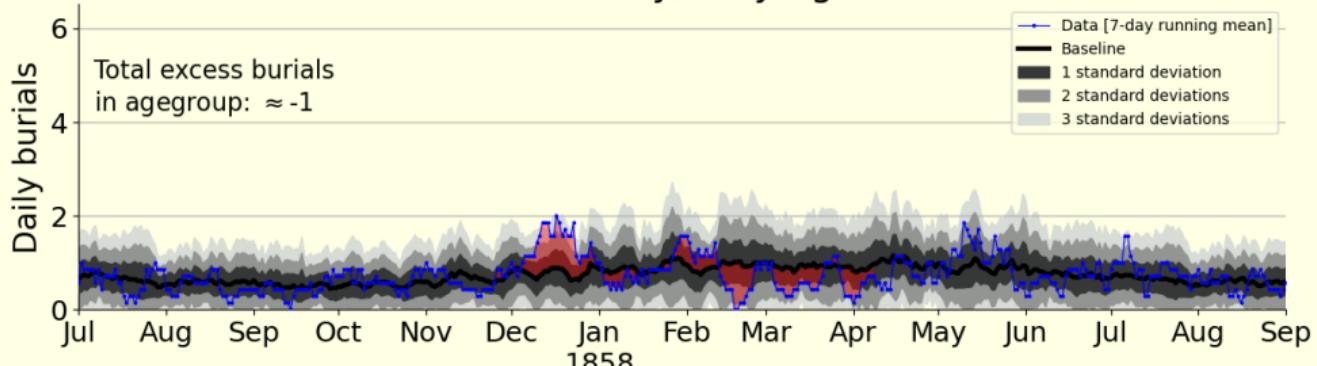
General discussion

Analyzing age-patterns

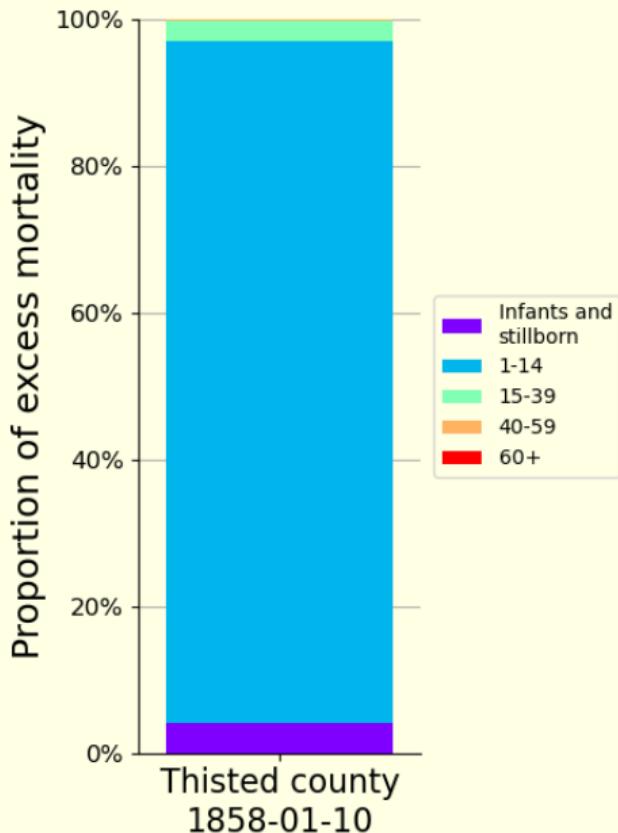
Thisted county, only ages 1-14



Thisted county, only ages 60+



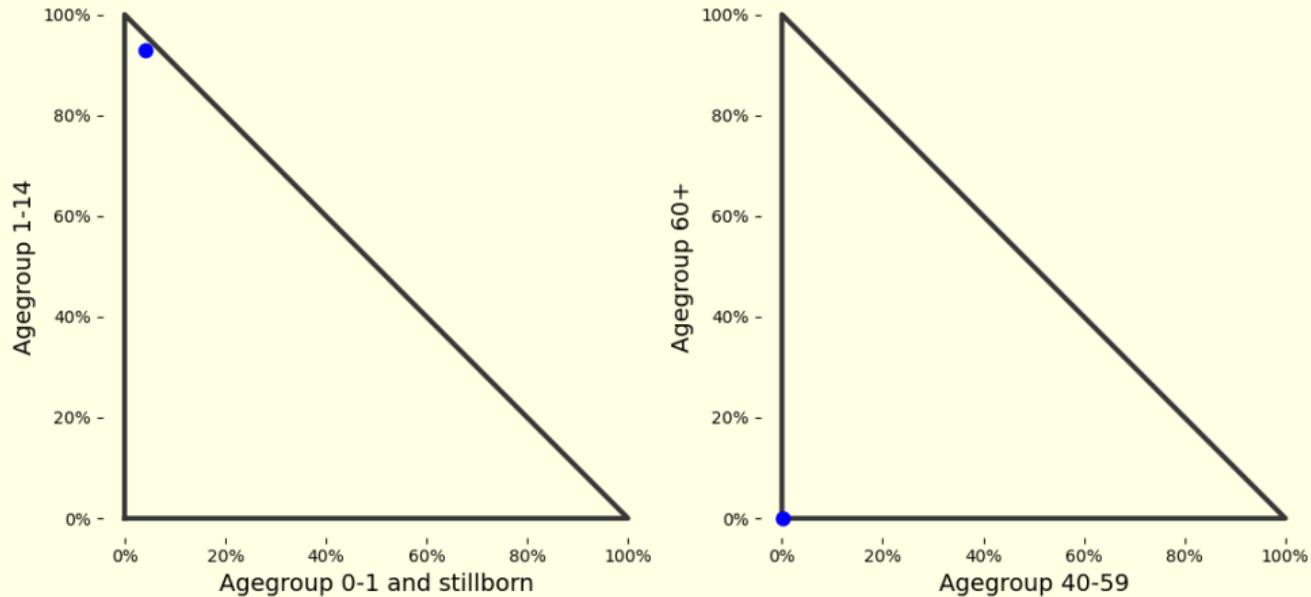
Analyzing age-patterns



Age-specific excess mortality in this period:

- ▶ Age group "60+": <1%
- ▶ Age group "40-59": <1%
- ▶ Age group "15-39": 3%
- ▶ Age group "1-14": 93%
- ▶ Age group "Below 1 year": 4%

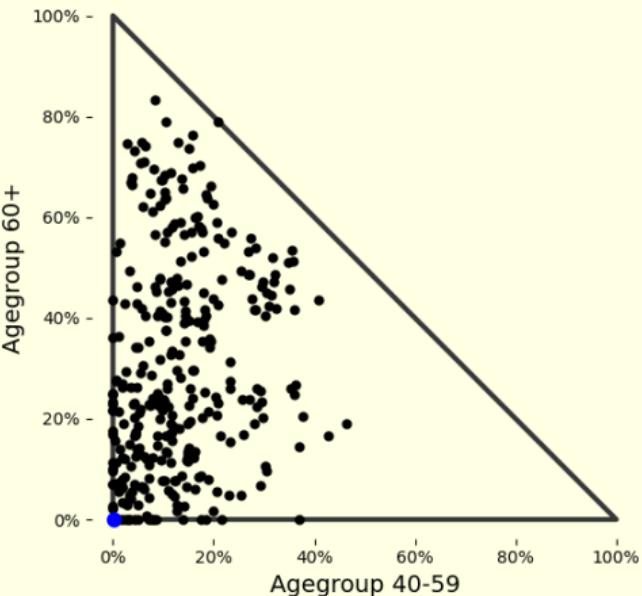
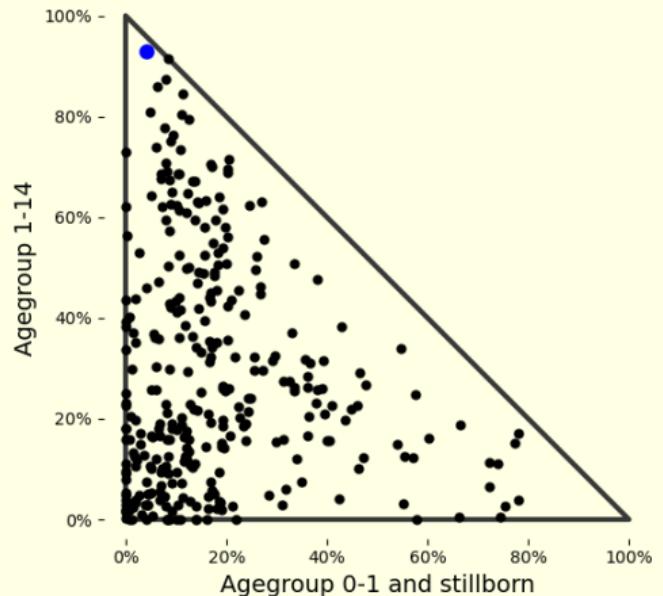




(Age group "15-39" not shown here)



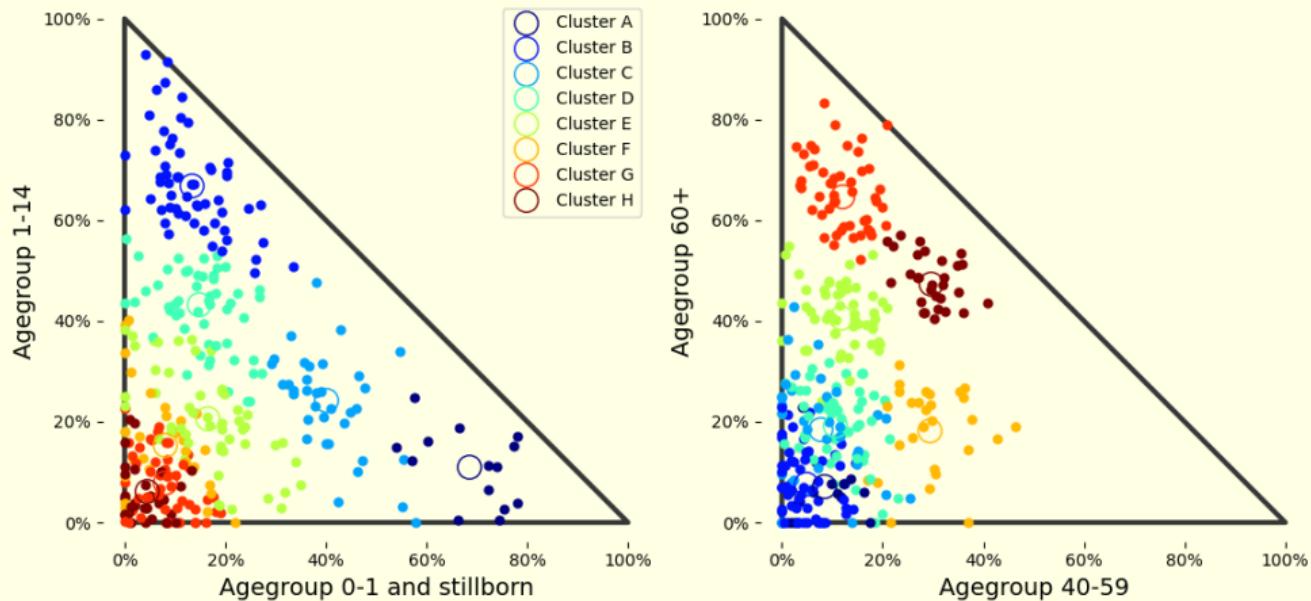
Mortality crises with comparable age patterns



Adding the other 319 mortality crises identified.

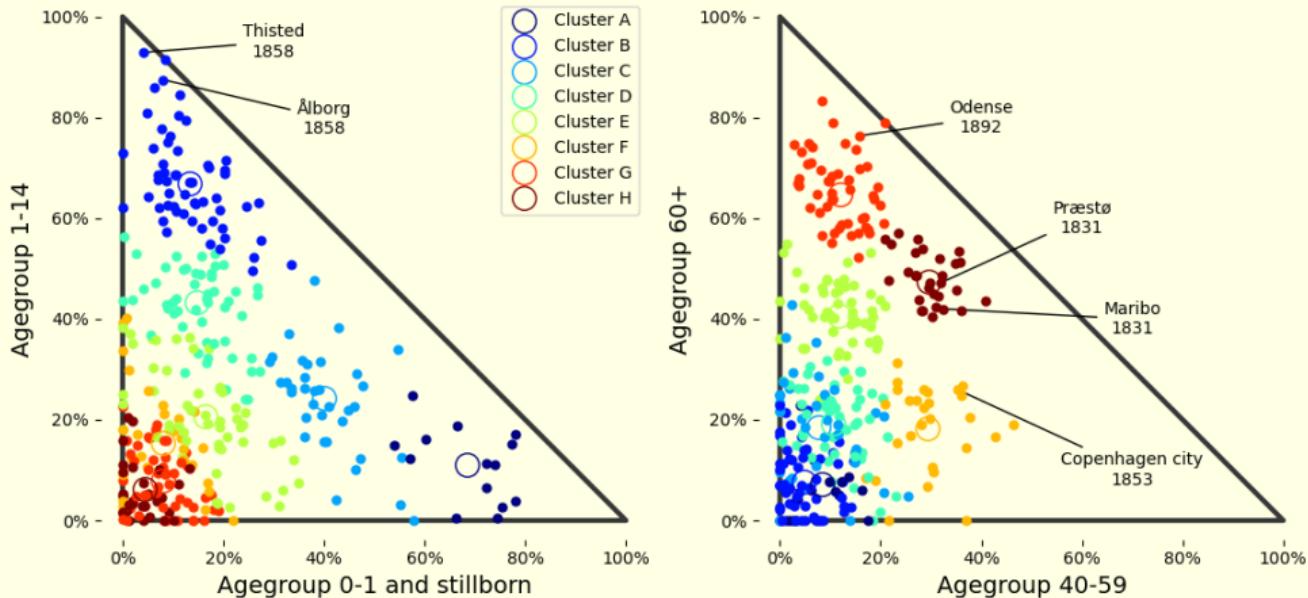


Mortality crises with comparable age patterns



Gaussian mixture modelling on full five-dimensional data.

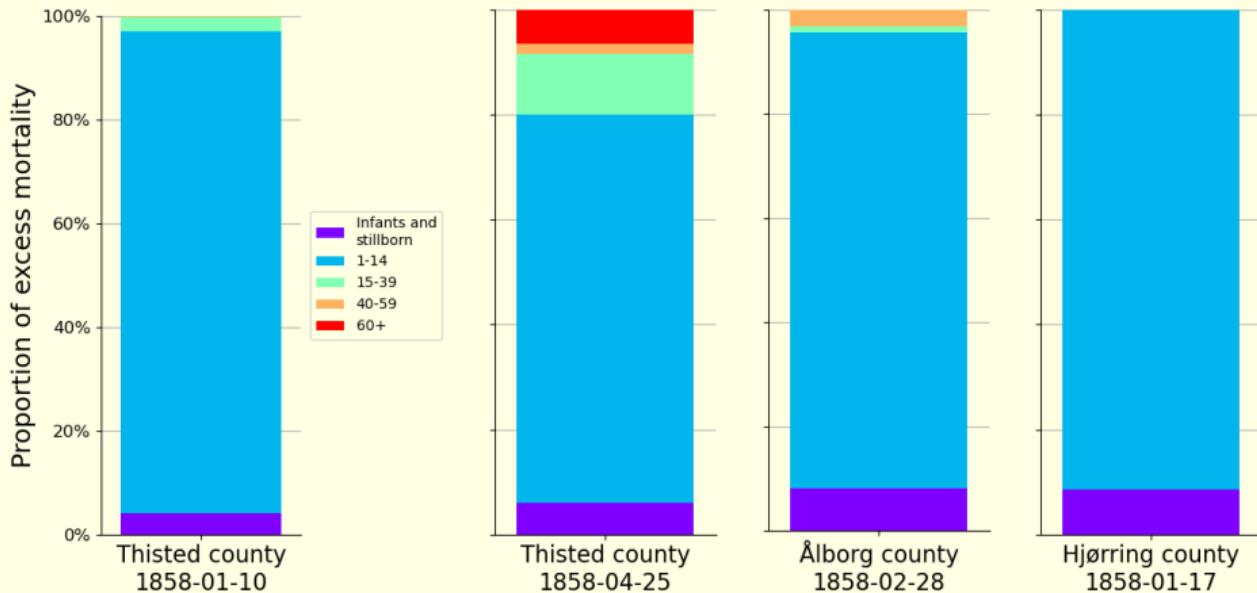
Mortality crises with comparable age patterns



Gaussian mixture modelling on full five-dimensional data.

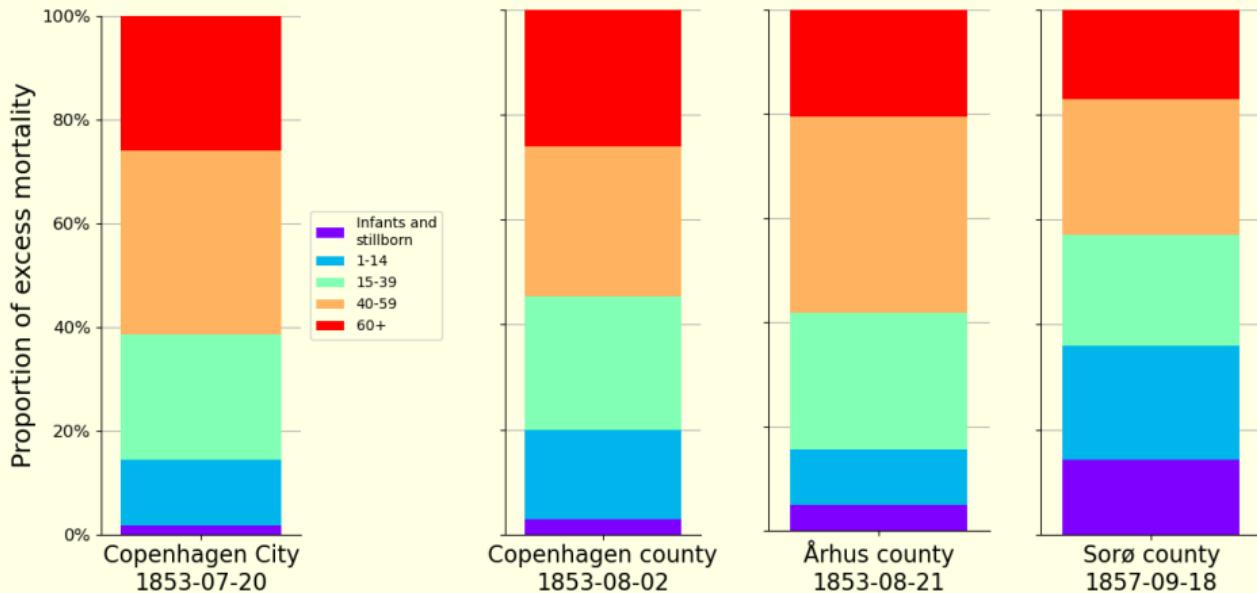
Mortality crises with comparable age patterns

Examples from "Cluster A"



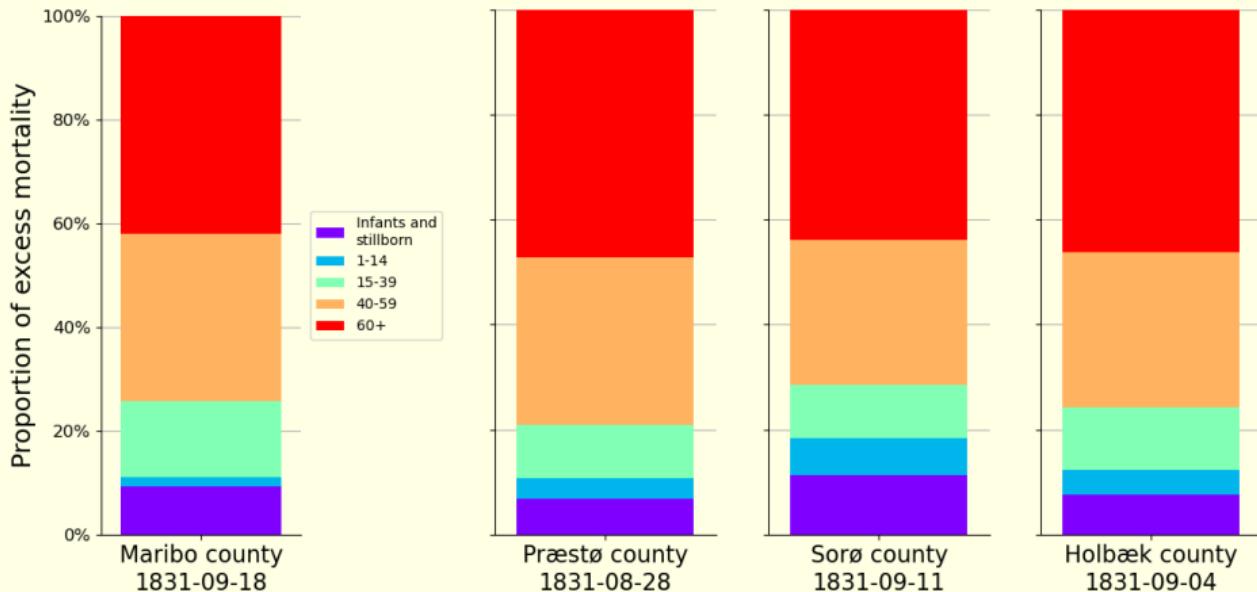
Mortality crises with comparable age patterns

Examples from "Cluster F"



Mortality crises with comparable age patterns

Examples from "Cluster H"



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

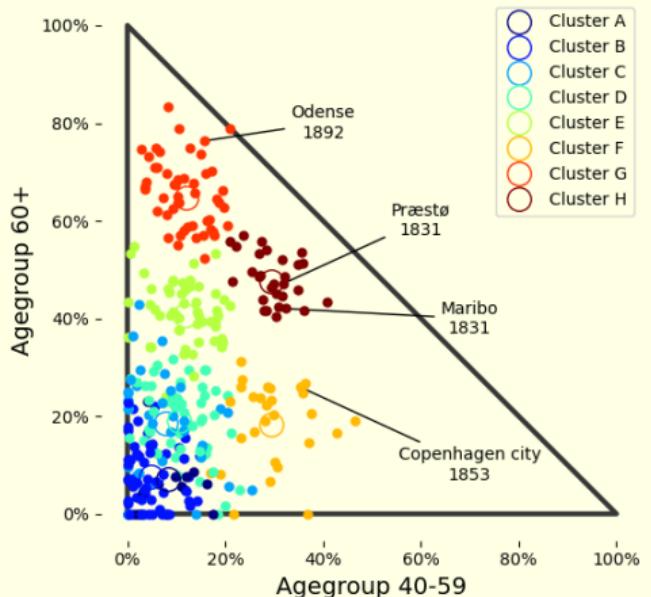
Comparing age patterns

Results and
discussionRepresentative signature
features

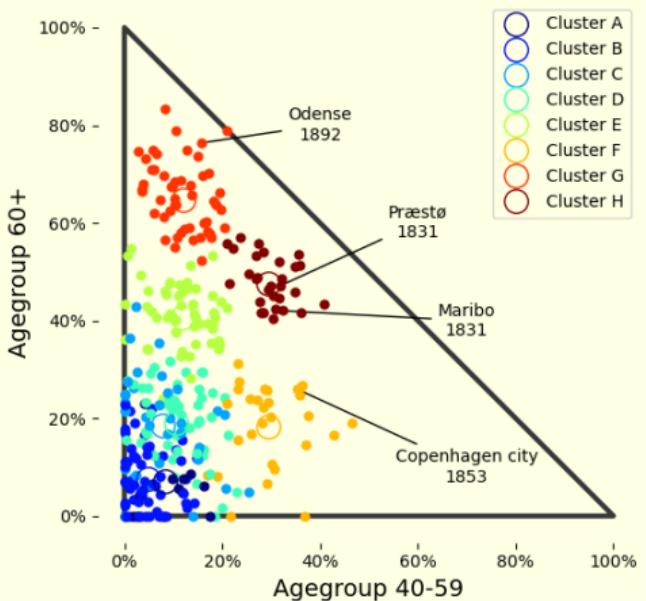
Grouping crises

General discussion

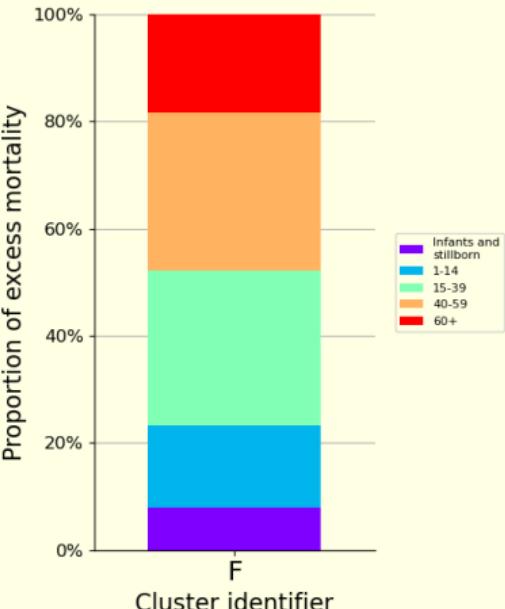
The signature features of certain diseases



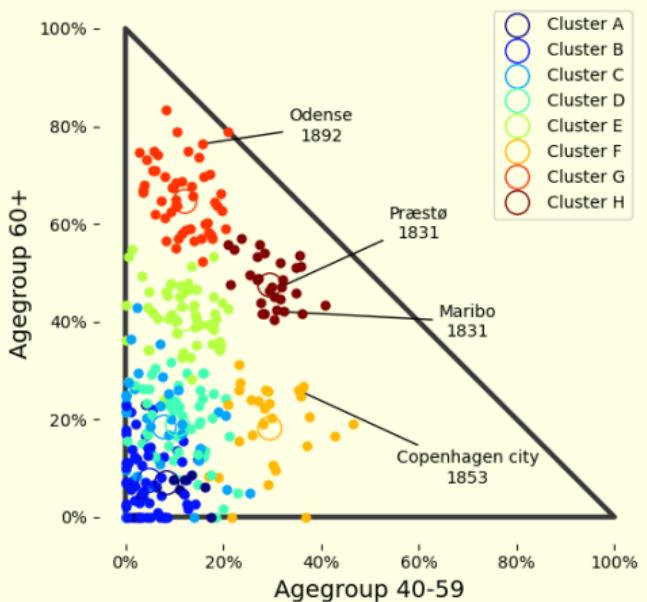
The signature features of certain diseases



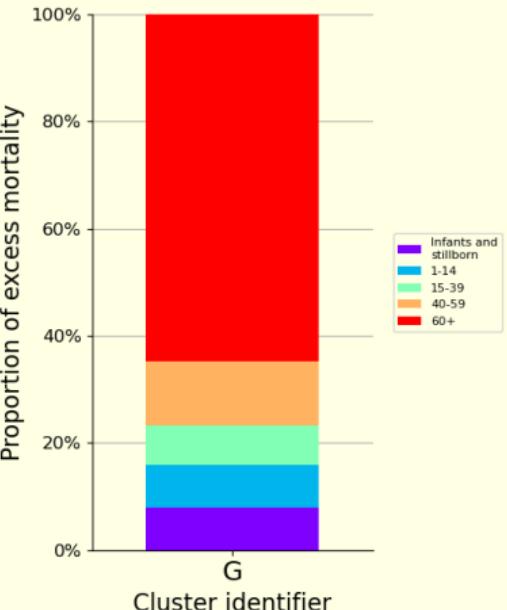
Cholera



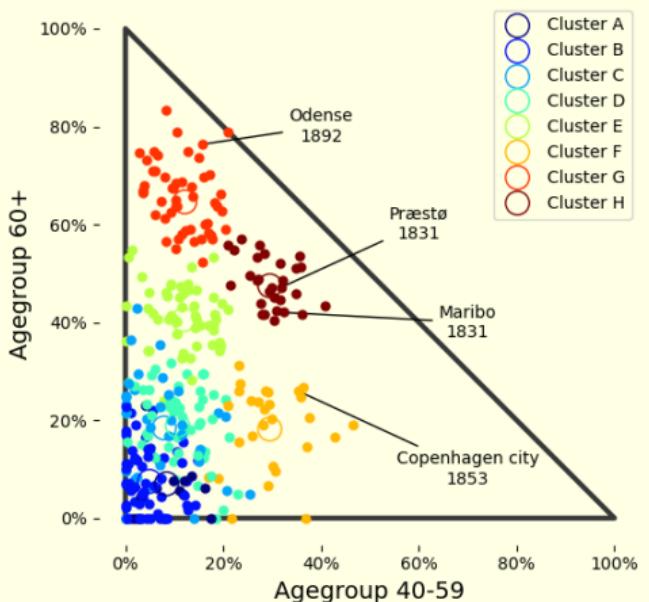
The signature features of certain diseases



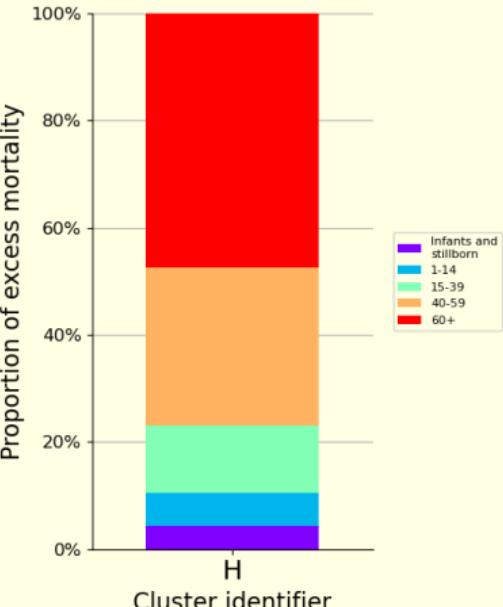
Pandemic influenza



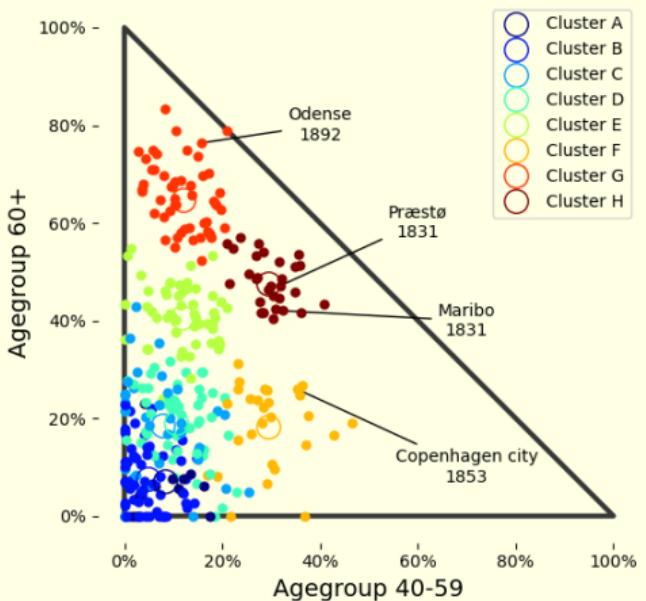
The signature features of certain diseases



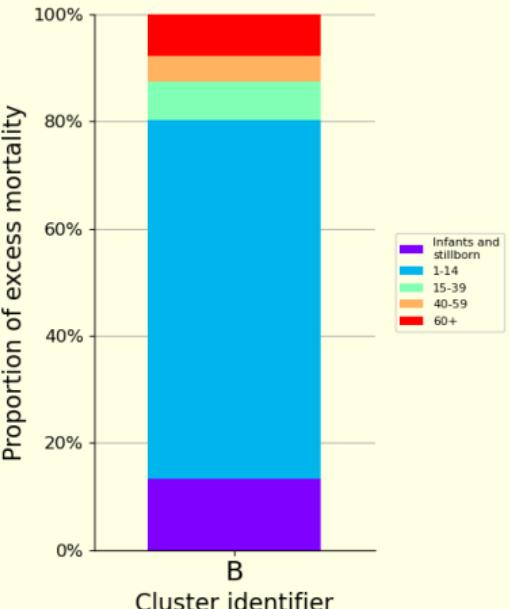
"Harvest epidemics"



The signature features of certain diseases



Scarlet fever



All epidemics and pandemics in 19th century Denmark

Going through all major mortality crises, and crossreferencing with historical records of epidemics:

Disease	Timing	Total excess	Age structure
Cholera	Late summer, 1853 and 1857	4588	Adults Cluster "F"
Scarlet fever	Winter 1857/1858	2323	Children (1-15) Cluster "B"
"Harvest" epidemic¹	Late summer, 1825-1831	11539	Adults Cluster "H" and "F"
Pandemic influenza	1892 and 1900	9532	Elderly Cluster "G"

And other epidemics as well as mortality crises unrelated to disease, e.g. war.

¹ The cause of this mortality crisis was probably a subsistence crisis as well as a range of diseases. Discussed in detail in Ingholt (2022) *Scandinavian Journal of History*



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

General discussion

- We determine mortality baselines on county-level and estimate excess mortality.



- We determine mortality baselines on county-level and estimate excess mortality.
- We identify 320 major mortality crises in 19th century Denmark.



- ▶ We determine mortality baselines on county-level and estimate excess mortality.
- ▶ We identify 320 major mortality crises in 19th century Denmark.
- ▶ For each crisis, we determine signature features:



- ▶ We determine mortality baselines on county-level and estimate excess mortality.
- ▶ We identify 320 major mortality crises in 19th century Denmark.
- ▶ For each crisis, we determine signature features:
 - ▶ Age-patterns.



- ▶ We determine mortality baselines on county-level and estimate excess mortality.
- ▶ We identify 320 major mortality crises in 19th century Denmark.
- ▶ For each crisis, we determine signature features:
 - ▶ Age-patterns.
 - ▶ Timing and seasonality.



- ▶ We determine mortality baselines on county-level and estimate excess mortality.
- ▶ We identify 320 major mortality crises in 19th century Denmark.
- ▶ For each crisis, we determine signature features:
 - ▶ Age-patterns.
 - ▶ Timing and seasonality.
 - ▶ Duration.



- ▶ We determine mortality baselines on county-level and estimate excess mortality.
- ▶ We identify 320 major mortality crises in 19th century Denmark.
- ▶ For each crisis, we determine signature features:
 - ▶ Age-patterns.
 - ▶ Timing and seasonality.
 - ▶ Duration.
 - ▶ Geography.



Summary

- ▶ We determine mortality baselines on county-level and estimate excess mortality.
- ▶ We identify 320 major mortality crises in 19th century Denmark.
- ▶ For each crisis, we determine signature features:
 - ▶ Age-patterns.
 - ▶ Timing and seasonality.
 - ▶ Duration.
 - ▶ Geography.
- ▶ By comparing these features and validating with historical sources, we are able to determine groups of mortality crises with the same etiology, and estimate the total number of excess deaths during specific epidemics.



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

General discussion

- Similar methods could be applied to modern data.



Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

General discussion

- ▶ Similar methods could be applied to modern data.
 - ▶ Clustering of age-patterns in modern all-cause mortality data.
A wide range of methods for clustering exists, see e.g. `scikit-learn` for python.



Final comments

- ▶ Similar methods could be applied to modern data.
 - ▶ Clustering of age-patterns in modern all-cause mortality data.
A wide range of methods for clustering exists, see e.g. `scikit-learn` for python.
 - ▶ Excess mortality calculation.
Available online soon, as both Python and R package.



Final comments

- ▶ Similar methods could be applied to modern data.
 - ▶ Clustering of age-patterns in modern all-cause mortality data.
A wide range of methods for clustering exists, see e.g. `scikit-learn` for python.
 - ▶ Excess mortality calculation.
Available online soon, as both Python and R package.
- ▶ Despite demographic differences between 19th century Denmark and modern times, the age patterns in the 19th century may be similar for modern epidemics.



Final comments

- ▶ Similar methods could be applied to modern data.
 - ▶ Clustering of age-patterns in modern all-cause mortality data.
A wide range of methods for clustering exists, see e.g. `scikit-learn` for python.
 - ▶ Excess mortality calculation.
Available online soon, as both Python and R package.
- ▶ Despite demographic differences between 19th century Denmark and modern times, the age patterns in the 19th century may be similar for modern epidemics.
- ▶ As more historical data becomes transcribed, e.g. thanks to improved computer vision, similar studies of other countries will become possible.



Thank you for your attention.



Feel free to email me with
questions or comments

Website: rasmuspedersen.com
Email: rakrpe@ruc.dk

"Identifying Signature Features of Epidemic Diseases in 19th Century All-cause Mortality Data"
Pedersen RK, Ingholt MM, van Wijhe M, Andreasen V & Simonsen L



- Introduction
- Background and data handling
 - Data source
 - Data cleaning
- Methodology
 - Mortality baseline
 - "Mortality crisis"
 - Age-specific mortality
 - Comparing age patterns
- Results and discussion
 - Representative signature features
 - Grouping crises
 - General discussion