

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

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

Summary & discussion

- ▶ 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

Summary & discussion

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



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

Summary & discussion



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

Summary & 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

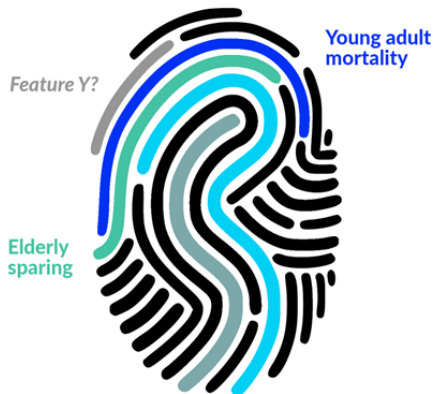
Grouping crises

Summary & 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

Summary & discussion

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

Summary & discussion

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

Summary & discussion

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

Summary & discussion

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.



505 1831. Døde. Mandkøn.

No.	Dødsdagen.	Begravelsesdagen.	Den Dødes Navn og Tilnavn.	Stand, Haandtering og Dypholdstid.	Alder.	Stort anført i det almindelige 4 ^{te} Jernførstes Register.	Bemærkninger.
35	29. August	4. Septbr.	Anders Jørgensen	Jæger i Højbjerg	58 Aar	692. 138	
36	31. August	4. Septbr.	Hans Carlén	Jæger i Højbjerg	65 Aar	692. 139	
37	30. August	3. Septbr.	Ole Jensen	Ablogermand i Højbjerg	89 Aar	692. 140	
38	3. Septbr.	7. Septbr.	Nelger Larsen	Jæger i Højbjerg	70 Aar	692. 141	
39	31. August	4. Septbr.	Hans Olsen	Jæger og Bøddler i Højbjerg	92 Aar	692. 142	
40	4. Septbr.	6. Septbr.	Niels Pedersen	Jæger i Højbjerg	60 Aar	692. 143	
41	5. Septbr.	9. Septbr.	Ole Hansen	Ablogermand i Højbjerg	61 Aar	693. 144	
42	4. Septbr.	9. Septbr.	Niels Christensen	Jæger i Højbjerg	57 Aar	693. 145	
43	7. Septbr.	12. Septbr.	Niels Larsen	Jæger i Højbjerg	84 Aar	693. 146	
44	6. Septbr.	12. Septbr.	Jens Andersen	Ablogermand i Højbjerg	72 Aar	693. 147	
45	8. Septbr.	13. Septbr.	Hans Mathsen	Jæger i Højbjerg	42 Aar	693. 148	
46	5. Septbr.	9. Septbr.	Lars Christoffersen	Ablogermand i Højbjerg	93 Aar	693. 149	
47	12. Septbr.	16. Septbr.	Niels Christensen	Jæger i Højbjerg	78 Aar	693. 150	
48	11. Septbr.	14. Septbr.	Hans Larsen	Jæger i Højbjerg	49 Aar	693. 151	

- Parish registers for Danish church parishes between 1815-1915

Døde. Dødsfald.					
No.	Kirkefoged.	Baptismand.	Den Døde Navn og Alder.	Sted, hvortil den er flyttet.	Sted, hvortil den er flyttet.
35	1. April	Anders Christensen	Anders Christensen	Anders Christensen	Anders Christensen
36	2. April	Anders Christensen	Anders Christensen	Anders Christensen	Anders Christensen
37	3. April	Anders Christensen	Anders Christensen	Anders Christensen	Anders Christensen
38	4. April	Anders Christensen	Anders Christensen	Anders Christensen	Anders Christensen
39	5. April	Anders Christensen	Anders Christensen	Anders Christensen	Anders Christensen
40	6. April	Anders Christensen	Anders Christensen	Anders Christensen	Anders Christensen
41	7. April	Anders Christensen	Anders Christensen	Anders Christensen	Anders Christensen
42	8. April	Anders Christensen	Anders Christensen	Anders Christensen	Anders Christensen
43	9. April	Anders Christensen	Anders Christensen	Anders Christensen	Anders Christensen
44	10. April	Anders Christensen	Anders Christensen	Anders Christensen	Anders Christensen
45	11. April	Anders Christensen	Anders Christensen	Anders Christensen	Anders Christensen
46	12. April	Anders Christensen	Anders Christensen	Anders Christensen	Anders Christensen
47	13. April	Anders Christensen	Anders Christensen	Anders Christensen	Anders Christensen
48	14. April	Anders Christensen	Anders Christensen	Anders Christensen	Anders Christensen

Scan of parish register for "Fakse" parish.

Data source

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

No.	Kirkefoged.	Begravelsesfoged.	Den Døde Børns Navn.	Stad, Bopæl og Æt.	Stad, Bopæl og Æt.	Stad, Bopæl og Æt.
35	J. L. Højgaard	A. Højgaard	Louise Christensen	Jens & Højgaard	Stad, Bopæl og Æt.	
36	J. L. Højgaard	A. Højgaard	Karl Carlén	Jens & Højgaard	Stad, Bopæl og Æt.	
37	J. L. Højgaard	A. Højgaard	Oluf Jensen	Højgaard & Højgaard	Stad, Bopæl og Æt.	
38	J. L. Højgaard	A. Højgaard	Nelger Larsen	Jens & Højgaard	Stad, Bopæl og Æt.	
39	J. L. Højgaard	A. Højgaard	Karl Olsen	Jens & Højgaard	Stad, Bopæl og Æt.	
40	J. L. Højgaard	A. Højgaard	Nelger Christensen	Jens & Højgaard	Stad, Bopæl og Æt.	
41	J. L. Højgaard	A. Højgaard	Oluf Christensen	Jens & Højgaard	Stad, Bopæl og Æt.	
42	J. L. Højgaard	A. Højgaard	Nelger Christensen	Jens & Højgaard	Stad, Bopæl og Æt.	
43	J. L. Højgaard	A. Højgaard	Nelger Christensen	Jens & Højgaard	Stad, Bopæl og Æt.	
44	J. L. Højgaard	A. Højgaard	Nelger Christensen	Jens & Højgaard	Stad, Bopæl og Æt.	
45	J. L. Højgaard	A. Højgaard	Nelger Christensen	Jens & Højgaard	Stad, Bopæl og Æt.	
46	J. L. Højgaard	A. Højgaard	Nelger Christensen	Jens & Højgaard	Stad, Bopæl og Æt.	
47	J. L. Højgaard	A. Højgaard	Nelger Christensen	Jens & Højgaard	Stad, Bopæl og Æt.	
48	J. L. Højgaard	A. Højgaard	Nelger Christensen	Jens & Højgaard	Stad, Bopæl og Æt.	

Scan of parish register for "Fakse" parish.

Identifying
Signature Features

RK Pedersen

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

Summary & discussion



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

No.	Beftjættelse.	Begravelsesdag.	Den Døde Børns Navn.	Kirke, hvortil den er begravet.	Den Døde Børns Alder.	Begravelsesdag.
35	1. April	1. April	Anders Christensen	St. Nicolai	58 Aar	1. April 1891
36	2. April	2. April	Hans Christensen	St. Nicolai	58 Aar	2. April 1891
37	3. April	3. April	Oluf Jensen	St. Nicolai	58 Aar	3. April 1891
38	4. April	4. April	Hilger Larsen	St. Nicolai	58 Aar	4. April 1891
39	5. April	5. April	Hans Olsen	St. Nicolai	58 Aar	5. April 1891
40	6. April	6. April	Anders Christensen	St. Nicolai	58 Aar	6. April 1891
41	7. April	7. April	Oluf Jensen	St. Nicolai	58 Aar	7. April 1891
42	8. April	8. April	Hilger Larsen	St. Nicolai	58 Aar	8. April 1891
43	9. April	9. April	Hans Olsen	St. Nicolai	58 Aar	9. April 1891
44	10. April	10. April	Anders Christensen	St. Nicolai	58 Aar	10. April 1891
45	11. April	11. April	Oluf Jensen	St. Nicolai	58 Aar	11. April 1891
46	12. April	12. April	Hilger Larsen	St. Nicolai	58 Aar	12. April 1891
47	13. April	13. April	Hans Olsen	St. Nicolai	58 Aar	13. April 1891
48	14. April	14. April	Anders Christensen	St. Nicolai	58 Aar	14. April 1891
49	15. April	15. April	Oluf Jensen	St. Nicolai	58 Aar	15. April 1891
50	16. April	16. April	Hilger Larsen	St. Nicolai	58 Aar	16. April 1891

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 of death, date of burial, gender, age and parish
- ▶ Property of the Danish National Archives, but digitized and transcribed by Ancestry



No.	Kirkefoged.	Begravelsesfoged.	Den Døde Børns Navn.	Stad, Bopæl og alder.	Stad, Bopæl og alder.	Begravelsesdag.
35	J. L. Højgaard	A. H. Højgaard	Louise Christensen	1831	1831	1831
36	J. L. Højgaard	A. H. Højgaard	Karl Christensen	1831	1831	1831
37	J. L. Højgaard	A. H. Højgaard	Carl Christensen	1831	1831	1831
38	J. L. Højgaard	A. H. Højgaard	Karl Christensen	1831	1831	1831
39	J. L. Højgaard	A. H. Højgaard	Karl Christensen	1831	1831	1831
40	J. L. Højgaard	A. H. Højgaard	Karl Christensen	1831	1831	1831
41	J. L. Højgaard	A. H. Højgaard	Karl Christensen	1831	1831	1831
42	J. L. Højgaard	A. H. Højgaard	Karl Christensen	1831	1831	1831
43	J. L. Højgaard	A. H. Højgaard	Karl Christensen	1831	1831	1831
44	J. L. Højgaard	A. H. Højgaard	Karl Christensen	1831	1831	1831
45	J. L. Højgaard	A. H. Højgaard	Karl Christensen	1831	1831	1831
46	J. L. Højgaard	A. H. Højgaard	Karl Christensen	1831	1831	1831
47	J. L. Højgaard	A. H. Højgaard	Karl Christensen	1831	1831	1831
48	J. L. Højgaard	A. H. Højgaard	Karl Christensen	1831	1831	1831

Scan of parish register for "Fakse" parish.

Identifying
Signature Features

RK Pedersen

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

Summary & discussion



Data cleaning and managing

Identifying
Signature Features

RK Pedersen

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

Summary & discussion

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 Tüylfang	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 Thjystrup	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)



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 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	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 Thjstrup	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



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	Vejlerv 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	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 Thjstrup	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)

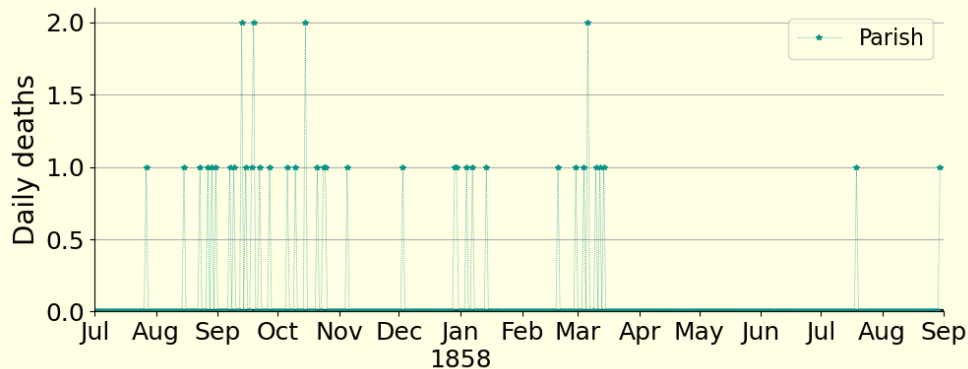
► Temporal resolution:

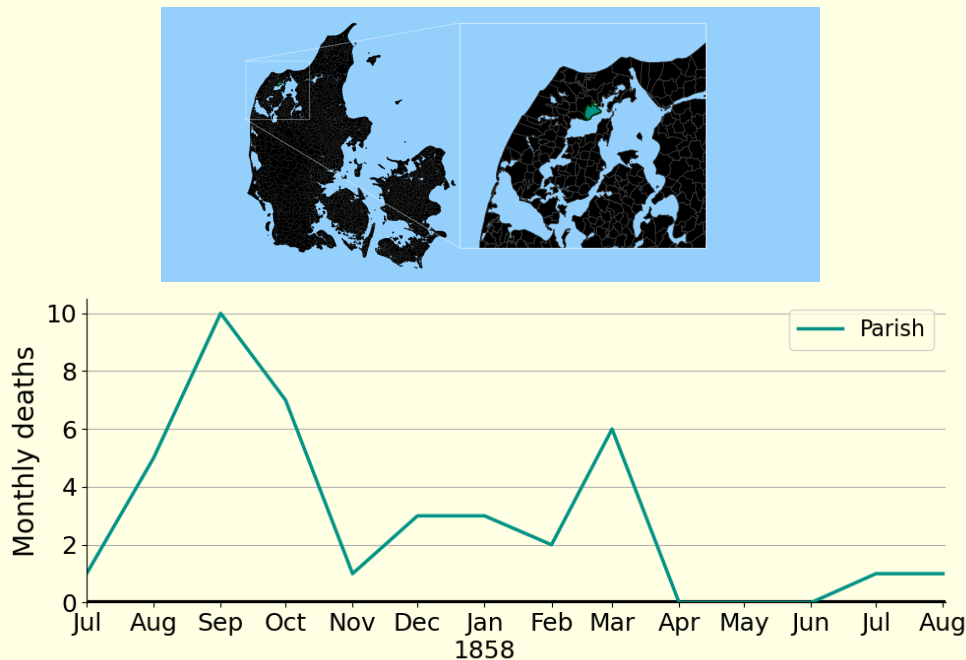
- Daily
- Weekly
- Monthly
- Yearly

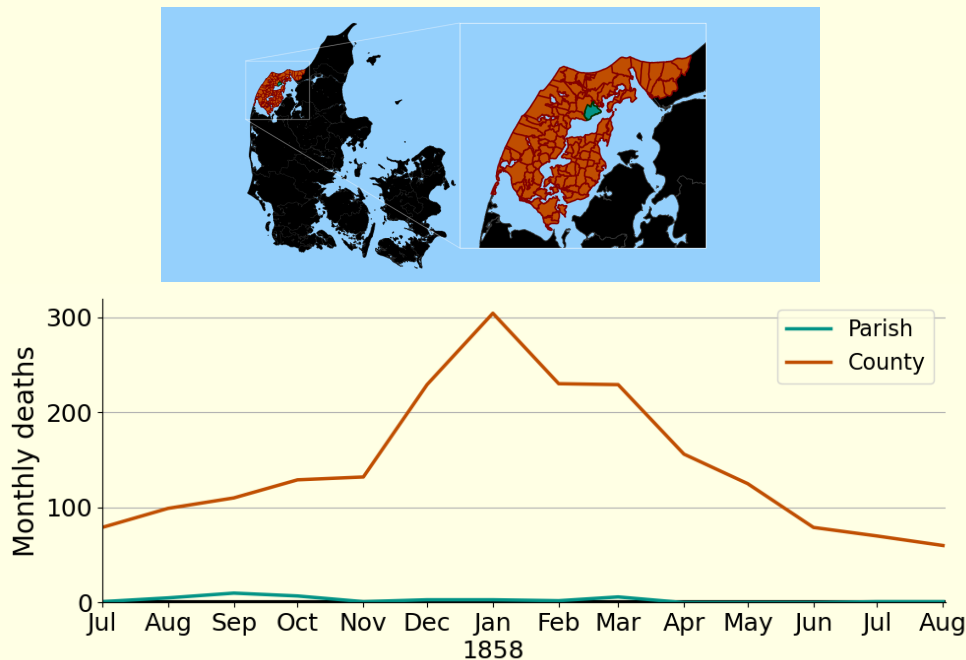
► Geographical resolution:

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









Calculating the mortality baseline

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

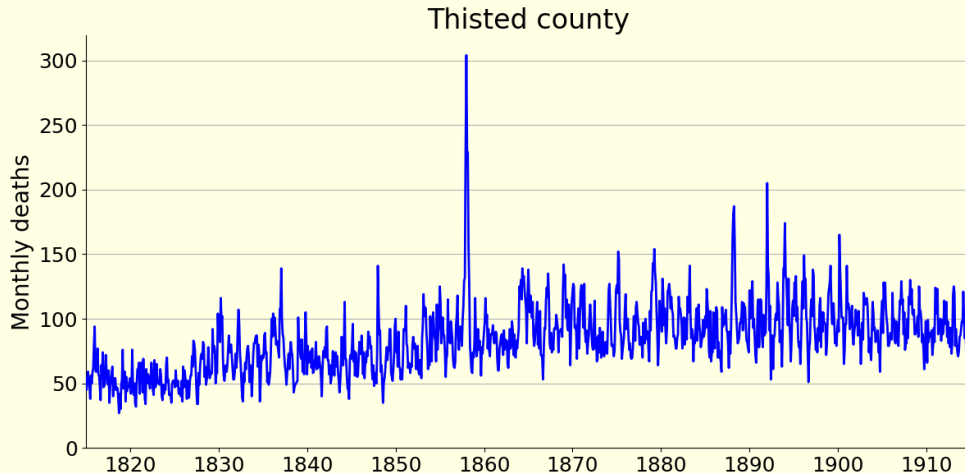
Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

Summary & discussion



RK Pedersen

Introduction

Background and data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

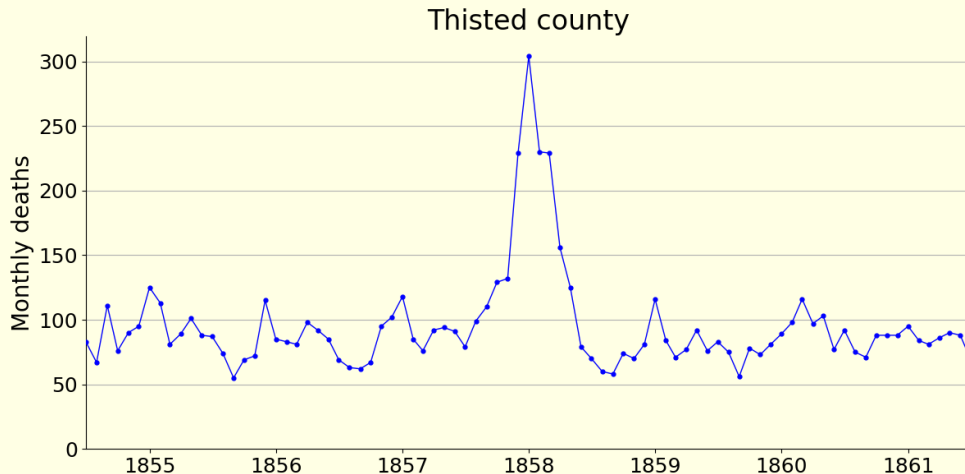
Comparing age patterns

Results and discussion

Representative signature features

Grouping crises

Summary & discussion



Calculating the mortality baseline

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

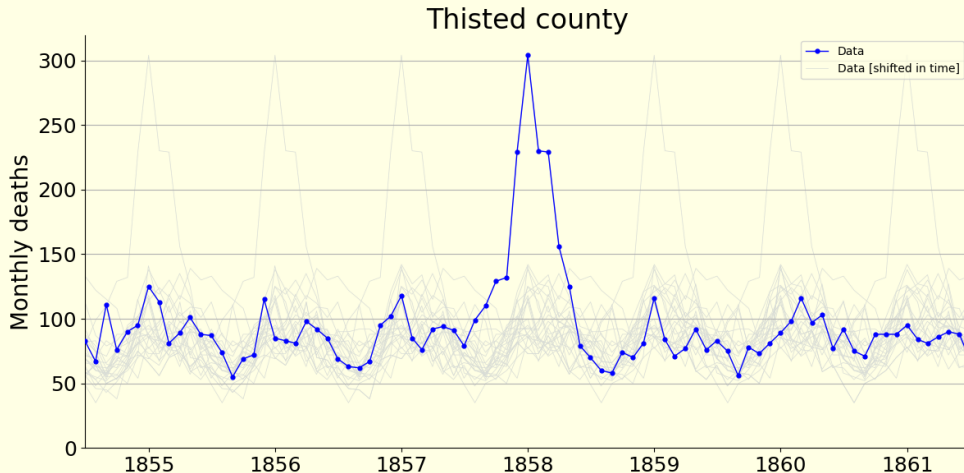
Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

Summary & discussion



Calculating the mortality baseline

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

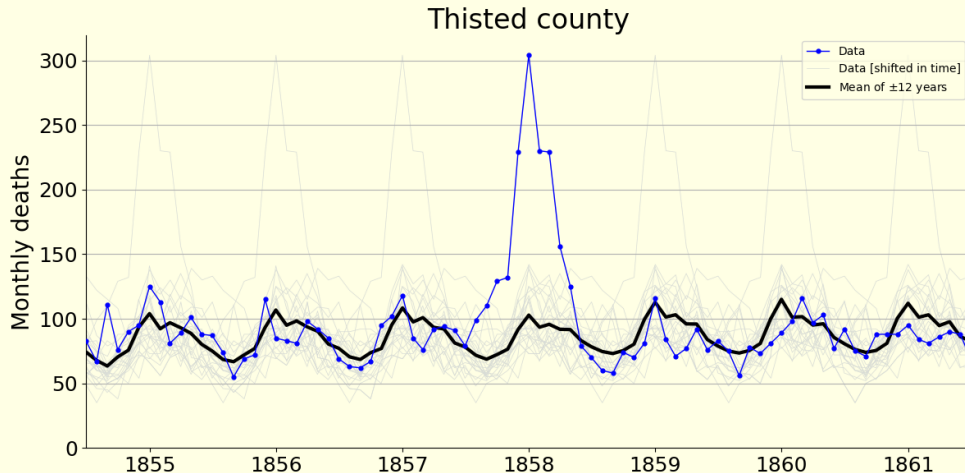
Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

Summary & discussion



Calculating the mortality baseline

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

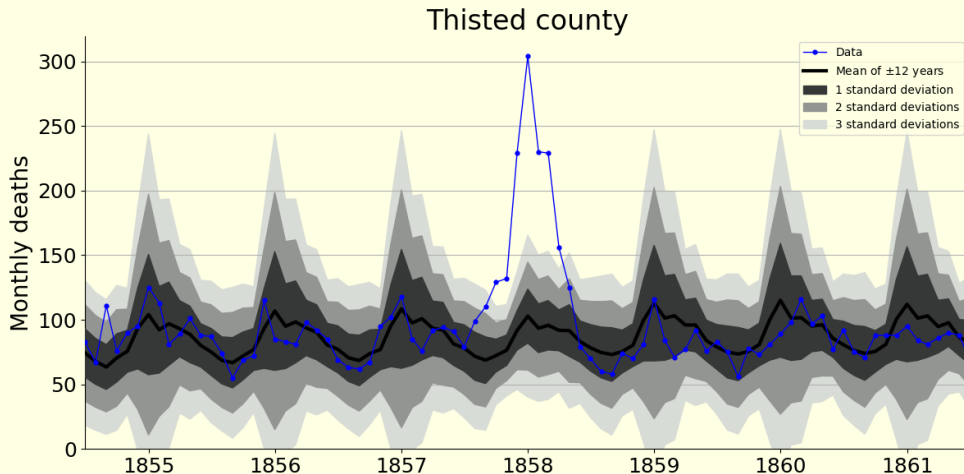
Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

Summary & discussion



Calculating the mortality baseline

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

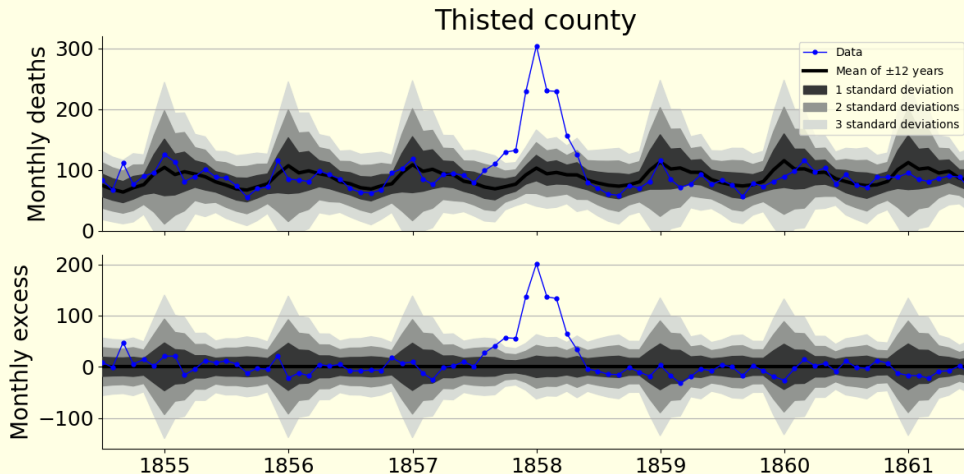
Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

Summary & discussion



Calculating the mortality baseline

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

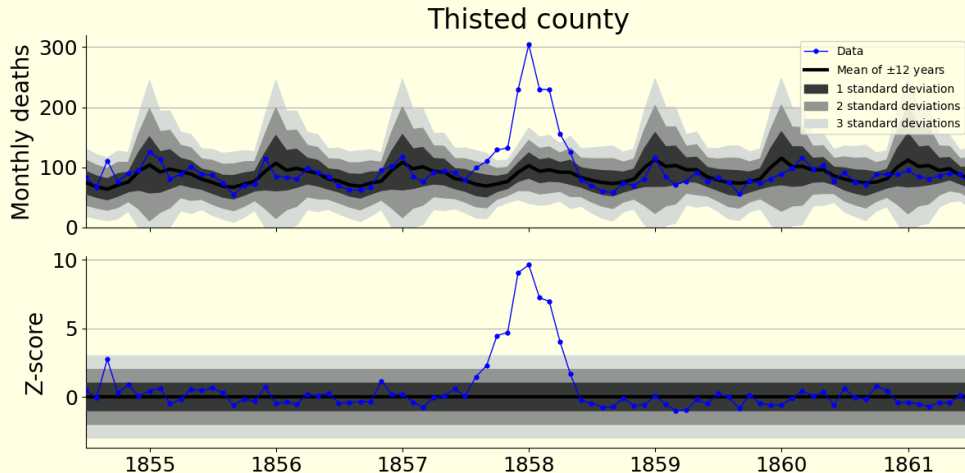
Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

Summary & discussion



Calculating the mortality baseline

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

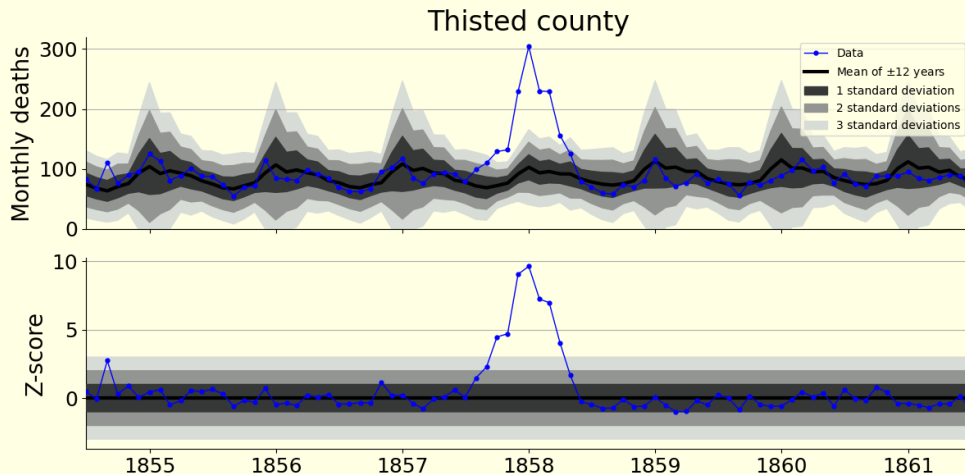
Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

Summary & discussion



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



Calculating the mortality baseline

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

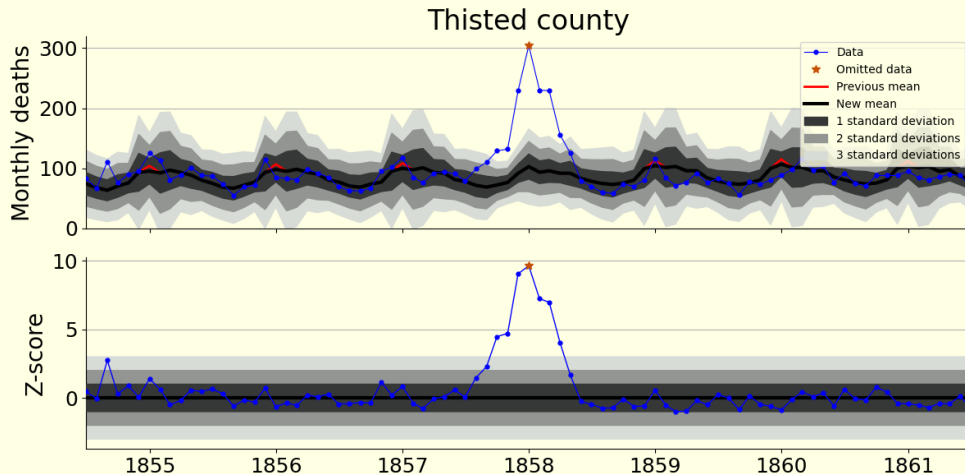
Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

Summary & discussion



Calculating the mortality baseline

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

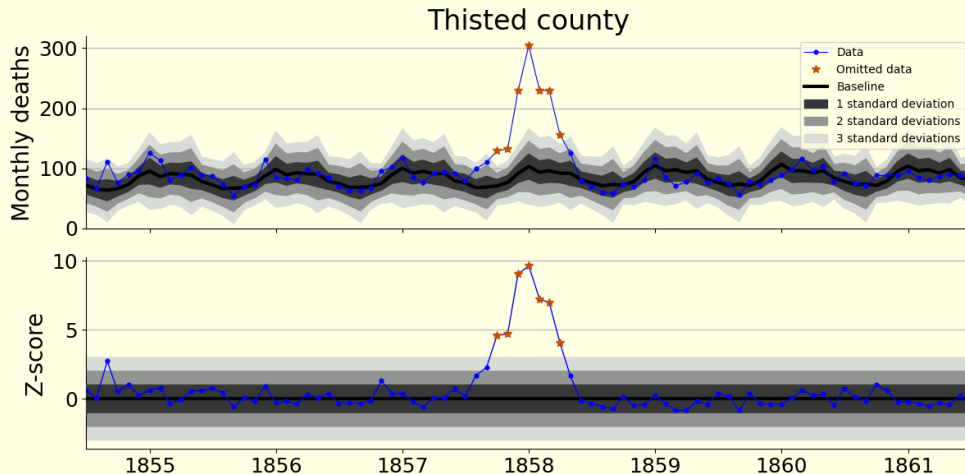
Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

Summary & discussion



Calculating the mortality baseline

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

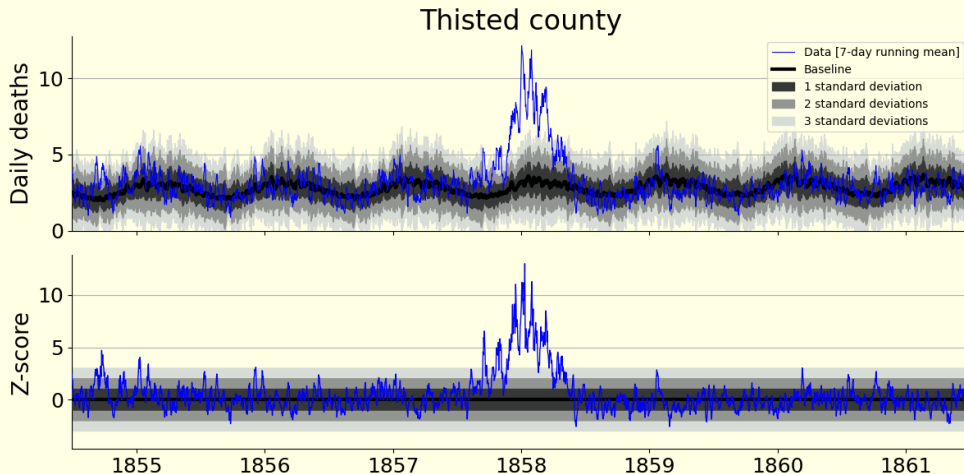
Comparing age patterns

Results and
discussion

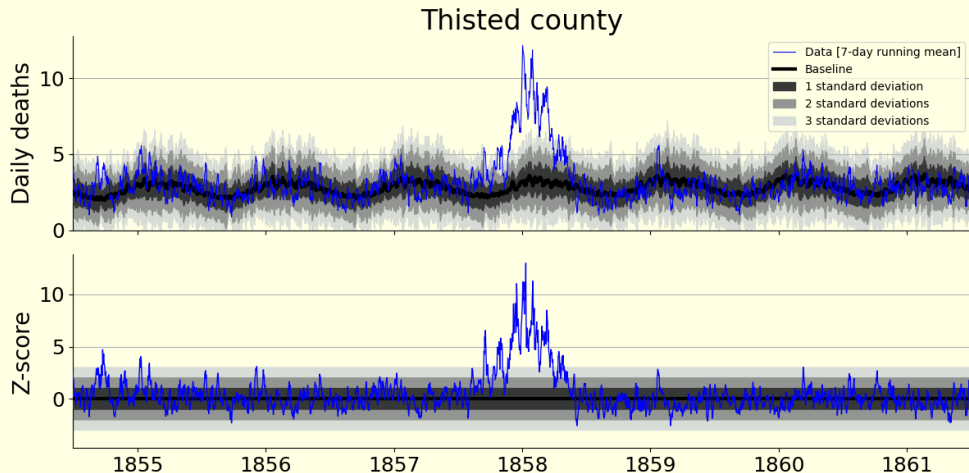
Representative signature
features

Grouping crises

Summary & discussion



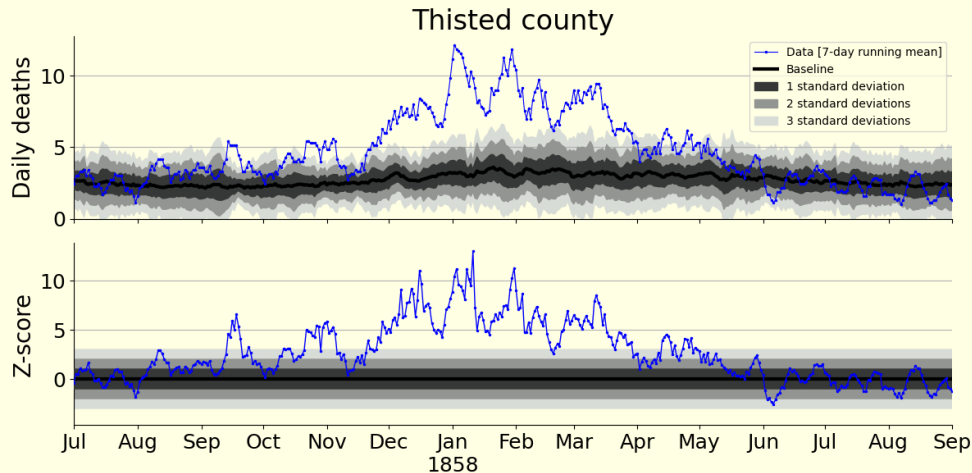
Identification of “mortality crises”



We wish to identify continuous periods of excess mortality.



Identification of “mortality crises”



Zooming in...



Identification of “mortality crises”

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

“Mortality crisis”

Age-specific mortality

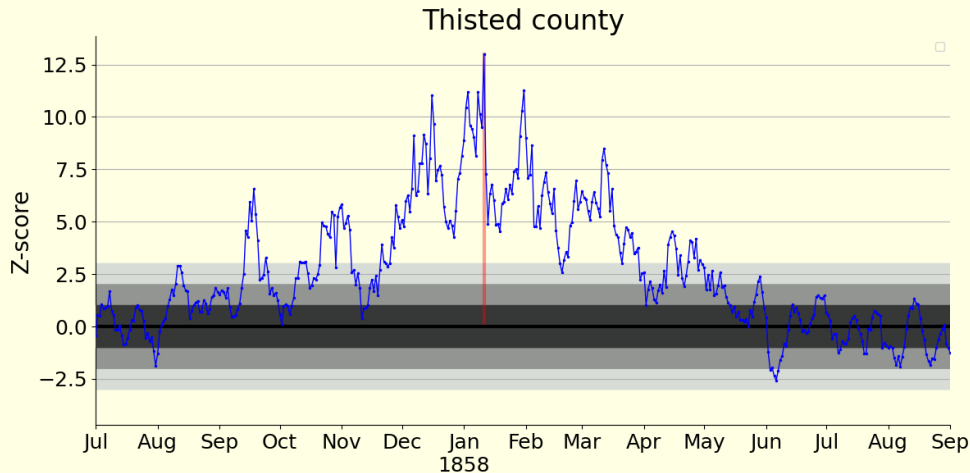
Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

Summary & discussion



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



Identification of “mortality crises”

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

“Mortality crisis”

Age-specific mortality

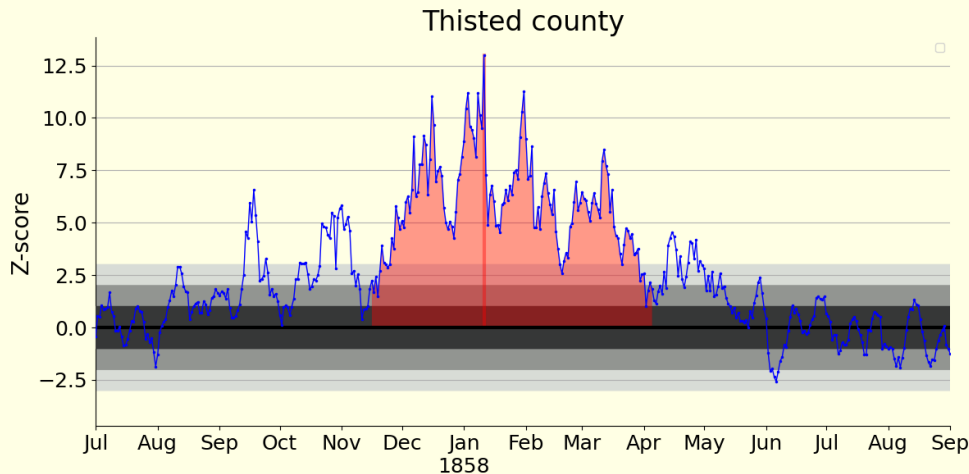
Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

Summary & discussion



... 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”

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

“Mortality crisis”

Age-specific mortality

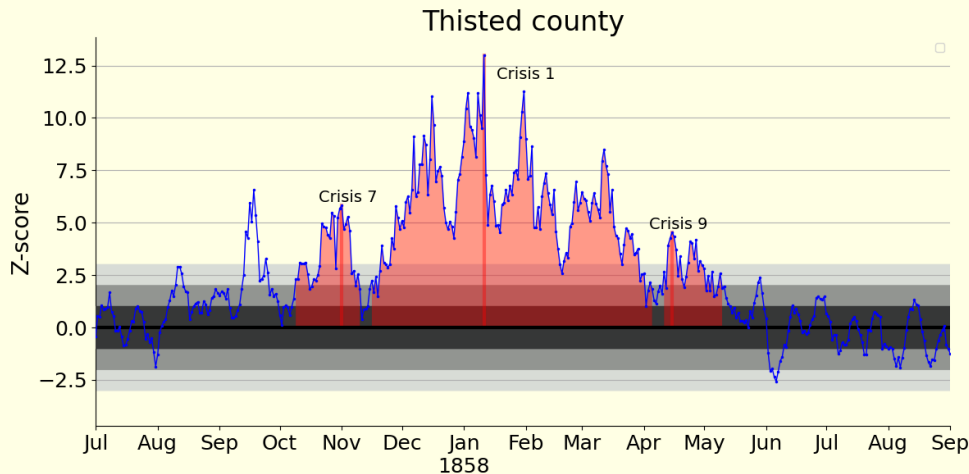
Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

Summary & discussion



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 319 mortality crises.

<i>County</i>	<i>#</i>			
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 319 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 319 mortality crises.

For each crisis,

- Peak-date.

<i>County</i>	<i>#</i>	<i>Peak-date</i>		
Thisted	1	1858-01-23		
Thisted	2	1892-01-23		
Thisted	3	1864-08-07		
⋮	⋮	⋮		
Copenhagen	1	1831-08-30		
Copenhagen	2	1891-12-13		
Copenhagen	3	1853-07-31		
⋮	⋮	⋮		
Aarhus	1	1892-01-06		
Aarhus	2	1853-08-24		
⋮	⋮	⋮		



Identifying main “signature features”

Using this methodology, we identify 319 mortality crises.

For each crisis,

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

<i>County</i>	<i>#</i>	<i>Peak-date</i>	<i>Total excess</i>	
Thisted	1	1858-01-23	655	
Thisted	2	1892-01-23	114	
Thisted	3	1864-08-07	105	
⋮	⋮	⋮	⋮	
Copenhagen	1	1831-08-30	592	
Copenhagen	2	1891-12-13	573	
Copenhagen	3	1853-07-31	552	
⋮	⋮	⋮	⋮	
Aarhus	1	1892-01-06	400	
Aarhus	2	1853-08-24	194	
⋮	⋮	⋮	⋮	



Identifying main “signature features”

Using this methodology, we identify 319 mortality crises.

For each crisis,

- Peak-date.
- Excess deaths.
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	1858-01-23	655	140 days
Thisted	2	1892-01-23	114	43 days
Thisted	3	1864-08-07	105	61 days
⋮	⋮	⋮	⋮	⋮
Copenhagen	1	1831-08-30	592	57 days
Copenhagen	2	1891-12-13	573	69 days
Copenhagen	3	1853-07-31	552	57 days
⋮	⋮	⋮	⋮	⋮
Aarhus	1	1892-01-06	400	61 days
Aarhus	2	1853-08-24	194	49 days
⋮	⋮	⋮	⋮	⋮



Identifying main “signature features”

Using this methodology, we identify 319 mortality crises.

For each crisis,

- ▶ Peak-date.
- ▶ Excess deaths.
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”

County	#	Peak-date	Total excess	Duration
Thisted	1	1858-01-23	655	140 days
Thisted	2	1892-01-23	114	43 days
Thisted	3	1864-08-07	105	61 days
⋮	⋮	⋮	⋮	⋮
Copenhagen	1	1831-08-30	592	57 days
Copenhagen	2	1891-12-13	573	69 days
Copenhagen	3	1853-07-31	552	57 days
⋮	⋮	⋮	⋮	⋮
Aarhus	1	1892-01-06	400	61 days
Aarhus	2	1853-08-24	194	49 days
⋮	⋮	⋮	⋮	⋮



Identifying main “signature features”

Using this methodology, we identify 319 mortality crises.

For each crisis,

- ▶ Peak-date.
- ▶ Excess deaths.
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	1858-01-23	655	140 days
Thisted	2	1892-01-23	114	43 days
Thisted	3	1864-08-07	105	61 days
⋮	⋮	⋮	⋮	⋮
Copenhagen	1	1831-08-30	592	57 days
Copenhagen	2	1891-12-13	573	69 days
Copenhagen	3	1853-07-31	552	57 days
⋮	⋮	⋮	⋮	⋮
Aarhus	1	1892-01-06	400	61 days
Aarhus	2	1853-08-24	194	49 days
⋮	⋮	⋮	⋮	⋮

But we also have data on age.



Analyzing age-patterns

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

Results and
discussion

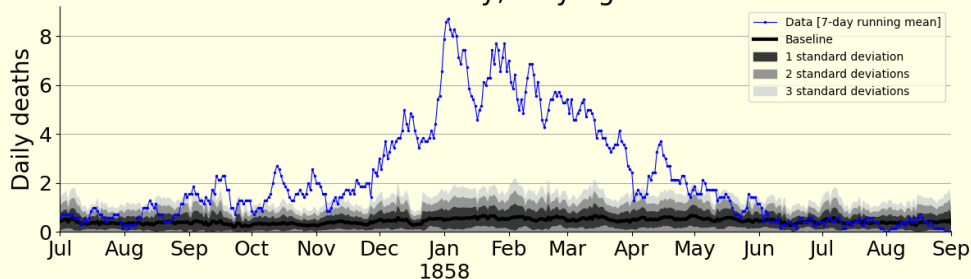
Representative signature
features

Grouping crises

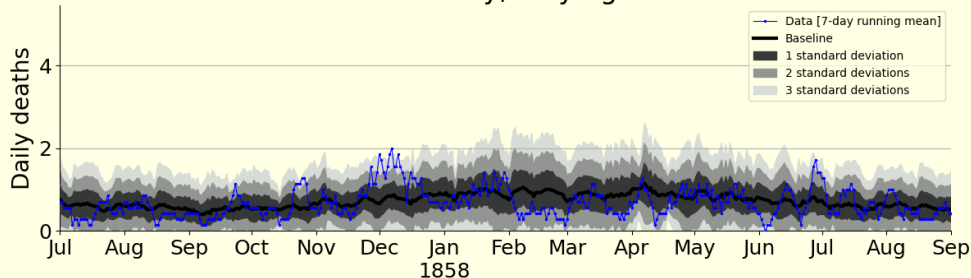
Summary & discussion



Thisted county, only ages 1-14



Thisted county, only ages 60+



Analyzing age-patterns

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

Results and
discussion

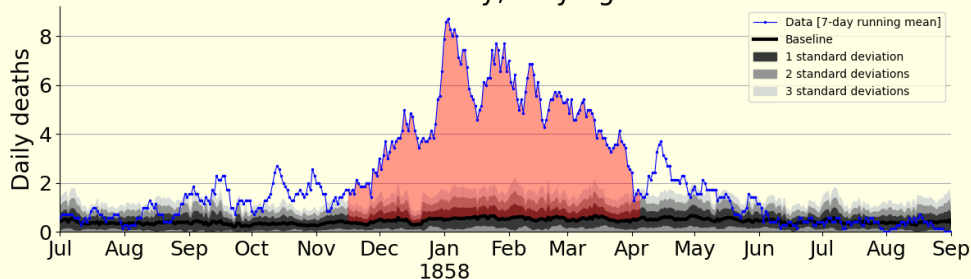
Representative signature
features

Grouping crises

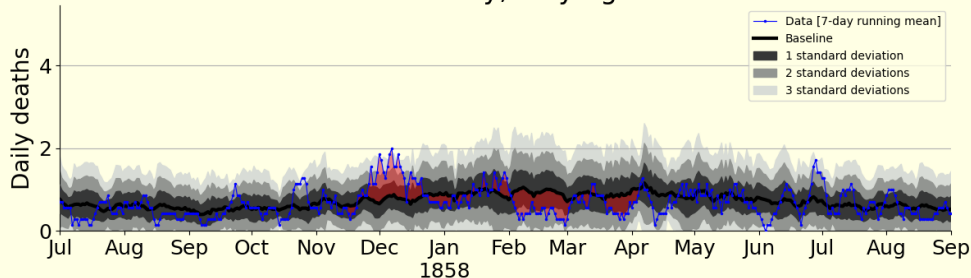
Summary & discussion



Thisted county, only ages 1-14



Thisted county, only ages 60+



Analyzing age-patterns

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

Results and
discussion

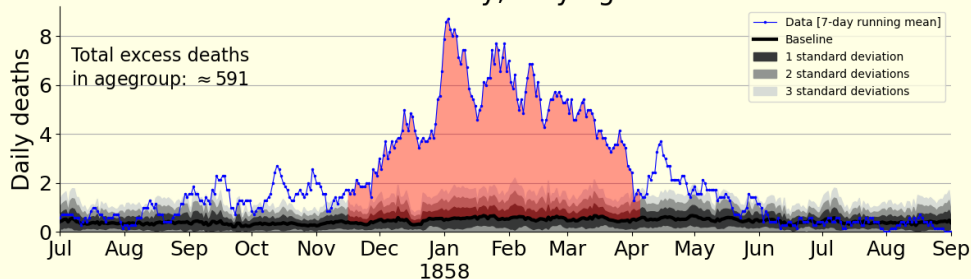
Representative signature
features

Grouping crises

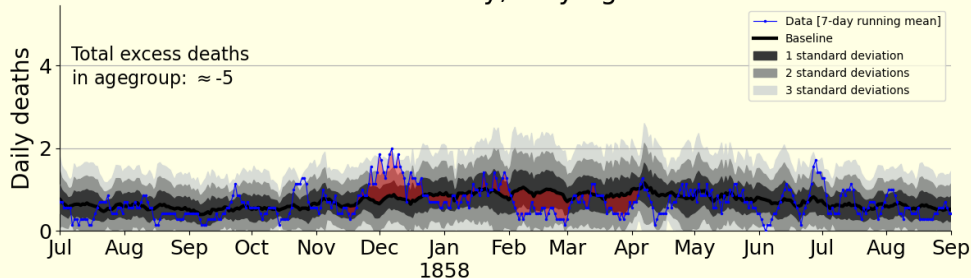
Summary & discussion



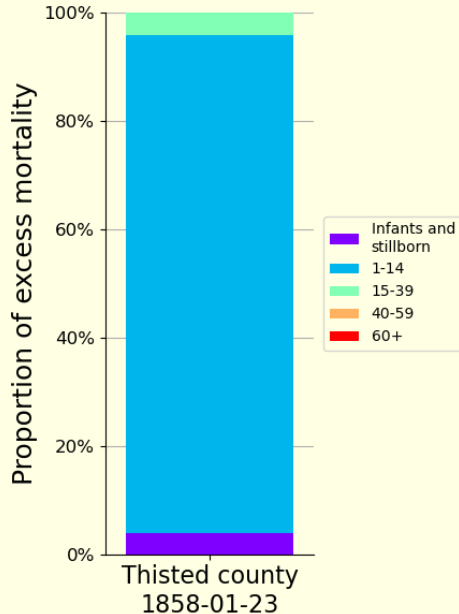
Thisted county, only ages 1-14



Thisted county, only ages 60+



Analyzing age-patterns



Age-specific excess mortality in this period:

- ▶ Age group "60+": 0%
- ▶ Age group "40-59": <1%
- ▶ Age group "15-39": 3%
- ▶ Age group "1-14": 93%
- ▶ Age group "Below 1 year": 4%

Mortality crises with comparable age patterns

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

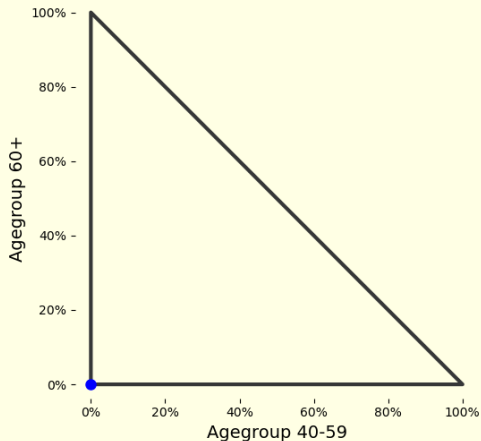
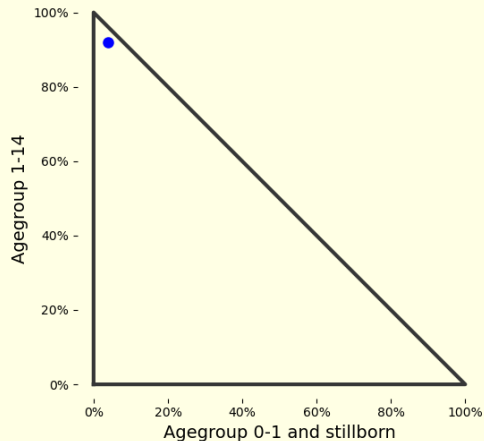
Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

Summary & discussion



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



Mortality crises with comparable age patterns

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

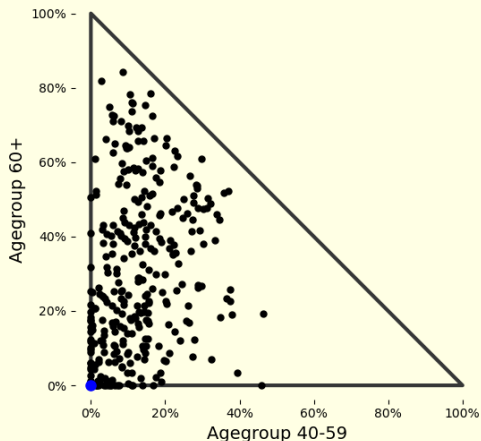
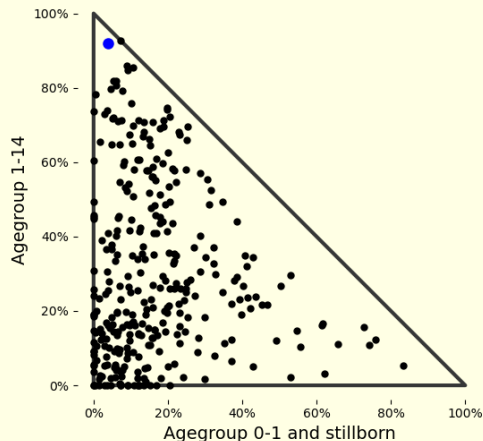
Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

Summary & discussion



Adding the other 318 mortality crises identified.



Mortality crises with comparable age patterns

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

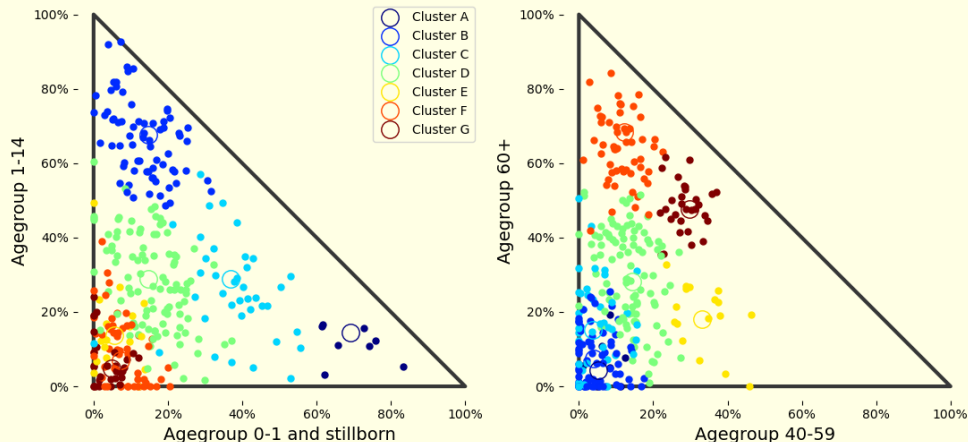
Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

Summary & discussion



Gaussian mixture modelling on full five-dimensional data.



Mortality crises with comparable age patterns

Identifying
Signature Features

RK Pedersen

Introduction

Background and
data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

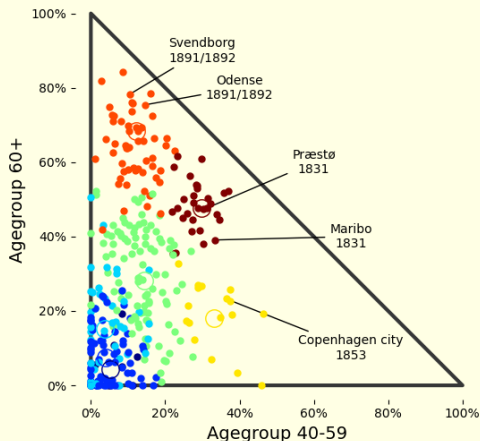
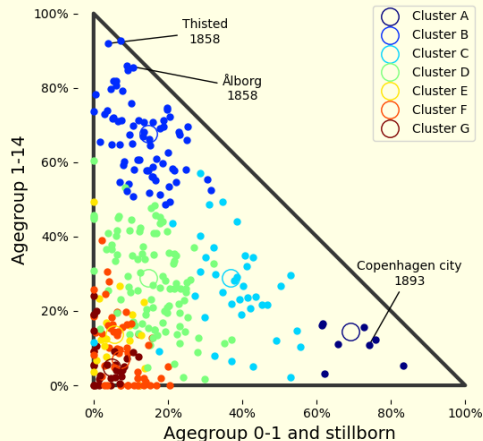
Comparing age patterns

Results and
discussion

Representative signature
features

Grouping crises

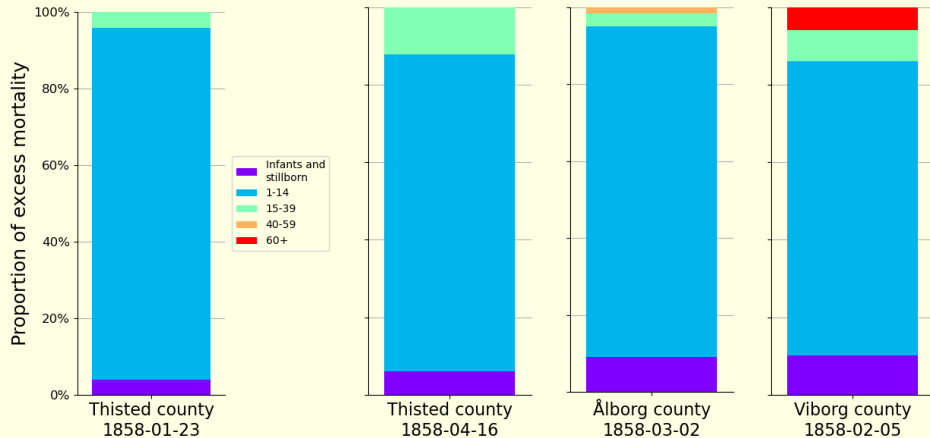
Summary & discussion



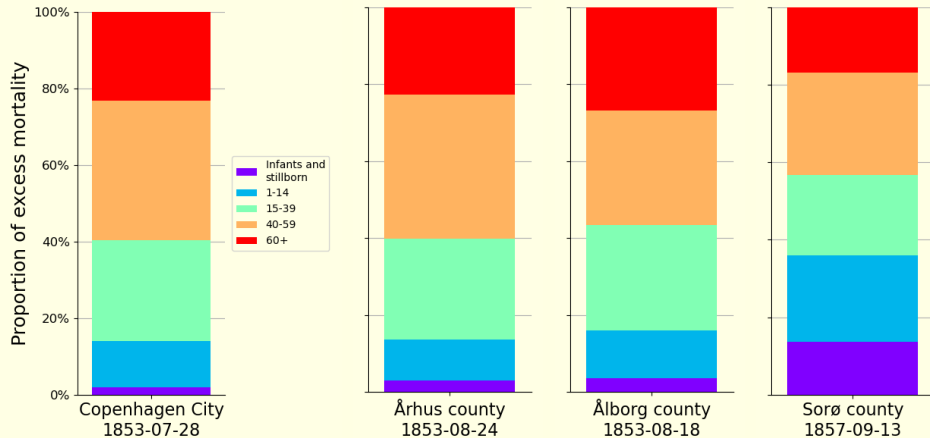
Gaussian mixture modelling on full five-dimensional data.



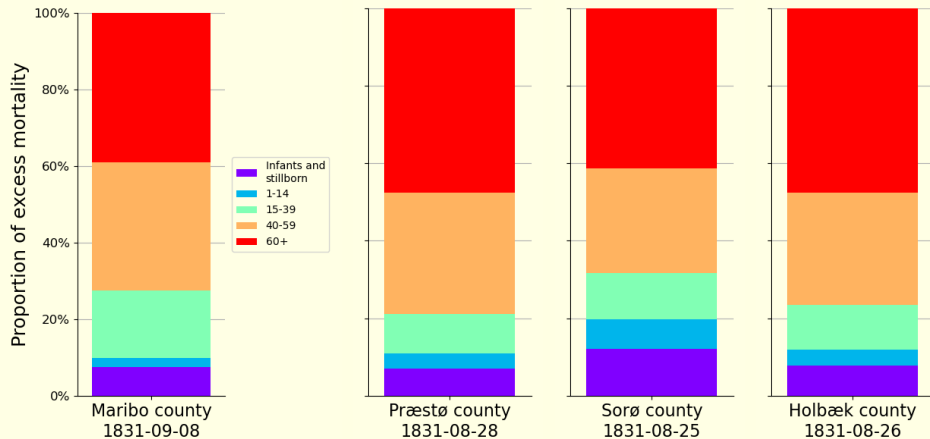
Examples from "Cluster A"



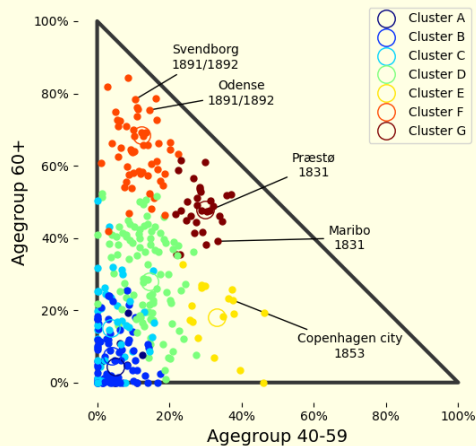
Examples from "Cluster E"



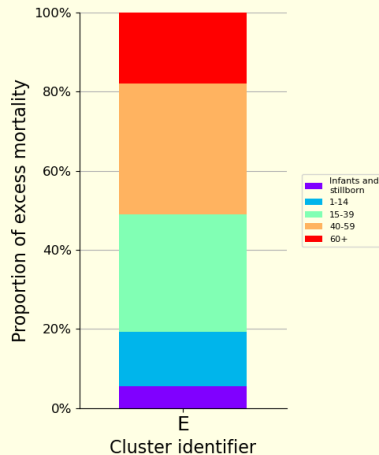
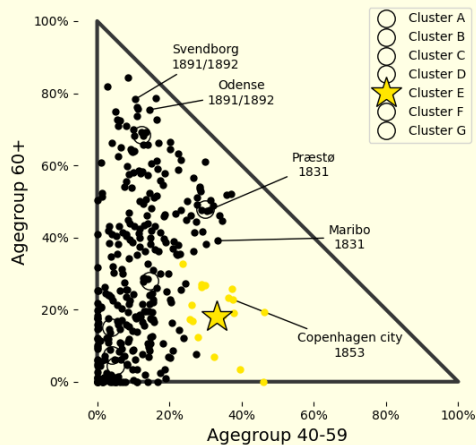
Examples from "Cluster G"



The signature features of certain diseases



The signature features of certain diseases



Multiple of the mortality crises in cluster E appear to be related to **cholera**.

Introduction

Background and data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

Results and discussion

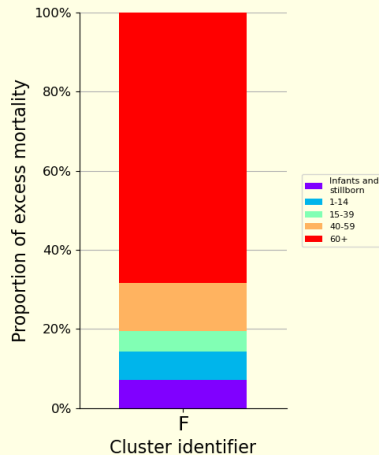
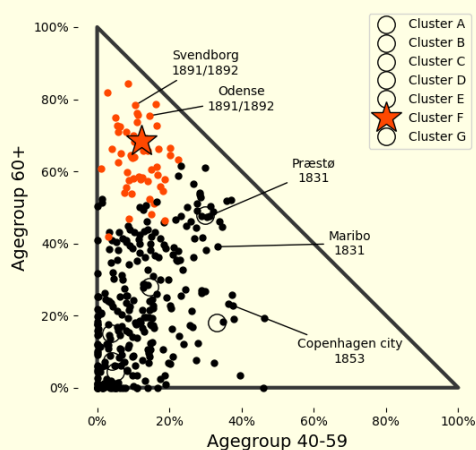
Representative signature features

Grouping crises

Summary & discussion



The signature features of certain diseases



Multiple of the mortality crises in cluster F appear to be related to **pandemic influenza**.

Introduction

Background and data handling

Data source

Data cleaning

Methodology

Mortality baseline

"Mortality crisis"

Age-specific mortality

Comparing age patterns

Results and discussion

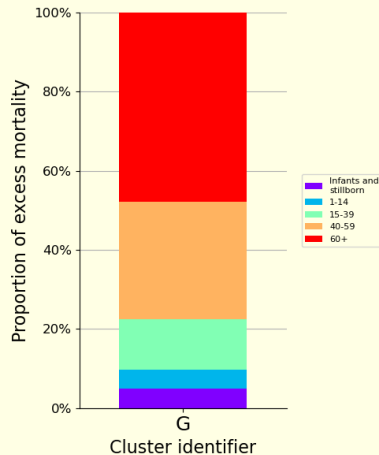
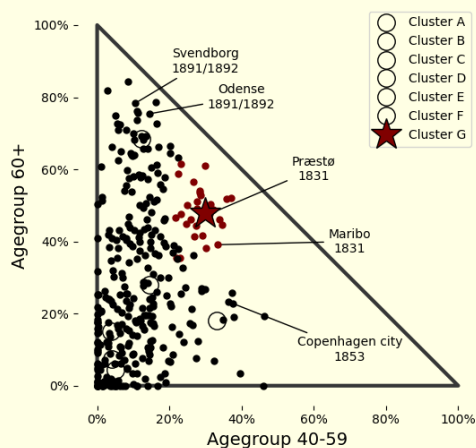
Representative signature features

Grouping crises

Summary & discussion



The signature features of certain diseases



Multiple of the mortality crises in cluster G appear to be related to **“the Harvest epidemics of 1826-1832”**.

Introduction

Background and data handling

Data source

Data cleaning

Methodology

Mortality baseline

“Mortality crisis”

Age-specific mortality

Comparing age patterns

Results and discussion

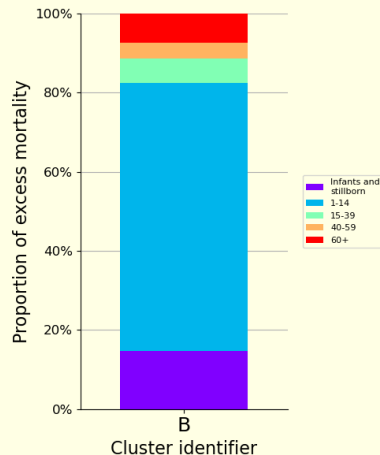
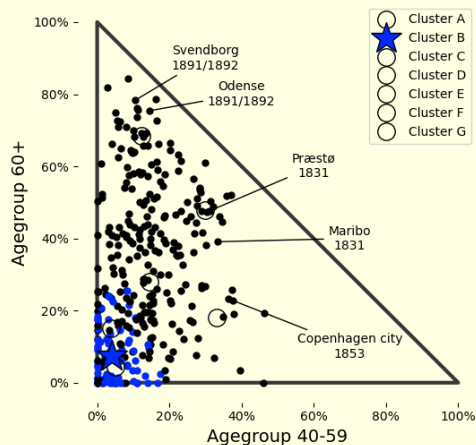
Representative signature features

Grouping crises

Summary & discussion



The signature features of certain diseases



Multiple of the mortality crises in cluster B appear to be related to **scarlet fever**.

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

Summary & discussion



All epidemics and pandemics in 19th century Denmark

Identifying
Signature Features

RK Pedersen

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

Summary & discussion

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

<i>Disease</i>	<i>Timing</i>	<i>Total excess</i>	<i>Age structure</i>
Cholera	Late summer, 1853 and 1857	5381	Adults Cluster "E"
Scarlet fever	Winter 1857/1858	2451	Children (1-15) Cluster "B"
"Harvest epidemics"¹	Late summer, 1826-1832	10818	Adults Cluster "G"
Pandemic influenza	1892 and 1900	8201	Elderly Cluster "F"

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

¹Various diseases, aggravated by a subsistence crisis. Discussed in detail by Ingholt (2022) *Scandinavian Journal of History*



- We determine mortality baselines on county-level, using an iterative process to omit outliers and estimate excess mortality.

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

Summary & discussion



- ▶ We determine mortality baselines on county-level, using an iterative process to omit outliers and estimate excess mortality.
- ▶ We identify 319 major mortality crises in 19th century Denmark.



- ▶ We determine mortality baselines on county-level, using an iterative process to omit outliers and estimate excess mortality.
- ▶ We identify 319 major mortality crises in 19th century Denmark.
- ▶ For each crisis, we determine signature features:



- ▶ We determine mortality baselines on county-level, using an iterative process to omit outliers and estimate excess mortality.
- ▶ We identify 319 major mortality crises in 19th century Denmark.
- ▶ For each crisis, we determine signature features:
 - ▶ Age-patterns.



- ▶ We determine mortality baselines on county-level, using an iterative process to omit outliers and estimate excess mortality.
- ▶ We identify 319 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, using an iterative process to omit outliers and estimate excess mortality.
- ▶ We identify 319 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, using an iterative process to omit outliers and estimate excess mortality.
- ▶ We identify 319 major mortality crises in 19th century Denmark.
- ▶ For each crisis, we determine signature features:
 - ▶ Age-patterns.
 - ▶ Timing and seasonality.
 - ▶ Duration.
 - ▶ Geography.



- ▶ We determine mortality baselines on county-level, using an iterative process to omit outliers and estimate excess mortality.
- ▶ We identify 319 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.



- Similar methods could be applied to modern data.



- ▶ 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.



- ▶ 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.



- ▶ 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.



- ▶ 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



Danmarks
Grundforskningsfond
Danish National
Research Foundation

This work was supported in part by funding from the Danish National Research Foundation (grant number DNRF170), the Carlsberg Foundation (grant number CF20-0046) and NordForsk (project number 104910).

Identifying
Signature Features

RK Pedersen

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

Summary & discussion

