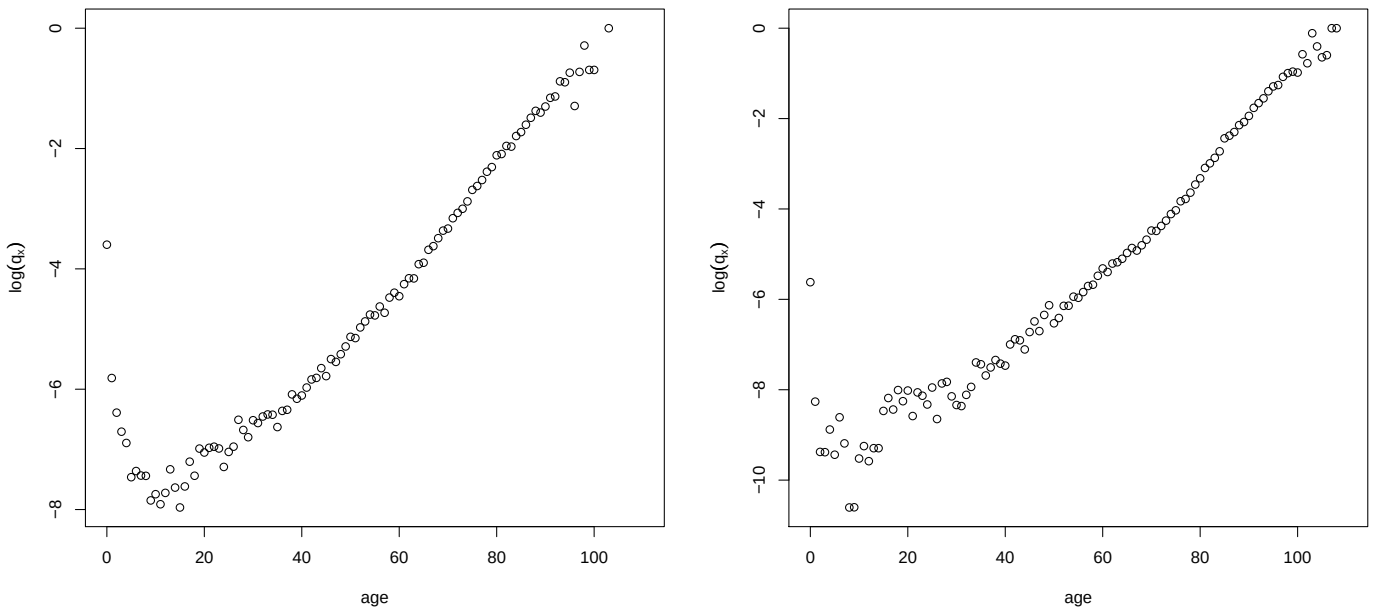
Forecasting Mortality Trends allowing for Cause-of-Death Mortality Dependence

Séverine Gaille & Michael Sherris

September 7, 2012

Table of contents

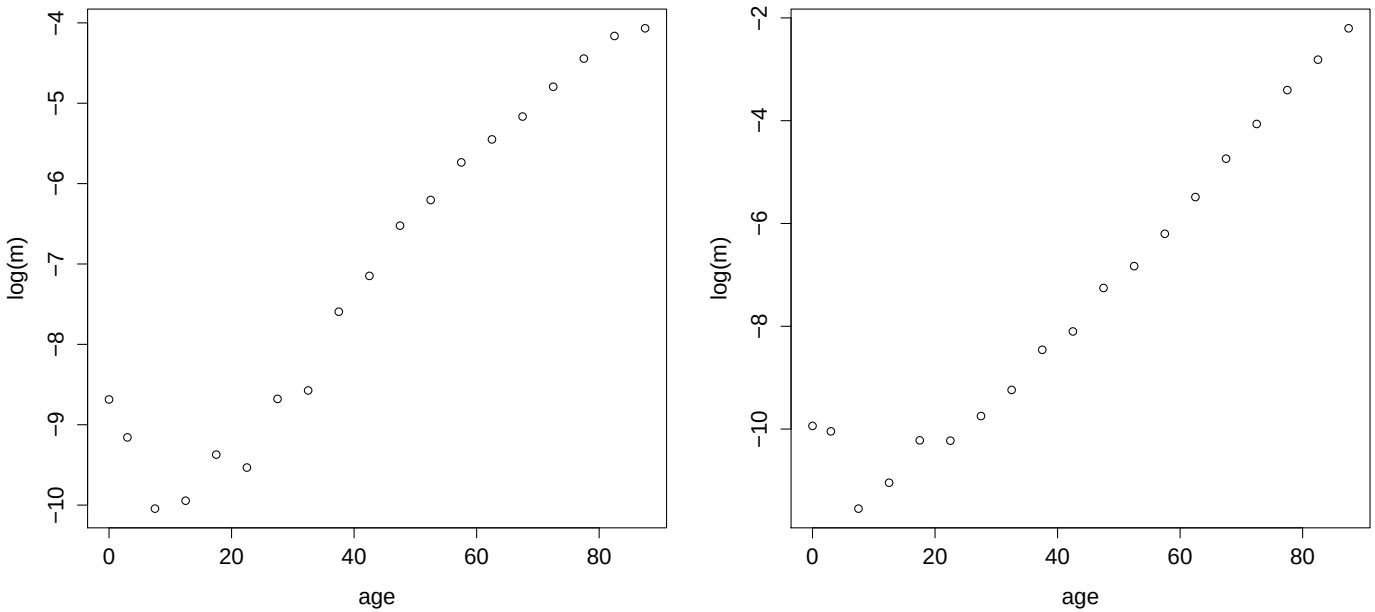
Why should we look at mortality by cause of death?



(a) 1950 (b) 2005

Figure: Log-mortality over ages, females, Switzerland

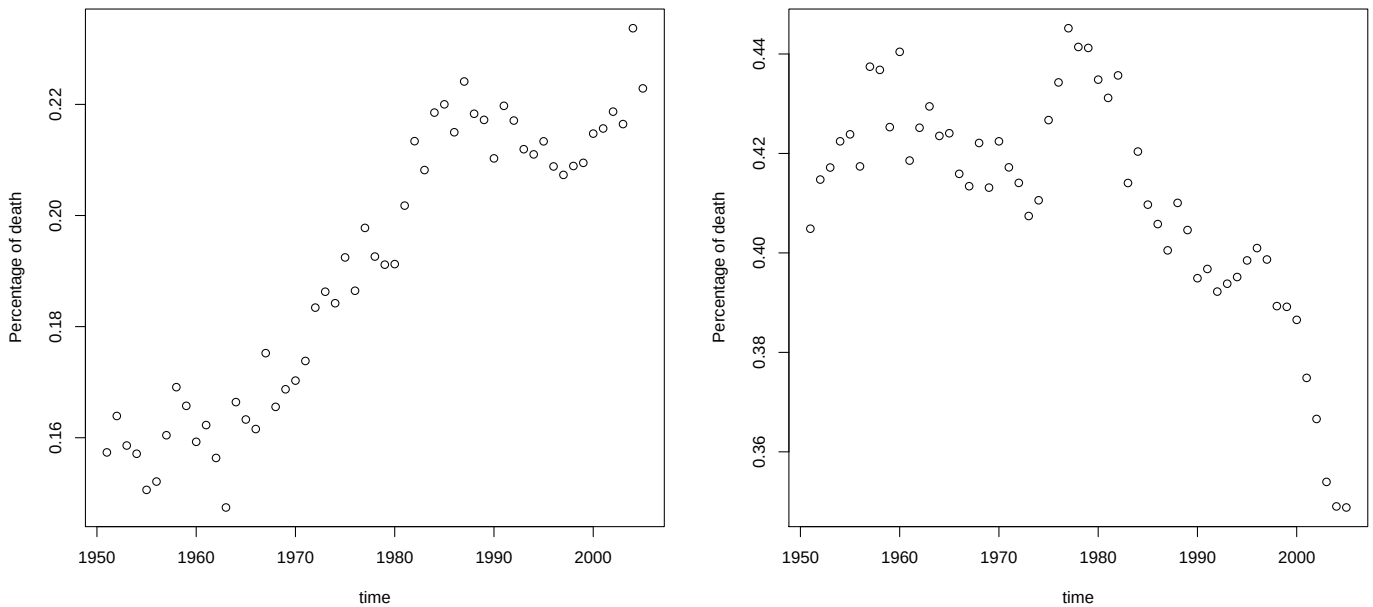
Why should we look at mortality by cause of death?



(a) Cancer (b) Circulatory system

Figure: Log-mortality over ages, females, Switzerland, 1955

Why should we look at mortality by cause of death?



(a) Cancer (b) Circulatory system

Figure: Percentage of deaths by cause, ages 65 and over, females, Switzerland

Why is nobody interested in causes of death?

Many problems arise:

- ...
- Causes of death = competing risks $\rightarrow$ a dependance exists.
- ...

Aim

What? Get a better understanding of mortality by cause of death
$\rightarrow$ especially the dependance

How? By developing a model which takes into account the main causes of death. The model will capture common features among countries and/or causes
$\rightarrow$ Improve the forecasting performance

Data

Countries:

- USA (1950 - 2007)
- Japan (1950 - 2009)
- France (1952-2008)
- E & W (1950 - 2009)
- Italy (1951 - 2003)
- Australia (1950 - 2004)
- Sweden (1951 - 2010)
- Switzerland (1951 - 2007)
- Singapore (1955 - 2009)
- Norway (1951 - 2009)

Causes of death:

- Diseases of the circulatory system
- Cancer
- Diseases of the respiratory system
- External causes (mainly: accidents)
- Infectious & parasitic diseases

The model

VAR(p)

$$\mathbf{y}_t = \mathbf{c} + \mathbf{\Phi}_1 \mathbf{y}_{t-1} + \mathbf{\Phi}_2 \mathbf{y}_{t-2} + \cdots + \mathbf{\Phi}_p \mathbf{y}_{t-p} + \epsilon_t$$

with

$$\begin{aligned} E(\epsilon_t) &= \mathbf{0} \\ E(\epsilon_t \epsilon_l) &= \begin{cases} \mathbf{\Omega} & \text{for } t = l \\ \mathbf{0} & \text{for } t \neq l \end{cases} \end{aligned}$$

$\rightarrow$ Need a stationary process

$\rightarrow$ What should we do with non-stationary variables?

Non-stationary variables

Possible answers

- Work on the first difference
- Use cointegrated relations
  - → Variables with common trends
  - → Long-run equilibrium relationships

Model selection

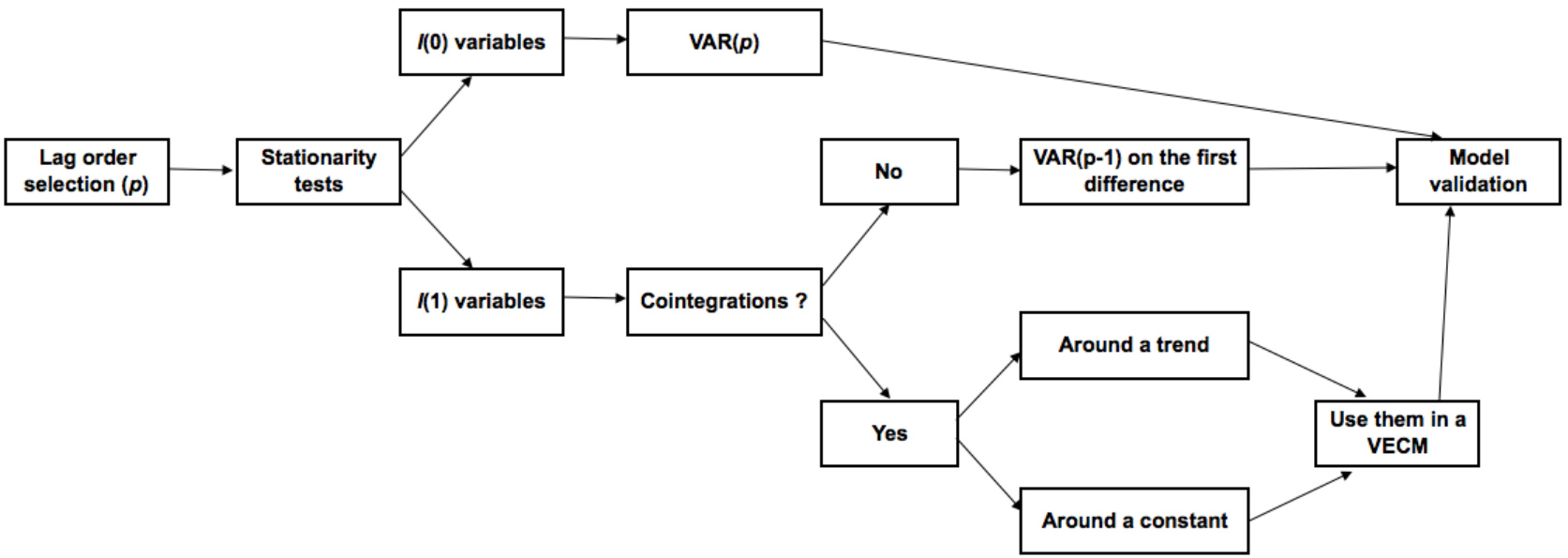


Figure: Steps to follow in a VECM analysis

Lag order

	Males	Females
USA	1	1
Japan	1	1
France	1	2
E&W	1	1
Italy	1	1
Australia	2	1
Sweden	1	1
Switzerland	1	2
Singapore	1	1
Norway	1	1

Figure: Number of past values to take into consideration in a VECM analysis

Stationarity

	Males	Females
USA	All causes: UR	All causes: UR
Japan	All causes: UR	All causes: UR
France	All causes: UR	All causes: UR
E&W	All causes: UR	Respiratory: S Other causes: UR
Italy	Respiratory: S Other causes: UR	Respiratory: S Other causes: UR
Australia	All causes: UR	All causes: UR
Sweden	All causes: UR	All causes: UR
Switzerland	Respiratory: S Other causes: UR	All causes: UR
Singapore	I&P: S Other causes: UR	I&P: S Other causes: UR
Norway	Respiratory: S Other causes: UR	(Cancer), Respiratory: S Other causes: UR

Figure: Stationarity of the five main causes of death in 10 countries

Cointegrated relations

	Males	Females
USA	1	1
Japan	1	1
France	1	1
E&W	1	1
Italy	1	1
Australia	1	1
Sweden	1	1
Switzerland	2	2
Singapore	2	2
Norway	2	1

Figure: Number of cointegrated relations among the five main causes of death in 10 countries

Cointegrated relations

	Males	Females
USA	Trend	Trend
Japan	Trend	Trend
France	Trend	Trend
E&W	Trend	Trend
Italy	Trend	Trend
Australia	Trend	Trend
Sweden	Trend	Trend
Switzerland	Trend	No trend
Singapore	Trend	Trend
Norway	Trend	No trend

Figure: Should we include a trend in the cointegrating relations?

Cointegrated relations

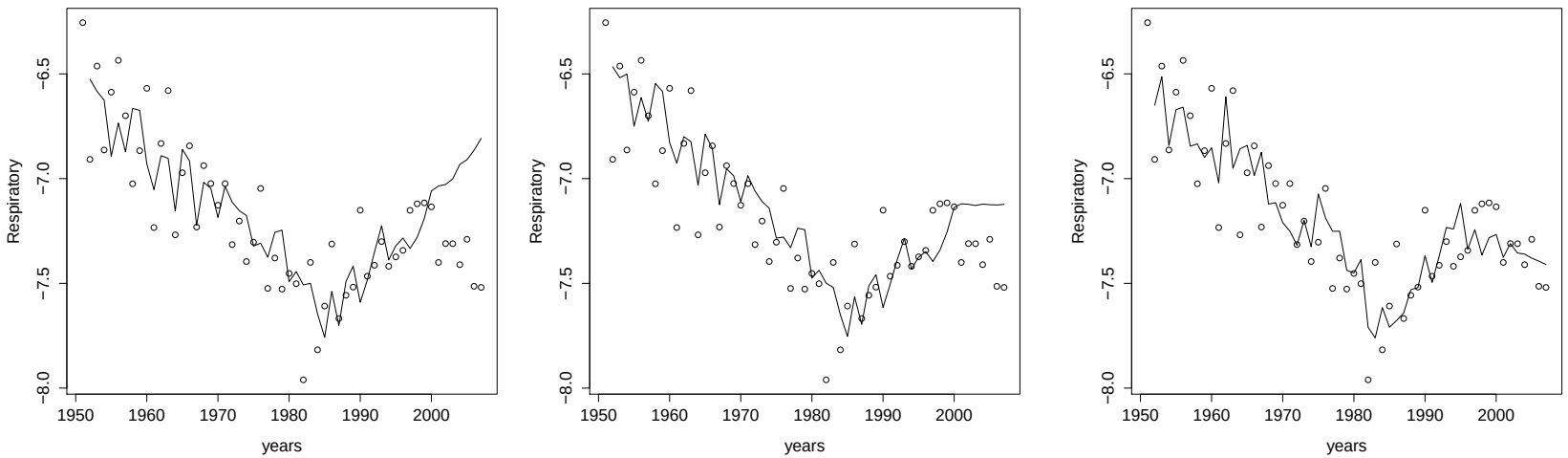
		Males					Females				
		Circulatory system	Cancer	Respiratory system	External causes	I&P	Circulatory system	Cancer	Respiratory system	External causes	I&P
USA	**1st**	10.69456	5.59757	-8.89295	-7.67594	-1.27201	-12.33722	21.23366	3.86850	9.11957	-2.44770
Japan	**1st**	7.88759	-18.19221	7.83418	1.91276	-5.31981	-7.04107	34.80843	-6.29159	-2.51984	2.79203
France	**1st**	3.61051	5.58839	9.92957	-8.57915	-9.35086	1.54265	-126.96155	13.87866	19.19450	-3.03437
E&W	**1st**	17.07335	16.15817	1.15538	19.21024	9.13450	23.66970	-41.64988	-8.28387	-0.97493	0.10822
Italy	**1st**	-4.65727	-15.23929	13.68104	10.70673	-2.51913	2.06643	7.74975	-11.33527	-0.84202	2.69301
Australia	**1st**	18.19951	-29.56920	-21.50232	10.31447	-1.79548	-6.94275	-13.93179	-9.76149	7.57787	3.14756
Sweden	**1st**	-10.08487	-3.35171	-5.54287	-6.00030	-7.55106	-23.67052	-3.53851	-2.54732	-1.99079	-4.38972
Switzerland	**1st**	-5.28092	9.94069	7.37395	-6.42822	-1.45477	-10.32052	15.33513	-1.79622	9.36036	1.06028
	2nd	-20.49919	28.31404	1.66893	5.96531	-1.34666	4.71468	-15.33328	-7.44481	0.59118	4.88562
Singapore	**1st**	15.18999	-8.63029	-5.62373	-4.32227	3.30489	-2.56778	8.37707	-4.68442	4.53691	-3.85025
	2nd	-6.06163	-0.48202	5.27013	-3.58569	3.53102	-12.32402	9.30955	4.75597	4.87606	0.64257
Norway	**1st**	12.75705	-8.84768	-13.00349	-7.90789	2.36239	10.63062	-16.28402	-11.00310	-9.68837	-0.54553
	2nd	15.80298	-4.81822	2.46397	10.22583	7.74822	-	-	-	-	-

Figure: Long-run equilibrium relationships

Next step

Can we use these models for forecasting?

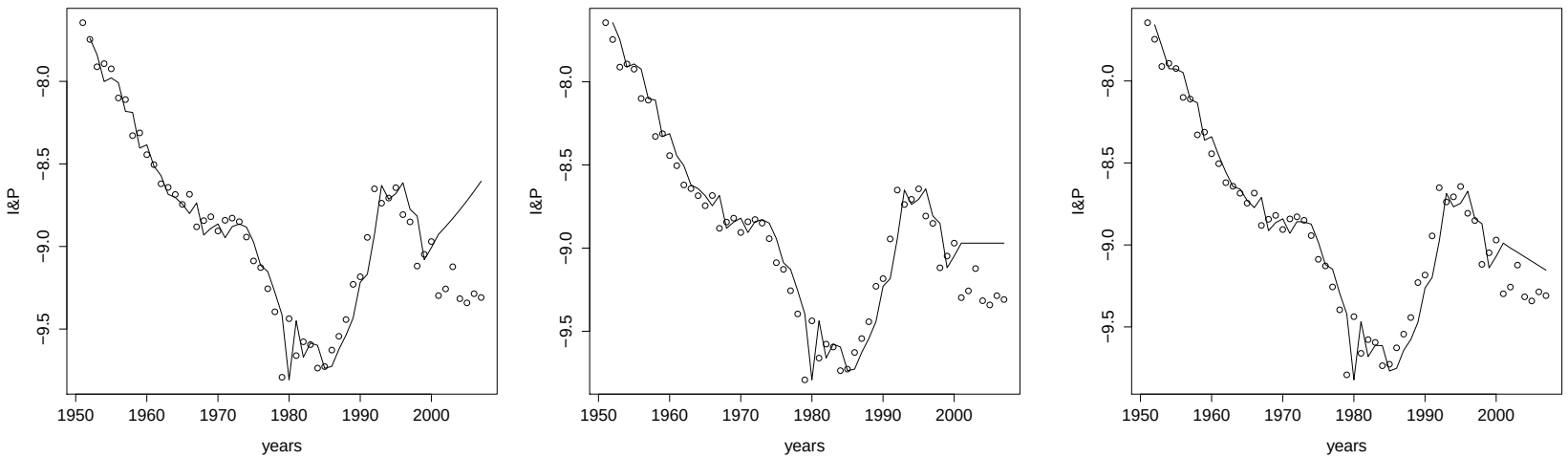
Respiratory diseases



(a) ARIMA (b) ARIMA, no trend (c) VECM

Figure: Death rate forecasts, females, Switzerland

Infectious and parasitic diseases



(a) ARIMA (b) ARIMA, no trend (c) VECM

Figure: Death rate forecasts, females, Switzerland

Summary statistics

	VECM	ARIMA
2001	8.78%	15.62%
2002	7.47%	12.96%
2003	4.77%	9.35%
2004	8.62%	17.82%
2005	10.75%	17.26%
2006	10.32%	20.55%
2007	9.13%	21.14%

(a) Females, Switzerland

	VECM	ARIMA
2001	4.81%	4.74%
2002	5.86%	5.87%
2003	7.99%	7.55%
2004	7.58%	7.68%
2005	5.63%	5.45%
2006	5.40%	5.96%
2007	5.28%	7.20%
2008	10.04%	13.33%
2009	15.79%	21.55%

(b) Males, E&W

Figure: Mean absolute percentage error

Concluding remarks

VECM capture key features of the cause-of-death data

- Improves our understanding of the dependance between these competing risks
  - → Long-run equilibrium relationships exist
- Forecasted cause-of-death mortality rates seem more reliable
  - → Becomes superior as the forecast interval is extended
  - → Data over longer time horizon are required

→ Models incorporating this information need to be further developed

Thank you for your attention!