

# Räumliche Epidemiologie und ihre methodischen Fallgruben

Thomas Waldhör

Abteilung für Epidemiologie  
Zentrum für Public Health  
Medizinische Universität Wien

Beispiele

Fallgruben

# Beispiele

## Type I diabetes

- BMI
- Trend

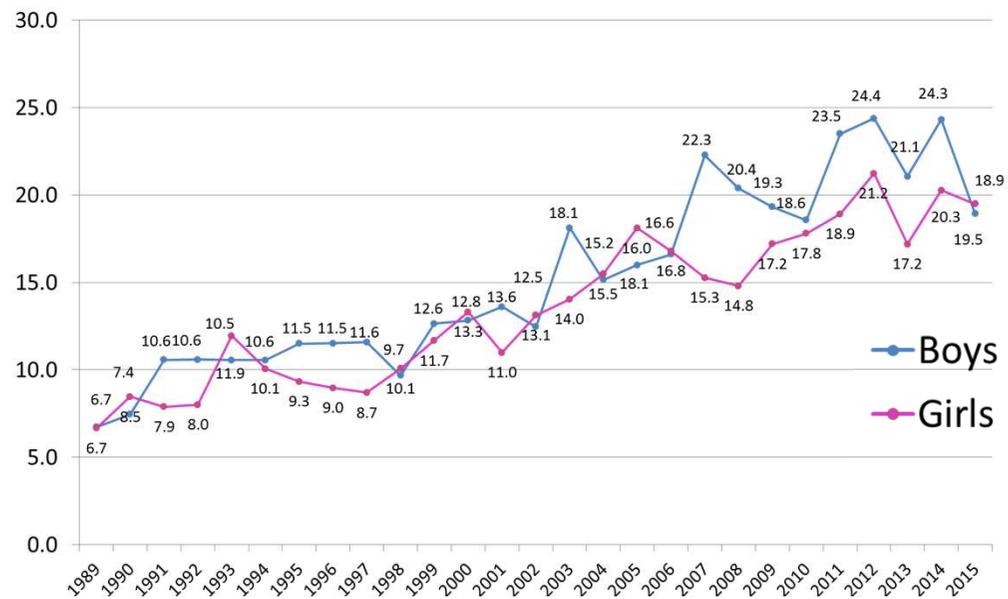
## Säuglingssterblichkeit

## Obduktion

## Methodische Arbeiten

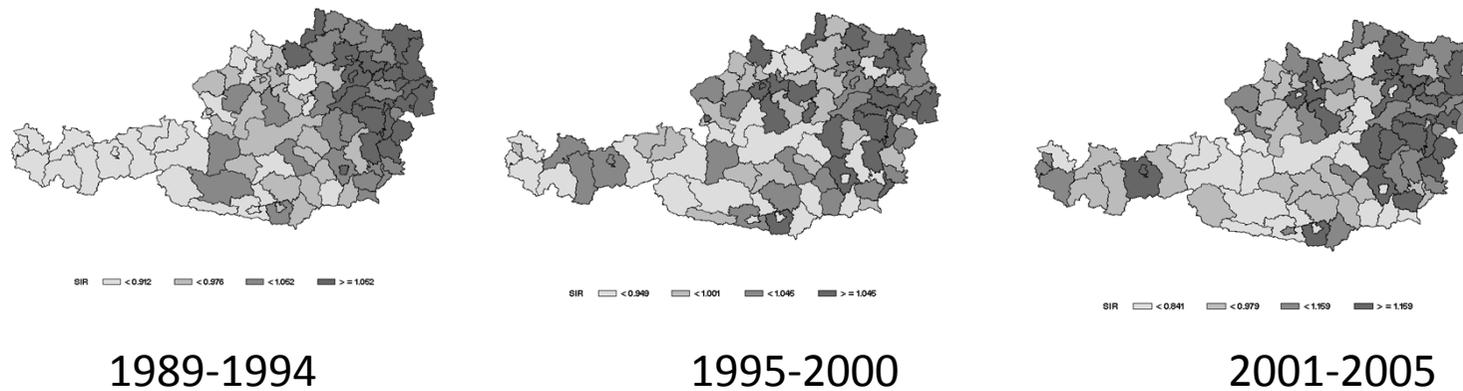
- Equivalence of rates
- (Type III error)

## Incidence of Type I Diabetes in children <15 yrs



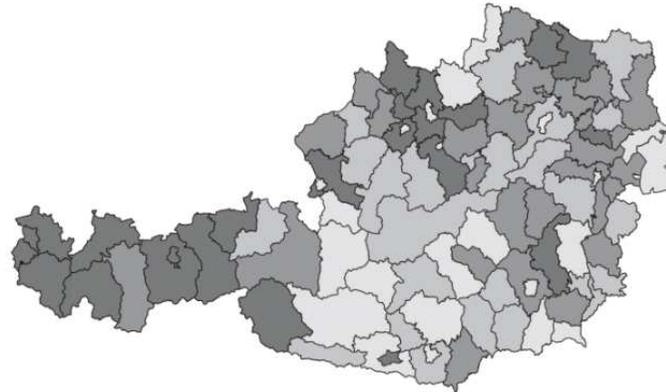
Rami-Merhar B, Hofer S, Fröhlich-Reiterer E, Fritsch M, Waldhör T, Group tA. Incidence trends of type 1 and type 2 diabetes in Austrian children < 15 years (1999–2015). Poster at the 42. ISPAD Meeting. Oct. 26.-29. 2016. Valcencia, Spain. Paediatric Diabetes, 2016: Suppl. 24(17)/(26.-29. Okt. 2016): 94-94.

## Changing geographical distribution of diabetes mellitus type 1 incidence in Austrian children 1989–2005

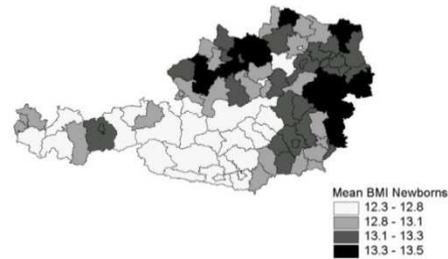


## Changing geographical distribution of diabetes mellitus type 1 incidence in Austrian children 1989–2005

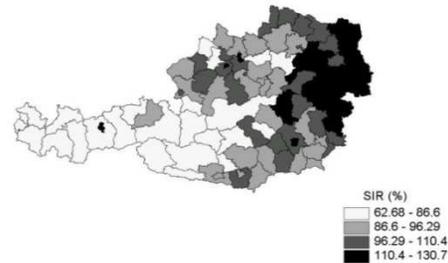
Estimated change in spatial distribution  
between 1989 to 2005



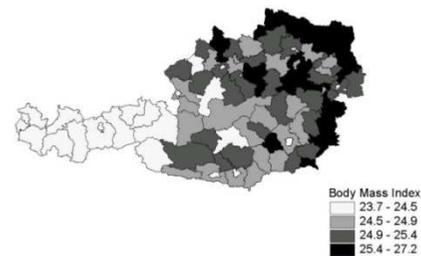
## Regional distribution of risk for childhood diabetes in Austria and possible association with body mass index



Mean BMI of newborns



Relative risk of Type I diabetes



Mean BMI of adults aged 20 yrs

## Analysis of the spatial distribution of infant mortality by cause of death in Austria in 1984 to 2006

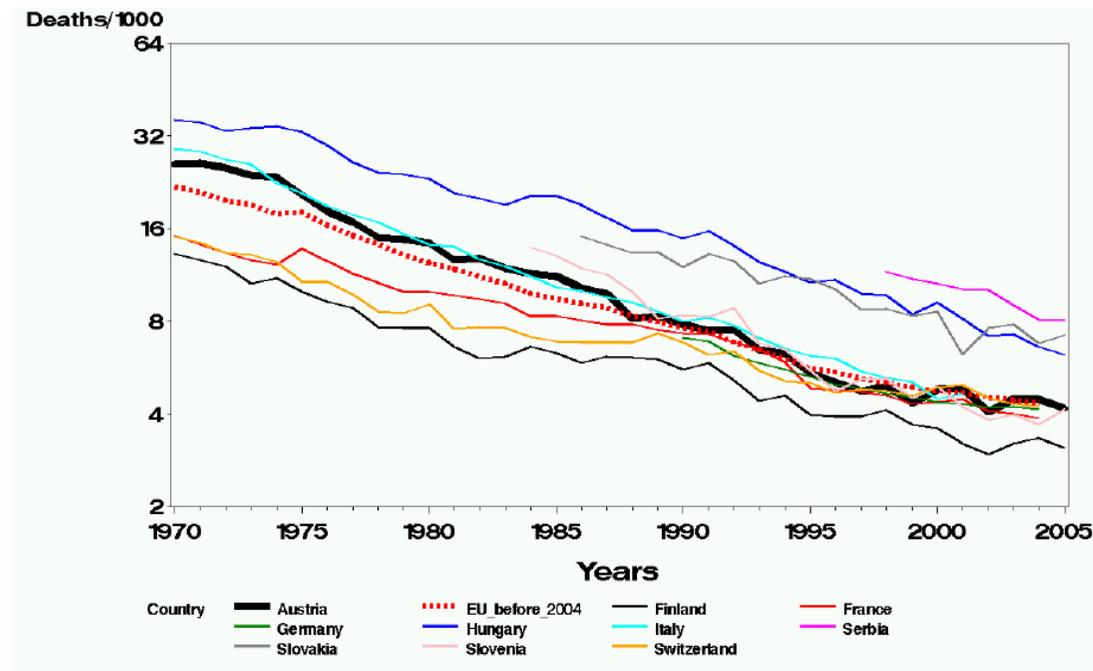
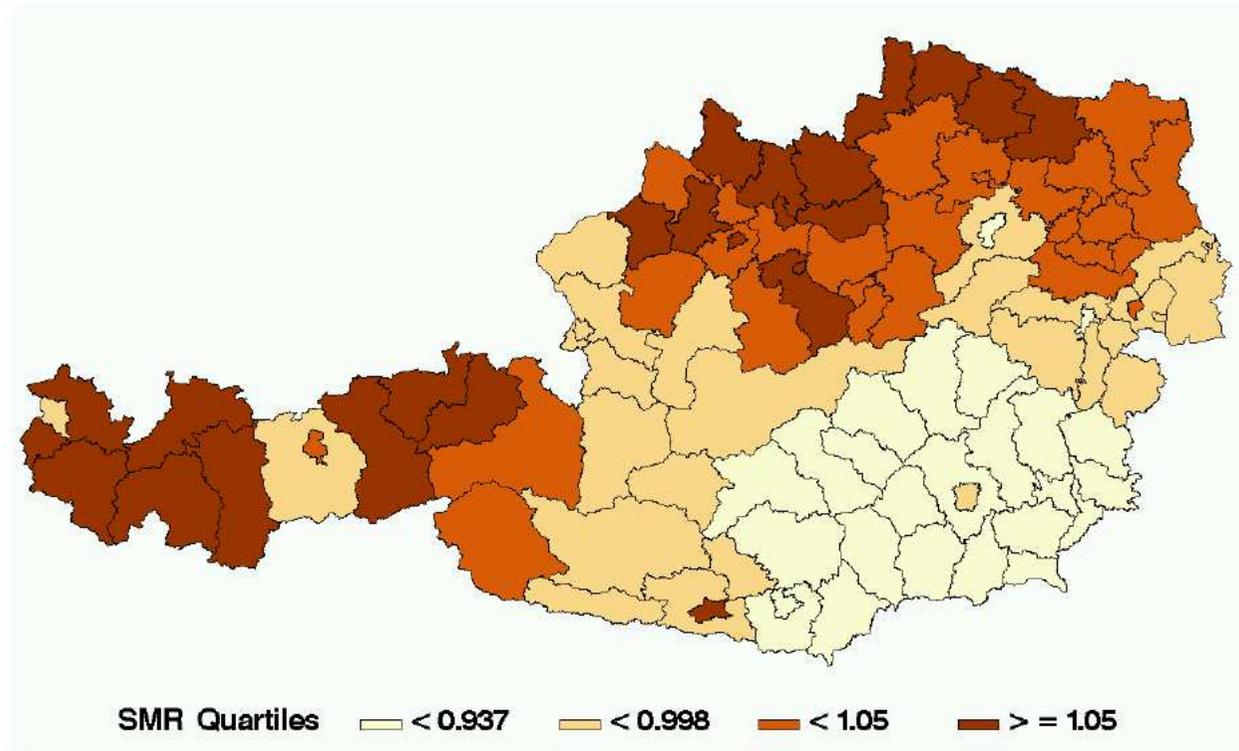


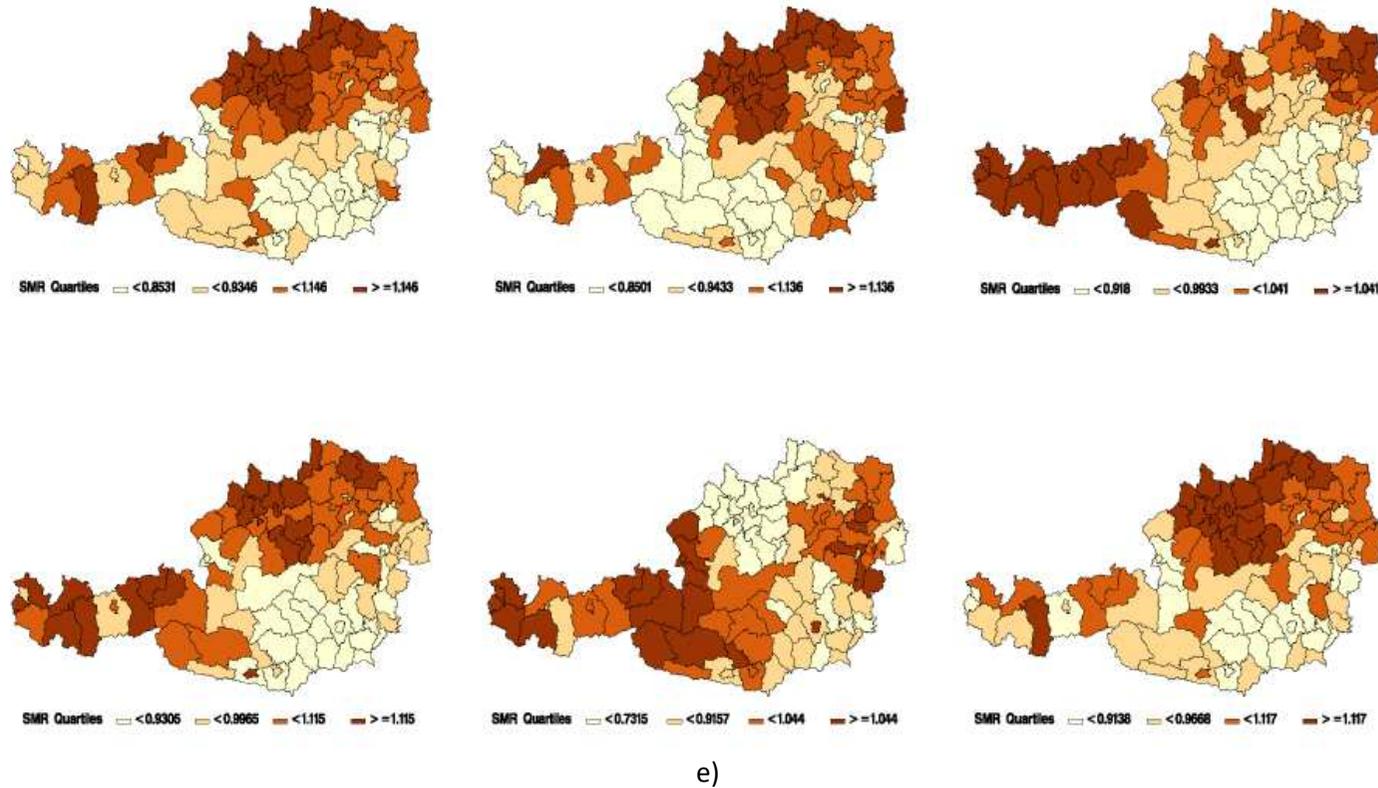
Figure 1  
Infant mortality rates per 1000 live births for some European countries.

## Standardized Mortality Ratios Infant Mortality 1984-2006



**Figure 2**  
The spatial distribution of SMR's for overall infant mortality.

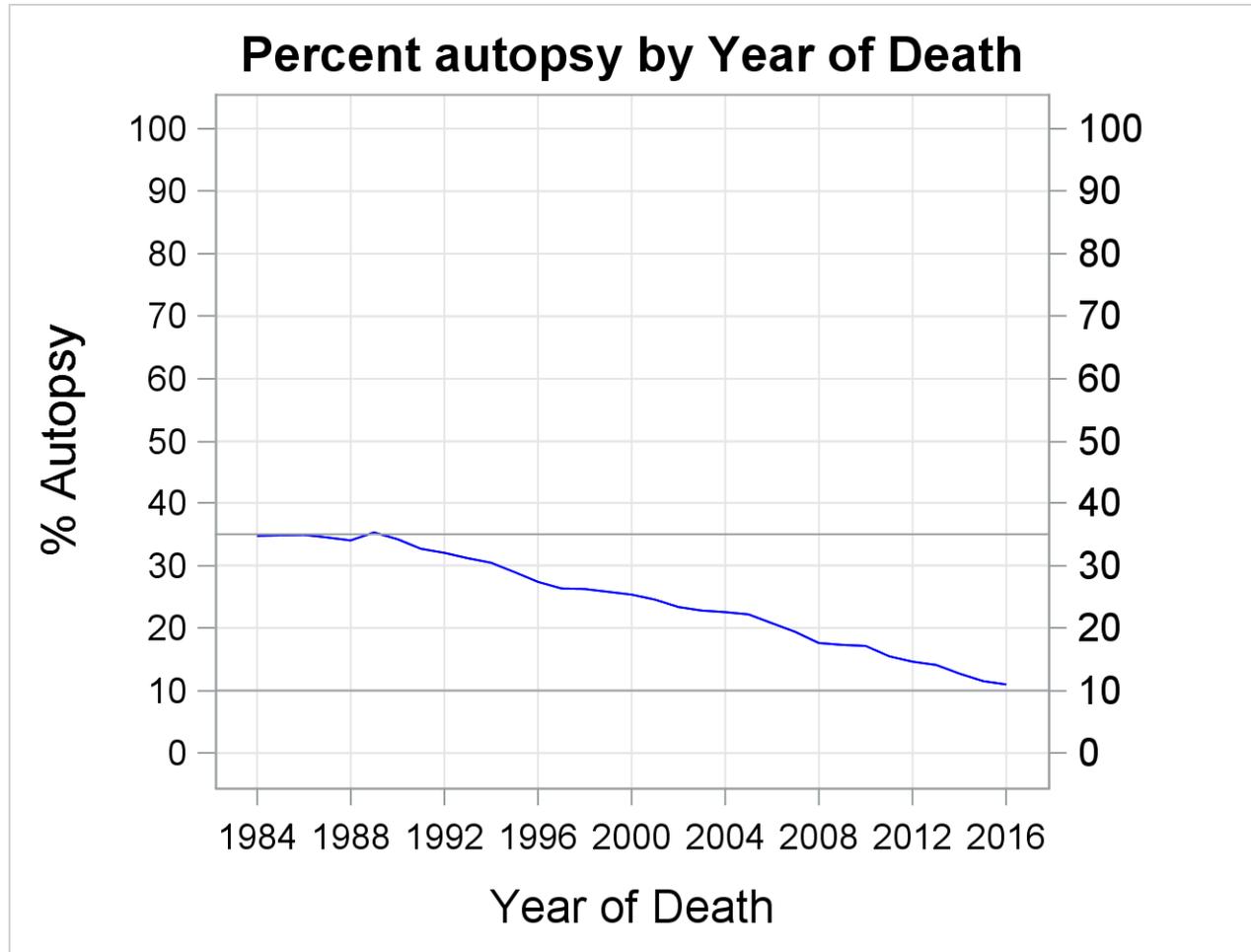
## Standardized Mortality Ratios Infant Mortality by Cause of Death



**Figure 3**

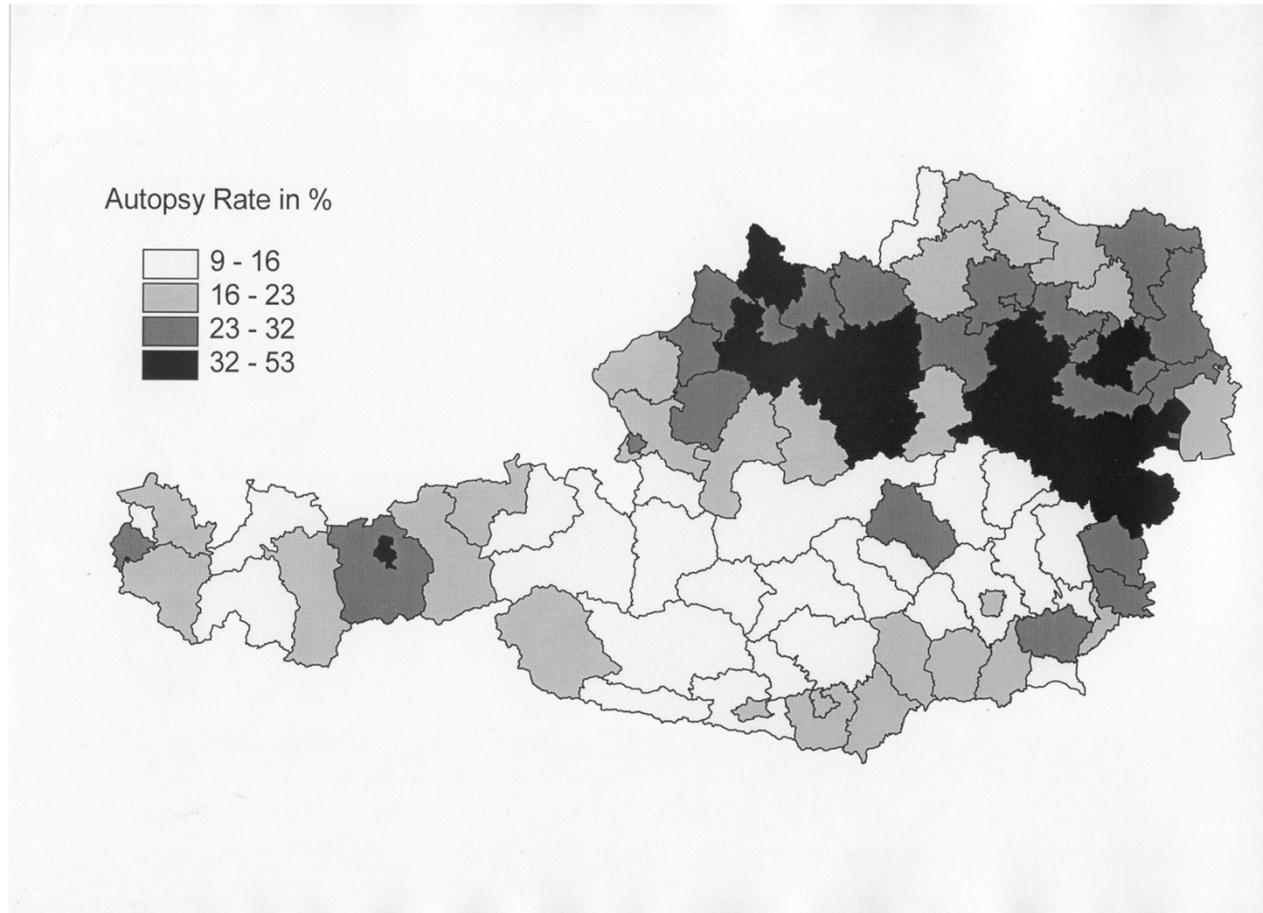
The spatial distribution of SMR's for infant mortality by groups of cause of death. a) "Infections, respiratory diseases"; b) "Peri-partal problems"; c) "Immaturity"; d) "Malformations"; e) "Sudden Infant Death Syndrome (SIDS)"; f) "All other".

# Autopsie



# Rates of postmortem examination in Austria

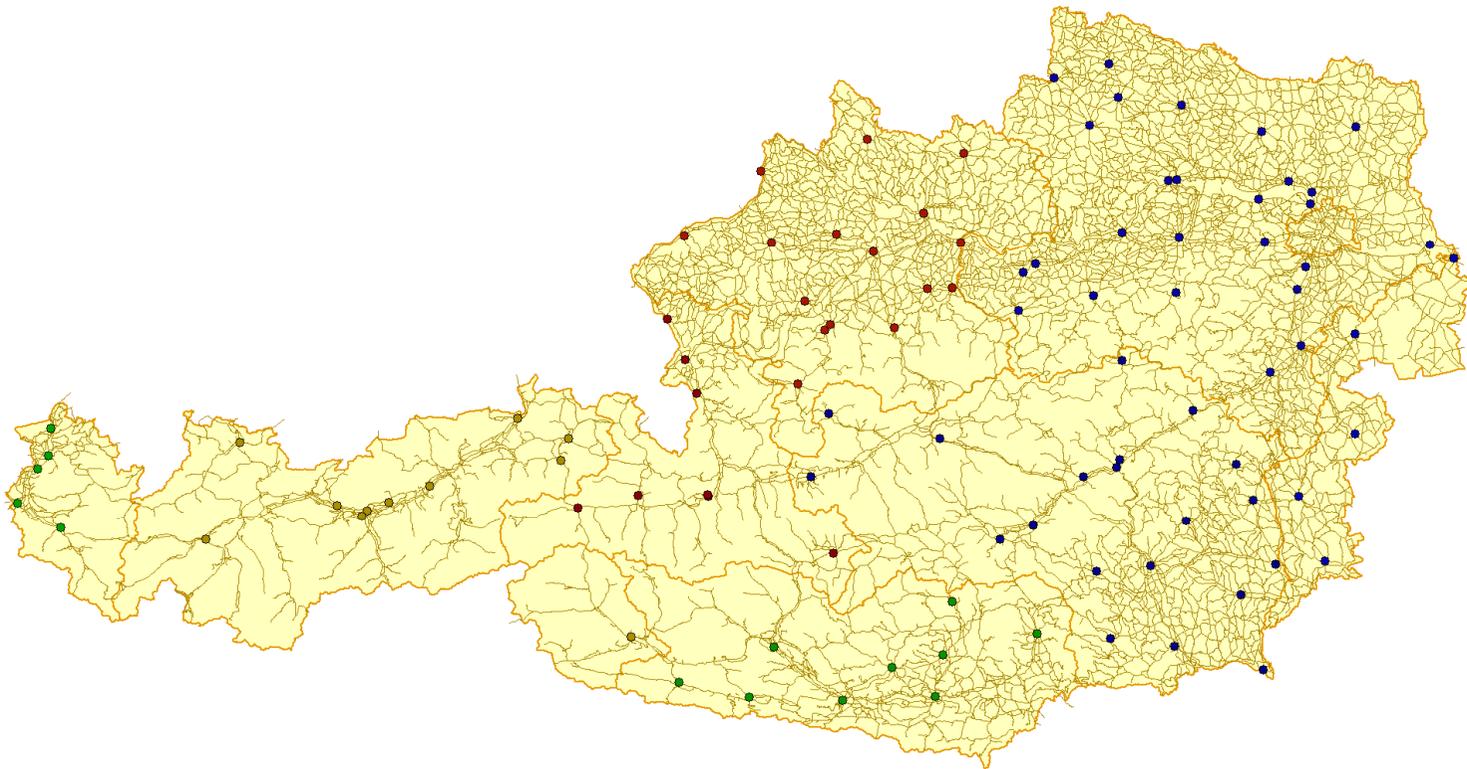
## The effect of distance between location of death and site of examination



T. Waldhoer<sup>a,\*</sup>, A. Berzlanovich<sup>b</sup>, C. Vutuc<sup>a</sup>, G. Haidinger<sup>a</sup>

Journal of Clinical Epidemiology 56 (2003) 891–895

## Road map and locations of autopsy facilities



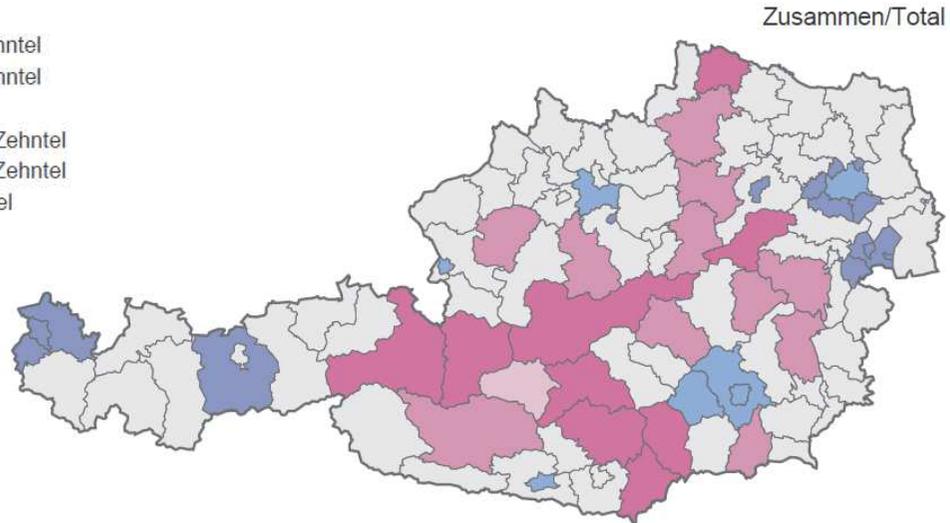
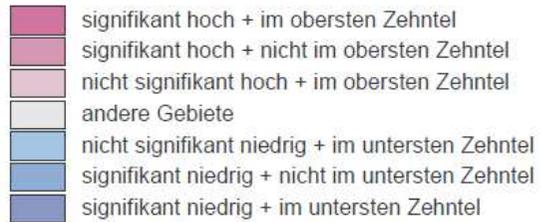
the more distant the autopsy facility is,  
the less likely is autopsy

# Methodische Arbeiten

Verletzungen und Vergiftungen 1998/2004  
Injury and poisoning 1998/2004

ICD-Code: V01-Y89

Standardisierte Sterbeziffer



Combining difference and equivalence test  
results in spatial maps

Zeige

vom Durchschnitt signifikant **verschiedene** Gebiete  
+  
dem Durchschnitt signifikant **äquivalente** Gebiete

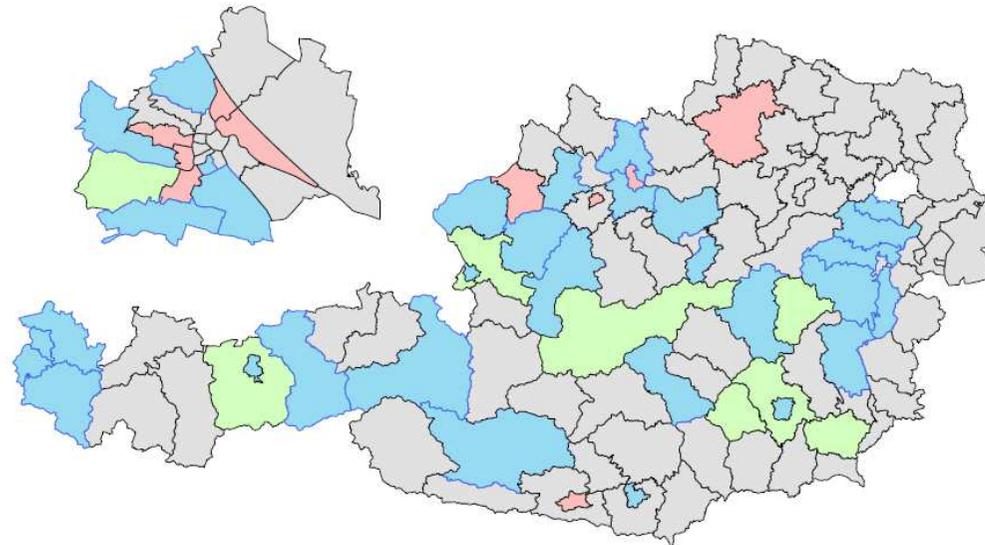


Figure 5 Difference/equivalence test results for infant mortality in Austria 1984-2007. Different colours refer to different results of a difference/equivalence test combination ("4 combined scenarios").

### Legend

#### combined test results

equivalent

not equivalent and significantly smaller

not equivalent and not significantly different

not equivalent and significantly larger

no decision possible at all



## Relevance of the type III error in epidemiological maps

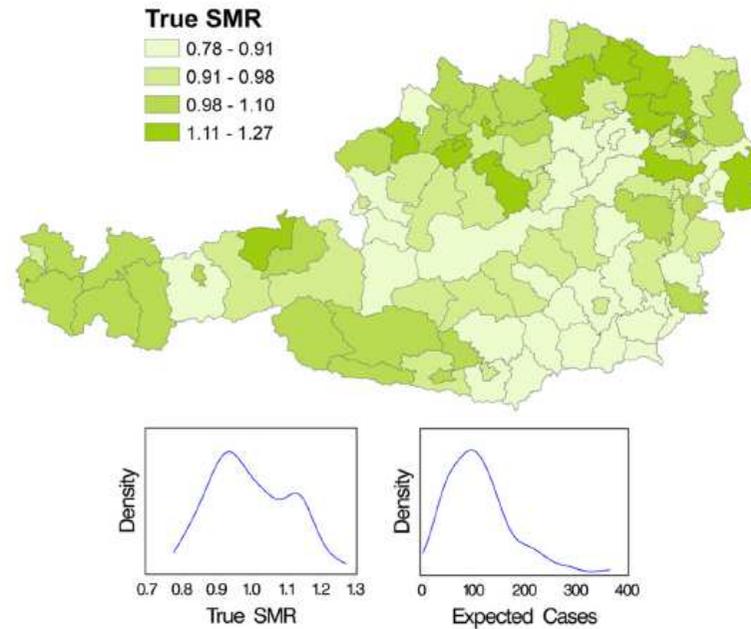
Fragestellung:

Falls ein signifikanter Unterschied beobachtet wird,  
wie groß ist Wahrscheinlichkeit, dass der wahre Effekt “verkehrt” ist?

z.B.: beobachtet: signifikant Relatives Risiko  $RR = 1.5$

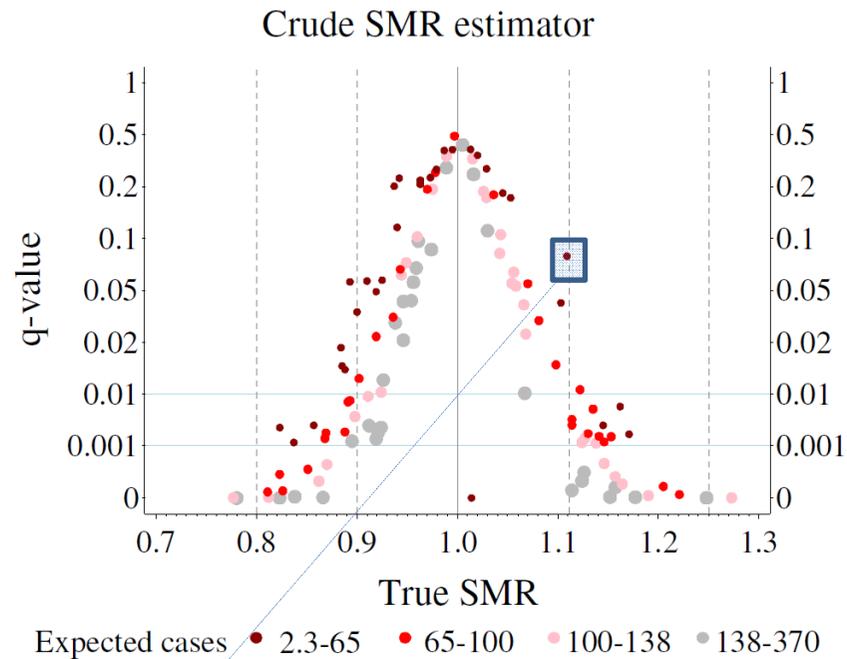
Wie wahrscheinlich, dass “wahres”  $RR < 1$ ?

# Relevance of the type III error in epidemiological maps



**Figure 1** Austrian infant mortality from 1984 to 2008 at the level of 121 administrative districts. SMRs have been estimated by the empirical Bayes procedure and have been grouped into quartiles. The expression "true SMR" refers to the use of these results as calibration reference points for type III error calculations in the example. Kernel estimates of the distributions of SMRs and expected cases are also shown.

## Observing significant result though true effect the opposite



For small units: If true RR~1.1  $\rightarrow$  Prob(observed RR<1)~ 0.1

# Fallgruben

**Tiefe Fallgruben**

**Technische Fallgruben**

## **Tiefe Fallgruben**

Handeln dann Denken  
Datenqualität

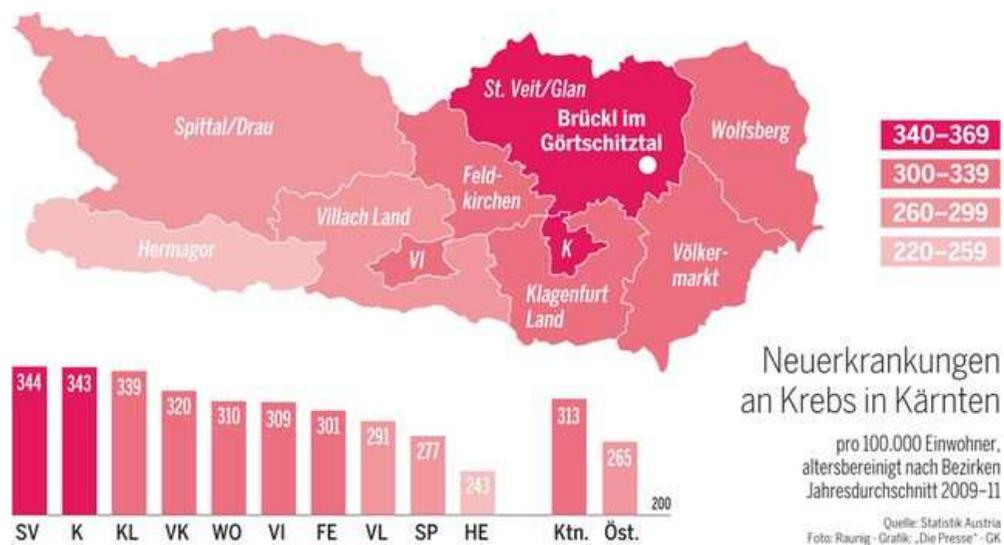
## **Technische Fallgruben**

Variabilität  
P-value fishing  
Räumliche Auflösung  
Größe der Region  
Farbwahl  
Ökologische Regression  
Auflösung

Fallgrube

„Handeln dann Denken“

## „Zusammenhang“ Krebsinzidenz im Görtschitztal und HCB-Belastung durch thermische Behandlung



März 2014 erhöhte HCB-Werte in einer Ricotta-Probe aus dem Görtschitztal

Zusammenhang Krebsrisiko mit Entsorgung des mit HCB-belastetem Blaukalks  
*Kalkdeponie Brückl I/liv*



pro 100.000 Einwohner,  
altersbereinigt nach Bezirken  
Jahresdurchschnitt 2009-11

Inzidenz



pro 100.000 Einwohner,  
altersbereinigt nach Bezirken  
Jahresdurchschnitt 2009-11

Inzidenz

Thermische Behandlung des mit HCB-belastetem Blaukalks ab 2012



„Beschreibung der Krebsinzidenzraten im Görtischtal im Vergleich zum Bundesland Kärnten „  
Masterarbeit, MPH, Katja Müller, 2017

## Fallgrube

Diagnoseungenauigkeit  
Vollzähligkeit von Daten

## Problem:

Es wird nicht die wahre räumliche Risikoverteilung abgebildet sondern eine verzerrte Verteilung bedingt durch mangelnde Datenqualität

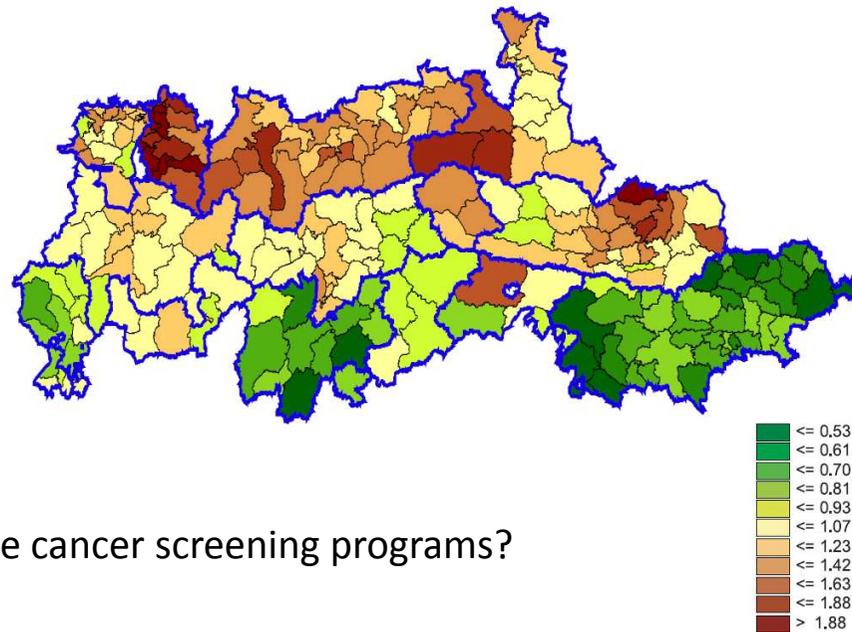
Nicky Best

Imperial College London, Faculty of Medicine, School of Public Health

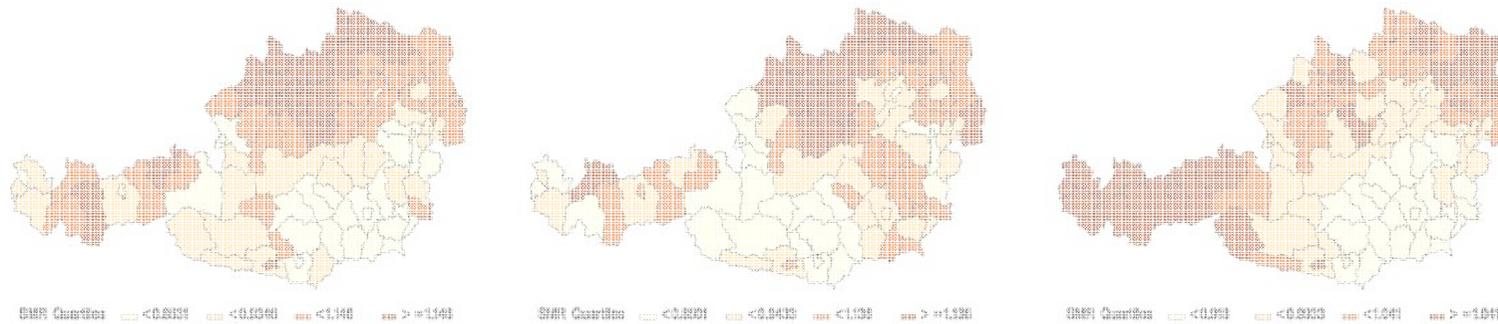
„...not showing spatial distribution of true risk  
but spatial distribution of data quality ....“

Often one does not model incidence rates but quality of diagnosis

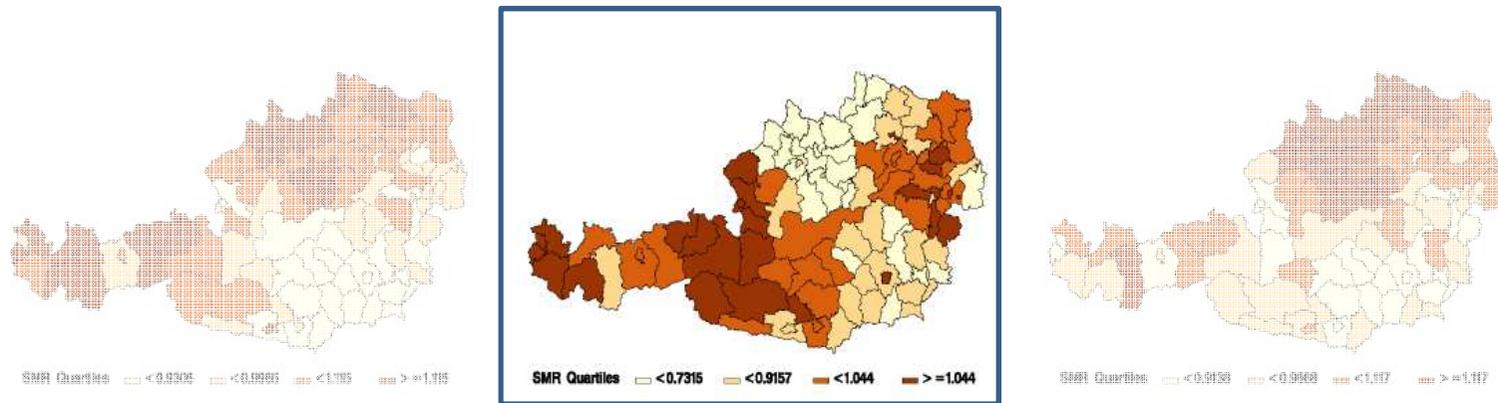
**Fig. 35: Prostate – Incidence – Smoothed Map**



Effect of prostate cancer screening programs?



## Plötzlicher Kindstod (SIDS) 1984-2006



The spatial distribution of SMR's for infant mortality by groups of cause of death. a) "Infections, respiratory diseases"; b) "Peripartal problems"; c) "Immaturity"; d) "Malformations"; e) "Sudden Infant Death Syndrom (SIDS)"; f) "All other".

### Analysis of the spatial distribution of infant mortality by cause of death in Austria in 1984 to 2006

Thomas Waldhoer\*<sup>1</sup>, Martin Wald<sup>2</sup> and Harald Heinzl<sup>3</sup>

# Vollzähligkeit

completeness of case ascertainment

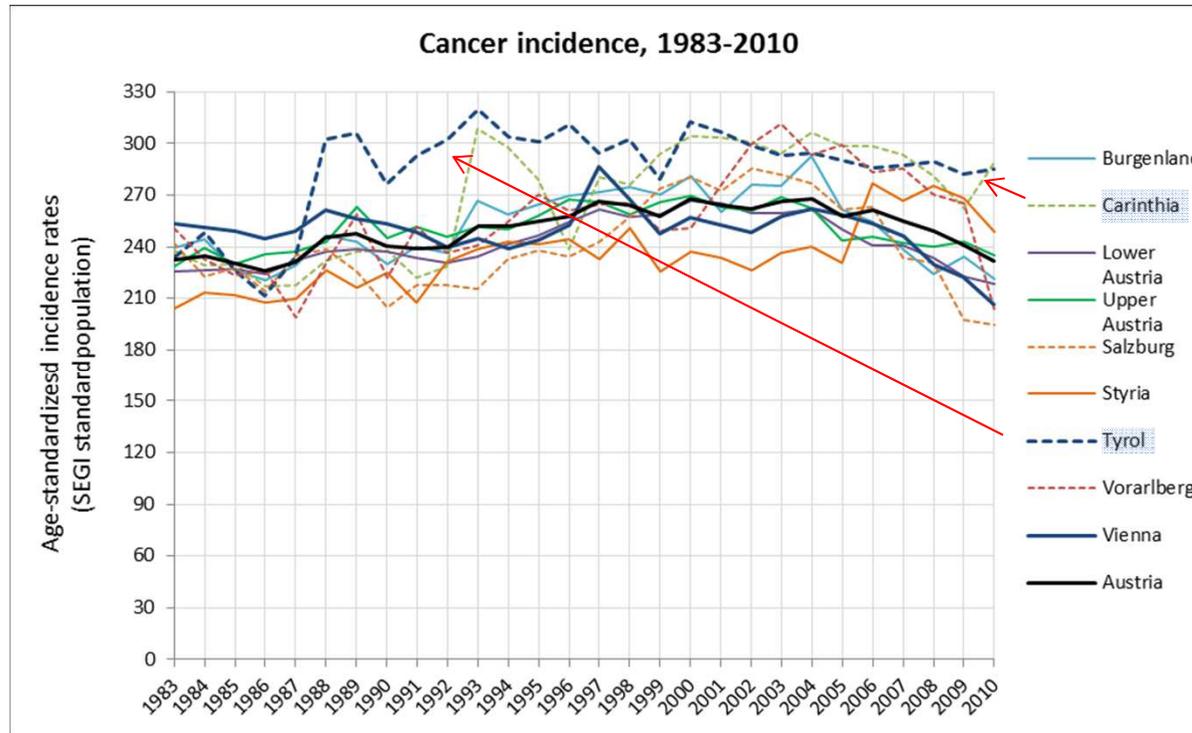
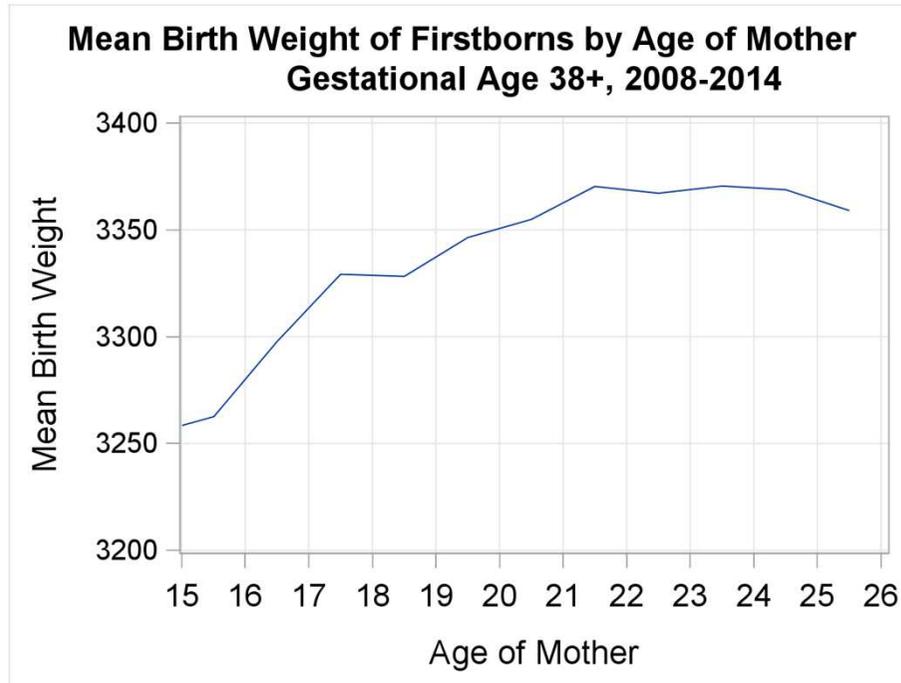


Figure 1: Age adjusted incidence rate for all malignant tumors, Austria and federal states, 1983-2010 (per 24.09.2012)

Evaluation of completeness estimation of Austrian cancer incidence data, PhD thesis, Monika Hackl, 2013

Technischen Fallgruben

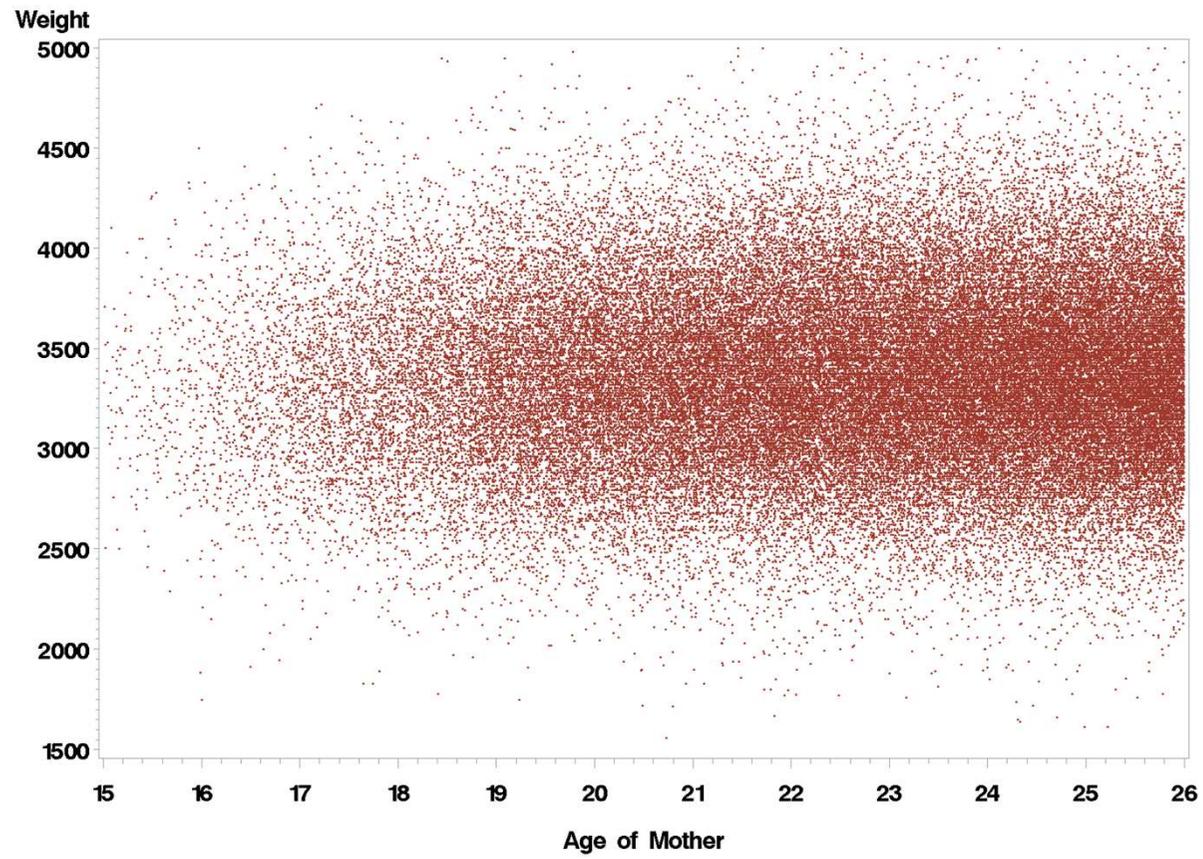
Ökologischer Regressionsbias



VIEWTABLE: Work.W			
	alterm	wohnbez	gebGew4
1	23.9	101	2630
2	22.3	101	2745
3	24.9	101	2885
4	23.7	101	2895
5	23	101	2965
6	20.7	101	2960
7	25.3	101	2960
8	23.9	101	3060
9	20.7	101	3085
10	25.6	101	3100
11	22.6	101	3100
12	25	101	3120
13	24.7	101	3180
14	25.2	101	3190

n ~ 79.000 **individual** observations

## Basierend auf individuellen Daten

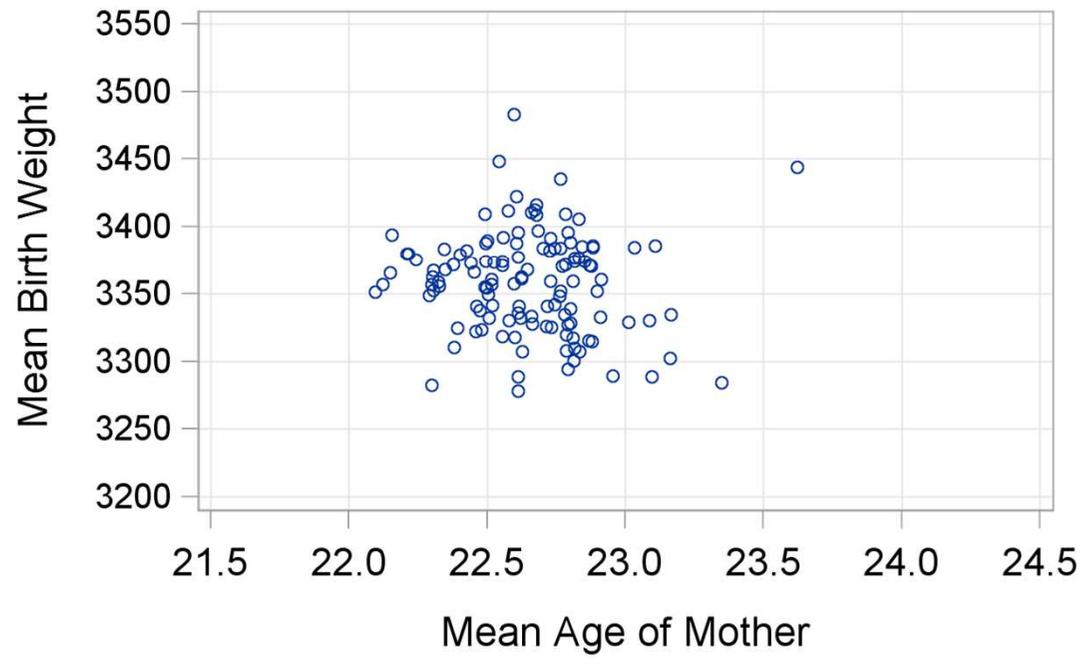


## Mean Birthweight and Mean Age by District

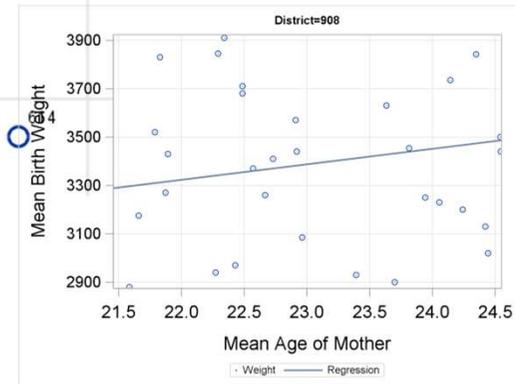
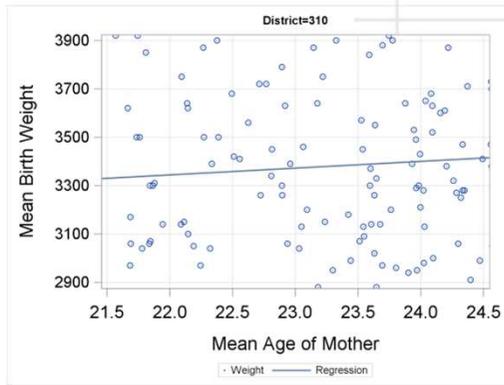
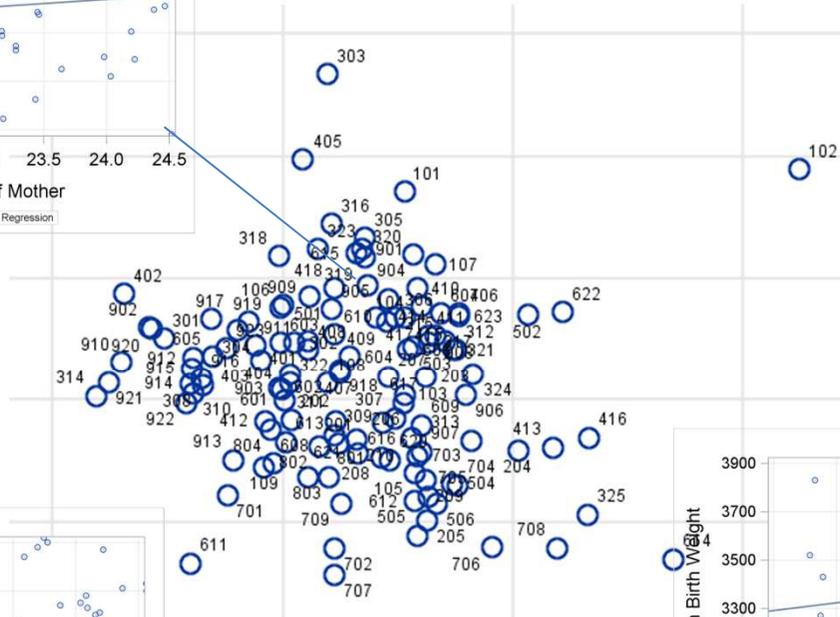
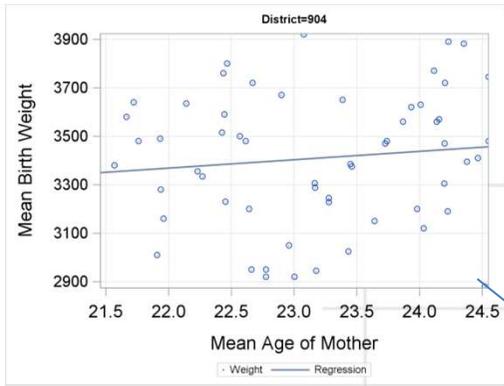
-> n=121 districts -> n=121 aggregated observations  
instead of n~79.000 individual observations

District	Mean Weight	Mean Age
101	3435.4	22.8
102	3444.3	23.6
103	3351.6	22.8
104	3397.1	22.7
105	3319.9	22.8
106	3387.7	22.5
...	...	...
...	...	...
919	3382.2	22.4
920	3379.6	22.2
921	3356.9	22.1
922	3348.5	22.3
923	3371.5	22.4

**Mean Birth Weight of Firstborns by Mean Age of Mother  
Gestational Age 38+, 2008-2014  
By Political District**



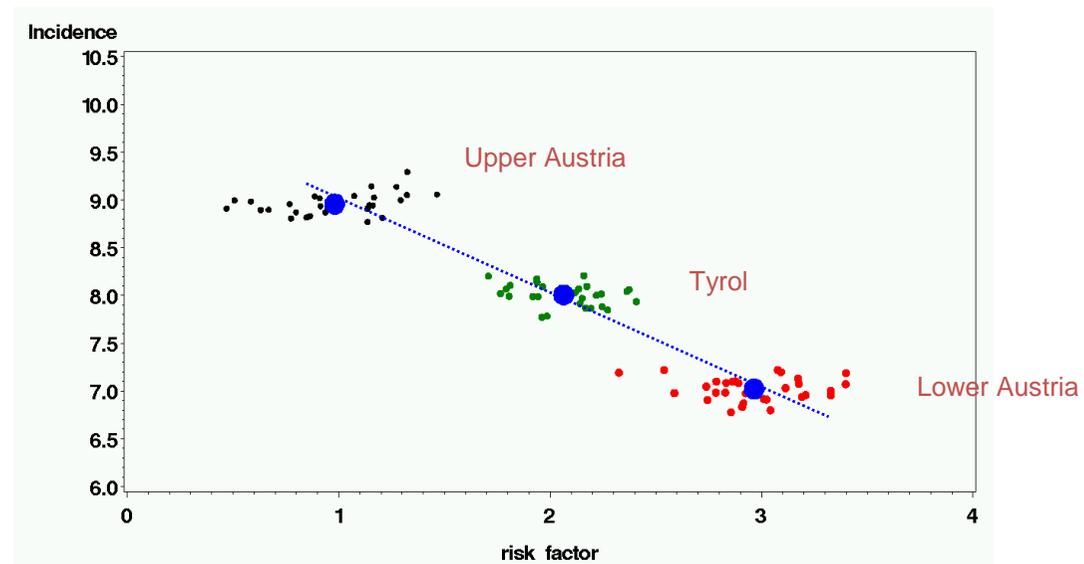
Pearson Correlation Coefficients Prob >  r  under H0: Rho=0	
	Mean Age
<b>mgew</b>	-0.09398
Mean Weight	0.2972



# Ecological bias

association

on individual level (municipalities)  $\neq$  on aggregated level (province)

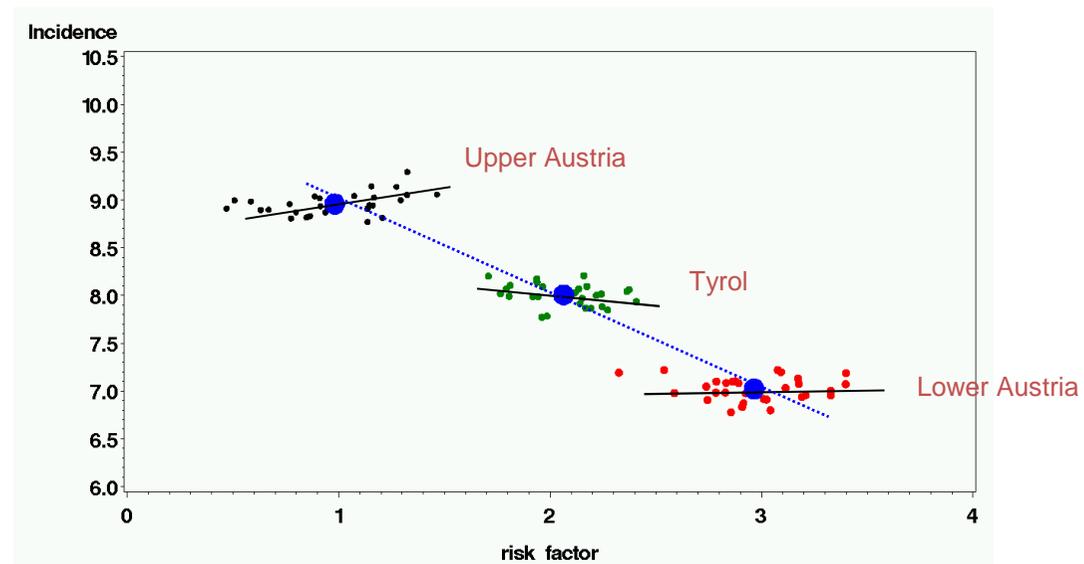


- ..... association based on aggregated data (province)
- association based on individual data (municipalities)

# Ecological bias

association

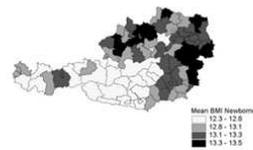
on individual level (municipalities)  $\neq$  on aggregated level (province)



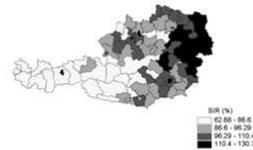
- ..... association based on aggregated data (province)
- association based on individual data (municipalities)

# Ökologischer Regressionsbias

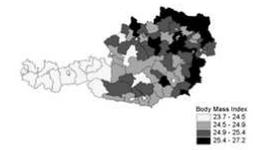
Analysen auf aggregiertem Datenniveau können zu unterschiedlichen Schlüssen führen im Vergleich zu Analysen auf individuellem Niveau



Mean BMI of newborns

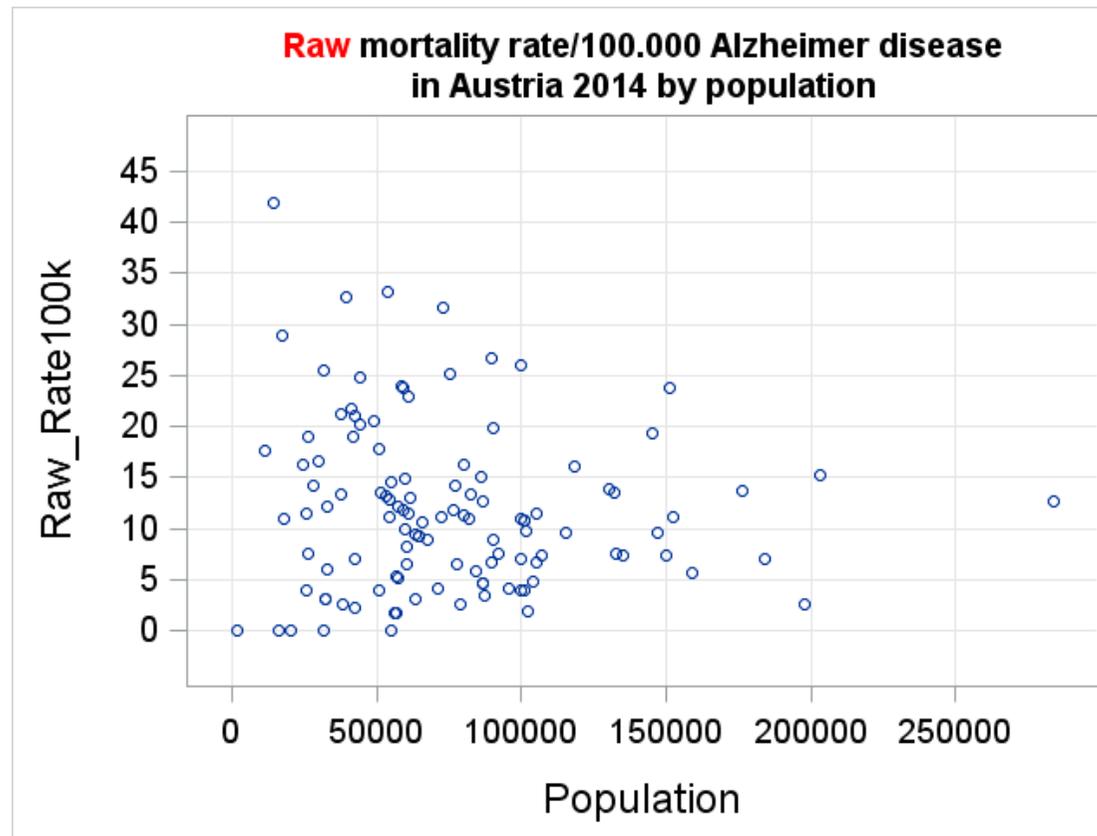


Relative risk of Type I diabetes



Mean BMI of adults aged 20 yrs

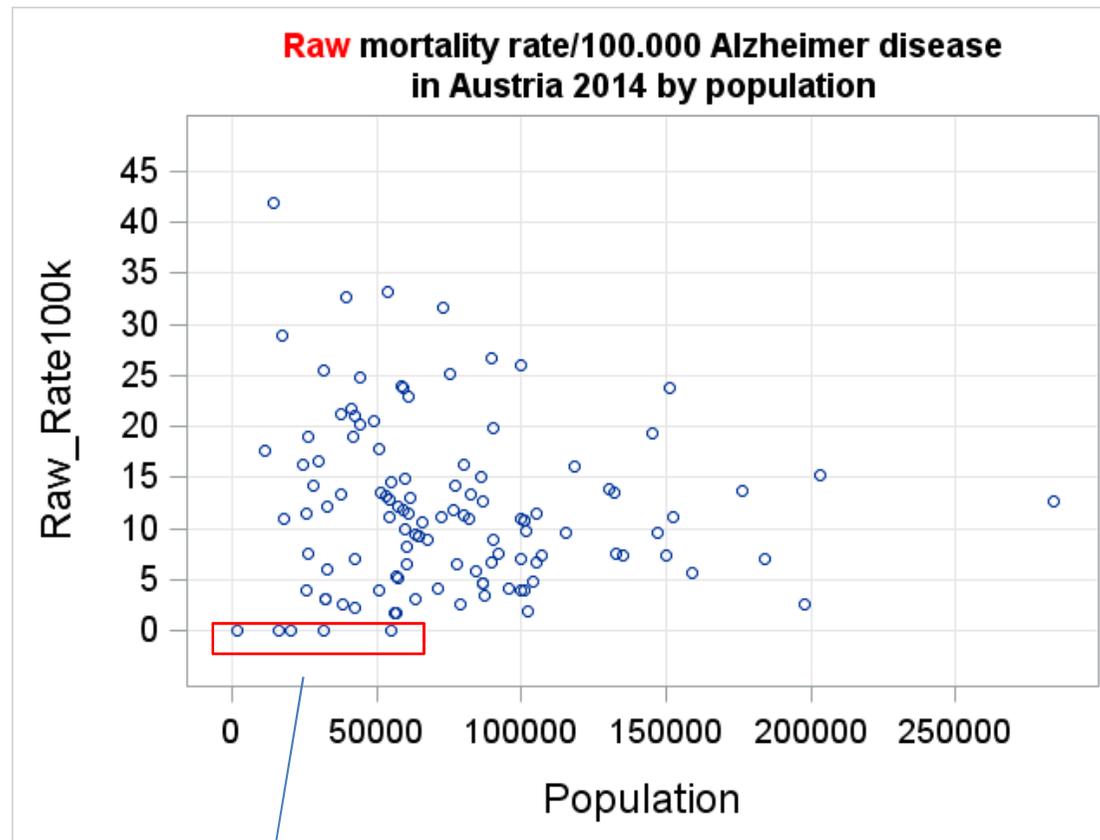
# Variabilität von Raten



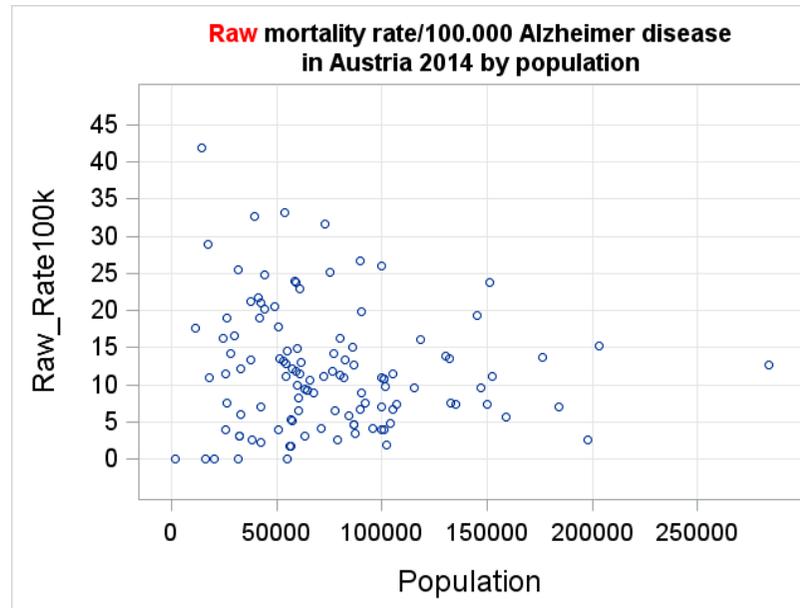
raw rates = # cases/population time\*

e.g. raw rate = 12 / 100000 yrs

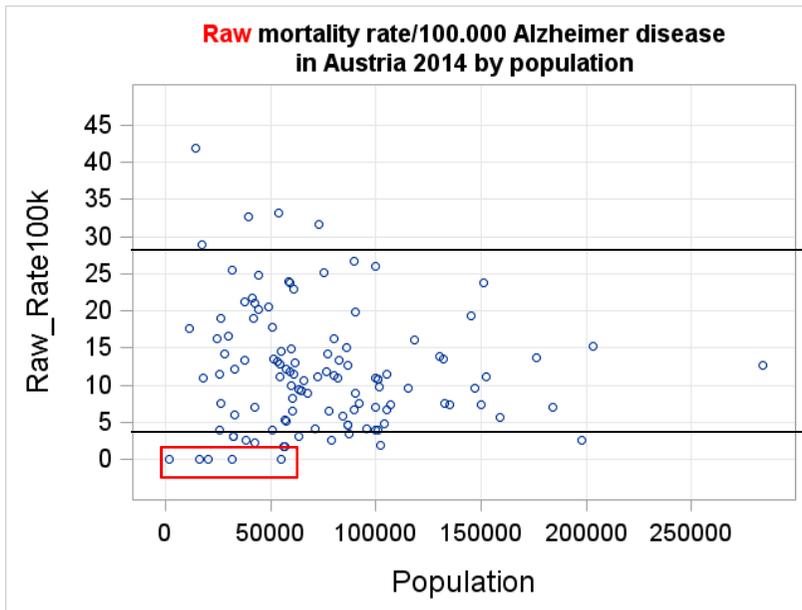
\*not standardized by age/sex



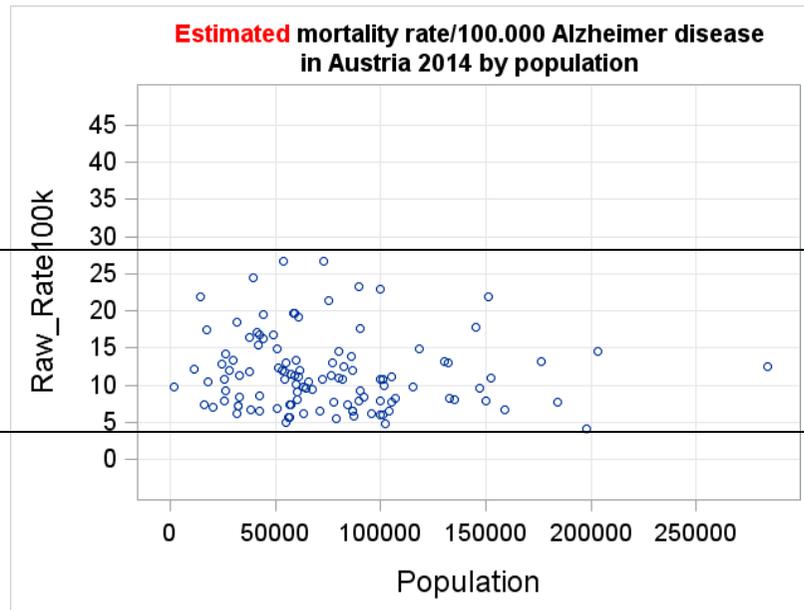
Zero rates of Alzheimer disease



Berechnung von „shrinkage rates“, welche die Variabilität  $\sigma^2$  der unterschiedlichen Bevölkerungsgrößen miteinbeziehen



raw rates



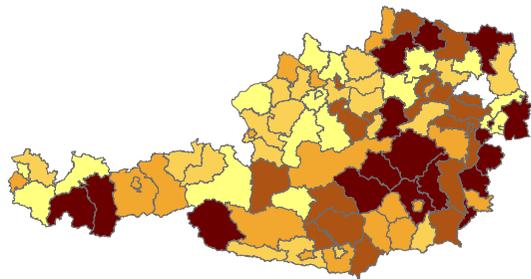
estimated rates

# Alzheimer disease 2014

raw rates

-

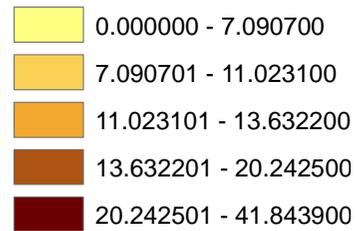
shrinkage rates



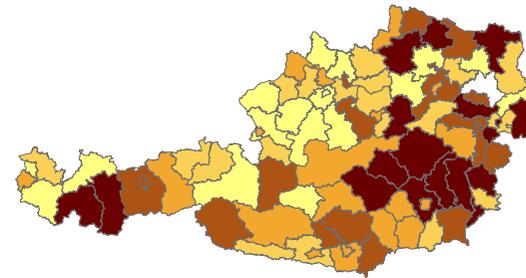
**Legend**

**BEZ**

**raw\_rate100k**



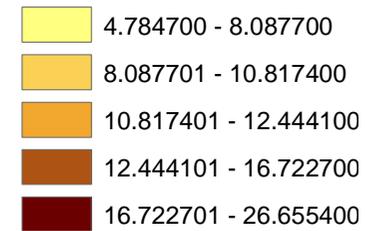
range: 41.8



**Legend**

**BEZ**

**predrate100k**



range: 21.9

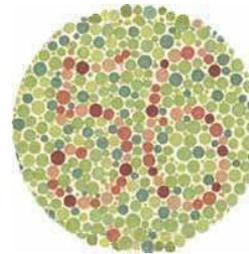
Rohe Raten sind oftmals wenig  
geeignet ein Risiko abzubilden

# Farben in Landkarten

Farbenblindheit

Genetische Krankheit,

Aberration auf dem X-Chromosom

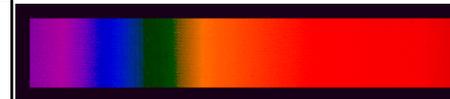


Quelle: <http://www.actoid.com/webdesign/farblichtsehen/Farbenblindheit.html>

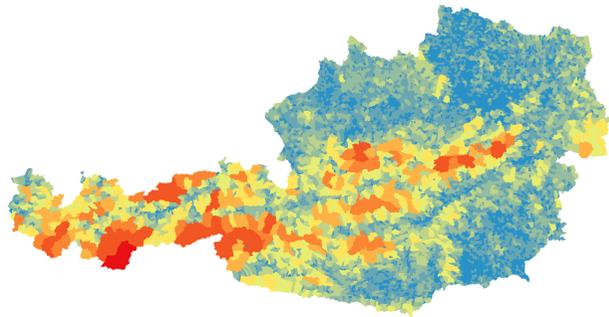
Trichromaten  
color spectrum seen by  
healthy people



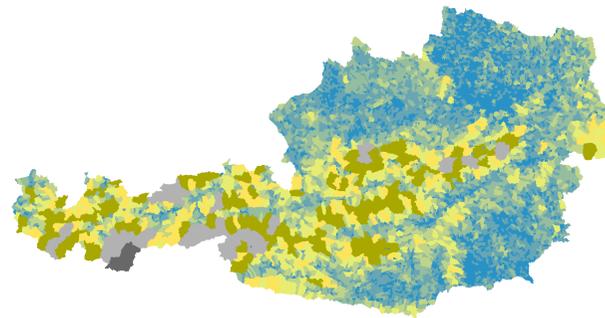
	Prevalence Men	Prevalence Women
Protanomalie (Rotschwäche)	1%	0.03%
Protanopie (red blindness)		
Deuteranomalie (Grünschwäche)	5%	0.35%
Deuteranopie (Grünblindheit)		
Tritanomalie (Blauschwäche)	0.005%	0.005%
Tritanopie (acyanopsia)		



Source: Color Vision Clinic, City University, London; [www.pboptik.ch/colorbr.html](http://www.pboptik.ch/colorbr.html)



color spectrum seen  
by healthy people



Protanopia	1%
(red blindness)	

## P-value fishing

P-Wert - Wahrscheinlichkeitsmaß um die Zufälligkeit eines beobachteten Effektes abzuschätzen

Der beobachtete Effekt wird mittels p-Wert  
als **signifikant** (=nicht zufällig)  
oder  
**nicht signifikant** (=zufällig)  
klassifiziert

## P-value fishing

P-Wert - Wahrscheinlichkeitsmaß um die Zufälligkeit eines beobachteten Effektes abzuschätzen

Der beobachtete Effekt wird mittels p-Wert  
als **signifikant** (=nicht zufällig)  
oder  
**nicht signifikant** (=zufällig)  
klassifiziert

signifikantes Relatives Risiko von 2 =  
wir glauben, dass das beobachtete Risiko nicht durch Zufall  
zustande kam sondern „wirklich besteht“

## P-value fishing

Entscheidung signifikant / nicht signifikant  
unsicher, mit Fehlerwahrscheinlichkeit behaftet

Saloppe Interpretation von 5% Signifikanz:

„Die Aussage über diesen auf signifikanten Effekts stimmt mit  
einer Wahrscheinlichkeit von  $(100\% - 5\%) = 95\%$ “

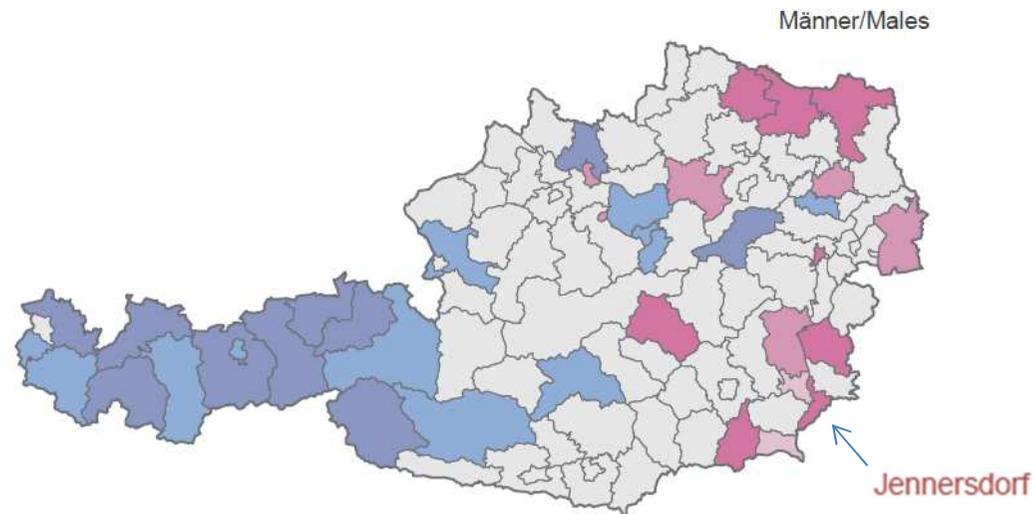
-> mit 5% stimmt die Aussage nicht

# Chronische Leberkrankheit und -zirrhose

*Chronic liver disease*

ICD-Code: K70, K73-K74

Bundesland, politischer Bezirk, Regionstyp	Männer				
	Index der Sterblichkeit (Österreich=100)			Standardsterbe- ziffer 1998/04	
	1998/04	1988/94	1978/84	1988/94 =100	1978/84 =100
Jennersdorf	196,7**	222,9***	97,1	62,6*	123,2



Standardisierte Sterbeziffer

<span style="color: red;">■</span> signifikant hoch + im obersten Zehntel	<span style="color: lightblue;">■</span> nicht signifikant niedrig + im untersten Zehntel
<span style="color: pink;">■</span> signifikant hoch + nicht im obersten Zehntel	<span style="color: blue;">■</span> signifikant niedrig + nicht im untersten Zehntel
<span style="color: lightpink;">■</span> nicht signifikant hoch + im obersten Zehntel	<span style="color: darkblue;">■</span> signifikant niedrig + im untersten Zehntel
<span style="color: lightgrey;">■</span> andere Gebiete	

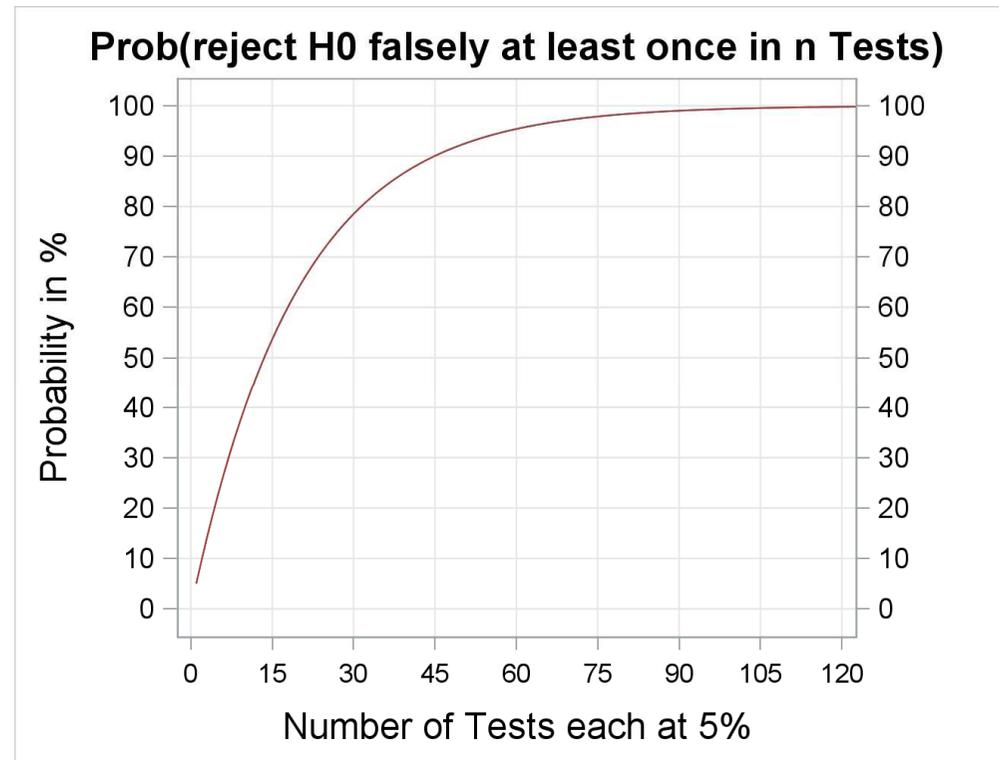
Österreichischer Todesursachenatlas 1998/2004, Statistik Austria

## P-value fishing

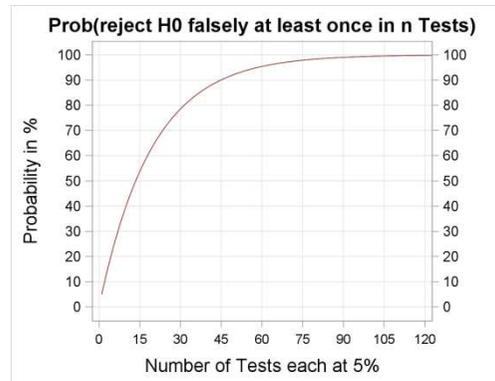
Fehlerwahrscheinlichkeit 5% gilt für jeden einzelnen  
Bezirksvergleich mit Durchschnitt Österreich

121 Bezirksvergleiche mit Österreich => 121 mal mit 5% einen  
falsche Aussage zu treffen

Wahrscheinlichkeit mindestens 1 mal  
eine falsch positiv Aussage zu treffen



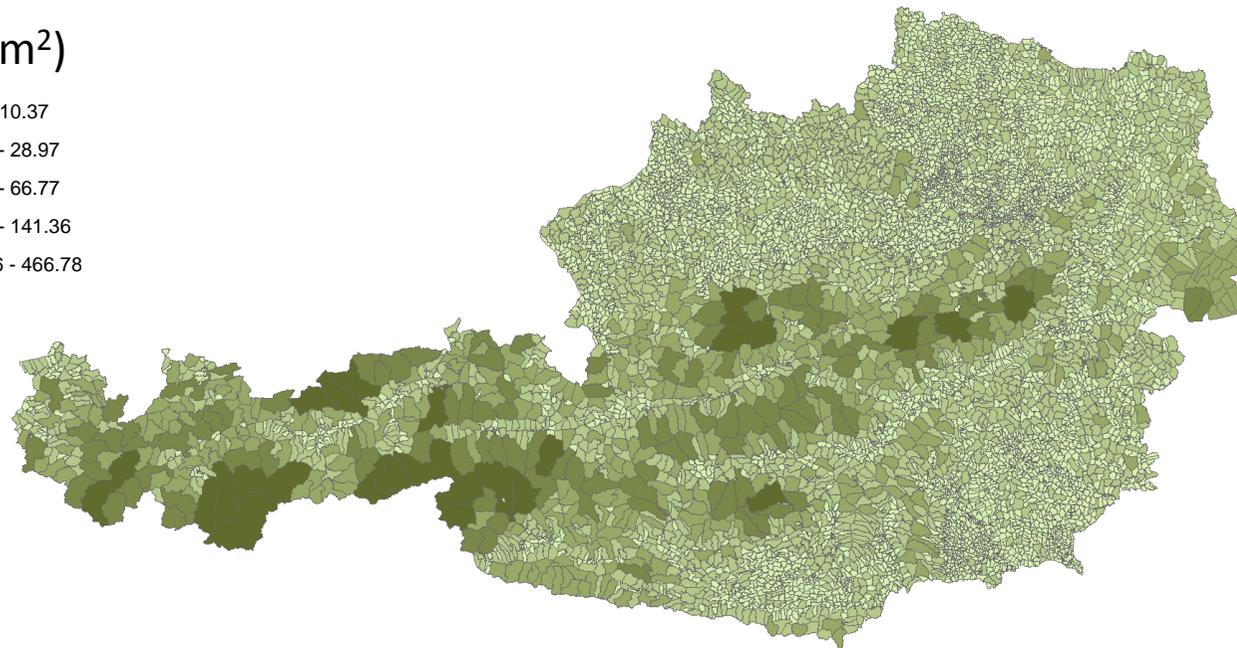
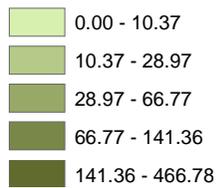
-> P(bei 121 Bezirksvergleichen mind. 1 falsch positive Aussage zu treffen) = ~100%



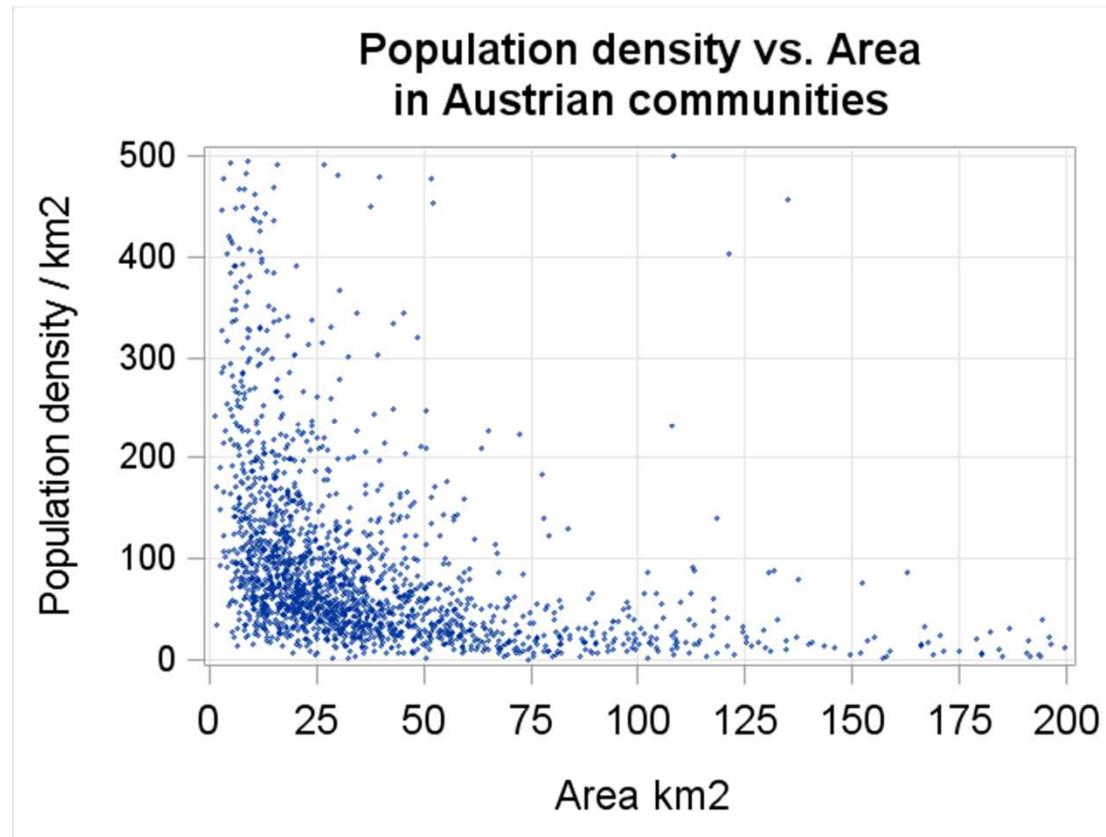
-> Durchsuchen einer Karte auf Auffälligkeiten mittels Signifikanz führt mit Sicherheit zu falsch positiven Aussagen

# Austrian communities hued by size of area

Area (km<sup>2</sup>)

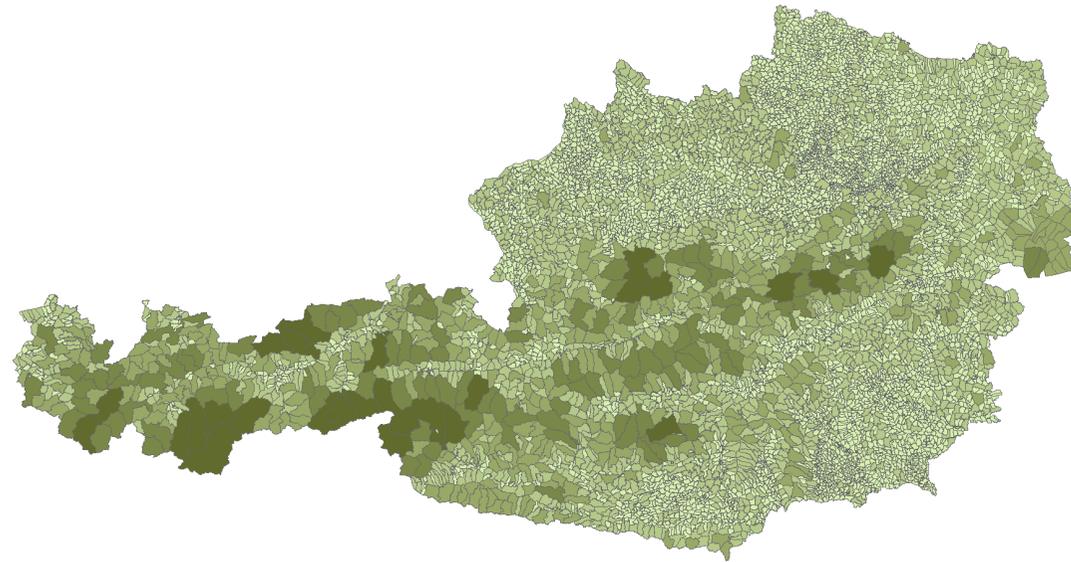


Few large areas dominate map



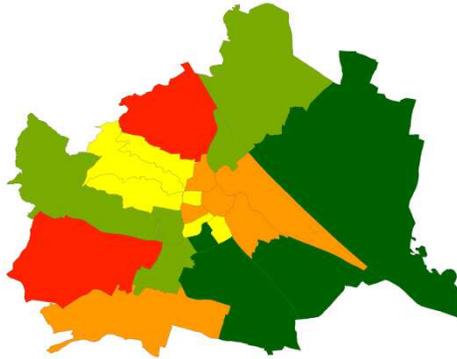
Larger areas -> smaller population density  
-> fewer cases -> larger variability in rates

## Austrian communities hued by size of area



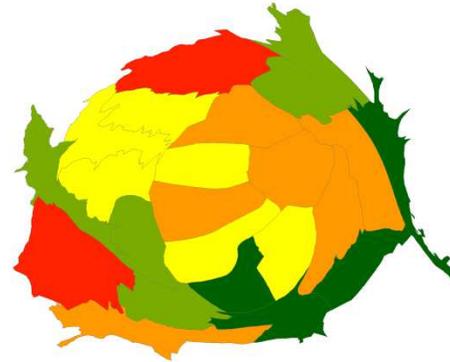
**Large but less populated areas**  
- showing increased variability in rates -  
**dominate map**

## Verwende Kartenanamorphote statt Flächenkartogramm

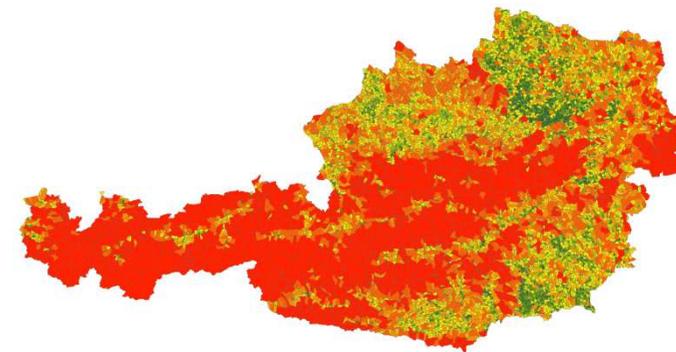
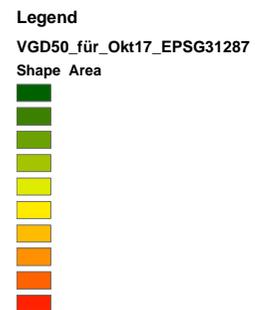
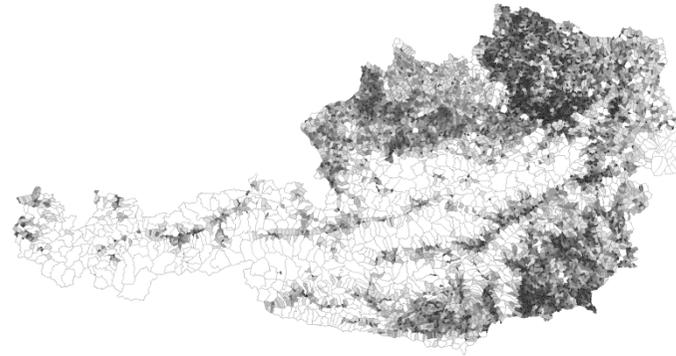
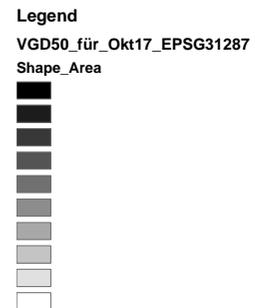


Choropleth (Flächenkartogramm)

Cartogram (Kartenanamorphote)

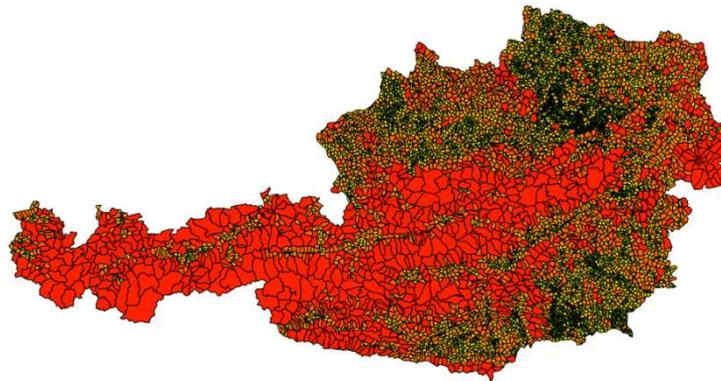
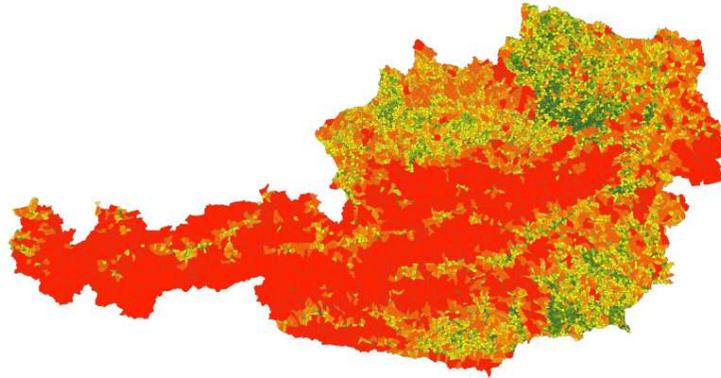


## Choice of color



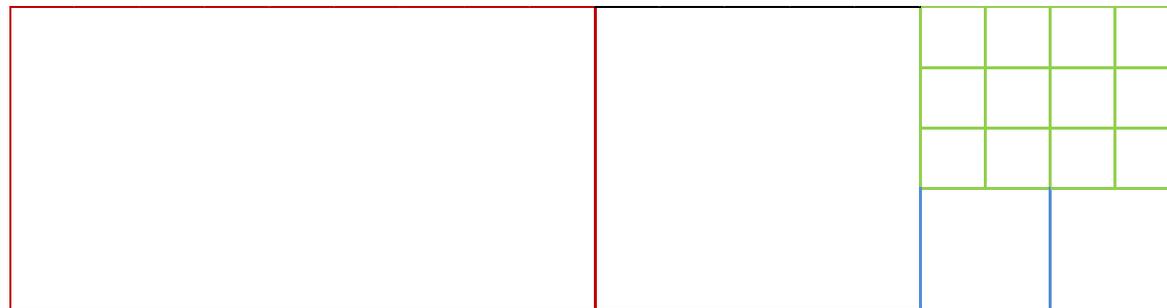
## Border colors

same legend color – different border color

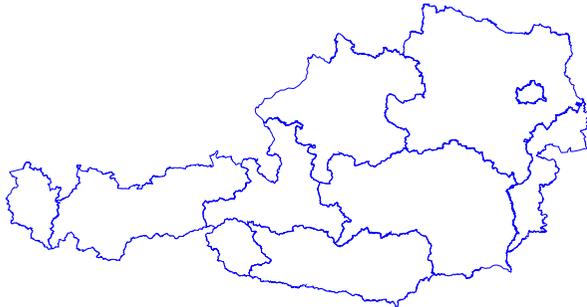


# Räumliche Auflösung

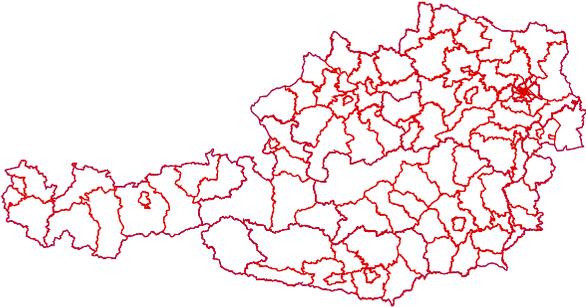
number of units influences results (resolution)



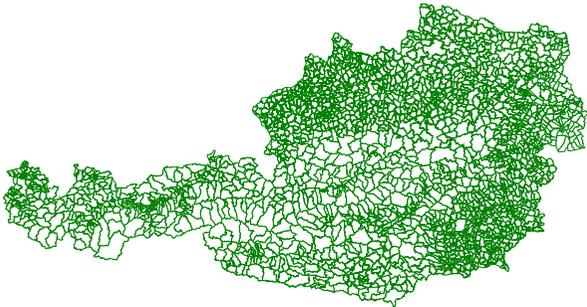
Province



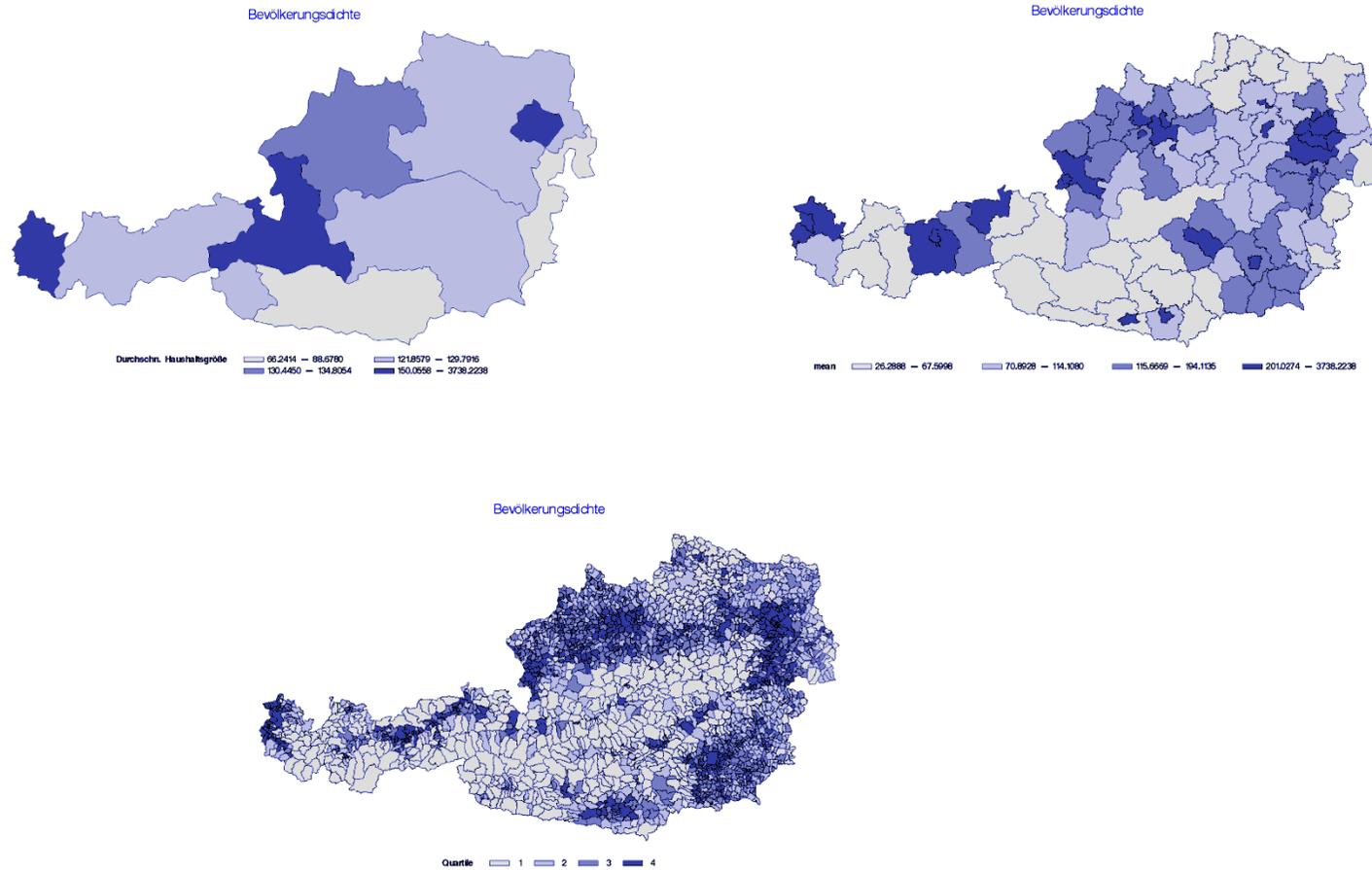
Districts



Municipalities

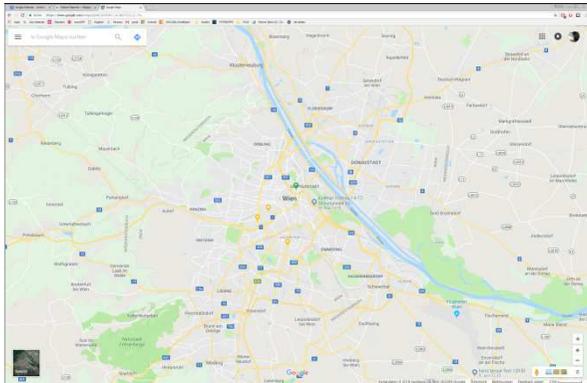
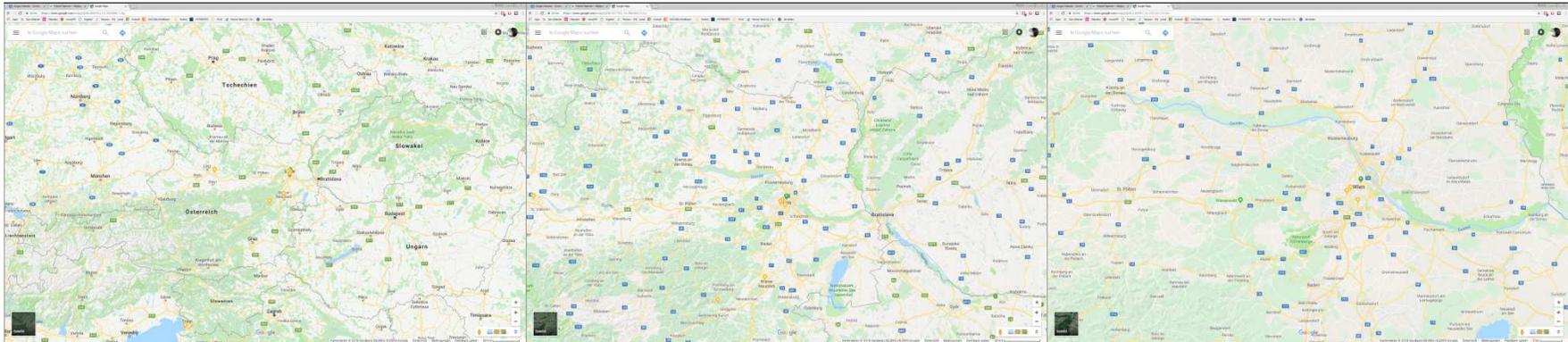


# Population density at different resolutions



## Conclusio:

Landkarten können wertvolle Werkzeuge sein um die räumliche Verteilung von Morbidität/Mortalität aufzuzeigen, führen aber in der Praxis manchmal zu verzerrten Ergebnissen



Danke für Ihre  
Aufmerksamkeit!

