On the combination of difference and equivalence tests in spatial maps

Waldhör Thomas and Harald Heinzl

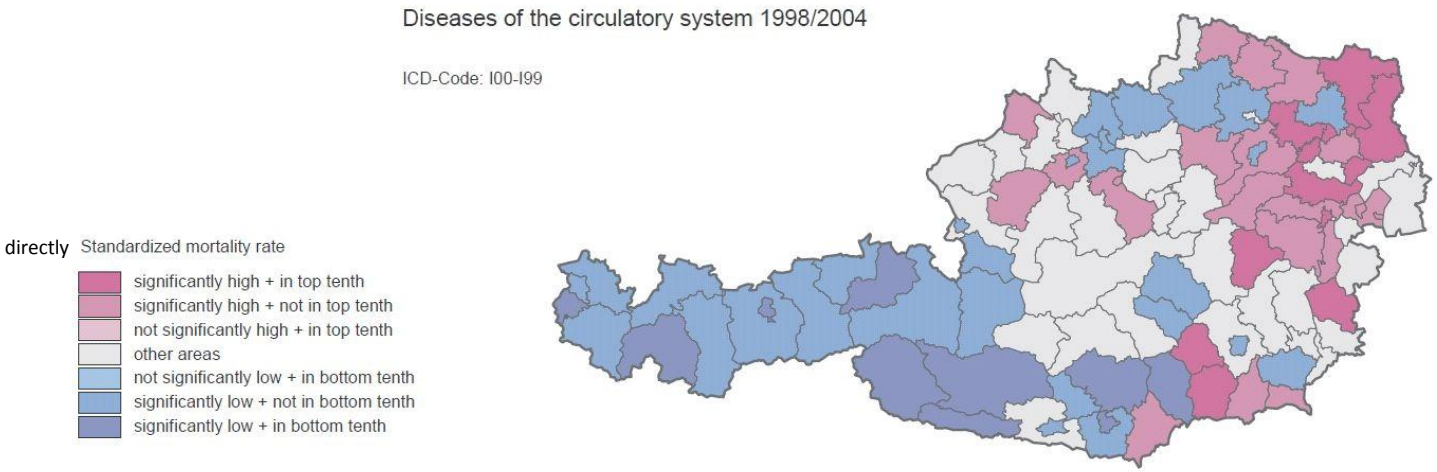
Department of Epidemiology; Center for Medical Statistics,

Informatics and Intelligent Systems

Medical University of Vienna

„Atlas of mortality in Austria by causes of death 1998/2004“*

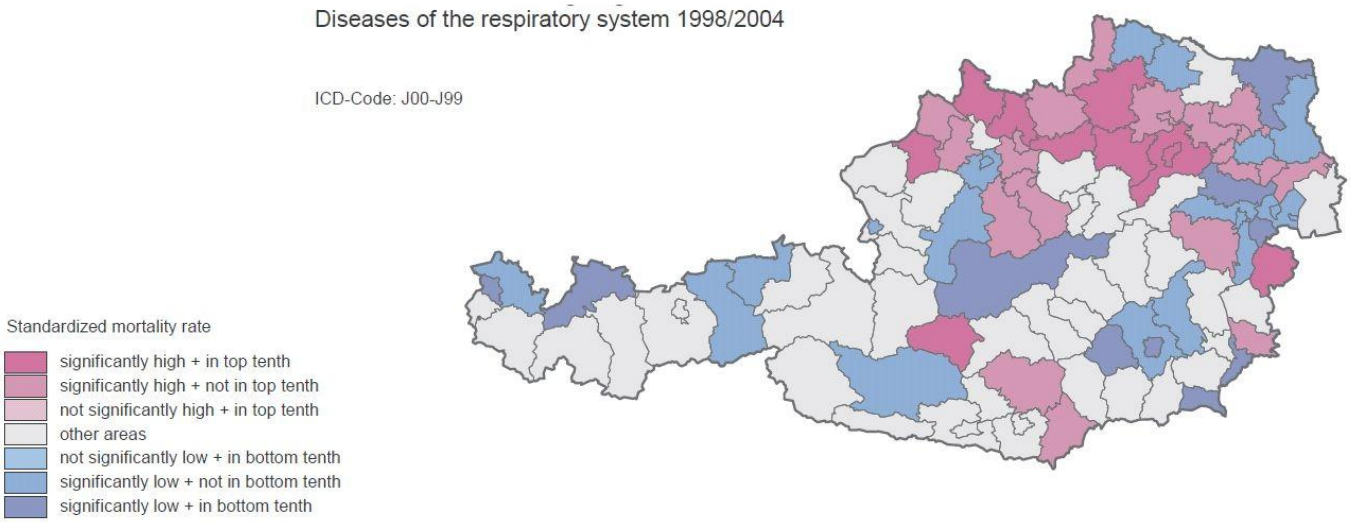
presents mortality rates and tests for difference



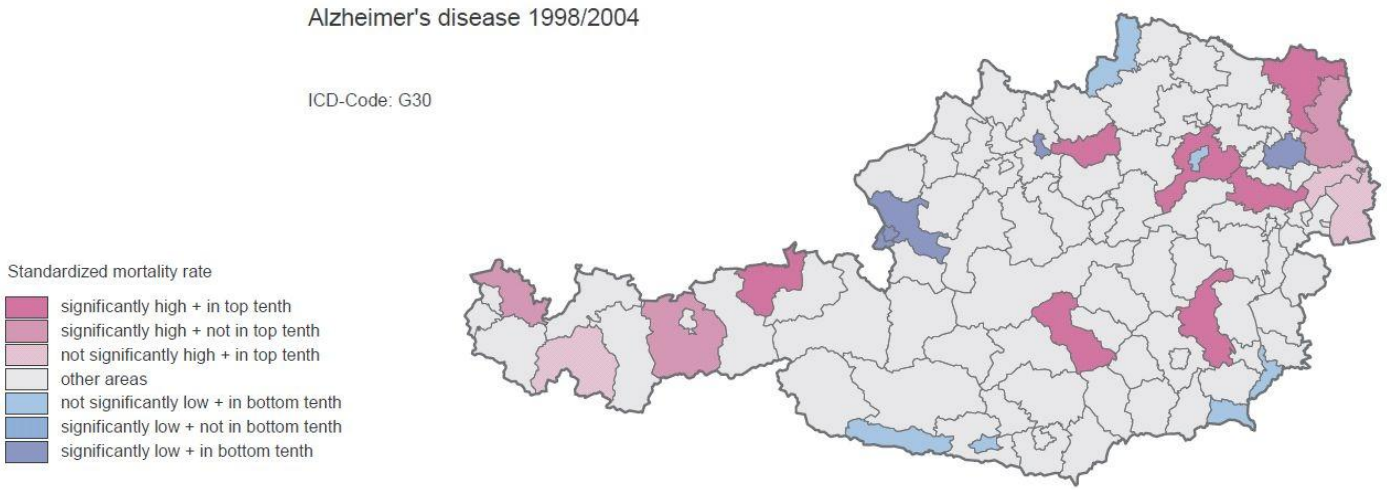
Circulatory system

*source: Statistik Austria, Bundesanstalt Österreich, 1110 Wien, Guglgasse 13

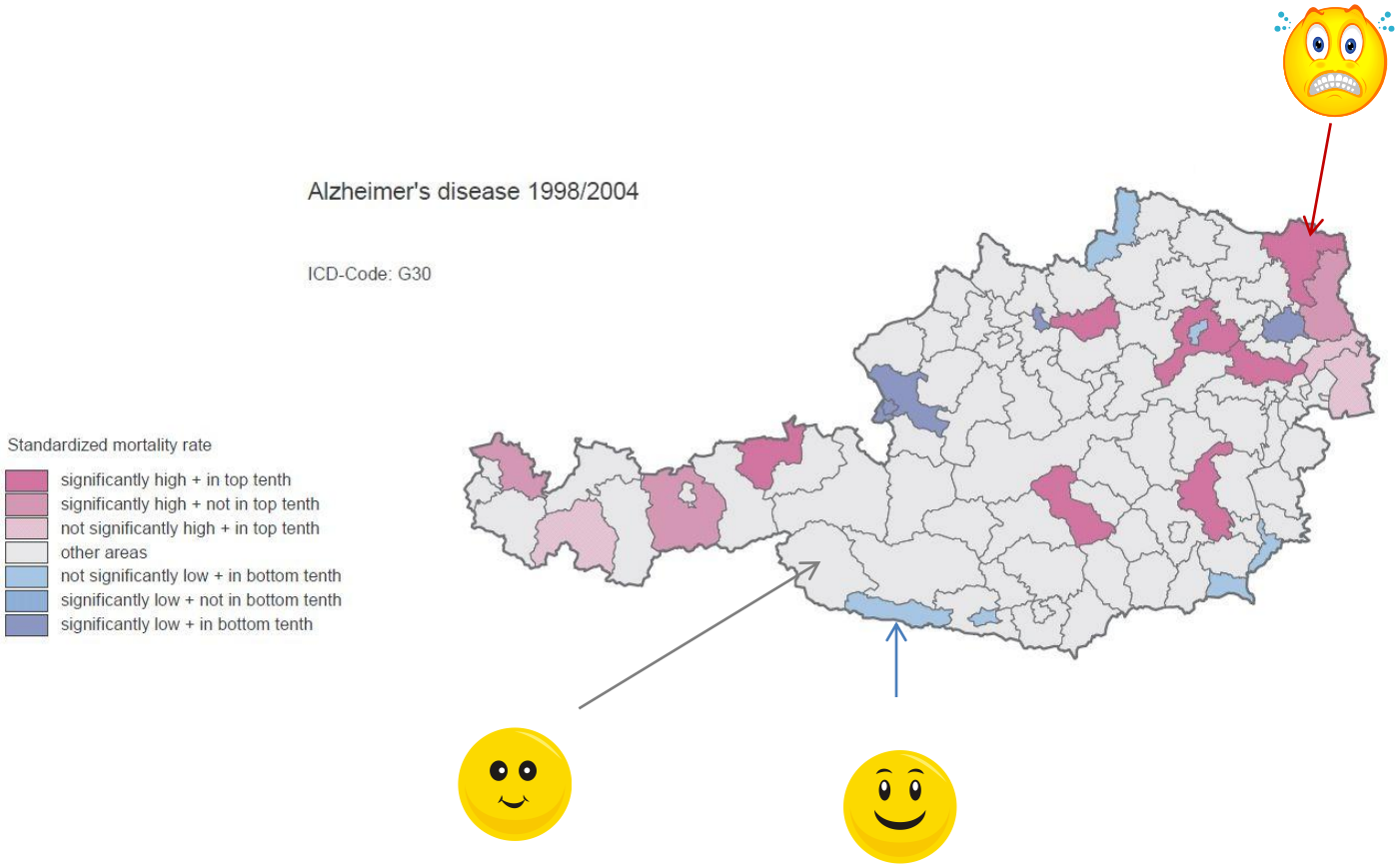
Respiratory system



Alzheimer disease



Alzheimer disease



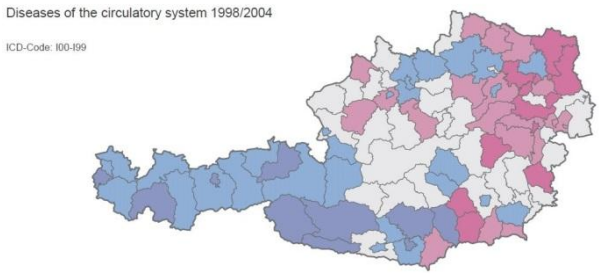
Readers mostly interprete non significant areas

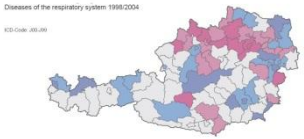
as being „equal“ or „equivalent“

to national average

- no need to worry

Numbers of deaths





Circulatory system

n= 267.457

Respiratory system

n= 29.285

Alzheimer disease

n= 2.416

We will observe a non significant difference –

because of small sample size

but **not**

because rate is „equivalent“ to average

In order to provide information for equivalence

we have to set up

equivalence instead of **difference** tests

In order to provide information for equivalence

we have to set up

equivalence instead of **difference** tests

Define „what“ means equivalent

Calculate corresponding equivalence test

In spatial epidemiology equivalence

may be defined by an SMR range

e.g. (0.8, 1.25) or (0.9,1.11)

True SMR‘s within equivalence range are

deemed equal to some comparison value c

e.g. SMR=1

Test for equivalence

two one-sided tests (TOST) approach

$H0_1$: SMR≤1-Δ_1 and $H0_2$: SMR≥1+Δ_2

Often Δ_1 , Δ_2 are set to

$1 - \Delta_1 = 1/(1 + \Delta_2)$

so e.g. Δ_1= 0.2 -> range = 0.8, 1.25

Test for equivalence

two one-sided tests (TOST) approach

$H0_1$: SMR≤1-Δ_1 and $H0_2$: SMR≥1+Δ_2

If both $H0_1$ and $H0_2$

are rejected at a significance level α each,

then the population SMR can be

declared *equivalent* to 1.

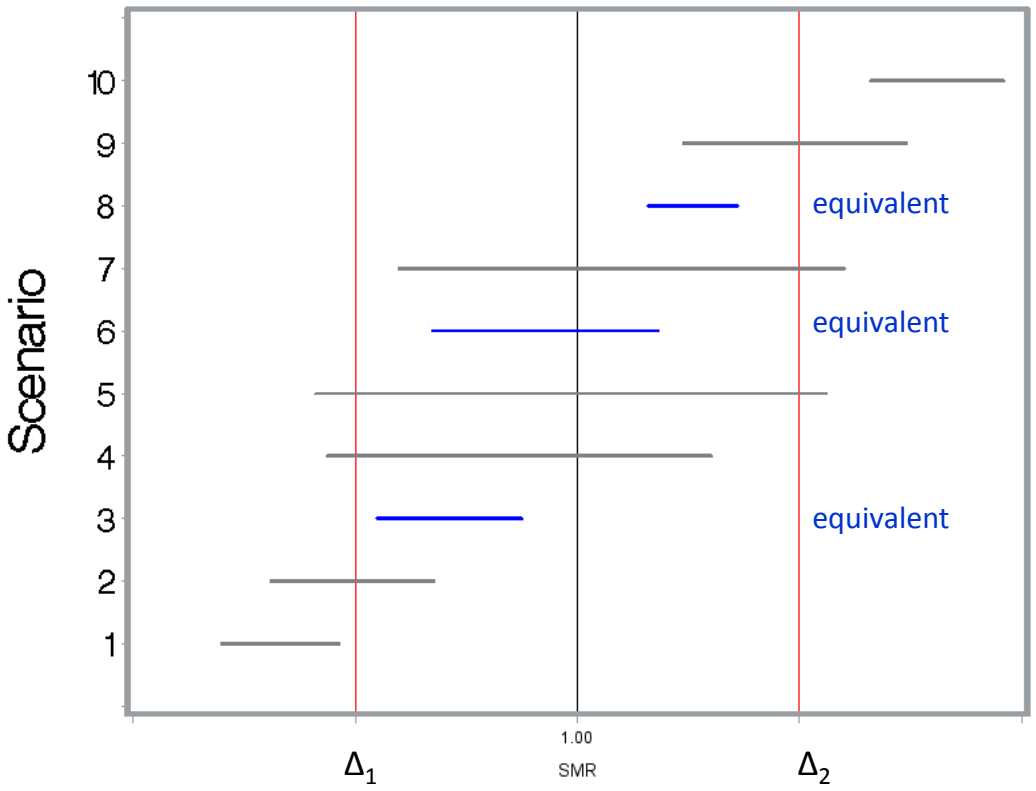
Combining difference and equivalence test results in spatial maps . Thomas Waldhoer and Harald Heinzl. *International Journal of Health Geographics* 2011, **10**:3 doi:10.1186/1476-072X-10-3

Rejection of both H0's:

two-sided (1 - 2α) - confidence interval for

estimated SMR is contained in equivalence range

10 possible scenarios for equivalence tests based on a 90% CI



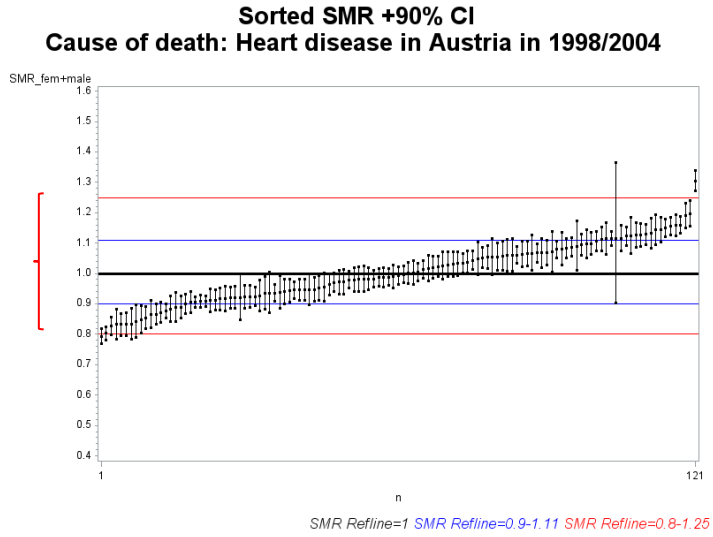
Equivalence is attained in scenarios E_3, E_6, E_8

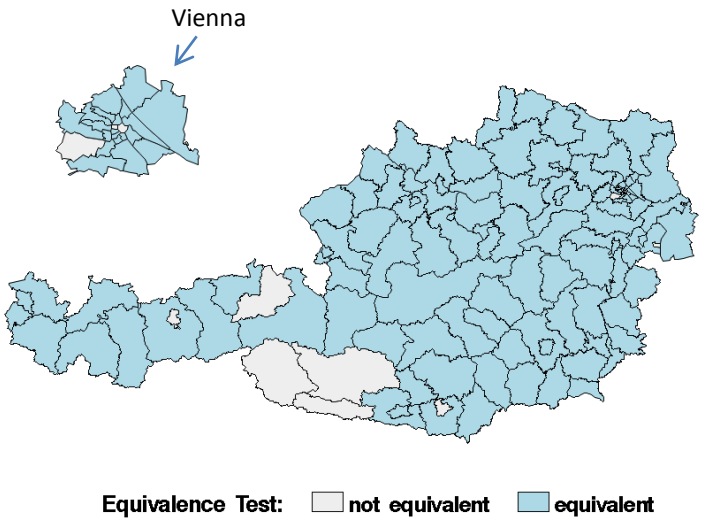
all other scenarios are not equivalent

Circulatory system

equivalence test for SMR range 0.8-1.25

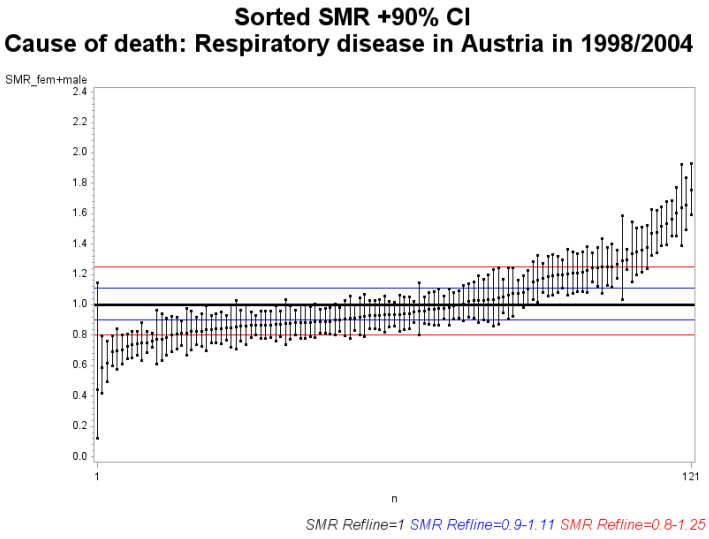
equiv.range
0.8-1.25

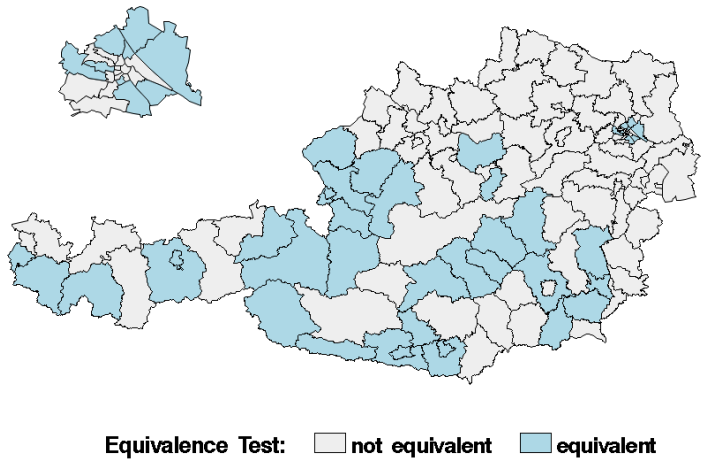




Respiratory diseases

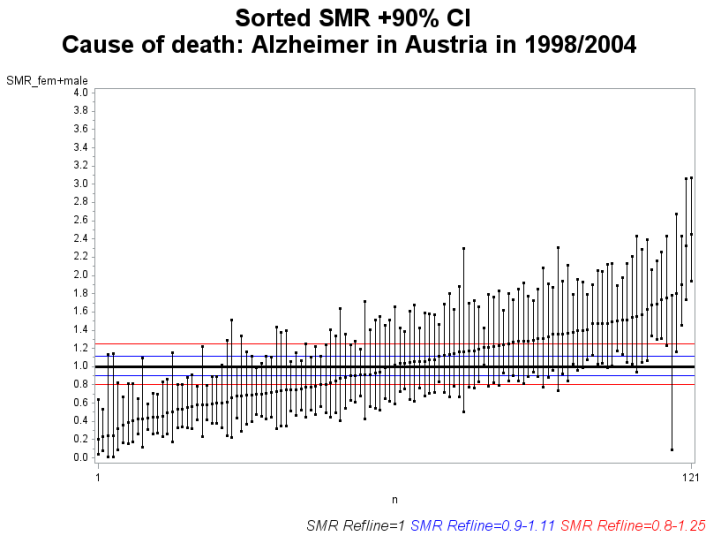
equivalence test for SMR range 0.8-1.25

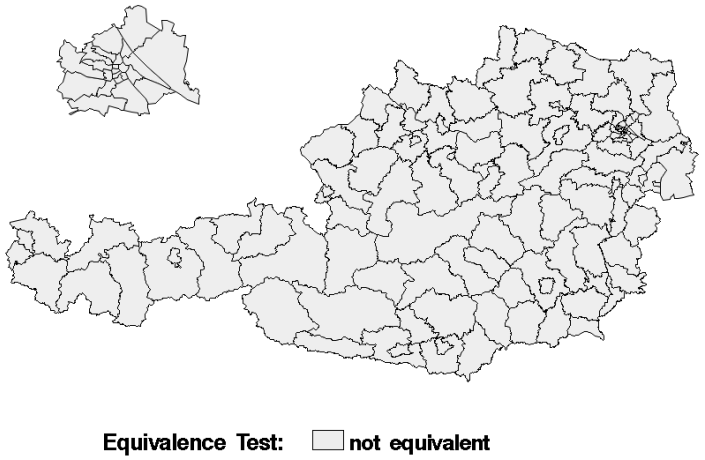




Alzheimer disease

equivalence test for SMR range 0.8-1.25

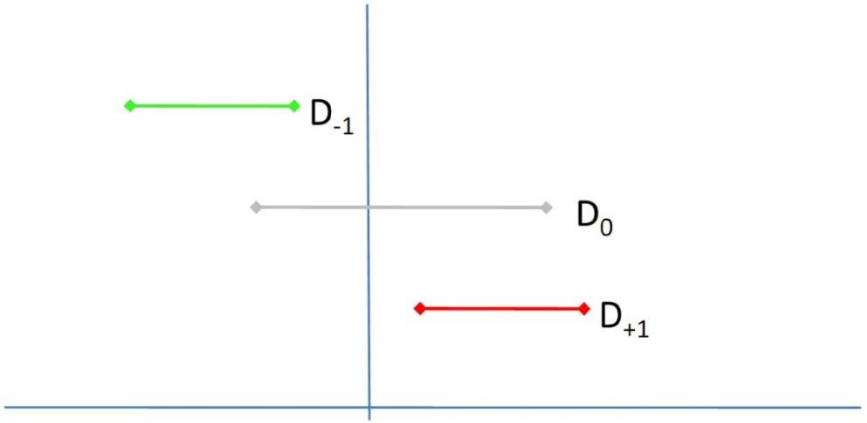




We may combine

equivalence and difference tests

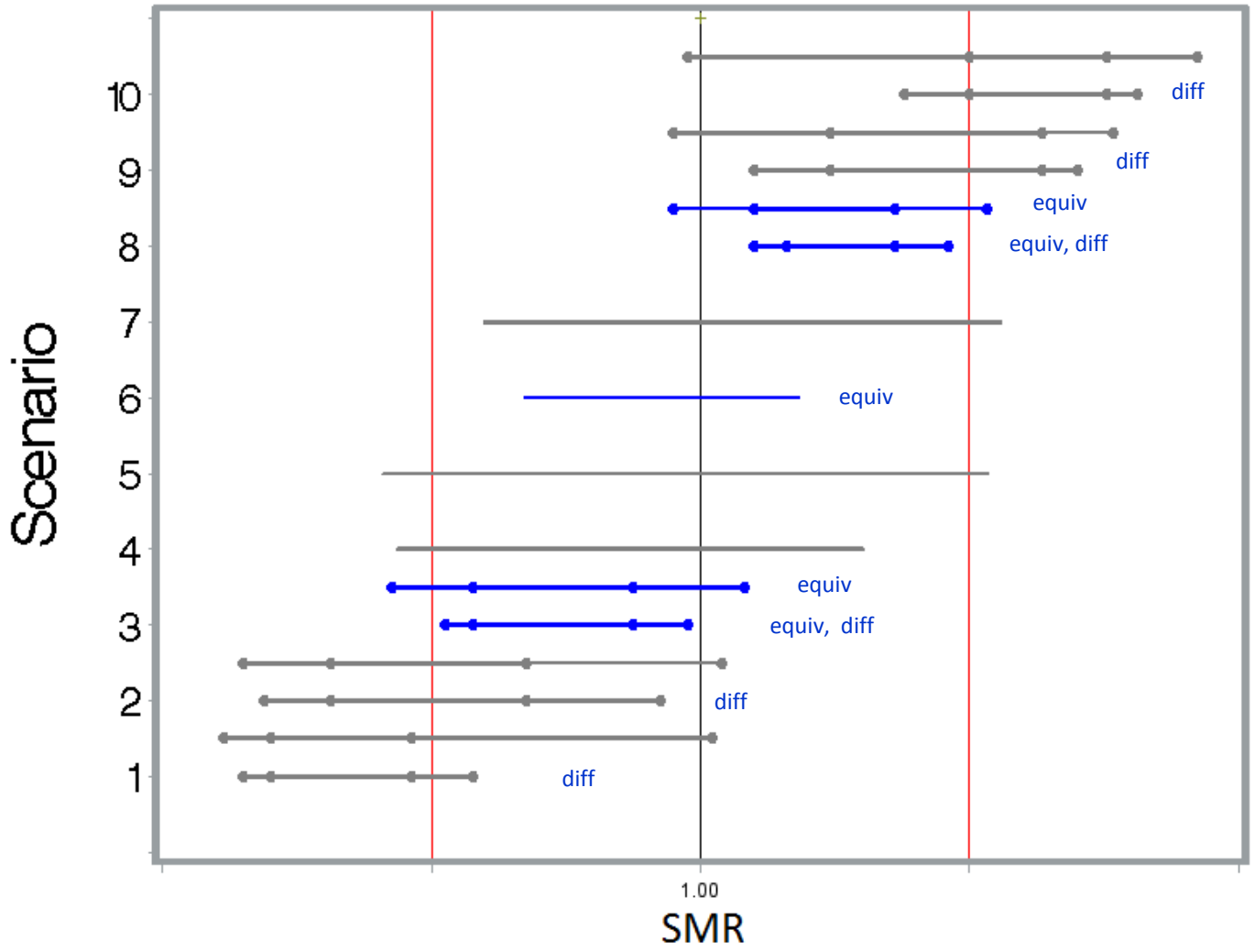
Test for difference based on a 95% confidence interval



SMR= 1

- Combination of equivalence and difference in map
- Use of 90% and 95% CI
- Group corresponding results for presentation in map

16 possible scenarios for combination of equivalence and difference tests based on a 90% and 95% CI



scenarios E4,E5,E6,E7 are not significantly different (D_0), only 90% CI shown

Two schemes to distinguish mutual difference and equivalence test results in choropleth maps

Six combined scenarios

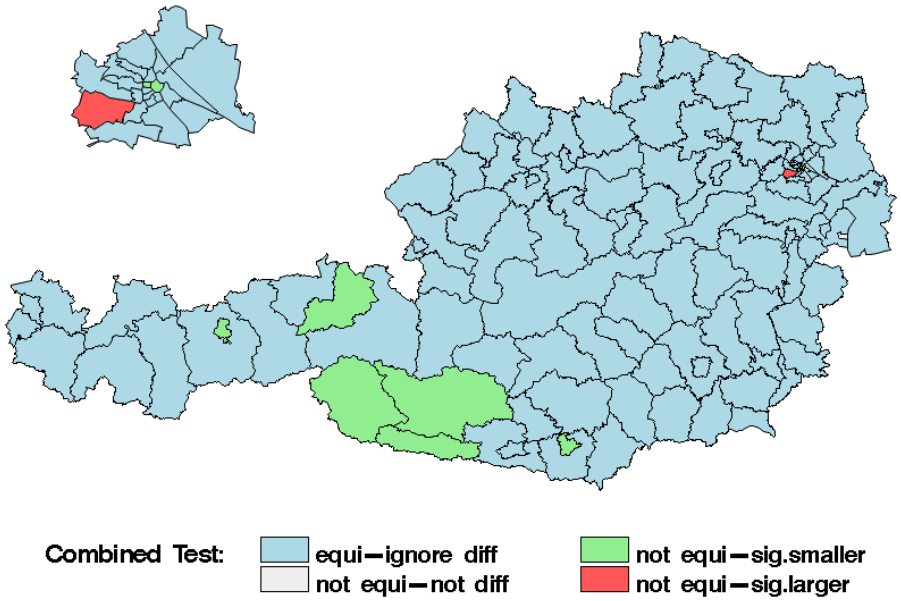
equivalent	sign. smaller
equivalent	not sign. different
equivalent	sign. larger
not equivalent	sign. smaller
not equivalent	not sign. different
not equivalent	sign. larger

Four combined scenarios

Equivalent (ignore difference test)	
not equivalent	sign. smaller
not equivalent	not sign. different
not equivalent	sign. larger

Circulatory system

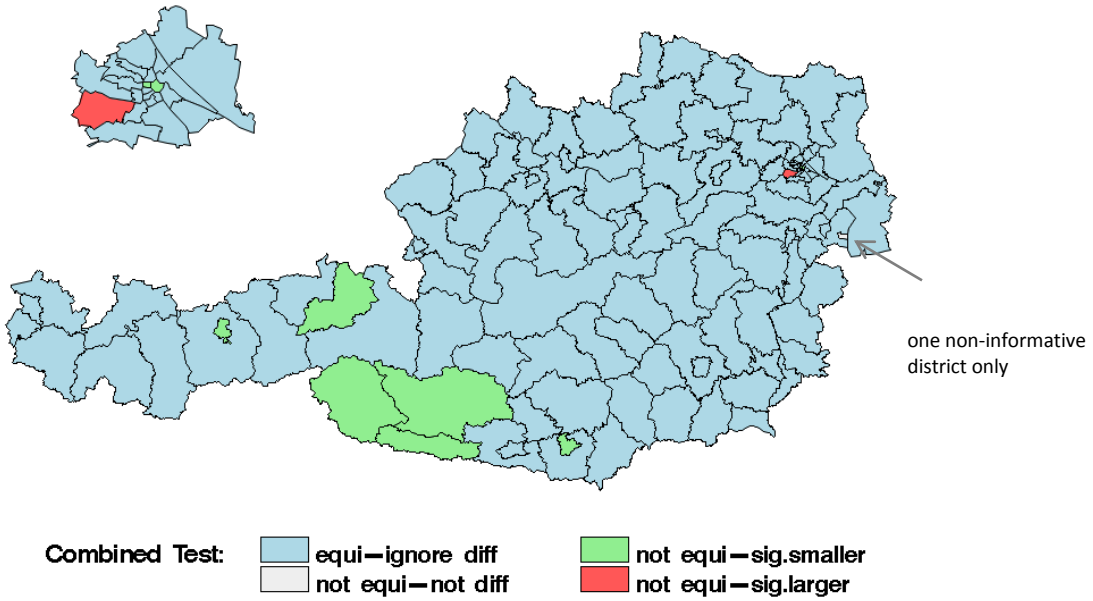
combination of equivalence and difference tests into 4 groups



equivalence range (0.8,1.25)

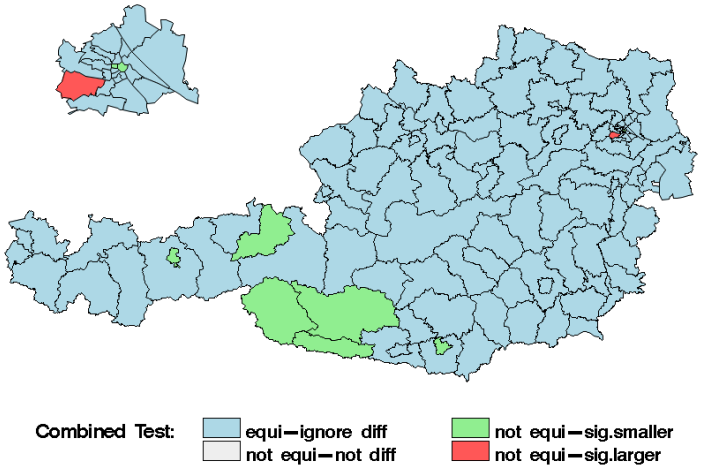
Circulatory system

combination of equivalence and difference tests into 4 groups

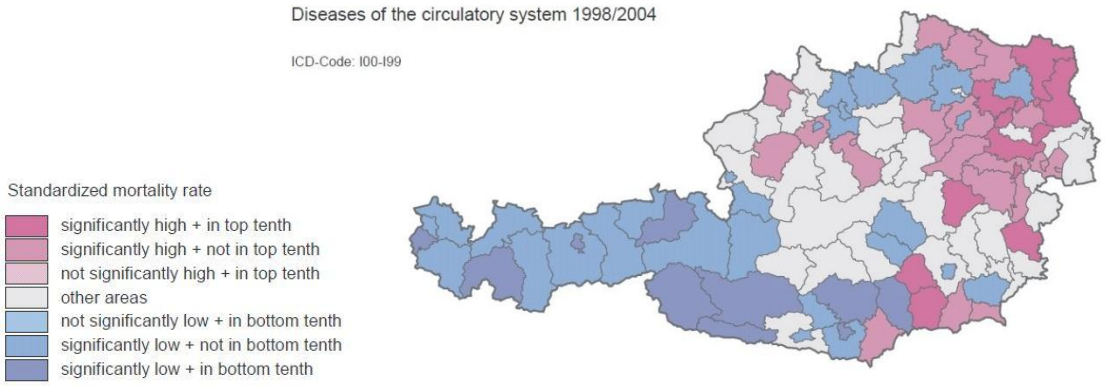


equivalence range (0.8,1.25)

Circulatory system



Combined test

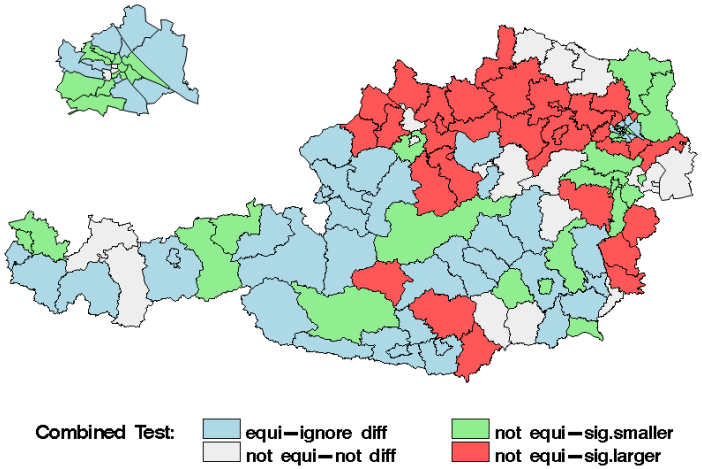


Difference test

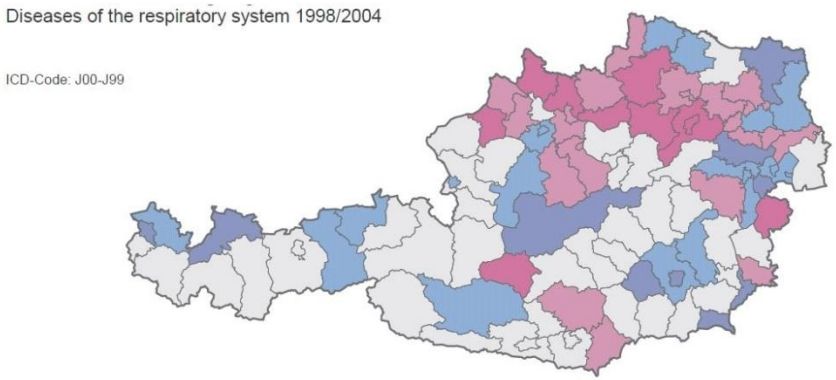
Atlas of mortality, Statistic Austria

Mind different color schemes!

Respiratory diseases

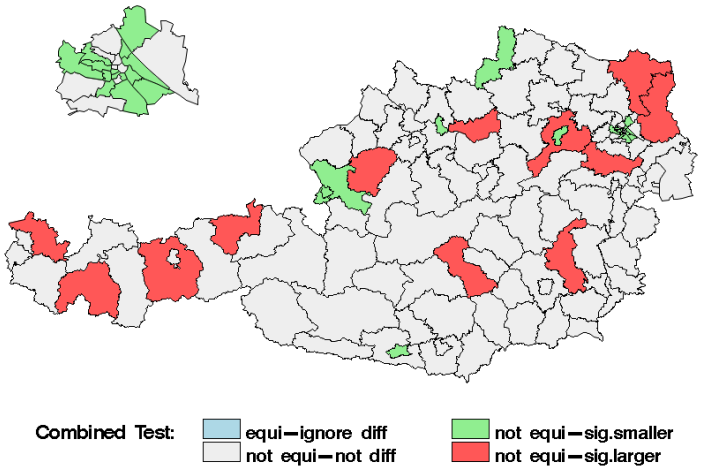


Combined test

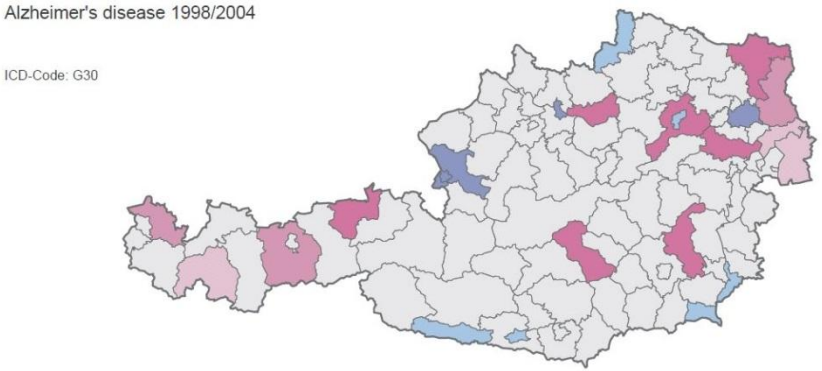


Difference test

Alzheimer disease



Combined test



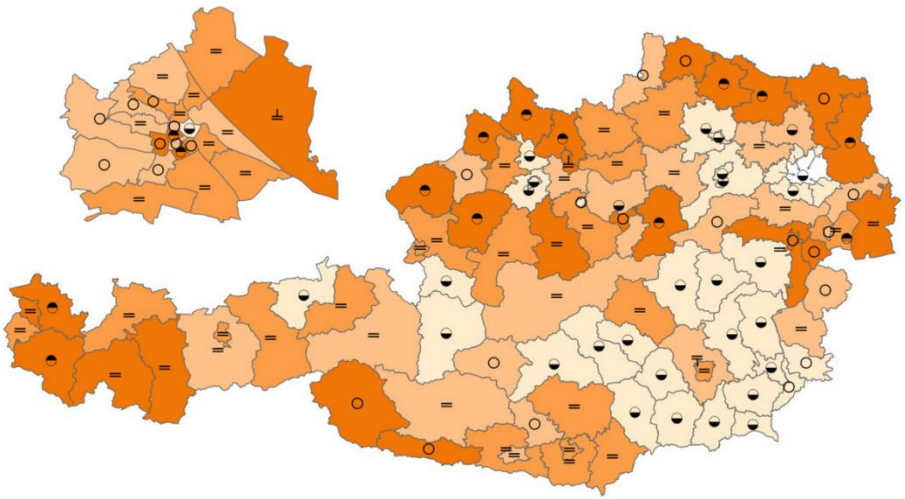
Difference test

Choropleth map of gestational age of newborns in Austria 2008

Legend 1 – Gestational age

mean gestational age per district

- 38.4 - 39.2 weeks
- 39.2 - 39.4 weeks
- 39.4 - 39.5 weeks
- 39.5 - 40.5 weeks



Legend 2 – Equivalence + Difference

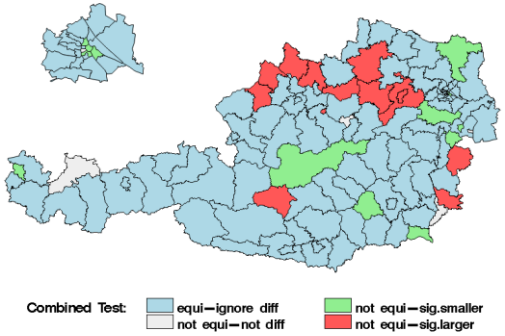
- ∓ equivalent and significantly smaller
- = equivalent and not significantly different
- ≐ equivalent and significantly larger
- ◒ not equivalent and significantly smaller
- ○ not equivalent and not significantly different
- ◓ not equivalent and significantly larger

Colors represent gestational age
Symbols represent combined test results (6 groups)

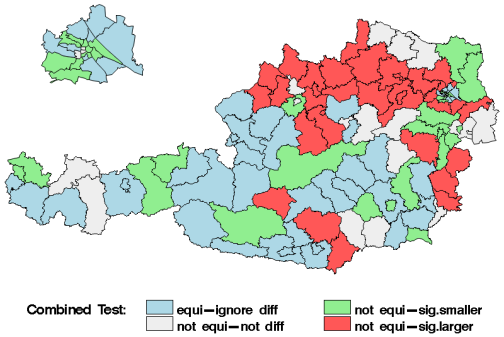
Results of equivalence tests depend strongly on equivalence range

Respiratory diseases

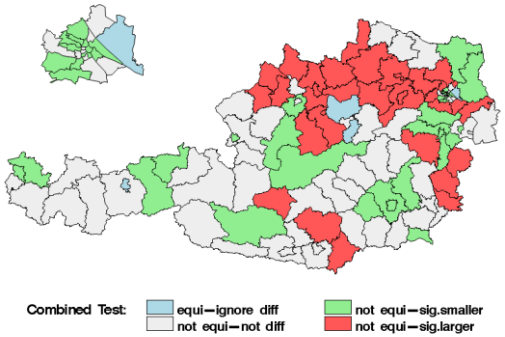
varying the width of range of equivalence



range 0.7 – 1.42
width: 0.72



range 0.8 – 1.25
width: 0.45

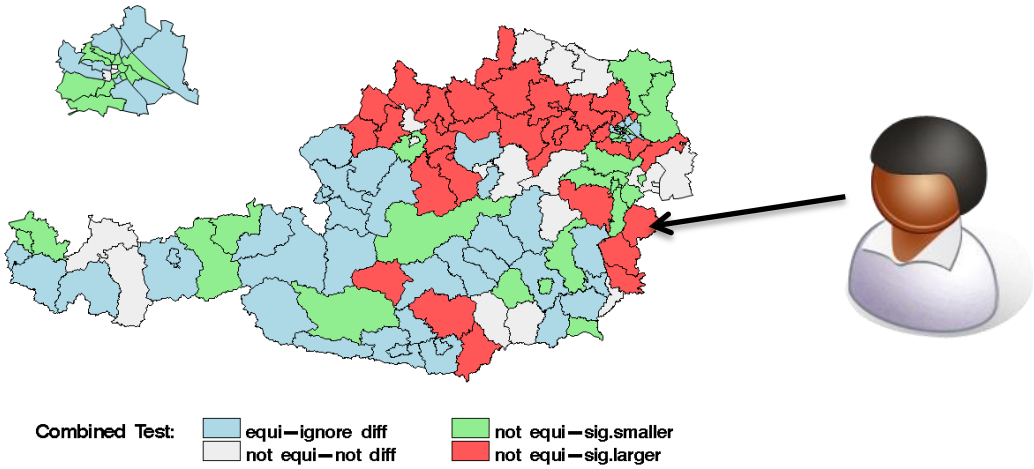


range 0.9 – 1.11
width: 0.21

Many spatial units lead to many statistical tests

so we have the problem of multiple tests

Seen from a person just interested in its own spatial unit, this multiple test problem does not apply



Readers interested in whole map:

We may use different models which

avoid the multiple test problem

Conclusion

Presentation of difference tests alone

may distort the perception of the reader

leading to many seemingly equivalent areas

Conclusion

Presentation of difference tests alone

may distort the perception of the reader

leading to many seemingly equivalent areas

Combination of difference and equivalence tests

provides more information than standard maps do

Conclusion

Setting up the equivalence range forces

the researcher to define

what is relevant and what is irrelevant