

Diagnostische Genauigkeit aggregierter Prävalenz- und Inzidenzdaten von 70 Millionen gesetzlich Versicherten am Beispiel des Typ-2-Diabetes



Brinks R^{1,2}, Tönnies T², Hoyer A³

¹Lehrstuhl für Medizinische Biometrie und Epidemiologie, Fakultät für Gesundheit, Universität Witten/Herdecke

²Institut für Biometrie und Epidemiologie, Deutsches Diabetes-Zentrum, Düsseldorf

³AG Biostatistik, Medizinische Fakultät OWL, Universität Bielefeld

Fragestellung

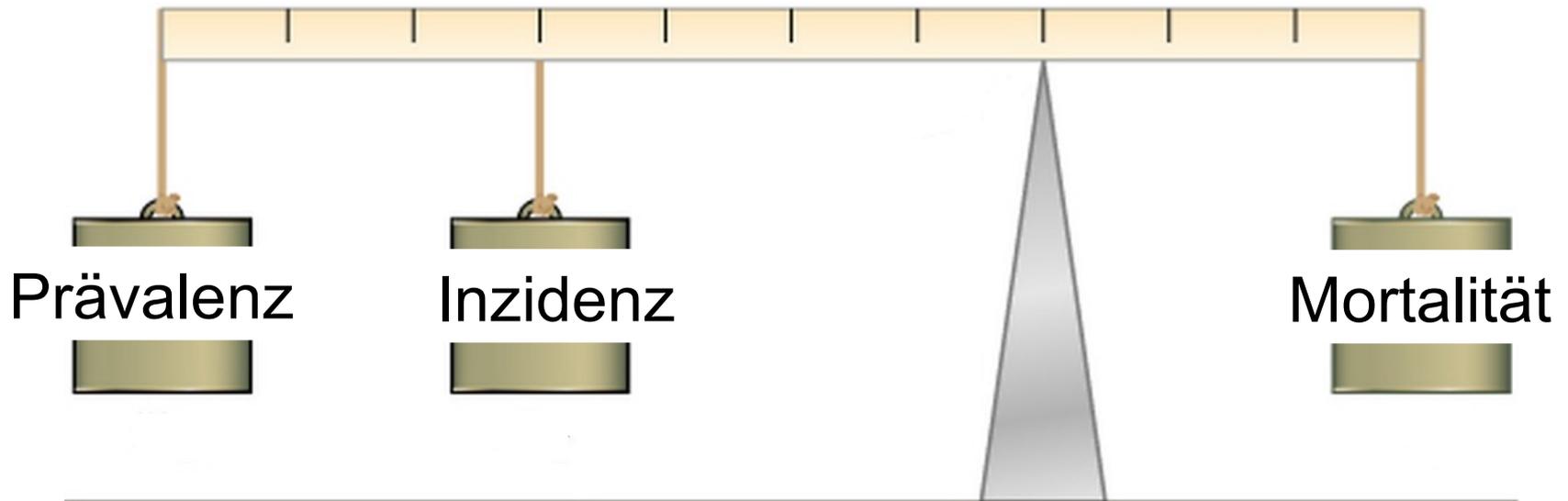
1. Daten zu **Prävalenz** und **Inzidenz** von chronischen Erkrankungen zunehmend erhältlich

Fragestellung

1. Daten zu **Prävalenz** und **Inzidenz** von chronischen Erkrankungen zunehmend erhältlich
2. Wenn zusätzlich **Mortalität** bekannt ist, was können wir über die diagnostische Qualität der Daten lernen?

Fragestellung

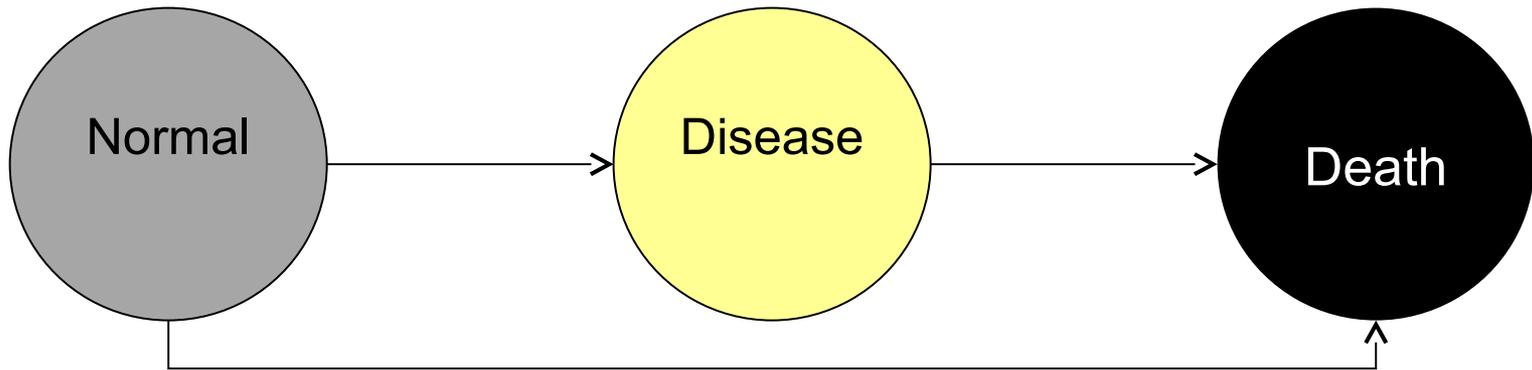
1. Daten zu **Prävalenz** und **Inzidenz** von chronischen Erkrankungen zunehmend erhältlich
2. Wenn zusätzlich **Mortalität** bekannt ist, was können wir über die diagnostische Qualität der Daten lernen?



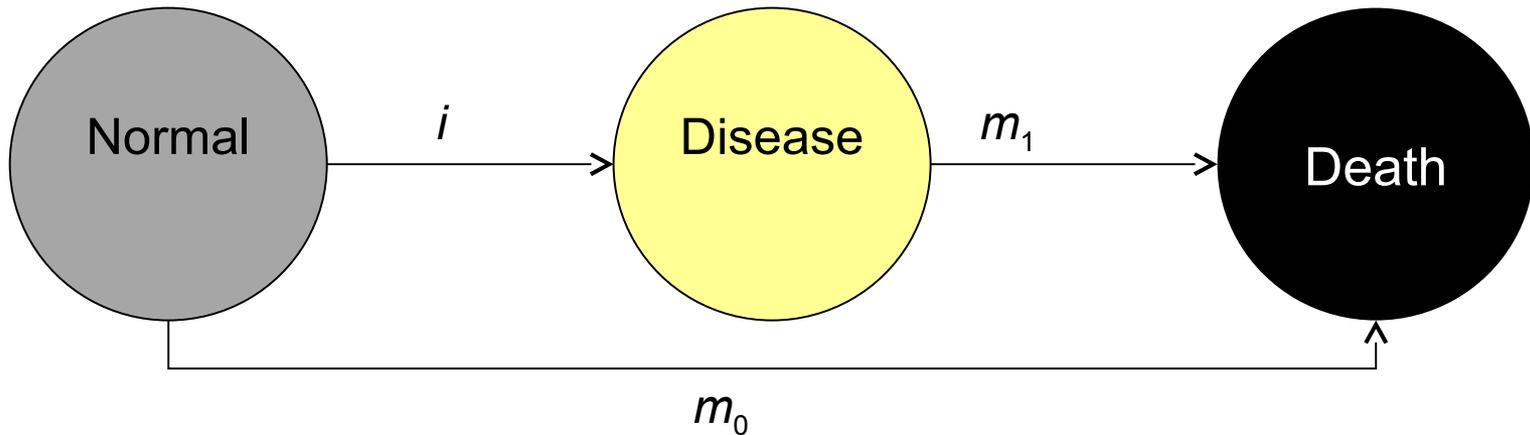
Gliederung

1. Illness-death model (IDM) für chronische Erkrankungen
2. Kassendaten: Diagnostische Genauigkeit
3. Aggregierte Prävalenz- und Inzidenzdaten zu Diabetes
4. Ergebnisse und Schlussfolgerung

Illness-death model für chronische Erkrankungen

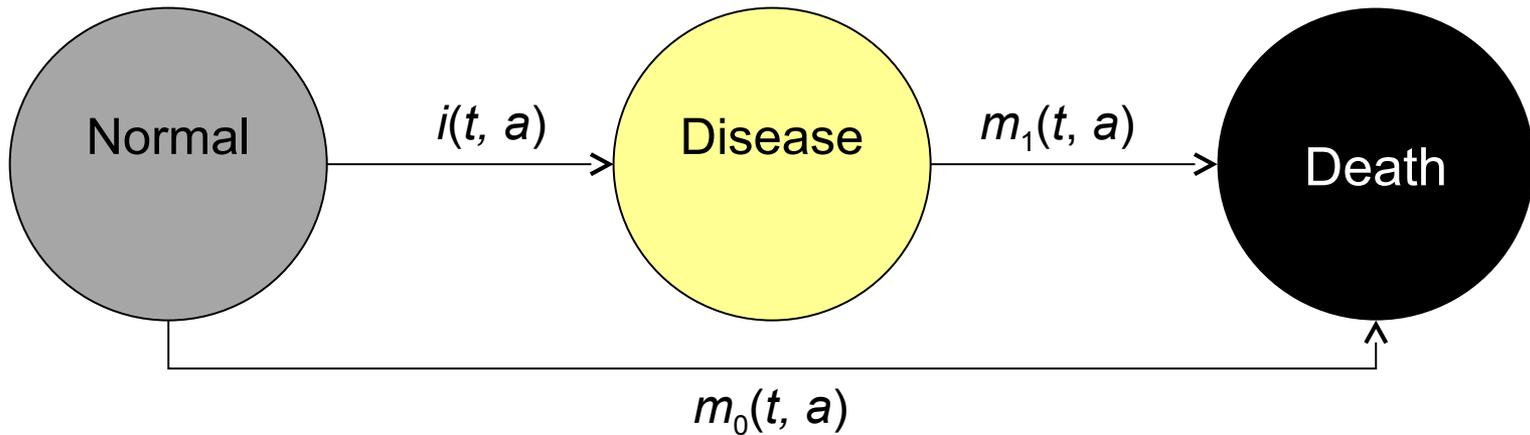


Illness-death model für chronische Erkrankungen



i Incidence rate
 m_0 Mortality rate of healthy people
 m_1 Mortality rate of diseased people

Illness-death model für chronische Erkrankungen



t Calendar time (period)

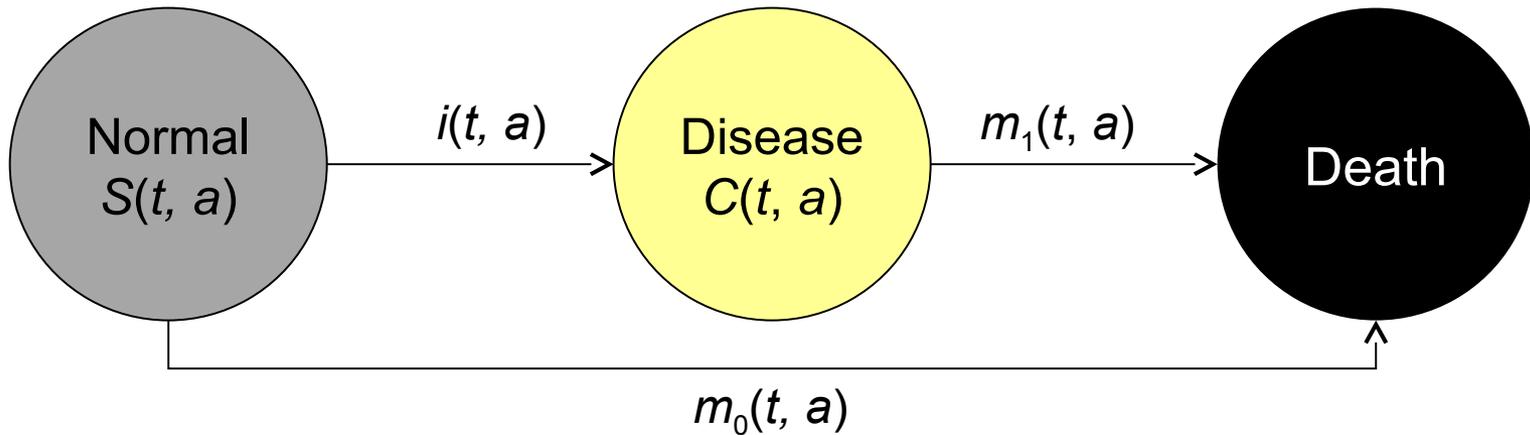
a Age

i Incidence rate

m_0 Mortality rate of healthy people

m_1 Mortality rate of diseased people

Illness-death model für chronische Erkrankungen



t Calendar time (period)

a Age

$S(t, a)$ Number of healthy people aged a at time t

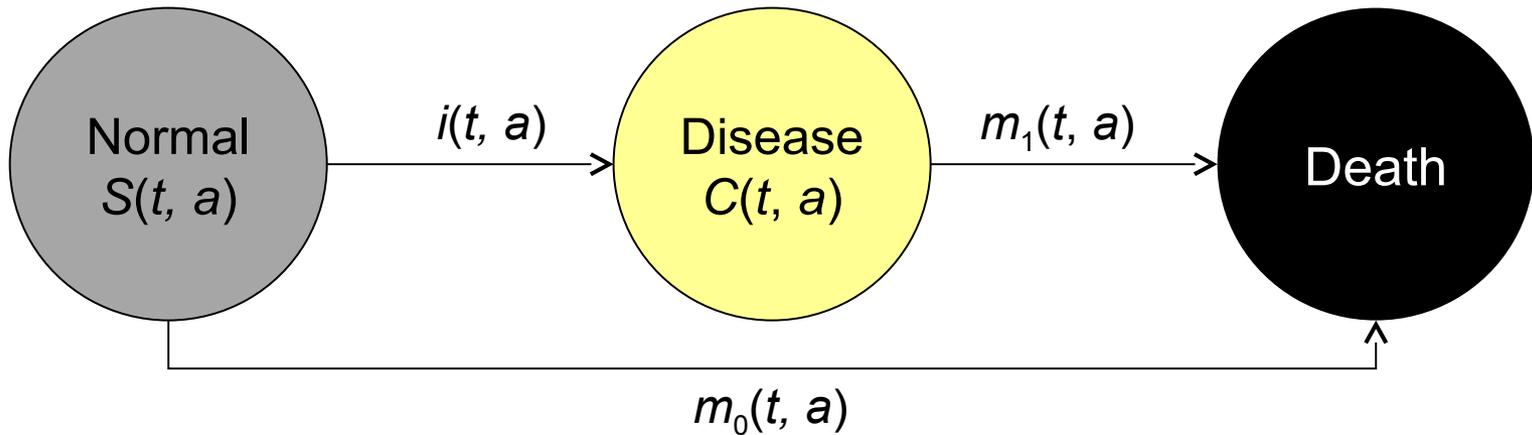
$C(t, a)$ Number of diseased people aged a at time t

i Incidence rate

m_0 Mortality rate of healthy people

m_1 Mortality rate of diseased people

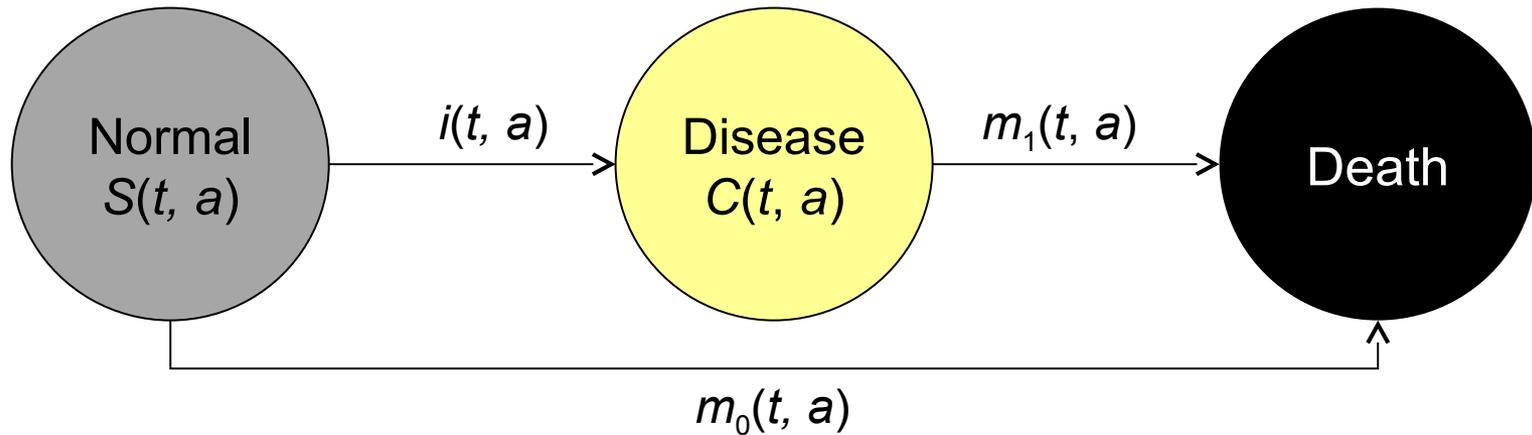
Illness-death model für chronische Erkrankungen



Age-specific prevalence

$$p(t, a) = \frac{C(t, a)}{S(t, a) + C(t, a)}$$

Illness-death model für chronische Erkrankungen

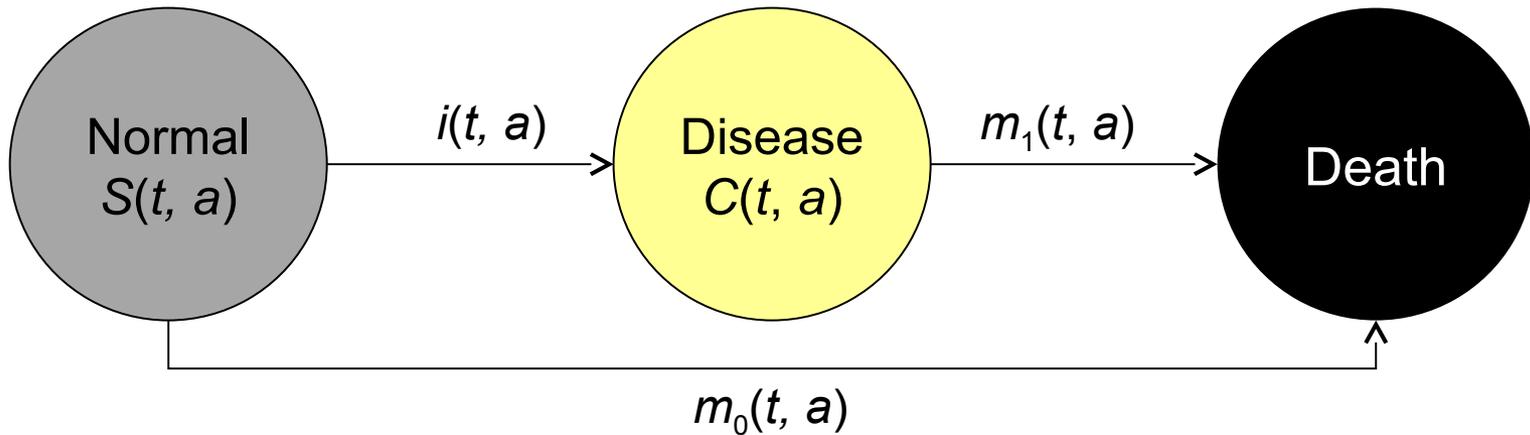


Age-specific prevalence

$$p(t, a) = \frac{C(t, a)}{S(t, a) + C(t, a)}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a} \right) p = (1 - p)[i - p(m_1 - m_0)]$$

Illness-death model für chronische Erkrankungen



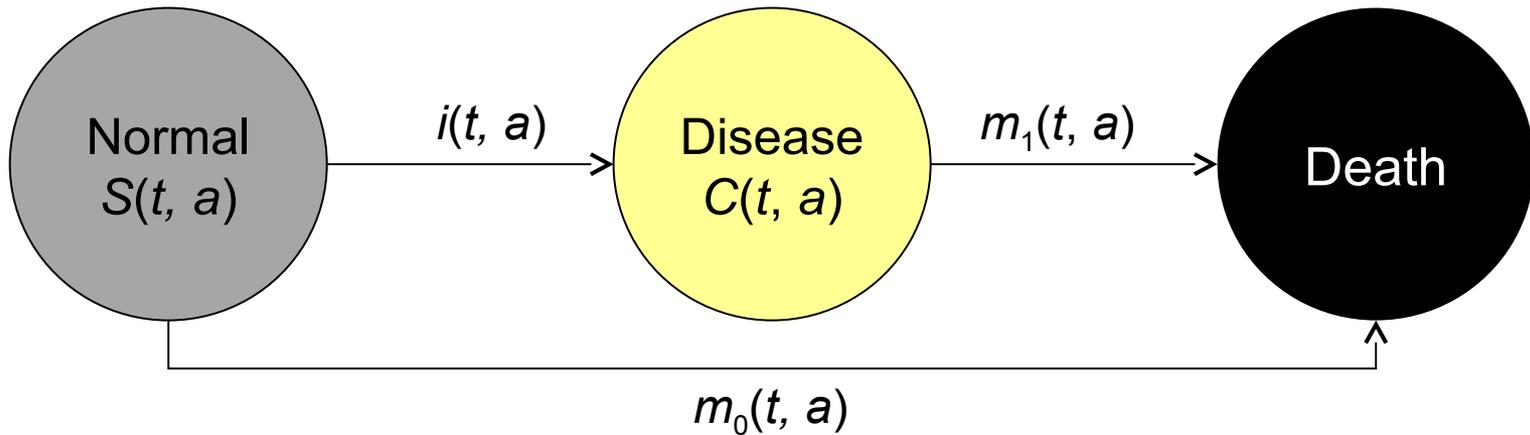
Age-specific prevalence

$$p(t, a) = \frac{C(t, a)}{S(t, a) + C(t, a)}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a} \right) p = (1 - p)[i - p(m_1 - m_0)]$$

Partial differential equation (PDE)

Illness-death model für chronische Erkrankungen

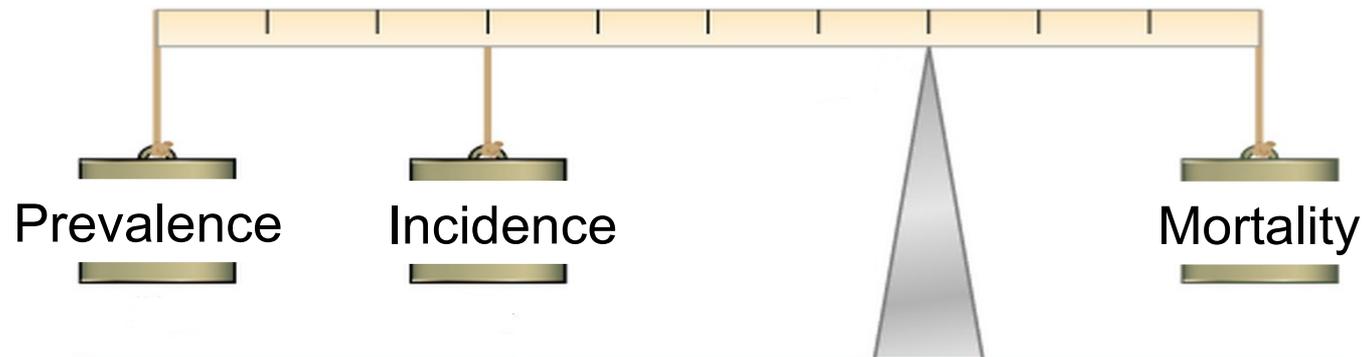


Age-specific prevalence

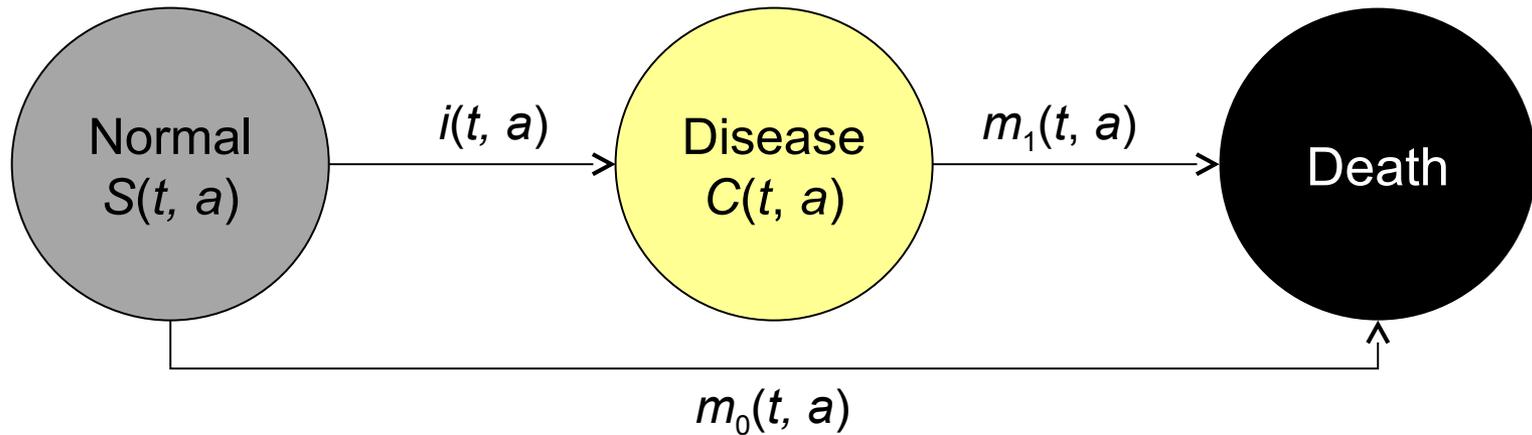
$$p(t, a) = \frac{C(t, a)}{S(t, a) + C(t, a)}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a} \right) p = (1 - p)[i - p(m_1 - m_0)]$$

Partial differential equation (PDE)



Illness-death model für chronische Erkrankungen



Age-specific prevalence

$$p(t, a) = \frac{C(t, a)}{S(t, a) + C(t, a)}$$

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a} \right) p = (1 - p)[i - p(m_1 - m_0)]$$

$$= (1 - p) \left[i - m \frac{p(HR - 1)}{1 + p(HR - 1)} \right]$$

m General mortality rate
 $HR = m_1/m_0$ Mortality rate ratio
 (Hazard ratio)

Illness-death model für chronische Erkrankungen



$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a} \right) p = (1 - p) \left[i - m \frac{p(HR - 1)}{1 + p(HR - 1)} \right]$$

Illness-death model für chronische Erkrankungen



$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial a} \right) p = (1 - p) \left[i - m \frac{p(HR - 1)}{1 + p(HR - 1)} \right]$$

Gegeben: allgemeine Sterberate (m) und Hazard-Ratio (HR)

Wie müssen **diagnostische Genauigkeit** der Prävalenz (p) und Inzidenz (i) sein, dass PDE erfüllt ist?

Was heißt diagnostische Genauigkeit?

Interpretation der Kassendaten als diagnostischer Test

Was heißt diagnostische Genauigkeit?

Interpretation der Kassendaten als diagnostischer Test

Gold standard	
Diseased	Not diseased

Was heißt diagnostische Genauigkeit?

Interpretation der Kassendaten als diagnostischer Test

Claims data	Gold standard	
	Diseased	Not diseased
Diagnosed		
Not diagnosed		

Was heißt diagnostische Genauigkeit?

Interpretation der Kassendaten als diagnostischer Test

Claims data	Gold standard	
	Diseased	Not diseased
Diagnosed	True positive	False positive
Not diagnosed	False negative	True negative

Was heißt diagnostische Genauigkeit?

Interpretation der Kassendaten als diagnostischer Test

Claims data	Gold standard	
	Diseased	Not diseased
Diagnosed	True positive	False positive
Not diagnosed	False negative	True negative

Wahre Prävalenz und Inzidenz bekommt man aus beobachteten (observed, ‚obs‘) Werten:

Was heißt diagnostische Genauigkeit?

Interpretation der Kassendaten als diagnostischer Test

Claims data	Gold standard	
	Diseased	Not diseased
Diagnosed	True positive	False positive
Not diagnosed	False negative	True negative

Wahre Prävalenz und Inzidenz bekommt man aus beobachteten (observed, ‚obs‘) Werten:

$$p = (p^{(\text{obs})} - FPR) / (se + FPR)$$

$$i = (i^{(\text{obs})} - FPR) / (se + FPR)$$



$$FPR = \Phi(se, p^{(obs)}, i^{(obs)}, m, R).$$

- Termumformungen liefern über PDE Zushg zwischen Sensitivität (se) und FPR



$$FPR = \Phi(se, p^{(obs)}, i^{(obs)}, m, R).$$

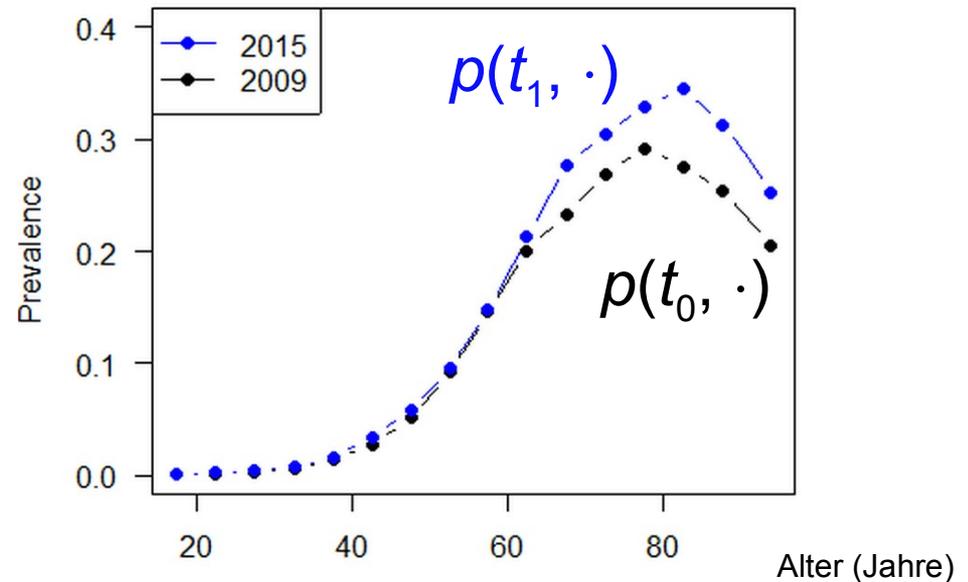
- Termumformungen liefern über PDE Zushg zwischen Sensitivität (se) und FPR
- se kennen wir nicht, wird aus Gleichverteilung ‚gezogen‘

Aggregierte Prävalenz- und Inzidenzdaten

Versorgungsatlas zu Prävalenz und Inzidenz des
Diabetes vom Zentralinstitut für Kassenärztliche
Versorgung (Zi)

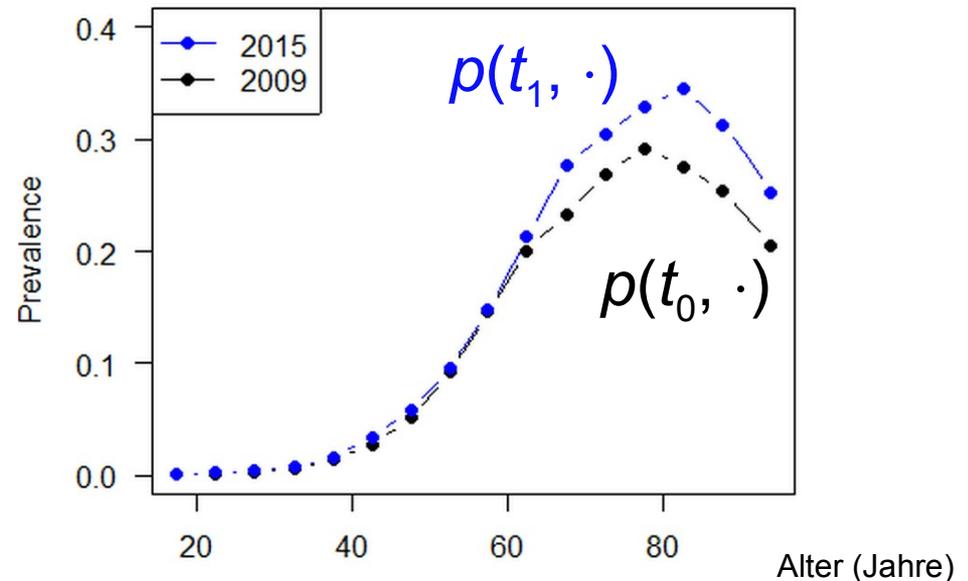
Aggregierte Prävalenz- und Inzidenzdaten

- Altersspezifische Prävalenz (p) von Typ 2 Diabetes bei Männern und Frauen in $t_0 = 2009$ und $t_1 = 2015$



Aggregierte Prävalenz- und Inzidenzdaten

- Altersspezifische Prävalenz (p) von Typ 2 Diabetes bei Männern und Frauen in $t_0 = 2009$ und $t_1 = 2015$

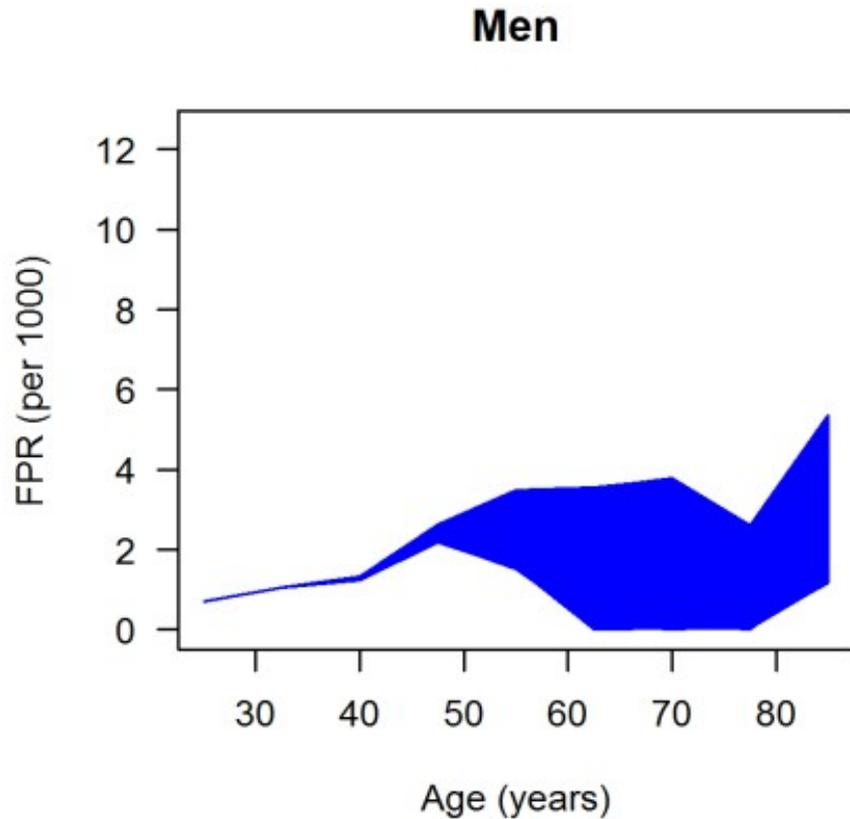


- Altersspezifische Inzidenzrate (i) 2012 bis 2014

	2012	2013	2014
0 bis unter 20 Jahre	0,02	0,02	0,02
20 bis unter 40 Jahre	0,18	0,19	0,19
40 bis unter 60 Jahre	1,26	1,25	1,21
60 bis unter 80 Jahre	2,52	2,47	2,26
80 Jahre und älter	2,20	2,14	1,95

Ergebnisse

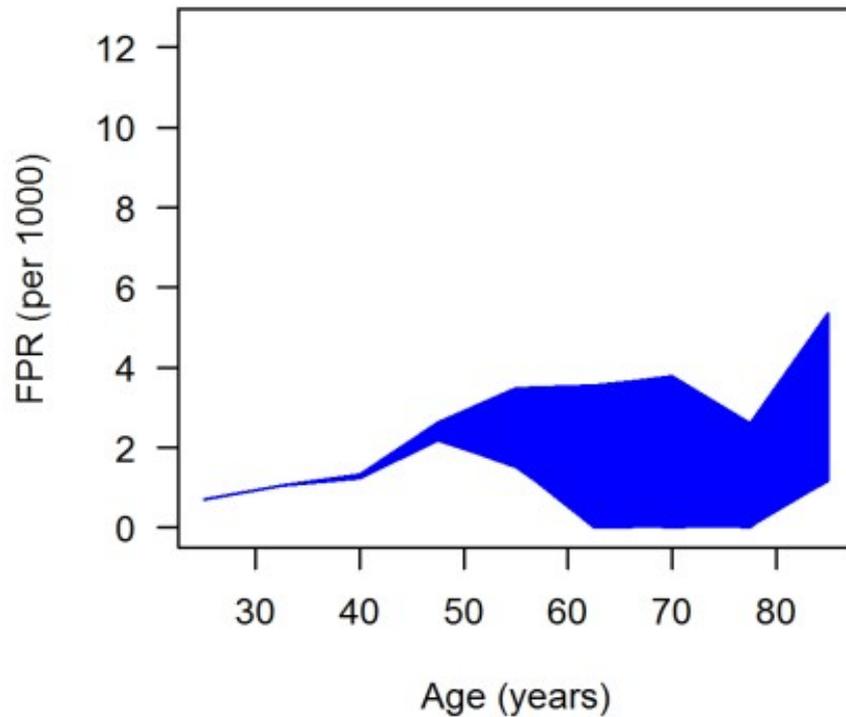
N = 100'000 Resamples der Sensitivität (se)



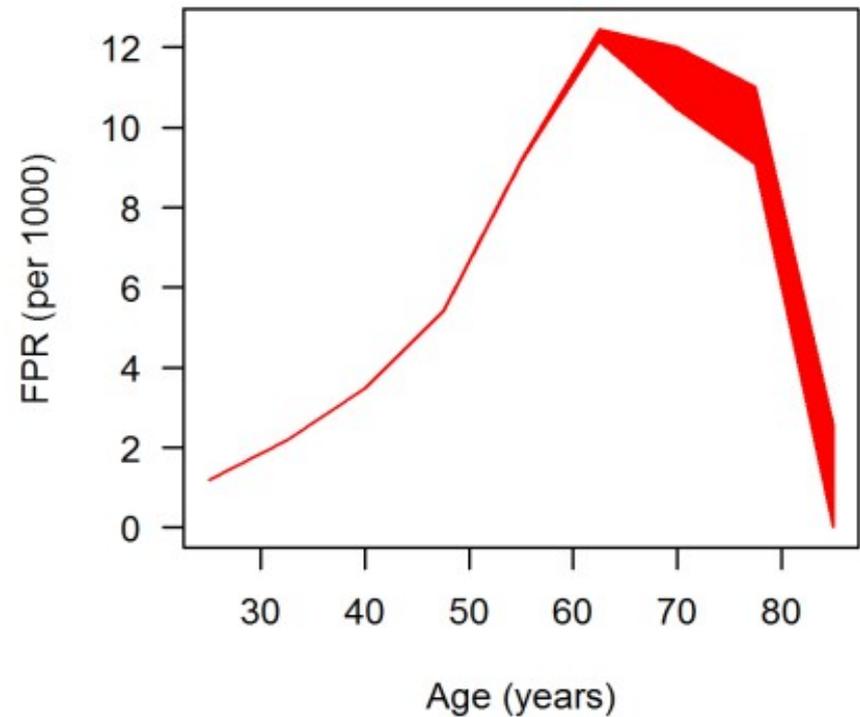
Ergebnisse

N = 100'000 Resamples der Sensitivität (se)

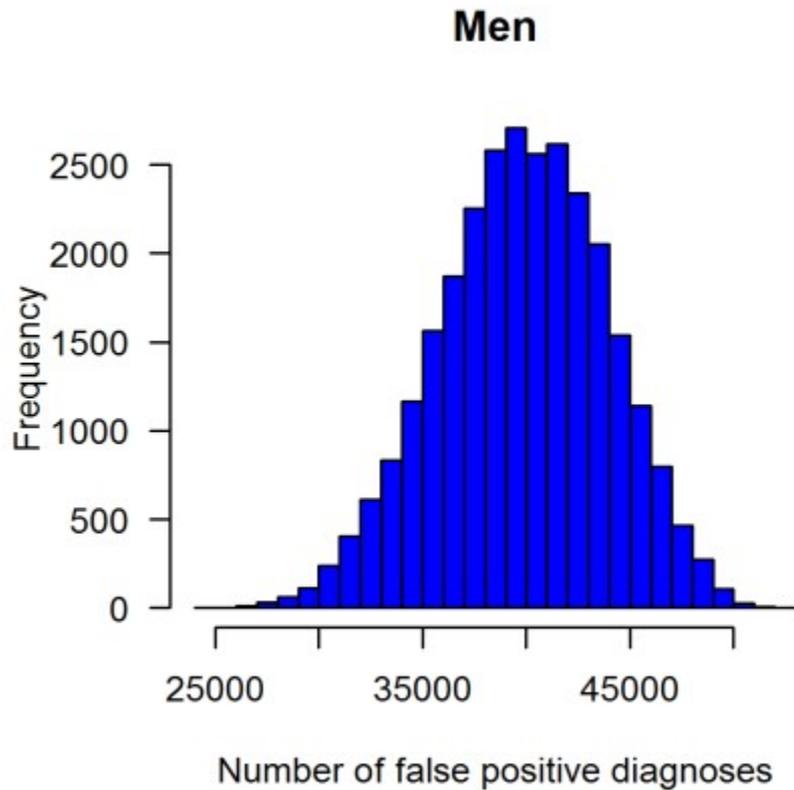
Men



Women



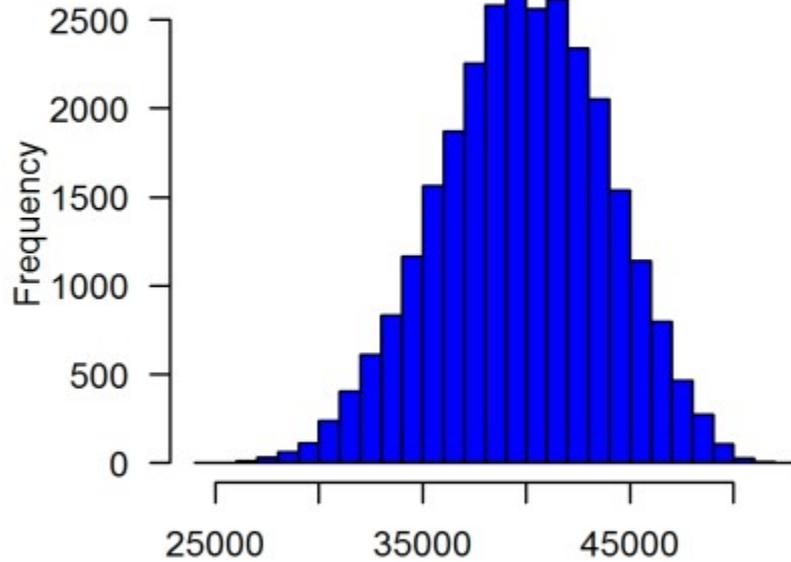
Ergebnisse



39.9	31.6 to 47.3
------	--------------

Ergebnisse

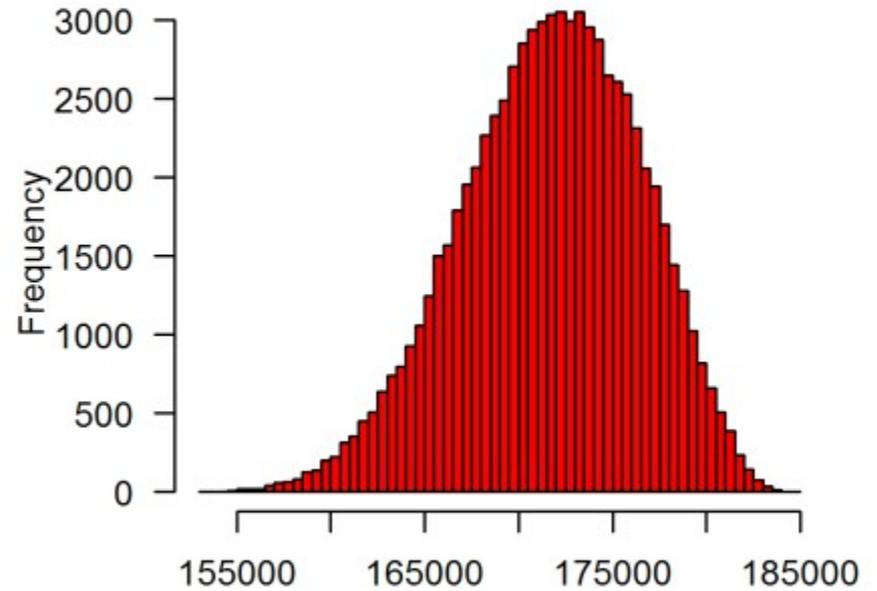
Men



39.9

31.6 to 47.3

Women



172

162 to 180

Zusammenfassung

- Damit PDE für Kassendaten erfüllt ist, müssen Fehlerterme (se , FPR) berücksichtigt werden.

Zusammenfassung

- Damit PDE für Kassendaten erfüllt ist, müssen Fehlerterme (se , FPR) berücksichtigt werden.
- Resultierender Zushg wird von FPR dominiert, was Abschätzung auch bei unbekannter se erlaubt

Zusammenfassung

- Damit PDE für Kassendaten erfüllt ist, müssen Fehlerterme (se , FPR) berücksichtigt werden.
- Resultierender Zushg wird von FPR dominiert, was Abschätzung auch bei unbekannter se erlaubt
- FPR bei Männer und Frauen unterschiedlich

Zusammenfassung

- Damit PDE für Kassendaten erfüllt ist, müssen Fehlerterme (se, *FPR*) berücksichtigt werden.
- Resultierender Zushg wird von *FPR* dominiert, was Abschätzung auch bei unbekannter se erlaubt
- *FPR* bei Männer und Frauen unterschiedlich
- Rund 4x so viele falsch-positive Diagnosen bei Frauen als bei Männern

Zusammenfassung

- Damit PDE für Kassendaten erfüllt ist, müssen Fehlerterme (se, *FPR*) berücksichtigt werden.
- Resultierender Zushg wird von *FPR* dominiert, was Abschätzung auch bei unbekannter se erlaubt
- *FPR* bei Männer und Frauen unterschiedlich
- Rund 4x so viele falsch-positive Diagnosen bei Frauen als bei Männern
- Insgesamt sind etwa 210'000 (von >6 Mio) Typ 2 Diagnosen in Zi-Daten falsch-positiv

Contact

eMail: `Ralph.Brinks@Uni-WH.de`



`@RBrinks`