

Standards der Daten- & Ergebnispräsentation

Virtueller ADT Workshop 22.4.2021

Prof. Dr. med. Andreas Stang, MPH

imibe.dir@uk-essen.de

**Institut für Medizinische Informatik,
Biometrie und Epidemiologie**

Medizinische Fakultät

UNIVERSITÄT
DUISBURG
ESSEN

Open-Minded

**Landeskrebsregister
Nordrhein-Westfalen**

LKR | NRW

Versorgungsforschung

Nutzung von evidenzbasiertem Wissen zu Diagnostik, Therapie und Prognostik

Evaluative Forschung zur Beurteilungen von Über/Unterlegenheit/Gleichwertigkeit verschiedener Ansätzen

Vergleichende Statistiken nach Versorger- oder regionalen Charakteristiken

Bei Unterschieden der Güte der Versorgung (Alarm): gibt es erklärende Variablen?

Sind nach Adjustierung für erklärende Variablen die Unterschiede plausibilisiert?

Vergleichende Statistiken in Bezug auf Diagnose, Therapie, Prognose

Confounding by indication & Selektionsbias

Sensitivitätsanalyse für Confounding & Selektionsbias

Evaluative Forschung mit KR-Daten

- (1) Welches Patientenkollektiv soll untersucht werden?
- (2) Welcher Einflussfaktor soll betrachtet werden?
- (3) Welches Outcome soll betrachtet werden?
- (4) Mit welchen Störgrößen ist zu rechnen?
- (5) Mit welchen Modifikatoren wird gerechnet?

Für die Punkte (1)-(5) bedarf es exakter Definitionen & Operationalisierungen

Beispiel:

Definition

Therapie: radikale Hysterektomie zur chirurgischen Primärbehandlung eines Endometriumkarzinoms

Operationalisierung: OPS-Codesim Zusammenhang mit...

Probleme der Forschung mit KR-Daten

Datensätze von Patienten

- Dubletten?
- Bevölkerungs- oder eher nur Klinikbezug?
- Umgang z.B. mit MV(-), DCO
- Register-abhängige Unterschiede der Kodierungen?
- Diagnose- & Follow-up-Zeiträume?

Variablen von Interesse (Exposition, Confounder, Modifikator, Outcome)

- Best-of Definitionen der Variablen
- Fehlende Werte & Missing-Mechanismen?
- Fehlerbehaftete Werte?
- Implausible Werte?
- Kategorisierung von Variablen
- Transformation von Variablen

Fluss-Diagramm: Brutto-, Netto- & Analyse-N

Brutto N

Ausschlüsse 1 [neutrale Ausfälle]

- Grund A: $n = \dots(\%)$
- Grund B: $n = \dots(\%)$
- Grund C: $n = \dots(\%)$

Netto N

Ausschlüsse 2 [nicht-neutrale Ausfälle]

- MV(-): $n = \dots(\%)$
- DCO: $n = \dots(\%)$
- Fehlendes Geschlecht: $n = \dots(\%)$

Analyse N

Gegebenfalls Vergleich von Merkmale ausgeschlossener und eingeschlossener Patienten in die Analysen bei „großen“ Mengen von Ausschlüssen

Statistik: Null Hypothesen Testen

Neyman-Pearson Orthodoxie

- (1) Festlegung H_0 & H_1
- (2) Festlegung Fehler 1. (α) und 2. Art (β)
- (3) Festlegung Test-Statistik (z.B. t-Test, Mann-Whitney-Test etc.)
- (4) Fallzahlkalkulation auf Grundlage (1)-(3)
- (5) Durchführung der Analysen mit der Fallzahl, Berechnung der Test-Statistik & p-Wert
- (6) Vergleich p-Wert mit α
- (7) Eiskalte Entscheidung: Ablehnung von H_0 oder nicht

Die Unterscheidung in $p \leq 0,001^{***}$, $p \leq 0,01^{**}$, $p \leq 0,05^{*}$ hat nichts mit Neyman-Pearson zutun

Notwendige Annahmen: Neyman-Pearson

- (1) Es wurde eine **geeignete Teststatistik** angewandt.
- (2) Es kam zu **keinem Selektionsbias** in die Studie oder zu einem Selektionsbias bei der Nachbeobachtung der Patienten.
- (3) Der Vergleich zwischen zwei Stichproben ist **nicht konfundiert** (keine Vermengung von Effekten).
- (4) Es gibt **keinen Informationsbias** bei der Messung der Variablen in der Studie.

6 Aussagen der American Statistical Association

- (1) P-Werte können angeben, wie unvereinbar die Daten mit einem bestimmten statistischen Modell sind.
- (2) P-Werte messen nicht die Wahrscheinlichkeit, dass die untersuchte Hypothese wahr ist, oder die Wahrscheinlichkeit, dass die Daten allein durch Zufall entstanden sind.
- (3) Wissenschaftliche Schlussfolgerungen und gesundheitspolitische Entscheidungen sollten nicht nur darauf beruhen, ob ein p-Wert einen bestimmten Schwellenwert überschreitet.
- (4) Eine korrekte Inferenz erfordert eine vollständige Berichterstattung und Transparenz.
- (5) Ein p-Wert, oder statistische Signifikanz, misst nicht die Größe eines Effekts oder die Wichtigkeit eines Ergebnisses.
- (6) Ein p-Wert an sich ist kein gutes Maß für die Evidenz eines Modells oder einer Hypothese.

Moderne (Krebs)Epidemiologie

“In vielen Fällen erfordert medizinische Literatur keine feste Entscheidung: sie trägt inkrementell zu einem bestehenden Wissensbestand bei” (Sterne & Davey Smith, 2001)

"Epidemiologische Forschung ist eine Übung im Messen. Ihr Ziel ist es, eine gültige und präzise Schätzung entweder des Auftretens von Krankheiten in einer Population oder der Wirkung einer Exposition auf das Auftreten von Krankheiten zu erhalten." (Lash, 2007)

Sterne JAC, Davey Smith G. Sifting the evidence – what’s wrong with significance tests? *BMJ* 2011;322:226-231

Lash TL: Heuristic thinking and inference from observational epidemiology. *Epidemiology* 2007;18:67-72

Notwendige Bedingung: Relatives Survival

Prinzip des relativen Survivals

Das beobachtete Survival wird der (alters-, geschlechts-, regionen- und kalenderjahrspezifischen) erwarteten Lebenserwartung gemäß der allgemeinen Sterbetafel gegenübergestellt.

Der Quotient aus beobachteten und erwartetem Survival entspricht dem relativen Survival und quantifiziert die Krebserkrankung-assoziierte Übersterblichkeit, ohne dass man Todesursachen des Leichenschauscheines verwenden musste.

Grundbedingung

Das Patientenkollektiv ist ein repräsentatives Abbild der Krebserkrankten der Bevölkerung, aus der die Sterbetafel-Information zur Berechnung des relativen Survivals gezogen wird.

Beispiel: Uvealmelanome, Universitätsklinikum Essen

Jan 2007 – Dez 2020: 5580 Uvealmelanome behandelt,
durchschnittlich 399 Patienten/Jahr (SD 13.9) behandelt

Zuweisung aus allen Bundesländern außer BE, BB & MV

Therapien

- Primäre Enukleation 20.4%
- Episklerale Plaquetherapie 67.2%
- Endoresektion oder transsklerale Resektion 7.1%
- Protonenbestrahlung 5.3%

Sterbetafel für BRD ?

männlich

Vollendetes Alter	Durchschnittliche Lebenserwartung in Jahren						
	2004/2006	2006/2008	2008/2010	2010/2012	2012/2014	2013/2015	2014/2016
0	76,17	76,71	77,08	77,40	77,81	77,88	78,00
1	75,57	76,09	76,43	76,74	77,13	77,21	77,33
2	74,60	75,12	75,46	75,76	76,16	76,23	76,36
3	73,62	74,14	74,48	74,78	75,17	75,25	75,37
4	72,64	73,15	73,49	73,79	74,19	74,26	74,38

Lever M et al. Demographic and clinical characteristics of uveal melanoma patients in a tertiary German ophthalmic oncology center between 2007 and 2020. DOG Jahrestagung 2021 (submitted)

Daten Modellierung

Kausale Modellierung

Def.: Optimierung der Schätzung der Änderung des Outcomes aufgrund der Veränderung einer Einflussvariablen

Q: Wie stark ist der Einfluss der Chemotherapie X auf die Überlebenswahrscheinlichkeit bei Brustkrebs?

Prädiktions-Modellierung

Def.: Optimierung der Vorhersage des Outcomes eines jeden Individuums

Q: Welche Krebspatienten haben die höchste/niedrigste Überlebenswahrscheinlichkeit?

Daten Modellierung

Kausale Modellierung

Variablenselektion

Minimal suffizientes
Adjustierungsset von
Konfoundern (DAG);
Vermeidung der Adjustierung
von Collidern oder Mediatoren

Modellevaluation

DAG impliziert bestimmte
statistische Unabhängigkeiten
zwischen Variablen, die
überprüft werden können

Prädiktions-Modellierung

Variablenselektion

automatisch bzw.
algorithmisch (Forwards,
Backwards, Stepwise,
Change in Estimate...)

Modellevaluation

Goodness of Fit Kriterien,
adj. R^2 , ROC

Daten Modellierung

Kausale und Prädiktions-Modellierungen unterscheiden sich in 5 Punkten:

- (1) Welche Covariaten überhaupt in Betrachtung gezogen werden sollten.
- (2) Wie ein geeignetes Set von Covariaten dann bestimmt wird.
- (3) Welche Covariaten dann ultimativ ins Modell kommen und in welcher Form.
- (4) Wie das Modell evaluiert wird.
- (5) Wie das Modell interpretiert wird.