

Nutzen diagnostischer Tests – vom Surrogat zur Patientenrelevanz

Dr. Sandra Janatzek

Fachbereich Evidenzbasierte Medizin

Medizinischer Dienst des Spitzenverbandes Bund
der Krankenkassen (MDS)

Essen

Untersuchung des Nutzens ist notwendig

➤ Deeks et al. (2001):

clinical value.^{23 24} Treatment interventions are recommended for use in health care only if they are shown on average to be of benefit to patients: the same criterion should also be applied for the use of a diagnostic test, and even the most accurate of tests can be clinically useless or do more harm than good. It should

(BMJ 2001; 323:157-162)

➤ Lijmer & Bossuyt (2002):

From the patient's perspective the contribution of diagnostic tests to a better health outcome is of primary importance, rather than their correspondence with the "truth".

(The Evidence Base of Clinical Diagnosis, BMJ Books 2002)

Phaseneinteilung für Diagnosestudien

z.B. von **Fryback & Thornbury** (1991)
(Übersetzung nach: Nachtnebel 2010)

Diagn. Validität

Level 1:

Technische Qualität

(→ Schärfe, Kontrast, Machbarkeit, Reliabilität, ...)

Level 2:

Diagnostische Validität

(→ Sensitivität, Spezifität, prädiktive Werte, ...)

Verändert der Testbefund Entscheidungen?

Level 3:

Diagnostischer Impact

(→ Auswirkungen des Testbefundes auf Diagnosefindung)

Level 4:

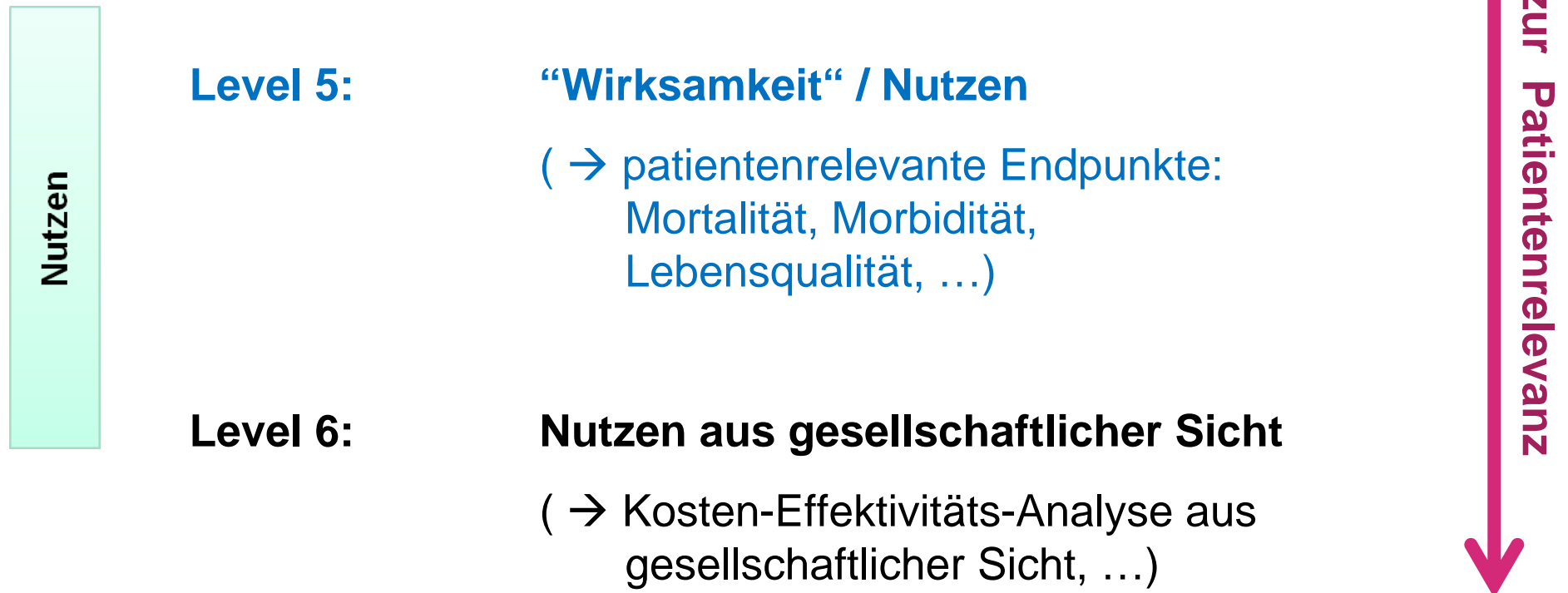
Therapeutischer Impact

(→ Einfluss des Testbefundes auf therapeutisches Vorgehen)

vom Surrogat zur ...



Phaseneinteilung für Diagnosestudien (Fortsetzung)



Fryback & Thornbury, Med Decis Making 1991;11: 88-94.

Nachtnebel, A. Evaluation diagnostischer Technologien – Hintergrund, Probleme, Methoden.
HTA-Projektbericht 2010; Nummer 36.

Dr. Sandra Janatzek, MDS

Verfahrensordnung des G-BA

Nutzen

(2) Bei der Klassifizierung der Unterlagen zu diagnostischen Methoden gelten folgende Evidenzstufen:

I a Systematische Übersichtsarbeiten von Studien der Evidenzstufe I b

I b Randomisierte kontrollierte Studien

I c Andere Interventionsstudien

Diagn. Validität

II a Systematische Übersichtsarbeiten von Studien zur diagnostischen Testgenauigkeit der Evidenzstufe II b

II b Querschnitts- und Kohortenstudien, aus denen sich alle diagnostischen Kenngrößen zur Testgenauigkeit (Sensitivität und Spezifität, Wahrscheinlichkeitsverhältnisse, positiver und negativer prädiktiver Wert) berechnen lassen

III Andere Studien, aus denen sich die diagnostischen Kenngrößen zur Testgenauigkeit (Sensitivität und Spezifität, Wahrscheinlichkeitsverhältnisse) berechnen lassen

IV Assoziationsbeobachtungen, pathophysiologische Überlegungen, deskriptive Darstellungen, Einzelfallberichte, u. ä.; nicht mit Studien belegte Meinungen anerkannter Expertinnen und Experten, Berichte von Expertenkomitees und Konsensuskonferenzen

(7) Bei der Bewertung sollen insbesondere auch die Verwendung patientenrelevanter Zielgrößen (wie z. B. Mortalität, Morbidität, Lebensqualität) ...

Konsequenz:

Für die Bewertung des Nutzens sind notwendig:

Das ist der
Regelfall

➤ **Level-5-Diagnostestudien**

direkte
Evidenz

oder

Das ist die
Ausnahme

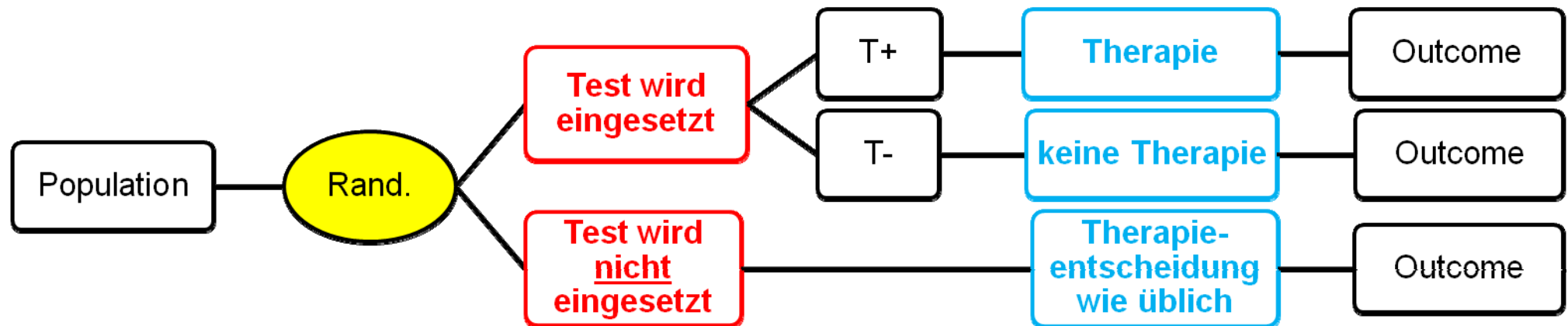
➤ **andere Wege, um von der
diagnostischen Validität
auf den Nutzen zu schließen
(„linked evidence“)**

indirekte
Evidenz

Designs für Level-5-Diagnostestudien

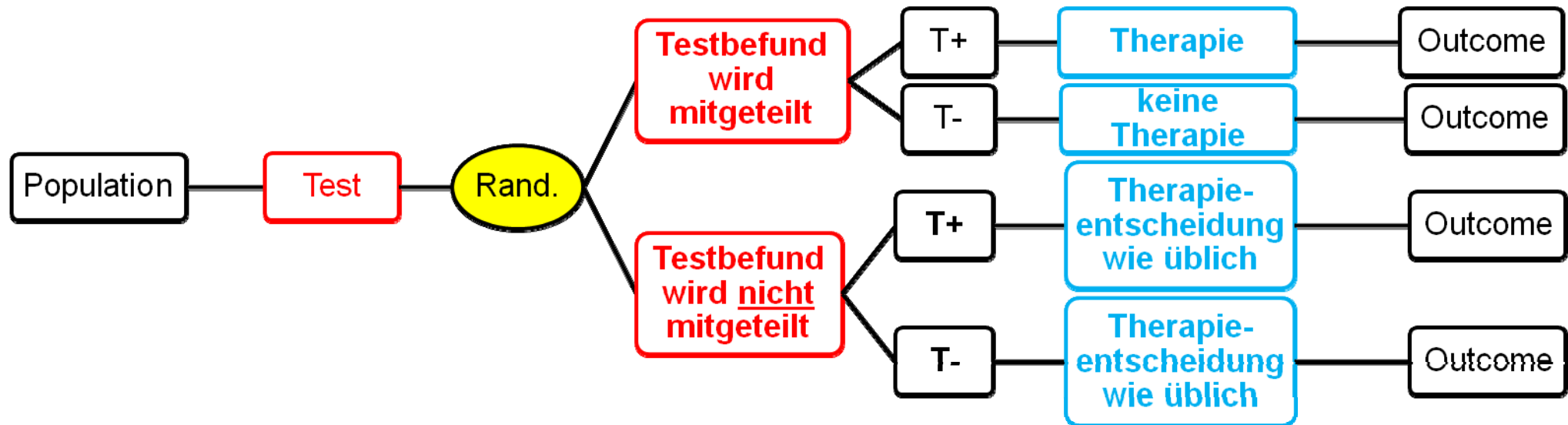
(Lijmer & Bossuyt 2009)

1) Am häufigsten verwendet: „RCT of testing“

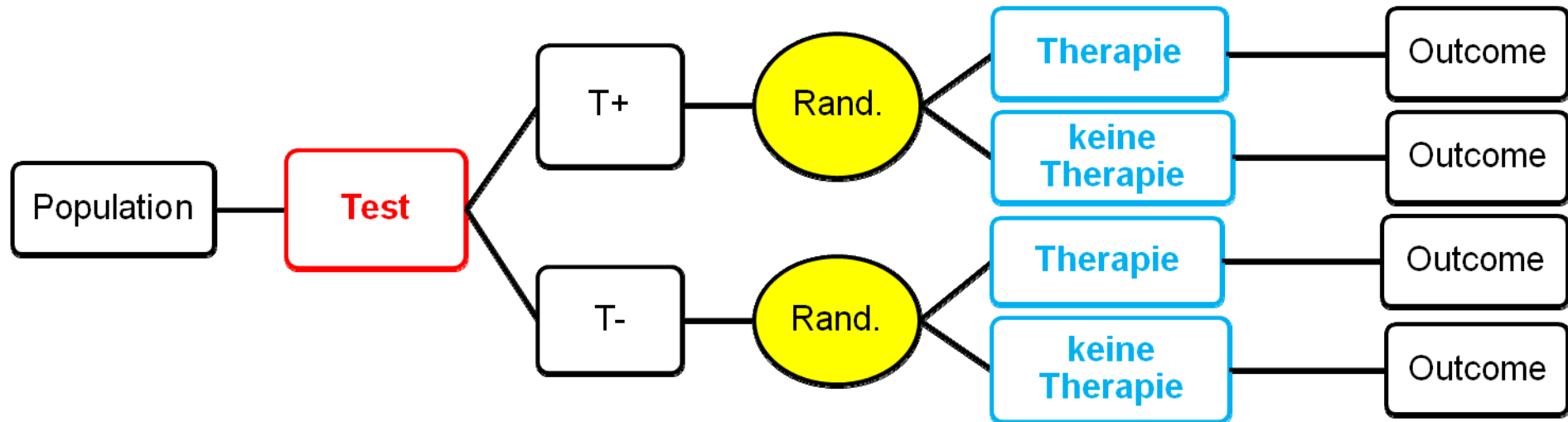


- Bewertung des Nutzens der „diagnostischen Kette“ (auf RCT-Ebene)
- Bei negativem Ausgang: unklar, ob Test insuffizient oder Therapie unwirksam
- Bei positivem Ausgang: unklar, ob auch T- von Therapie profitieren könnten

2) „Random disclosure design“ (Modifikation von 1)



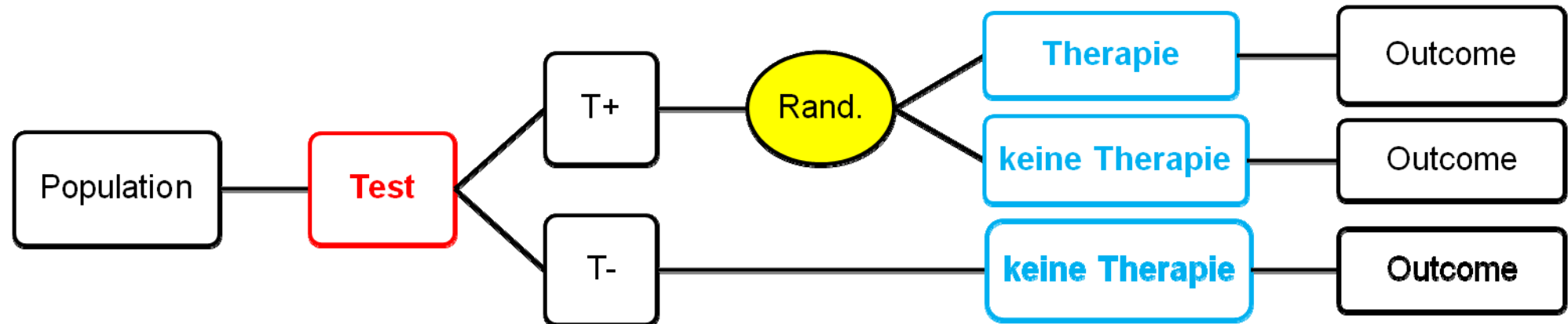
3) „Testing at baseline in RCT“



- Antwort auf Frage,
 - ob T+ stärker von Therapie profitieren als T- (Wechselwirkung)
 - ob auch T- von Therapie profitieren
 - wie viel die Therapie bei T+ bringt
- Keine Bewertung des Nutzens der „diagnostischen Kette“ (auf RCT-Ebene)

4) „RCT of test positives“

Modifikation von 3 für Situation: unstrittig, dass T- unbehandelt bleiben sollen



- Antwort auf Frage, ob und wie viel T+ von Therapie profitieren
- Keine Bewertung des Nutzens der „diagnostischen Kette“ (auf RCT-Ebene)

- ⇒ Level-5-Diagnostestudien sind machbar.
- ⇒ Es gibt mehrere RCT-Designs. Geeignete Wahl hängt vom Kontext ab.

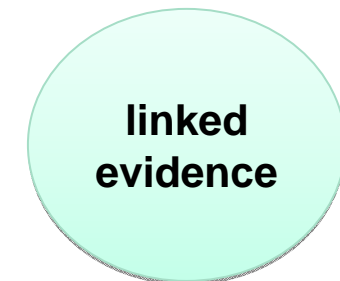
Werden immer Level-5-Studien gebraucht oder kann die diagnostische Validität ausreichen?

JA, sie kann unter gewissen Voraussetzungen ausreichen:

- Wenn neuer Test **nur Modifikation** eines etablierten Tests mit nachgewiesenem Nutzen
- **Diagn. Validität** ist durch Level-2-Studie(n) belegt



Nutzen der Therapie **bei den T+**
(Referenzverfahren oder zu prüfender Test)
ist durch Therapiestudie(n) belegt



Linked Evidence

- Kann häufig – aufgrund der Voraussetzungen – nicht angewendet werden (Nachtnebel 2010)
- Aussagekraft geringer als bei direkter Evidenz aus RCT (Lord et al. 2009)
- Schönemann et al. (2008):

Nevertheless, in the GRADE system, valid studies of test accuracy also start as high quality in the diagnostic framework. Such studies are, however, vulnerable to limitations and often provide low quality evidence for recommendations as a result of the indirect evidence they usually offer on impact on patient-important outcomes.

Zusammenfassung

- **Nutzen-Beleg** muss gefordert werden, um Schaden von Patienten abzuwenden
- Dazu werden i.R. **Level-5-Studien** benötigt, aber nicht immer.
- In Ausnahmefällen können Level-2-Studien (in Kombination mit weiteren Informationen) genügen.
- Methodik für Ausnahmefälle (**linked evidence**) muss weiterentwickelt werden.
- **Wünschenswert:** Studienmacher von Level-2-Studien legen dar, ob (und falls ja, warum) aus diagnostischer Validität – OHNE Level-5-Studie – auf Nutzen geschlossen werden kann.
- Beim Design von Level-5-Studien kann **mehr Kreativität** effizientere Studien ermöglichen und mehr Erkenntnis liefern.