

KiMeKo: KI-Med-Kollaborationsplattform

Große Sprachmodelle für die Extraktion von
kausalem Wissen aus medizinischen Texten

Robin Nicolay, Sebastian Bader, Thomas Kirste

Mobile Multimediale Informationssysteme
Institut für Visual & Analytic Computing
Fakultät für Informatik und Elektrotechnik

KiMeKo: KI-Med-Kollaborationsplattform

- **Projektziel:**

- Aufbau eines KI-Ökosystems zur Unterstützung der Entwicklung KI-gestützter Medizintechnologien in Norddeutschland.

- **Hintergrund:** KI SIGS

- Vorgängerprojekt KI SIGS: Konzeption eines „KI-Space“ für intelligente Gesundheitssysteme.
- Bisherige Partner: Norddeutsche KI-Institute in Kooperation mit führenden medizinischen Unternehmen und Universitätskliniken.

- **Unsere Rolle als neuer Partner im KiMeKo-Projekt**

- Extraktion kausaler Zusammenhänge aus medizinischen Texten (Leitlinien, Berichte, ...)
- Umsetzung eines kausalen Alignments KI-basierter Expertensysteme an medizinischen Erkenntnisse
- Schließen einer Lücke zwischen LLM-basierter KI-Technologie und der praktischen Anwendung in der Medizin.



Bundesministerium
für Bildung
und Forschung

Universität
Rostock



Traditio et Innovatio

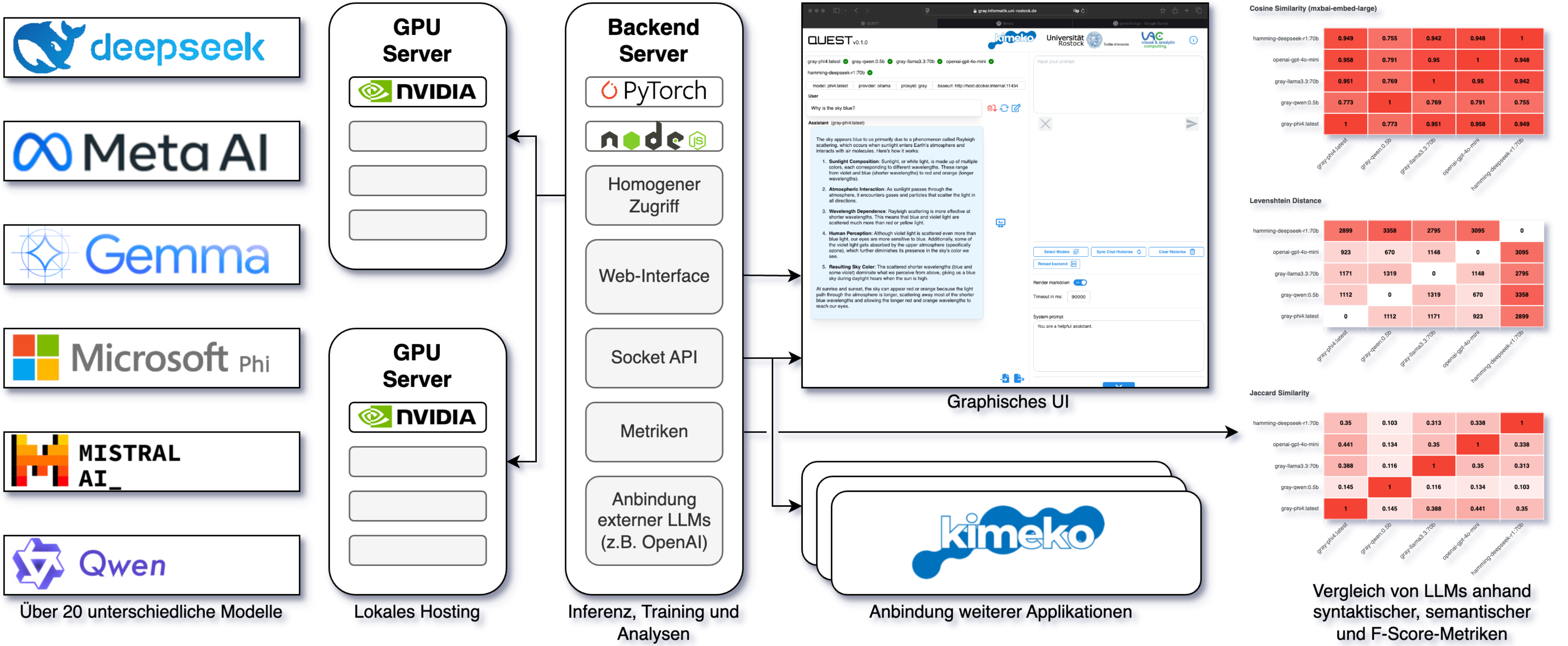
TUHH
Technische
Universität
Hamburg



Universität
Bremen



QUEST: Lokale Trainingsarchitektur und Evaluationsframework zur Quantitativen Untersuchung und Evaluation von Sprachmodellen auf Textbasis

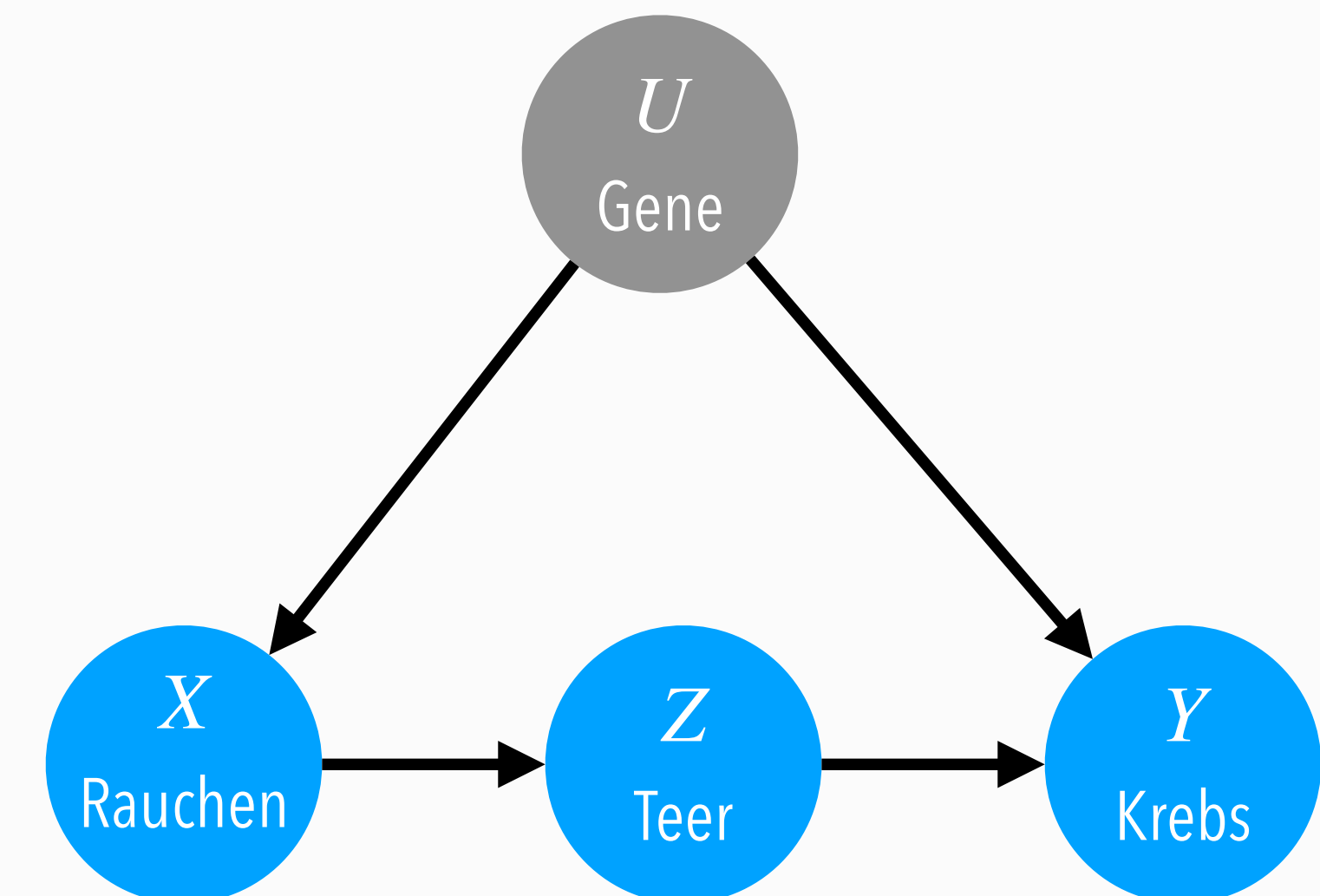
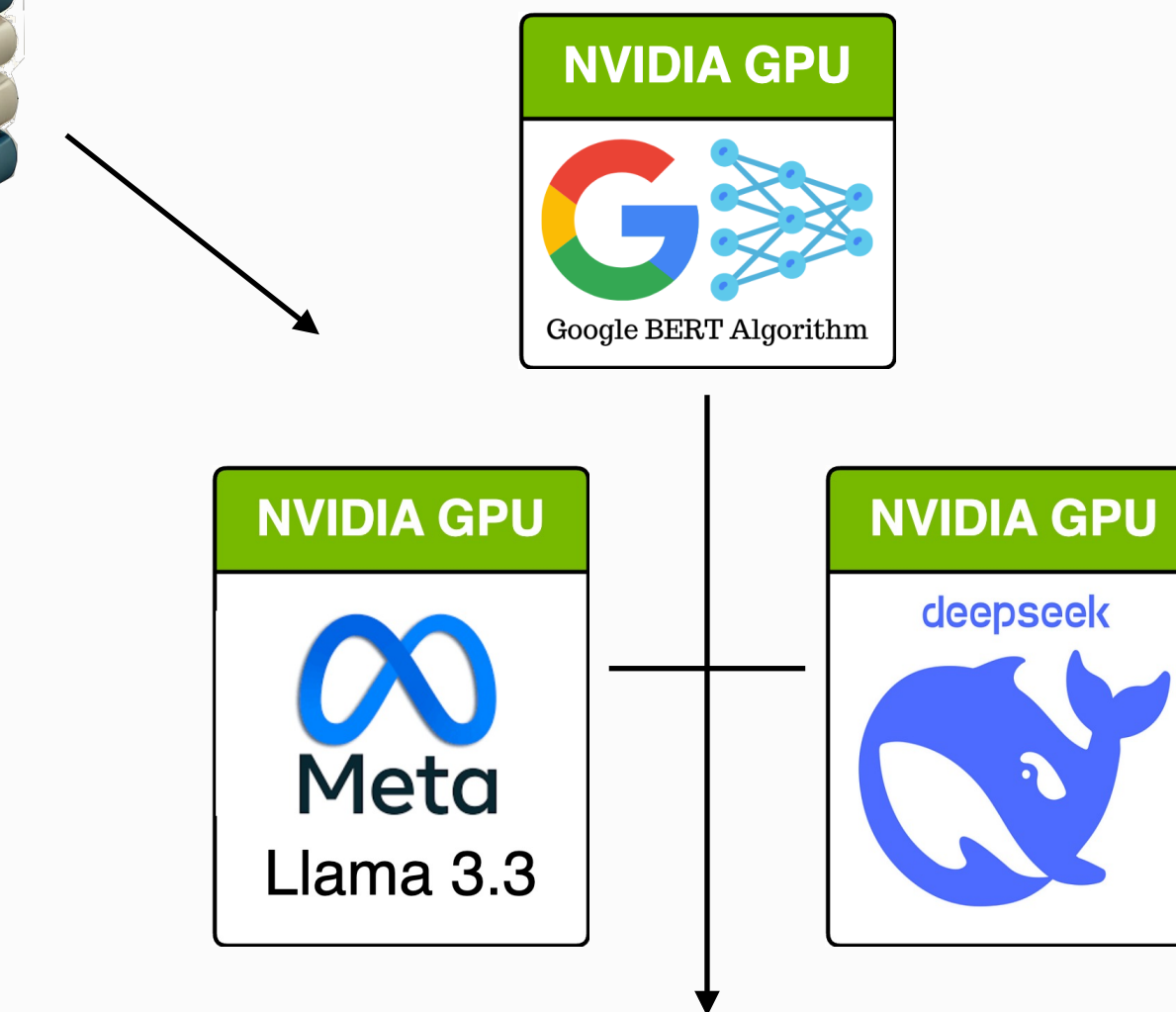


• KiMeKo Projektziele

- Bestimmung der Fähigkeit von LLMs zum kausalen Verständnis medizinischer Fachliteratur
- Training von LLMs zur Extraktion kausaler Graphen aus med. Texten (Leitlinien, Berichten, ...)
- Kausales Alignment medizinischer KI-Expertensysteme anhand kausaler Modelle.
- Integration von Methoden ins KI-Med Ökosystem

• Unser Beitrag

- Know-how zu Aufbau, Installation, lokalem Hosting und Training verschiedener Schlüsselmodelle wie Llama3.3 und DeepSeek-R1.
- Konzeption und Umsetzung von Werkzeugen und Pipelines zum Training und zur Bestimmung von LLM-Fähigkeiten (z. B. Cos-Embedding in Quest, F1-Scores von BERT-Trainings)



Warum eigentlich?

- LLM-Unterstützung in der medizinischen Versorgung

- Unterstützt bei schnellen und präzisen Diagnosen
- Unterstützt bei objektiven Therapieentscheidungen
- Entlastet das medizinischen Personal
- **Aber:** Im unbeaufsichtigten Einsatz können fehlinterpretierte Kontexte und Halluzinationen zu kritischen Fehlern führen.



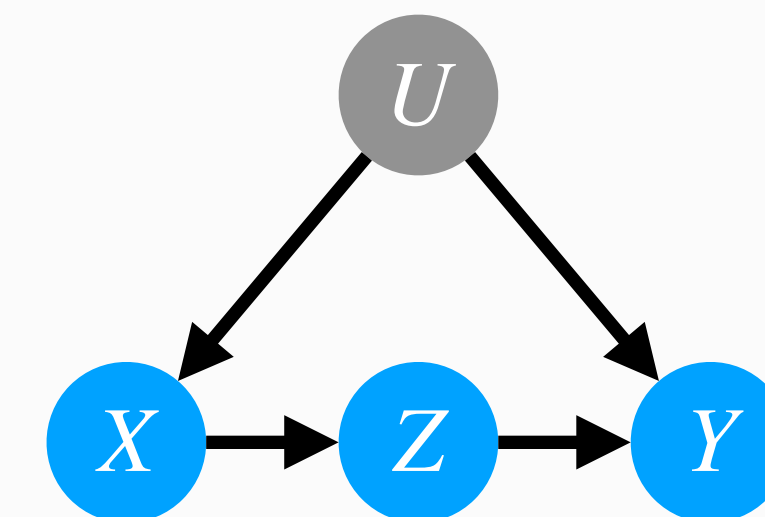
Warum eigentlich?

- LLM-Unterstützung in der medizinischen Versorgung

- Unterstützt bei schnellen und präzisen Diagnosen
- Unterstützt bei objektiven Therapieentscheidungen
- Entlastet das medizinischen Personal
- **Aber:** Im unbeaufsichtigten Einsatz können fehlinterpretierte Kontexte und Halluzinationen zu kritischen Fehlern führen.

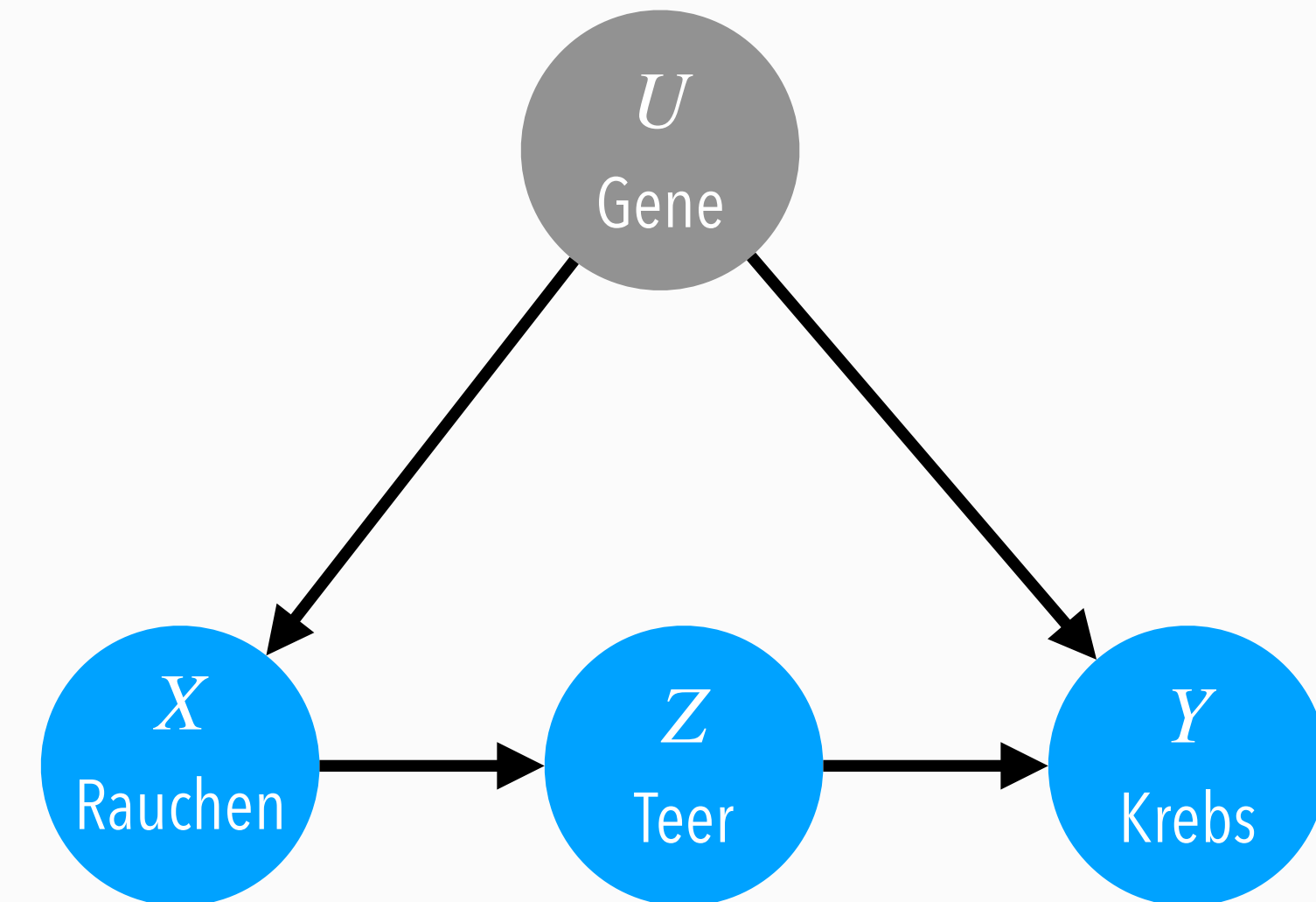
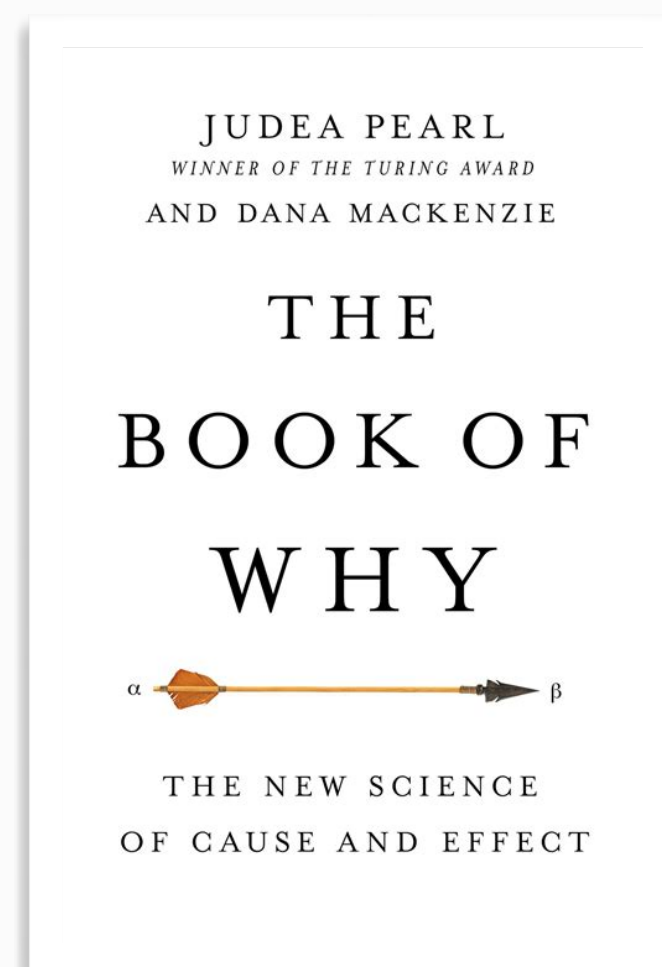
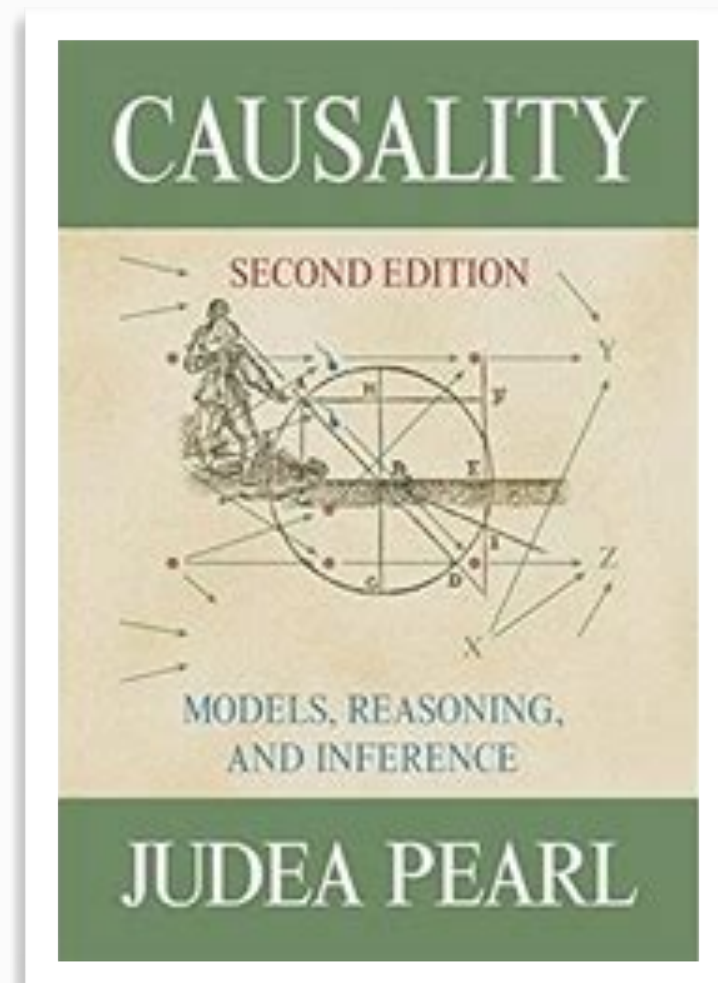
- Extraktion von Kausalbeziehungen aus medizinischen Texten könnte:

- die quellenübergreifende automatische Entdeckung von Kausalzusammenhängen ermöglichen,
- die Synthese neuer Hypothesen über Krankheitsursachen und Interventionswirkungen erlauben,
- durch Generierung kausaler Diagramme den Nutzen von Beobachtungsstudien und vorhandenen Daten sichern.



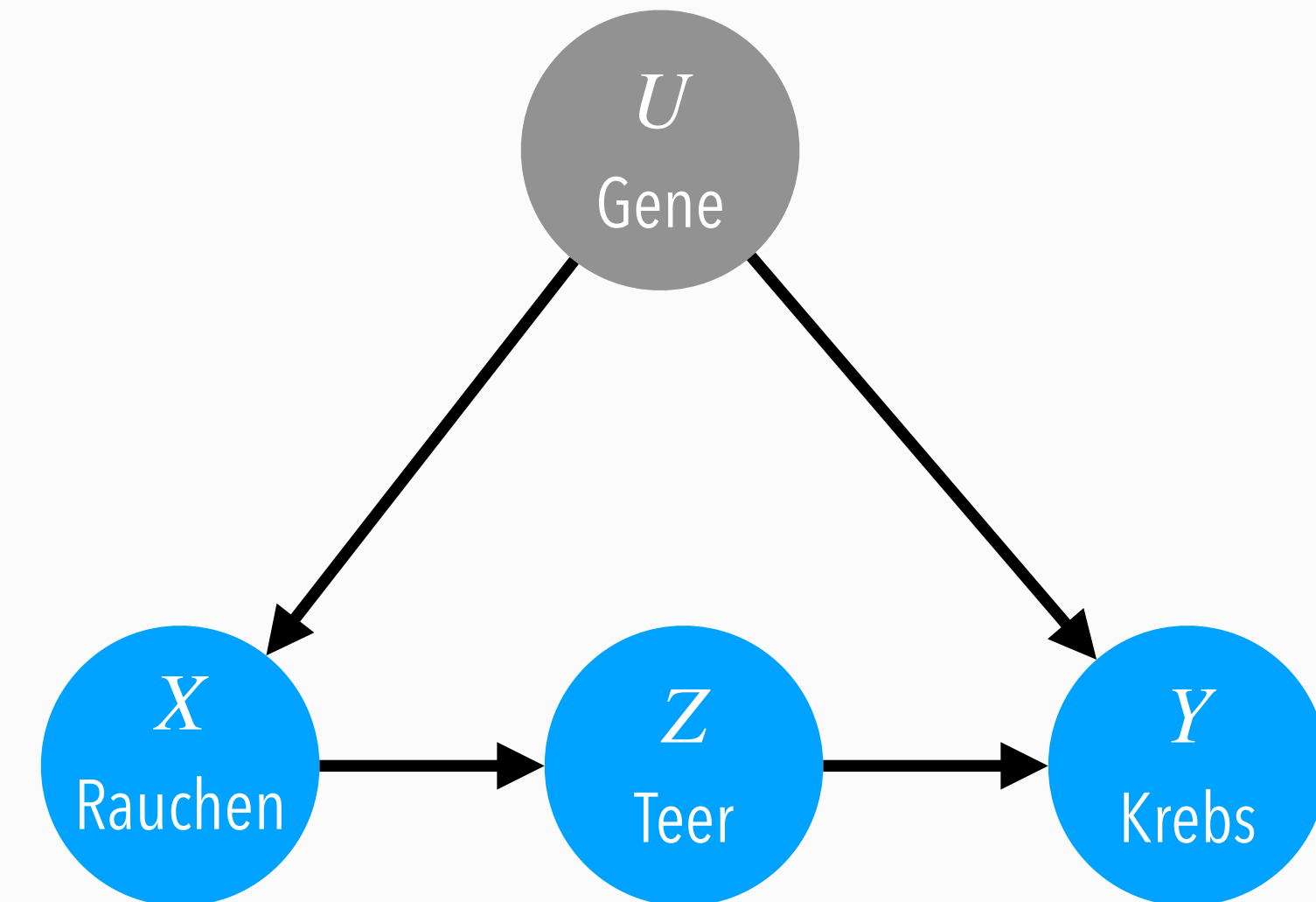
Übergewicht bewirkt **hormonelle Ungleichgewichte**, die kausal zu **erhöhter Insulinresistenz** und damit zu **Schwangerschaftsdiabetes** führen.

- Kausales Diagramm:
Wie wirken Variablen aufeinander ein?
- Hintergrund:
 - *Do-Kalkül*: Judea Pearl „Causality“, 2nd ed., 2009.
(Gewinner des Turing-Awards 2011,
Erfinder der Bayesschen Netze.)

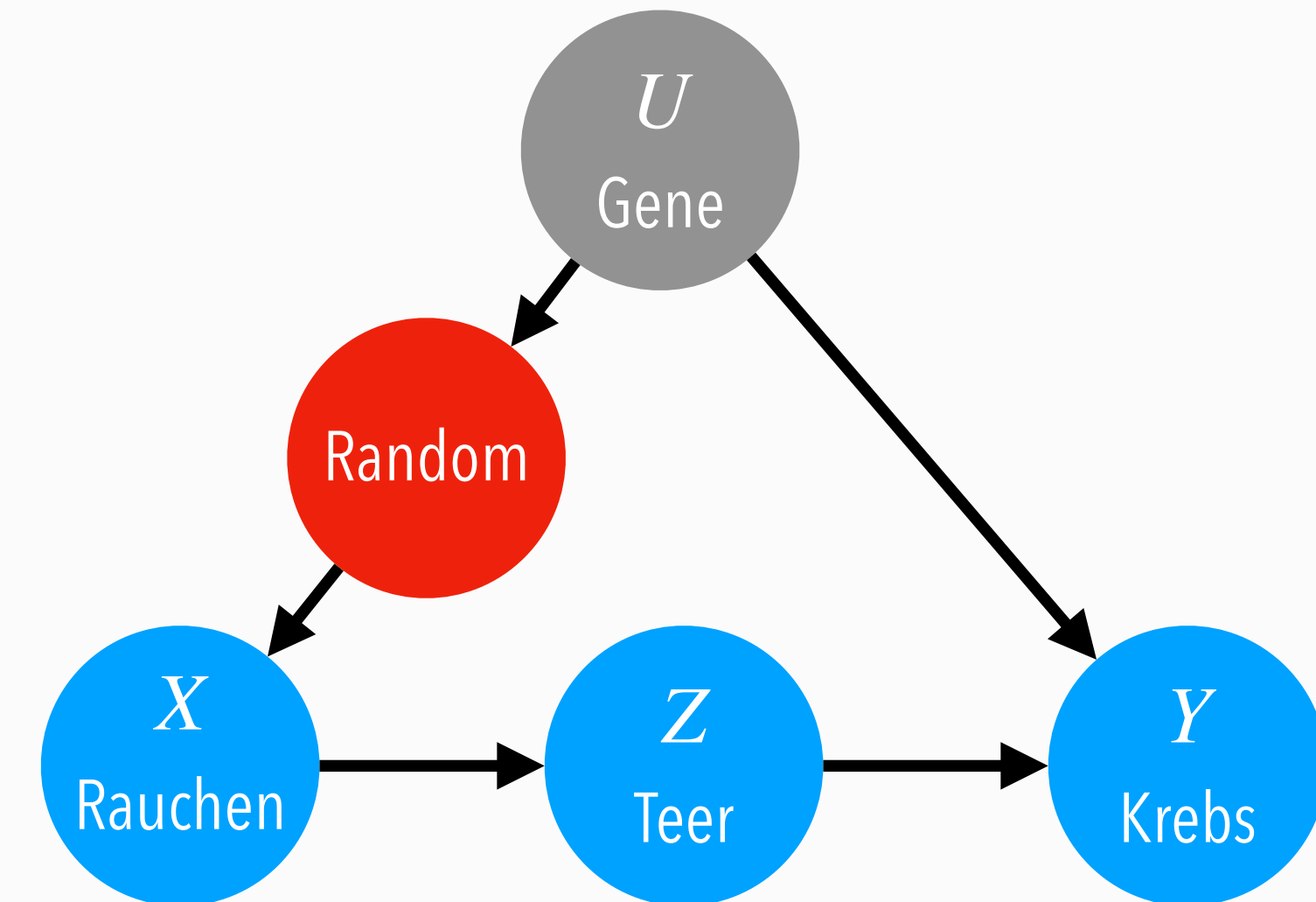


Kausalität

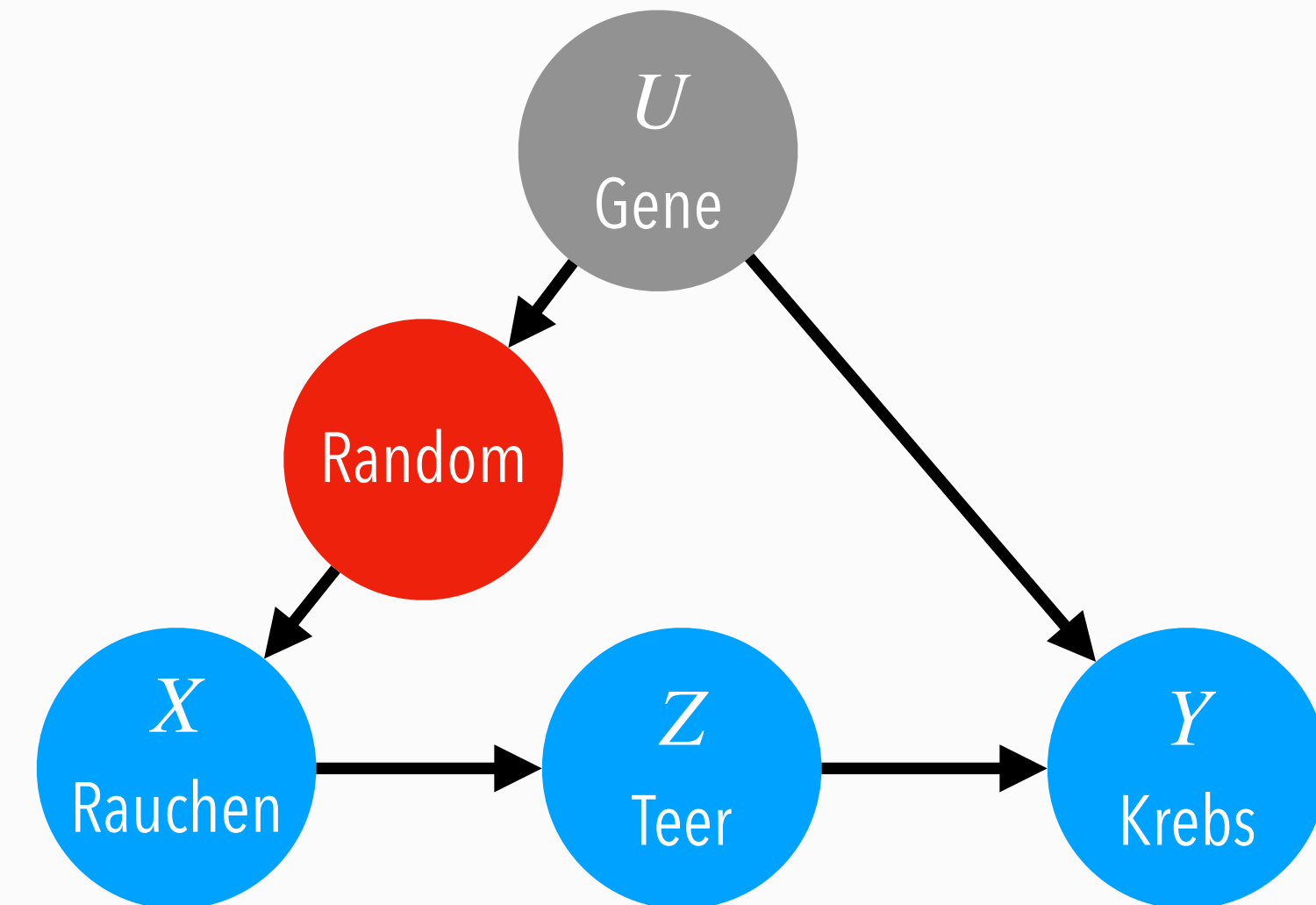
- Frage: wie wird Y von X beeinflusst?
- Beobachtungsstudie: X, Y, Z .
- Unbeobachtet: U .
- Frage lässt sich nicht beantworten.
- Der Einfluss von U ist nicht bekannt.



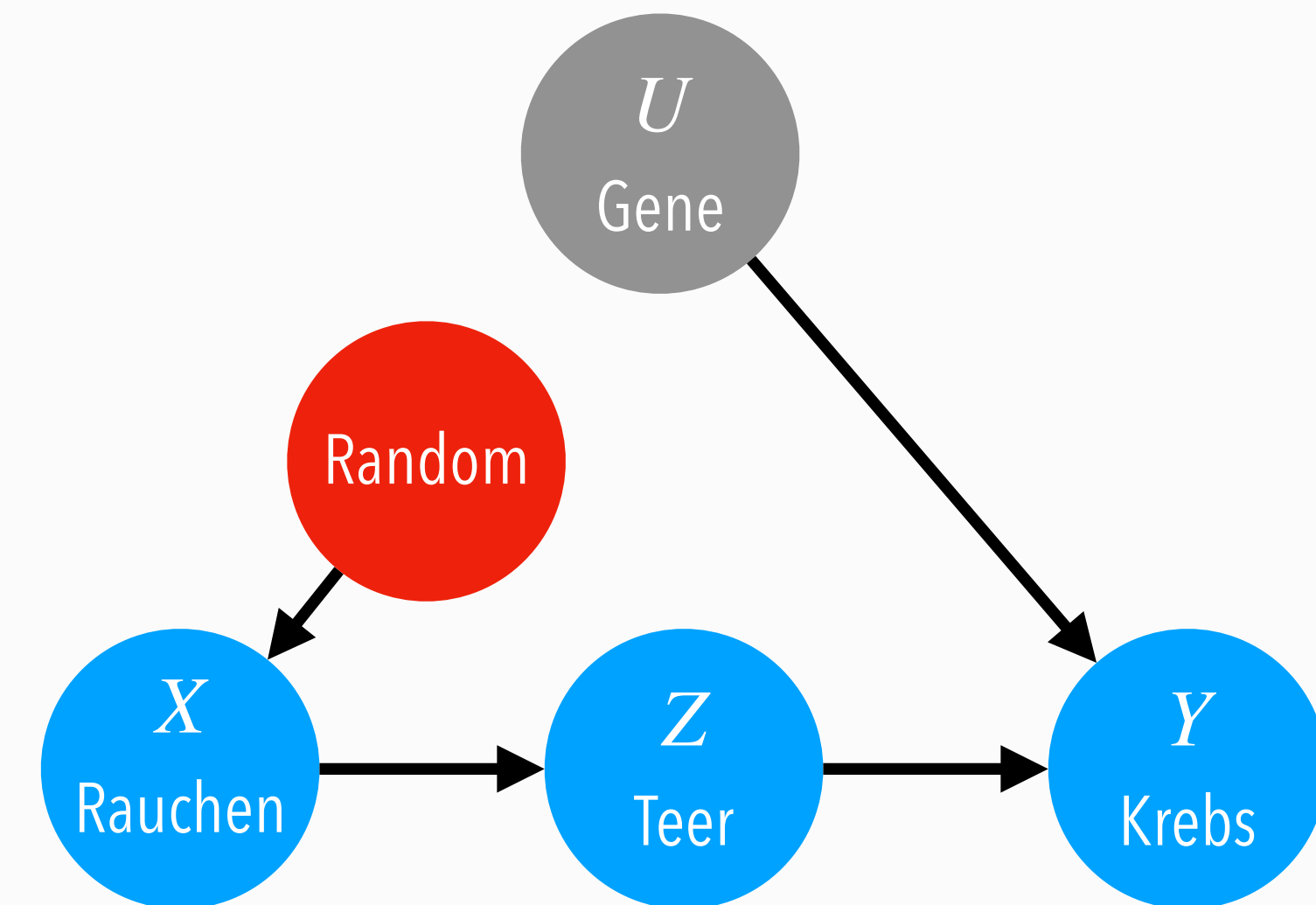
- Frage: wie wird Y von X beeinflusst?
- Goldstandard:
Randomisierte kontrollierte Studie (RCT).



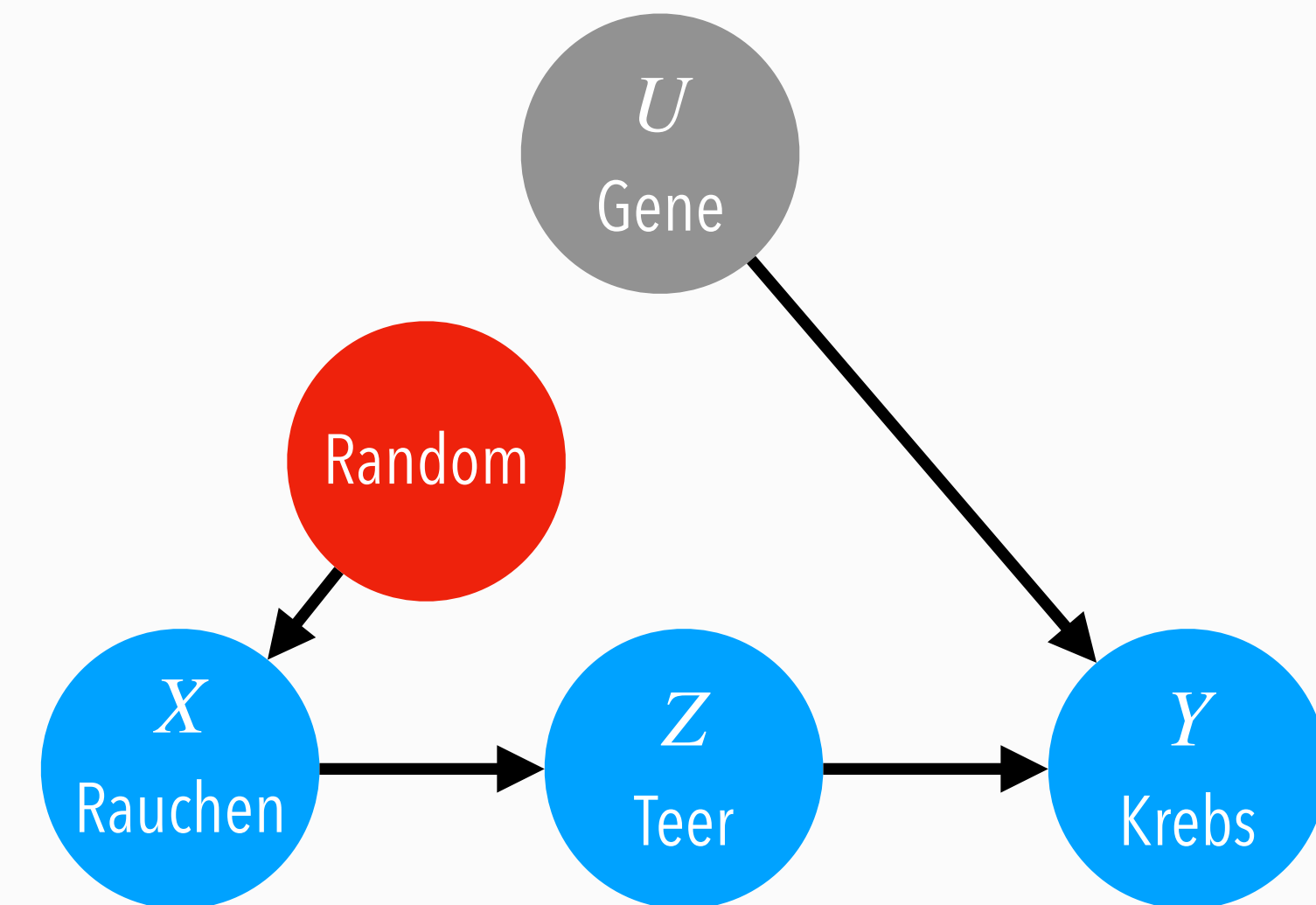
- Frage: wie wird Y von X beeinflusst?
- Goldstandard:
Randomisierte kontrollierte Studie (RCT).
 - Die randomisierte Manipulation von X wird als $P(y | Do(x))$ geschrieben.



- Frage: wie wird Y von X beeinflusst?
- Goldstandard:
Randomisierte kontrollierte Studie (RCT).
 - Die randomisierte Manipulation von X wird als $P(y | Do(x))$ geschrieben.
 - Randomisierung unterbricht den Einfluss von U auf X .
 - (und marginalisiert über $U \dots$;)
 - X, Y, Z werden gesammelt.
 - Berechnung von $P(y | x)$ beantwortet die Frage.



- Frage: wie wird Y von X beeinflusst?
- Goldstandard:
Randomisierte kontrollierte Studie (RCT).
 - Die randomisierte Manipulation von X wird als $P(y | Do(x))$ geschrieben.
 - Randomisierung unterbricht den Einfluss von U auf X .
 - (und marginalisiert über $U \dots$;)
 - X, Y, Z werden gesammelt.
 - Berechnung von $P(y | x)$ beantwortet die Frage.
- Nachteile:
 - Aufwändig und teuer.
 - Ethisch problematisch.
 - Daten der vorhandenen Beobachtungsstudie sind wertlos.

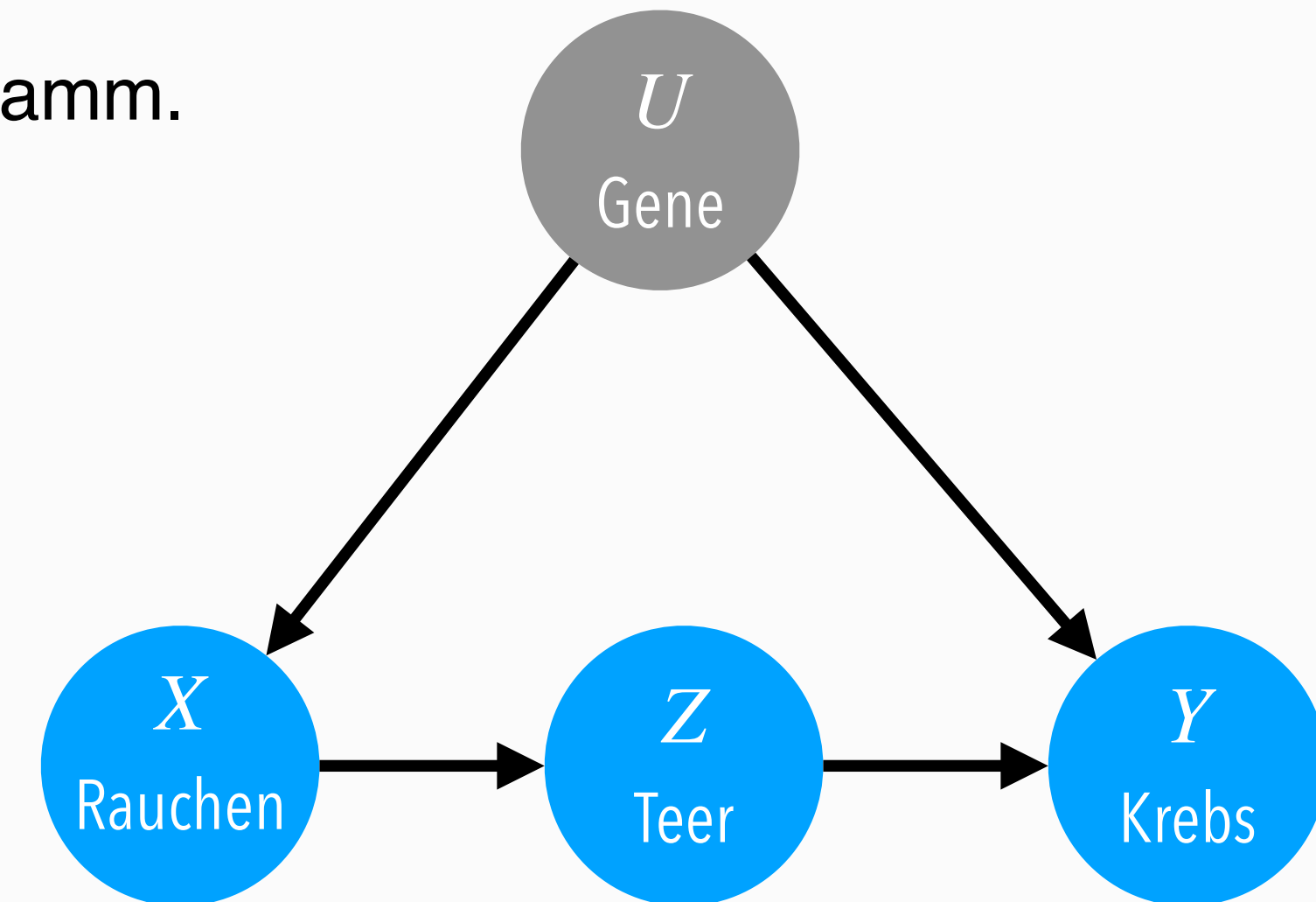


- Alternative: Ausnutzung der kausalen Struktur

- $P(y | Do(x)) =$ „Formel ohne *Do*“ $\Rightarrow$ kann mit Beobachtungsstudie beantwortet werden.
- Das *Do*-Kalkül berechnet die „Formel ohne *Do*“ aus dem kausalen Diagramm.

- Für das Beispieldiagramm gilt:

- $P(y | Do(x)) = \sum_z P(z | x) \cdot \sum_{x'} p(y | x', z) \cdot P(x')$



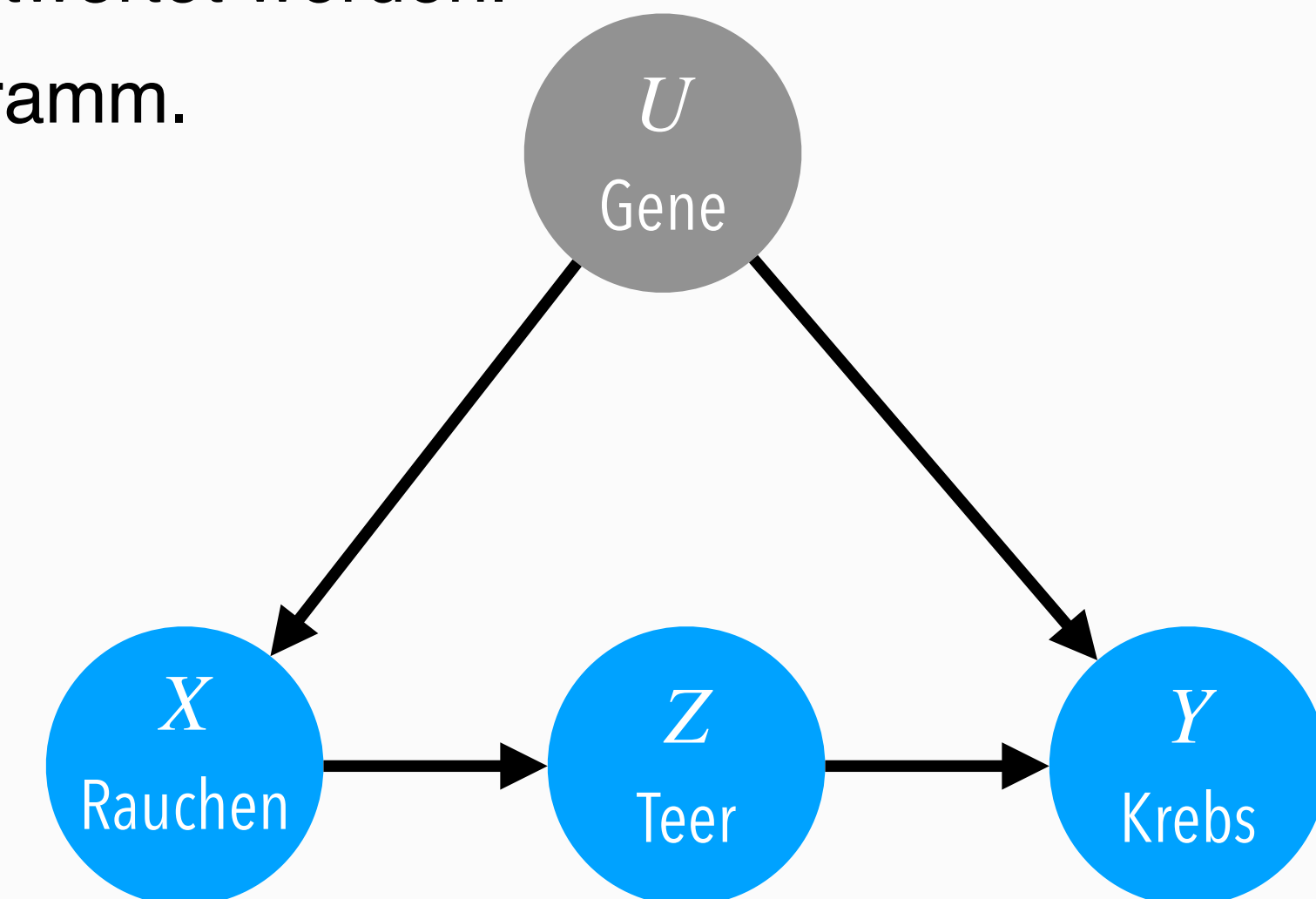
- Alternative: Ausnutzung der kausalen Struktur

- $P(y | Do(x)) =$ „Formel ohne Do “ $\Rightarrow$ kann mit Beobachtungsstudie beantwortet werden.
- Das Do -Kalkül berechnet die „Formel ohne Do “ aus dem kausalen Diagramm.

- Für das Beispieldiagramm gilt:

- $P(y | Do(x)) = \sum_z P(z | x) \cdot \sum_{x'} p(y | x', z) \cdot P(x')$

- Der rechte Teil der Gleichung kann aus der vorhandenen Beobachtungsstudie beantwortet werden!



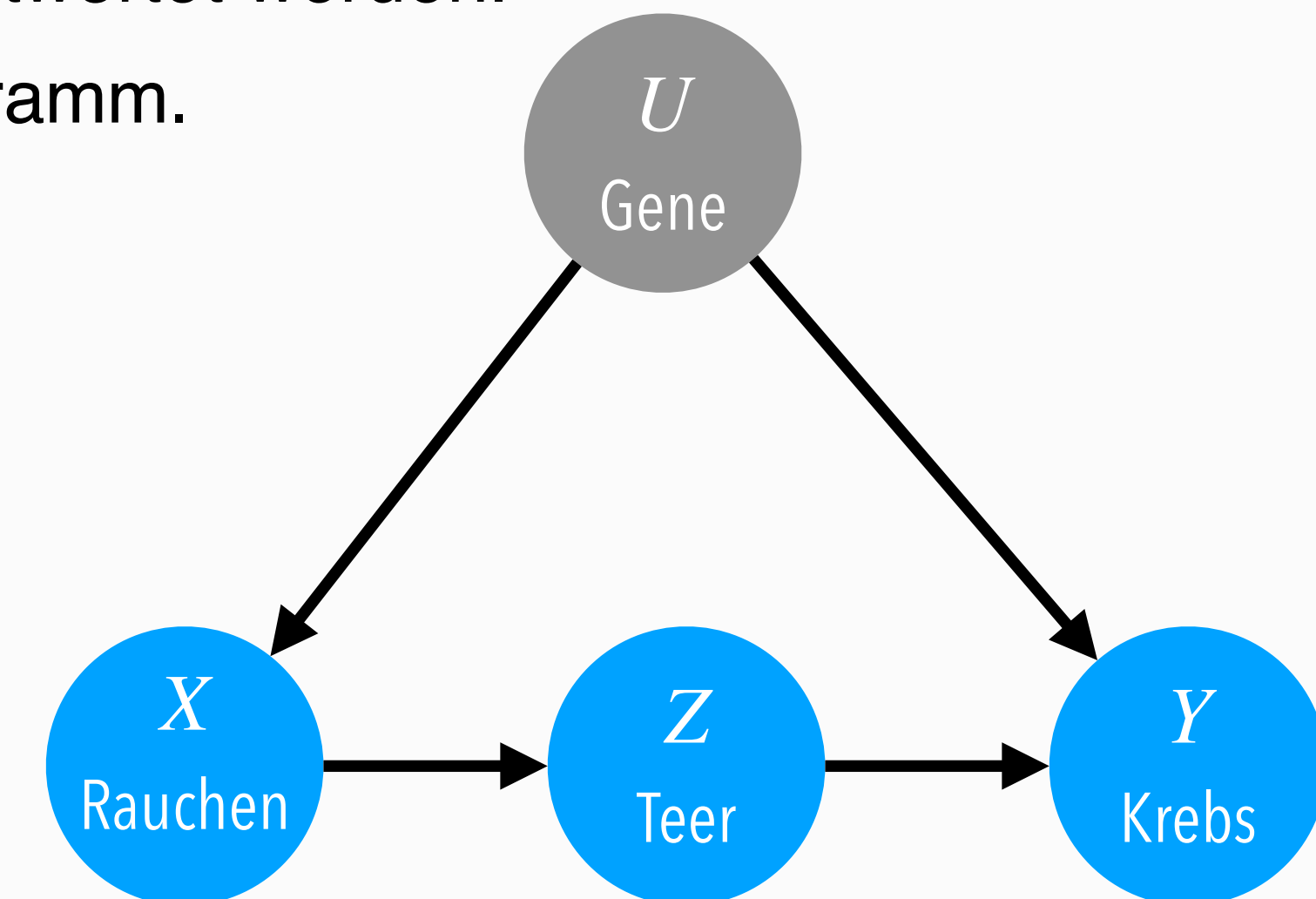
- Alternative: Ausnutzung der kausalen Struktur

- $P(y | Do(x)) =$ „Formel ohne *Do*“ $\Rightarrow$ kann mit Beobachtungsstudie beantwortet werden.
- Das *Do*-Kalkül berechnet die „Formel ohne *Do*“ aus dem kausalen Diagramm.

- Für das Beispieldiagramm gilt:

- $P(y | Do(x)) = \sum_z P(z | x) \cdot \sum_{x'} p(y | x', z) \cdot P(x')$

- Der rechte Teil der Gleichung kann aus der vorhandenen Beobachtungsstudie beantwortet werden!



- Die Extraktion potentieller kausalen Strukturen aus Texten ermöglicht es, die Datensätze vorhandener Beobachtungsstudien zu heben.

Integration von LLMs und kausalen Modellen

- Modellierung und Datenintegration für hybride KI-gestützte Medizinsysteme
 - Untersuchung, wie gut große Sprachmodelle (LLMs) kausale Zusammenhänge aus medizinischen Texten erkennen und extrahieren können
 - Untersuchung von Methoden zur Integration großer Sprachmodelle mit kausalen Modellen medizinischen Fachwissens, um robuste und erklärbare Lösungen für den Einsatz im medizinischen Bereich zu entwickeln.

Clinical Review & Education

US Preventive Services Task Force

USPSTF Recommendation: Screening for Gestational Diabetes

Screening strategy	Glucose load of OGTT, g	Threshold, mg/dL			
		Fasting	1-Hour	2-Hour	3-Hour
Two-step screening: an initial screening 50-g OGCT is administered. If the OGCT is positive (≥ 130 –140 mg/dL at 1 h), then proceed with OGTT. Diagnosis of gestational diabetes if ≥ 2 thresholds met on OGTT					
Carpenter and Coustan ³⁵	100	95	180	155	140
National Diabetes Data Group (NDDG) ³⁶	100	105	190	165	145
One-step screening: diagnosis of gestational diabetes if ≥ 1 thresholds met on OGTT					
International Association of Diabetes and Pregnancy Study Group (IADPSG) ³⁸	75	92	180	153	NA

Abbreviations: NA, not applicable; OGCT, oral glucose challenge test; OGTT, oral glucose tolerance test.

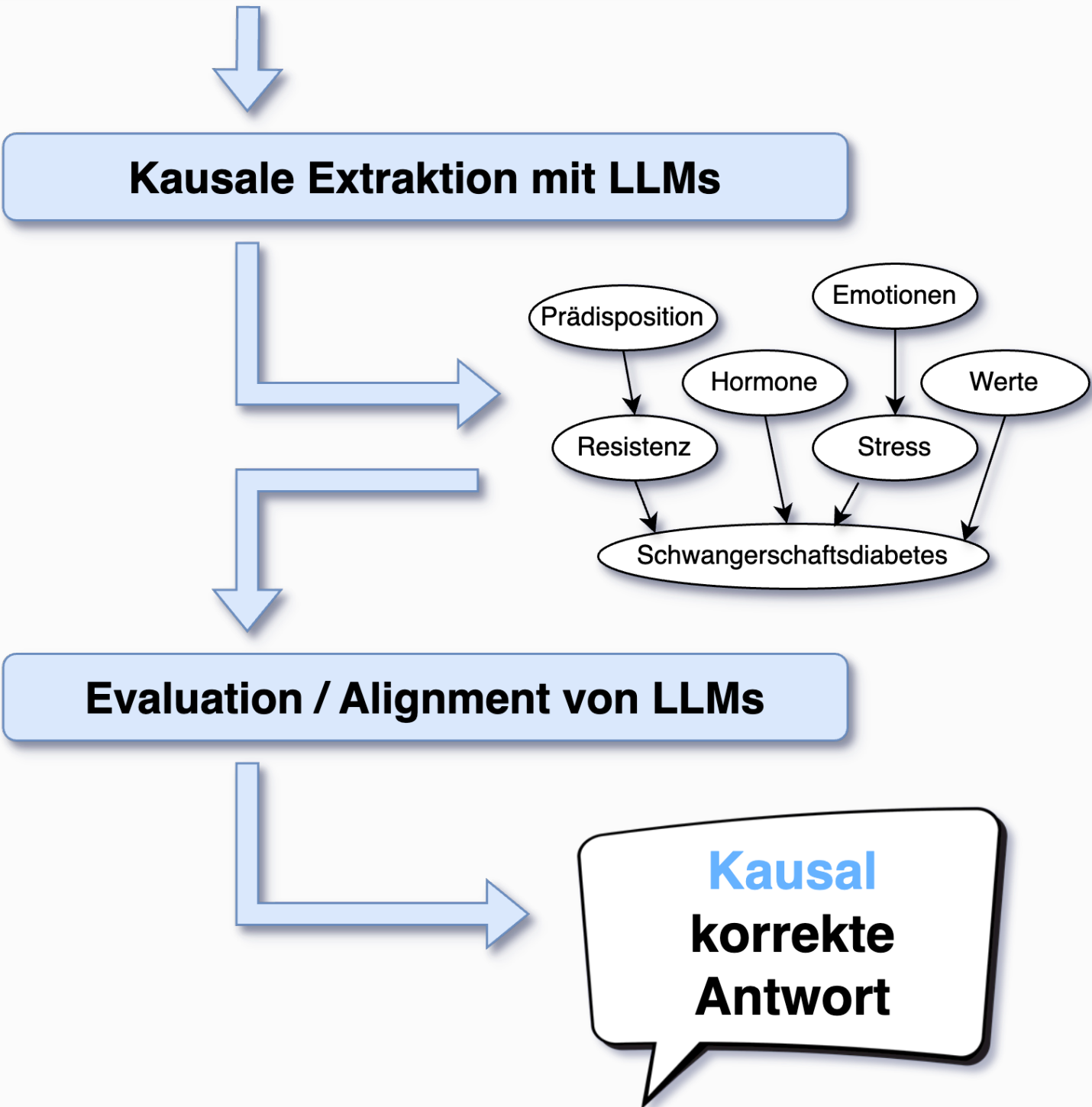
SI conversion factor: To convert glucose values to mmol/L, multiply by 0.0555.

Current Practice

Although current data are limited, a 2014–2015 survey found that universal screening is the most common practice in the US, with 90% of obstetricians reporting routinely screening for gestational diabetes using a 2-step approach.²⁶ Other potential (although not widely used) approaches to screening include fasting plasma glucose level, glycosylated hemoglobin (HbA_{1c}) concentration, and risk-based screening tools. Some pregnant persons are screened earlier than 24 weeks of gestation because they have risk factors for type 2 diabetes, such as obesity, family history of type 2 diabetes, or fetal macrosomia during a previous pregnancy. If a pregnant person presents in the first trimester or in early pregnancy with risk factors for type 2 diabetes, clinicians should use their clinical judgment to determine what is appropriate screening for that individual patient, given the patient's health needs.

tational diabetes (undiagnosed type 2 diabetes) is not in the scope of this recommendation.

Accuracy of Screening Tests and Risk Assessment
The USPSTF reviewed 45 prospective studies of fair or good quality that assessed the accuracy of various screening tests for gestational diabetes, including the 50-g OGCT, fasting plasma glucose level, HbA_{1c} concentration, and screening based on risk factors. The reference standard in these studies was the diagnostic OGTT, but the cutoff thresholds used for gestational diabetes varied, with most using criteria from Carpenter and Coustan, the International Association of Diabetes and Pregnancy Study Group (IADPSG), and the National Diabetes Data Group (NDDG) (Table 2). In all studies, the entire study population that had a screening test (regardless of test result) was offered the diagnostic OGTT reference standard.



Stand der Technik: Kausale Extraktion



- Aktuelle Best Practices

Clinical Review & EducationUS Preventive Services Task ForceUSPSTF Recommendation: Screening for Gestational Diabetes

Table 2. Common Screening Strategies for Gestational Diabetes

Screening strategy	Glucose load of OGTT, g	Threshold, mg/dL			
		Fasting	1-Hour	2-Hour	3-Hour
Two-step screening: an initial screening 50-g OGCT is administered. If the OGCT is positive (≥130–140 mg/dL at 1 h), then proceed with OGTT. Diagnosis of gestational diabetes if ≥2 thresholds met on OGTT					
Carpenter and Coustan ³⁵	100	95	180	155	140
National Diabetes Data Group (NDDG) ³⁶	100	105	190	165	145
One-step screening: diagnosis of gestational diabetes if ≥1 thresholds met on OGTT					
International Association of Diabetes and Pregnancy Study Group (IADPSG) ²⁸	75	92	180	153	NA

Abbreviations: NA, not applicable; OGCT, oral glucose challenge test; OGTT, oral glucose tolerance test. SI conversion factor: To convert glucose values to mmol/L, multiply by 0.0555.

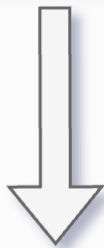
Current Practice

Although current data are limited, a 2014-2015 survey found that universal screening is the most common practice in the US, with 90% of obstetricians reporting routinely screening for gestational diabetes using a 2-step approach.²⁶ Other potential (although not widely used) approaches to screening include fasting plasma glucose level, glycosylated hemoglobin (HbA_{1c}) concentration, and risk-based screening tools. Some pregnant persons are screened earlier than 24 weeks of gestation because they have risk factors for type 2 diabetes, such as obesity, family history of type 2 diabetes, or fetal macrosomia during a previous pregnancy. If a pregnant person presents in the first trimester or in early pregnancy with risk factors for type 2 diabetes, clinicians should use their clinical judgment to determine what is appropriate screening for that individual patient, given the patient's health needs.

tational diabetes (undiagnosed type 2 diabetes) is not in the scope of this recommendation.

Accuracy of Screening Tests and Risk Assessment

The USPSTF reviewed 45 prospective studies of fair or good quality that assessed the accuracy of various screening tests for gestational diabetes, including the 50-g OGCT, fasting plasma glucose level, HbA_{1c} concentration, and screening based on risk factors. The reference standard in these studies was the diagnostic OGTT, but the cutoff thresholds used for gestational diabetes varied, with most using criteria from Carpenter and Coustan, the International Association of Diabetes and Pregnancy Study Group (IADPSG), and the National Diabetes Data Group (NDDG) (Table 2). In all studies, the entire study population that had a screening test (regardless of test result) was offered the diagnostic OGTT reference standard.



Schwangerschaftsdiabetes

- Aktuelle Best Practices

- Zero-Shot, Few-Shot and Finetuning und Post-Hoc Verification [SWC+24].

F1-Score 86% – Geeignet für explizite Kausalitäten.

Clinical Review & Education US Preventive Services Task Force

USPSTF Recommendation: Screening for Gestational Diabetes

Table 2. Common Screening Strategies for Gestational Diabetes

Screening strategy	Glucose load of OGTT, g	Threshold, mg/dL			
		Fasting	1-Hour	2-Hour	3-Hour
Two-step screening: an initial screening 50-g OGCT is administered. If the OGCT is positive (≥ 130 – 140 mg/dL at 1 h), then proceed with OGTT. Diagnosis of gestational diabetes if ≥ 2 thresholds met on OGTT					
Carpenter and Coustan ³⁵	100	95	180	155	140
National Diabetes Data Group (NDDG) ³⁶	100	105	190	165	145
One-step screening: diagnosis of gestational diabetes if ≥ 1 thresholds met on OGTT					
International Association of Diabetes and Pregnancy Study Group (IADPSG) ²⁸	75	92	180	153	NA

Abbreviations: NA, not applicable; OGCT, oral glucose challenge test; OGTT, oral glucose tolerance test. SI conversion factor: To convert glucose values to mmol/L, multiply by 0.0555.

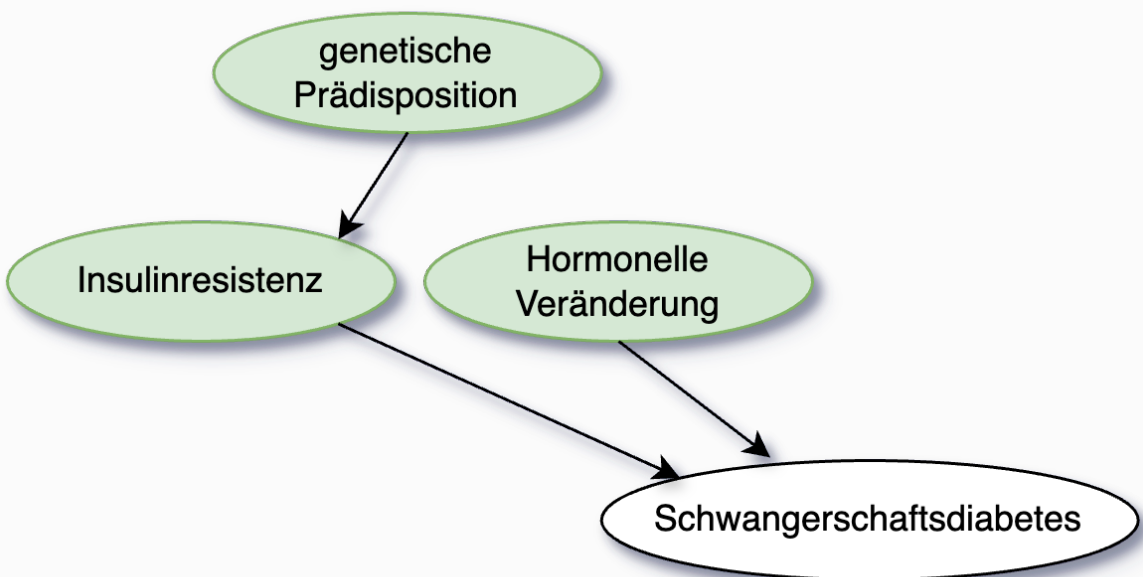
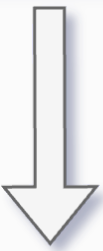
Current Practice

Although current data are limited, a 2014-2015 survey found that universal screening is the most common practice in the US, with 90% of obstetricians reporting routinely screening for gestational diabetes using a 2-step approach.²⁶ Other potential (although not widely used) approaches to screening include fasting plasma glucose level, glycosylated hemoglobin (HbA_{1c}) concentration, and risk-based screening tools. Some pregnant persons are screened earlier than 24 weeks of gestation because they have risk factors for type 2 diabetes, such as obesity, family history of type 2 diabetes, or fetal macrosomia during a previous pregnancy. If a pregnant person presents in the first trimester or in early pregnancy with risk factors for type 2 diabetes, clinicians should use their clinical judgment to determine what is appropriate screening for that individual patient, given the patient's health needs.

tational diabetes (undiagnosed type 2 diabetes) is not in the scope of this recommendation.

Accuracy of Screening Tests and Risk Assessment

The USPSTF reviewed 45 prospective studies of fair or good quality that assessed the accuracy of various screening tests for gestational diabetes, including the 50-g OGCT, fasting plasma glucose level, HbA_{1c} concentration, and screening based on risk factors. The reference standard in these studies was the diagnostic OGTT, but the cutoff thresholds used for gestational diabetes varied, with most using criteria from Carpenter and Coustan, the International Association of Diabetes and Pregnancy Study Group (IADPSG), and the National Diabetes Data Group (NDDG) (Table 2). In all studies, the entire study population that had a screening test (regardless of test result) was offered the diagnostic OGTT reference standard.



- Aktuelle Best Practices

- Zero-Shot, Few-Shot and Finetuning und Post-Hoc Verification [SWC+24].
F1-Score 86% – Geeignet für explizite Kausalitäten.
- Kombinationen von BioBERT mit graphenbasierten & heuristischen Ansätzen [NHF+25].
F1-Score 77% – Komplexer Aufbau, bedingt geeignet für implizite Kausalitäten in klaren Grammatiken (DPT-Errors).

Clinical Review & Education US Preventive Services Task Force

USPSTF Recommendation: Screening for Gestational Diabetes

Table 2. Common Screening Strategies for Gestational Diabetes

Screening strategy	Glucose load of OGTT, g	Threshold, mg/dL			
		Fasting	1-Hour	2-Hour	3-Hour
Two-step screening: an initial screening 50-g OGCT is administered. If the OGCT is positive (≥130–140 mg/dL at 1 h), then proceed with OGTT. Diagnosis of gestational diabetes if ≥2 thresholds met on OGTT					
Carpenter and Coustan ³⁵	100	95	180	155	140
National Diabetes Data Group (NDDG) ³⁶	100	105	190	165	145
One-step screening: diagnosis of gestational diabetes if ≥1 thresholds met on OGTT					
International Association of Diabetes and Pregnancy Study Group (IADPSG) ²⁸	75	92	180	153	NA

Abbreviations: NA, not applicable; OGCT, oral glucose challenge test; OGTT, oral glucose tolerance test. SI conversion factor: To convert glucose values to mmol/L, multiply by 0.0555.

Current Practice

Although current data are limited, a 2014-2015 survey found that universal screening is the most common practice in the US, with 90% of obstetricians reporting routinely screening for gestational diabetes using a 2-step approach.²⁶ Other potential (although not widely used) approaches to screening include fasting plasma glucose level, glycosylated hemoglobin (HbA_{1c}) concentration, and risk-based screening tools. Some pregnant persons are screened earlier than 24 weeks of gestation because they have risk factors for type 2 diabetes, such as obesity, family history of type 2 diabetes, or fetal macrosomia during a previous pregnancy. If a pregnant person presents in the first trimester or in early pregnancy with risk factors for type 2 diabetes, clinicians should use their clinical judgment to determine what is appropriate screening for that individual patient, given the patient's health needs.

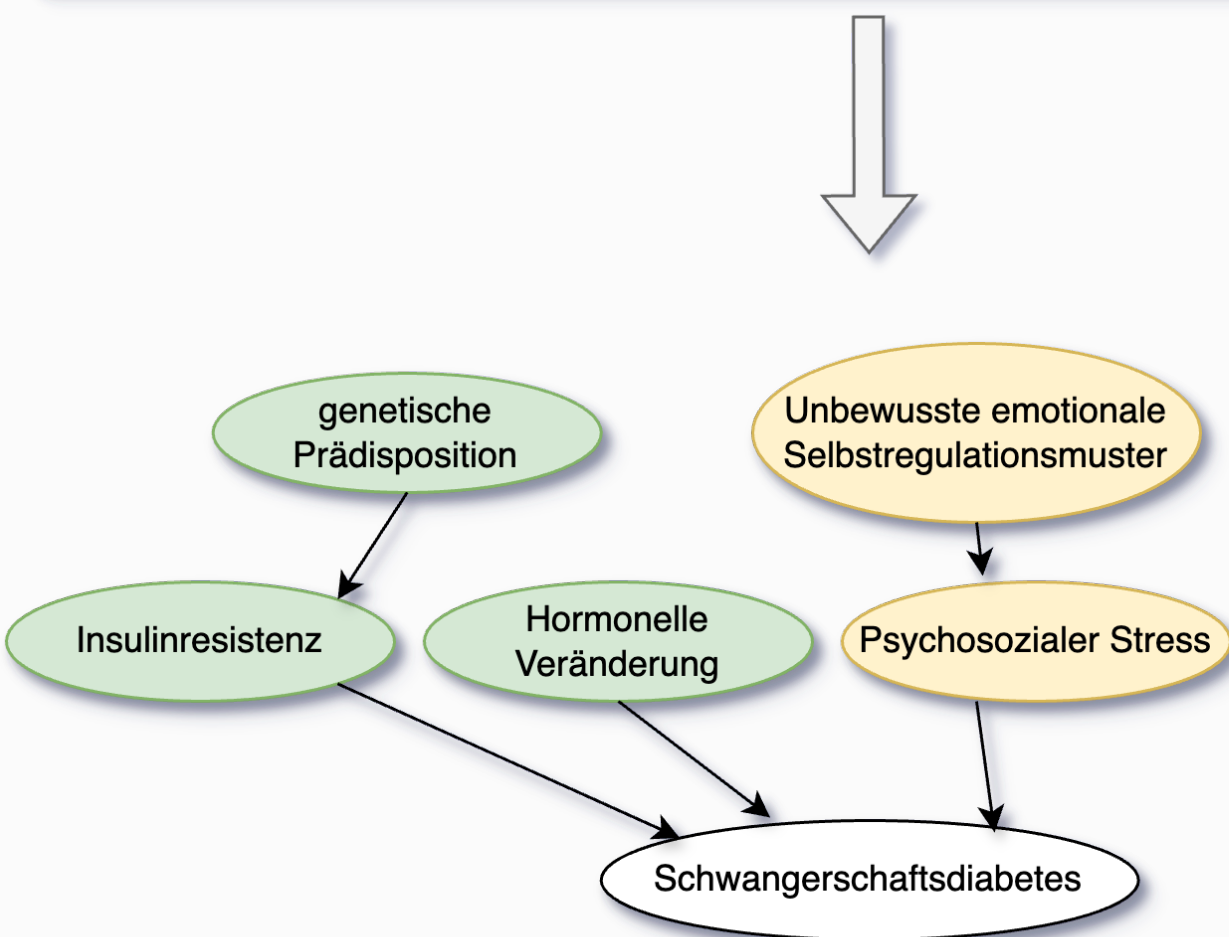
tational diabetes (undiagnosed type 2 diabetes) is not in the scope of this recommendation.

Accuracy of Screening Tests and Risk Assessment

The USPSTF reviewed 45 prospective studies of fair or good quality that assessed the accuracy of various screening tests for gestational diabetes, including the 50-g OGCT, fasting plasma glucose level, HbA_{1c} concentration, and screening based on risk factors. The reference standard in these studies was the diagnostic OGTT, but the cutoff thresholds used for gestational diabetes varied, with most using criteria from Carpenter and Coustan, the International Association of Diabetes and Pregnancy Study Group (IADPSG), and the National Diabetes Data Group (NDDG) (Table 2). In all studies, the entire study population that had a screening test (regardless of test result) was offered the diagnostic OGTT reference standard.

Explizite Kausalität

Implizite Kausalität



- Aktuelle Best Practices

- Zero-Shot, Few-Shot and Finetuning und Post-Hoc Verification [SWC+24].
F1-Score 86% – Geeignet für explizite Kausalitäten.
- Kombinationen von BioBERT mit graphenbasierten & heuristischen Ansätzen [NHF+25].
F1-Score 77% – Komplexer Aufbau, bedingt geeignet für implizite Kausalitäten in klaren Grammatiken (DPT-Errors).
- Prompting GPT und Decomposed Reasoning mit Multi Agent Framework für Event Causality Identification (ECI) [CYZ+25]
F1-Score 84% – Geeignet für implizite Kausalitäten, jedoch Abhängig von externen LLMs wie ChatGPT-3.5-Turbo und PaLM2.

Clinical Review & Education US Preventive Services Task Force

USPSTF Recommendation: Screening for Gestational Diabetes

Table 2. Common Screening Strategies for Gestational Diabetes

Screening strategy	Glucose load of OGTT, g	Threshold, mg/dL			
		Fasting	1-Hour	2-Hour	3-Hour
Two-step screening: an initial screening 50-g OGCT is administered. If the OGCT is positive (≥130–140 mg/dL at 1 h), then proceed with OGTT. Diagnosis of gestational diabetes if ≥2 thresholds met on OGTT					
Carpenter and Coustan ³⁵	100	95	180	155	140
National Diabetes Data Group (NDDG) ³⁶	100	105	190	165	145
One-step screening: diagnosis of gestational diabetes if ≥1 thresholds met on OGTT					
International Association of Diabetes and Pregnancy Study Group (IADPSG) ³⁷	75	92	180	153	NA

Abbreviations: NA, not applicable; OGCT, oral glucose challenge test; OGTT, oral glucose tolerance test. SI conversion factor: To convert glucose values to mmol/L, multiply by 0.0555.

Current Practice

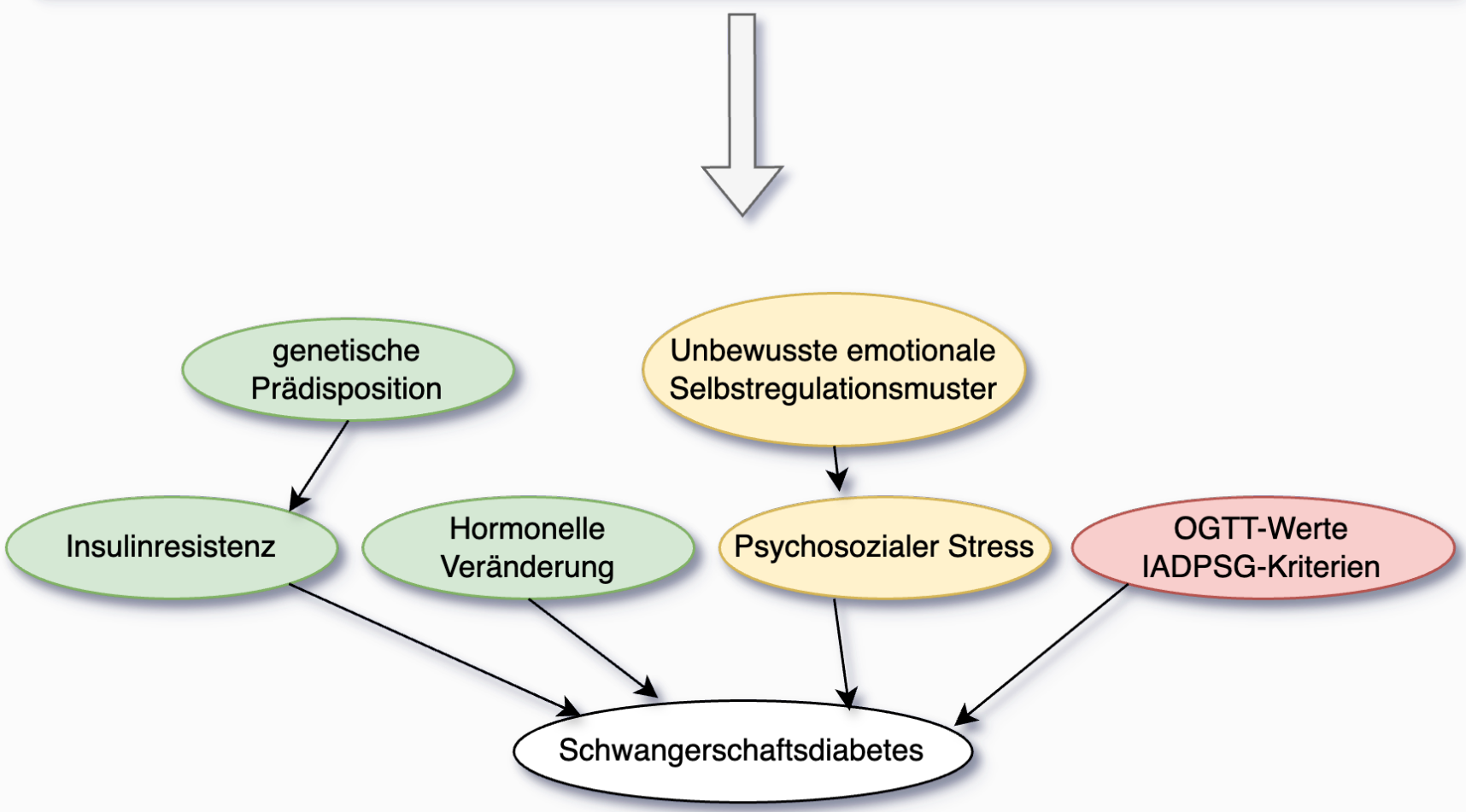
Although current data are limited, a 2014-2015 survey found that universal screening is the most common practice in the US, with 90% of obstetricians reporting routinely screening for gestational diabetes using a 2-step approach.²⁶ Other potential (although not widely used) approaches to screening include fasting plasma glucose level, glycosylated hemoglobin (HbA_{1c}) concentration, and risk-based screening tools. Some pregnant persons are screened earlier than 24 weeks of gestation because they have risk factors for type 2 diabetes, such as obesity, family history of type 2 diabetes, or fetal macrosomia during a previous pregnancy. If a pregnant person presents in the first trimester or in early pregnancy with risk factors for type 2 diabetes, clinicians should use their clinical judgment to determine what is appropriate screening for that individual patient, given the patient's health needs.

tational diabetes (undiagnosed type 2 diabetes) is not in the scope of this recommendation.

Accuracy of Screening Tests and Risk Assessment

The USPSTF reviewed 45 prospective studies of fair or good quality that assessed the accuracy of various screening tests for gestational diabetes, including the 50-g OGCT, fasting plasma glucose level, HbA_{1c} concentration, and screening based on risk factors. The reference standard in these studies was the diagnostic OGTT, but the cutoff thresholds used for gestational diabetes varied, with most using criteria from Carpenter and Coustan, the International Association of Diabetes and Pregnancy Study Group (IADPSG), and the National Diabetes Data Group (NDDG) (Table 2). In all studies, the entire study population that had a screening test (regardless of test result) was offered the diagnostic OGTT reference standard.

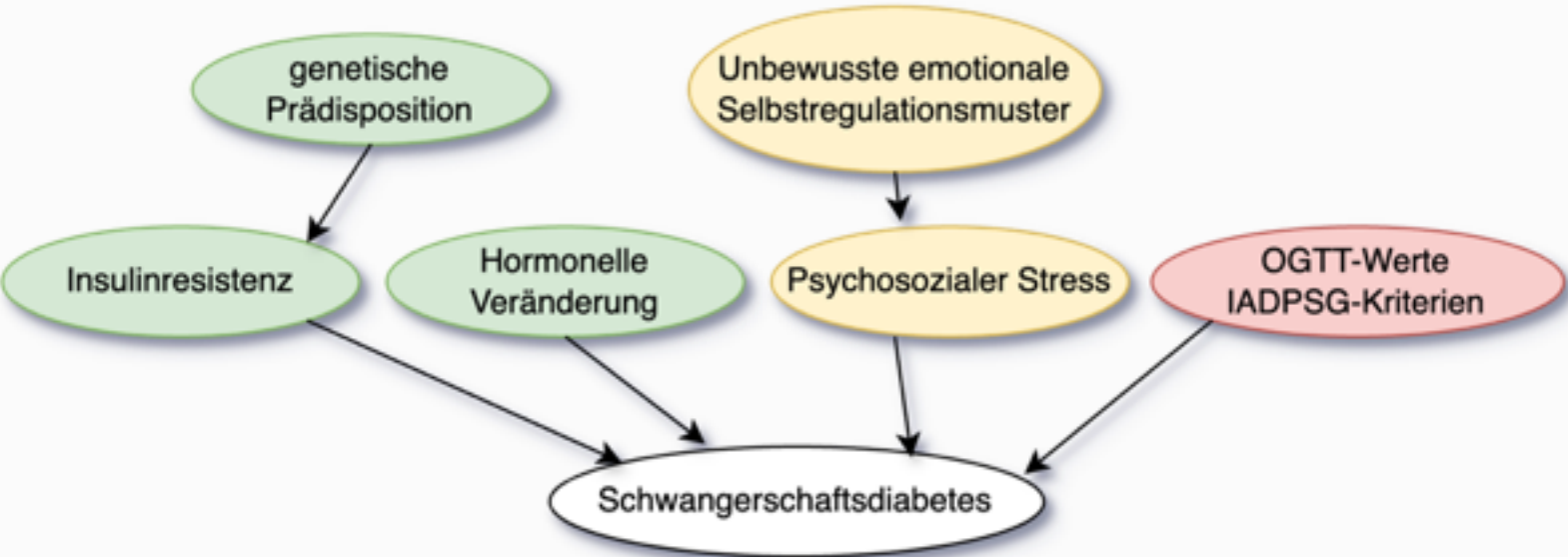
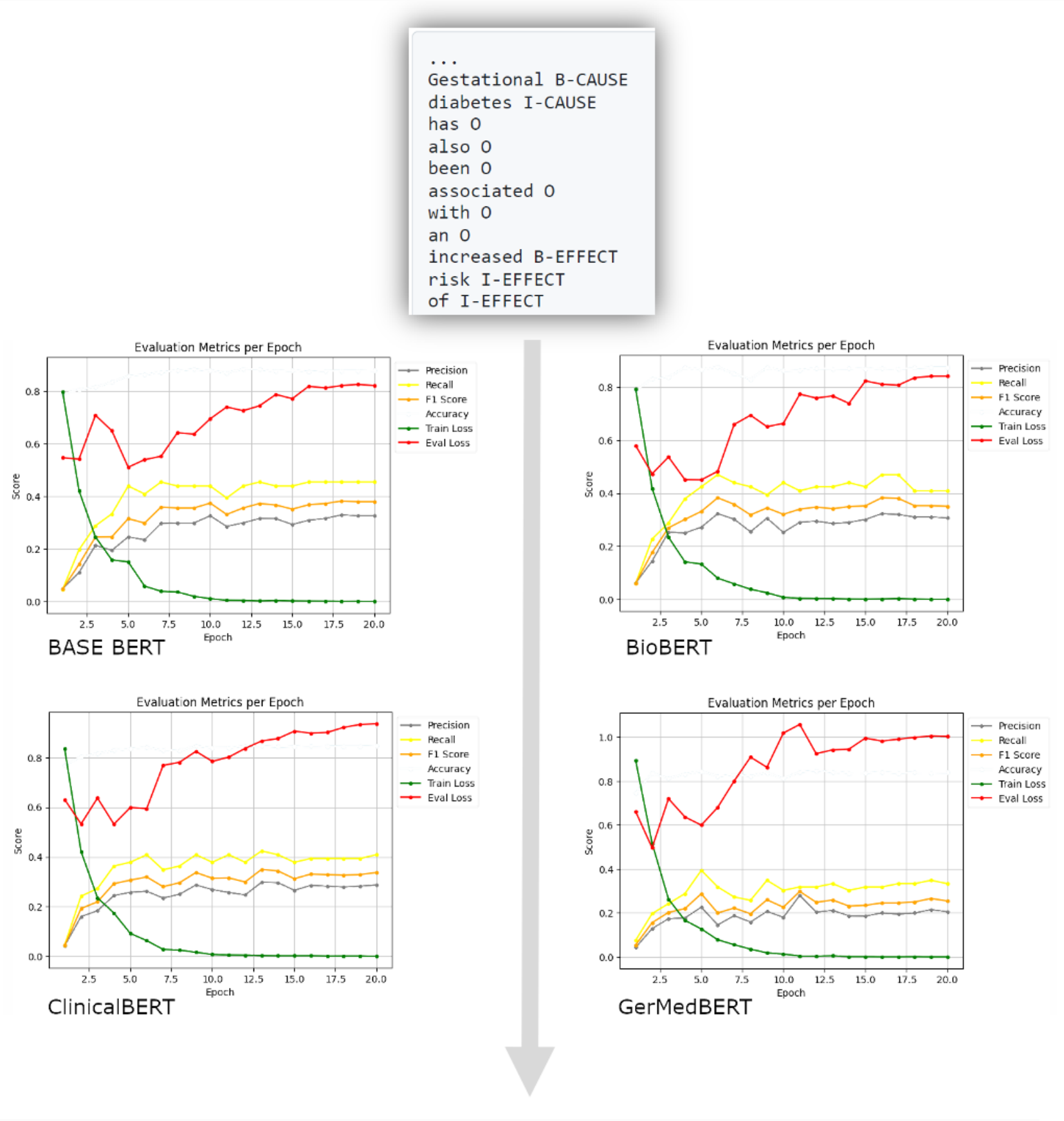
- Explizite Kausalität
- Implizite Kausalität
- Multimodal Grammatik



Aktuelle & Nächste Schritte



- Aktuelle Schritte
 - Bewertung der Leistungsfähigkeit von Encoder-Modellen wie BioBERT, ClinicalBERT und GerMedBERT anhand Precision, Recall, F1-Score und Loss mit aktuellen Datensätzen aus medizinischen Leitlinien [GGZ25].



Aktuelle & Nächste Schritte



- Aktuelle Schritte

- Bewertung der Leistungsfähigkeit von Encoder-Modellen wie BioBERT, ClinicalBERT und GerMedBERT anhand Precision, Recall, F1-Score und Loss mit aktuellen Datensätzen aus medizinischen Leitlinien [GGZ25].
- Bewertung der Leistungsfähigkeit von Decoder-Modellen wie DeepSeek-R1, Llama3.3, Gemma, ChatGPT anhand syntaktischer und semantischer Metriken mit QUEST.

QUEST v0.1.0

gray-phi4:latest

gray-qwen:0.5b

gray-llama3.3:70b

openai-gpt-4o-mini

hamming-deepseek-r1:70b

model: phi4:latest

provider: ollama

proxyid: gray

baseurl: http://host.docker.internal:11434

User

Welche Ursachen existieren für Schwangerschaftsdiabetes?

Assistant (gray-phi4:latest)

Schwangerschaftsdiabetes, auch als Gestationsdiabetes bekannt, ist eine Form von Diabetes, die während der Schwangerschaft auftritt. Es gibt mehrere Faktoren und Mechanismen, die zur Entwicklung von Schwangerschaftsdiabetes beitragen können:

- Hormonelle Veränderungen:** Während der Schwangerschaft produzieren hormonell bedingte Veränderungen im Körper der Frau Insulinresistenz. Hormone wie das Plazentahormon und Cortisol hemmen die Wirkung von Insulin, was bedeutet, dass mehr Insulin benötigt wird, um den Blutzuckerspiegel zu regulieren.
- Insulinproduktion:** Obwohl die meisten Schwangeren in der Lage sind, genügend Insulin zu produzieren, können einige Frauen mit einem verminderten Insulinsignalgefüge oder einer ineffizienteren Insulinsekretion konfrontiert werden, was den Blutzuckerspiegel erhöht.
- Genetische Faktoren:** Ein familiäres Vorliegen von Diabetes kann das Risiko für Schwangerschaftsdiabetes erhöhen. Auch genetische Prädispositionen können eine Rolle spielen.
- Alter der Mutter:** Frauen, die ein höheres Alter bei der Empfängnis haben (z.B. über 25 Jahre), sind einem größeren Risiko ausgesetzt.
- Übergewicht und Adipositas:** Ein erhöhter Body-Mass-Index (BMI) vor oder während der Schwangerschaft kann das Risiko für Schwangerschaftsdiabetes steigern, da Fettgewebe Insulinresistenz fördert.
- Frühere Schwangerschaften mit Gestationsdiabetes:** Frauen, die in einer früheren Schwangerschaft bereits an Schwangerschaftsdiabetes erkrankt sind, haben ein erhöhtes Risiko für eine Wiederholung der Erkrankung in zukünftigen Schwangerschaften.
- Ethnische Zugehörigkeit:** Bestimmte ethnische Gruppen haben ein höheres Risiko für die Entwicklung von Gestationsdiabetes. Dazu gehören Afroamerikaner, Hispanics, Asiaten und Pazifikinsulaner.
- Ernährungsgewohnheiten:** Eine unausgewogene Ernährung mit hohem Anteil an verarbeiteten Lebensmitteln und einfachen Kohlenhydraten kann das Risiko erhöhen.
- Sedentärer Lebensstil:** Ein Mangel an körperlicher Aktivität kann ebenfalls zur Entwicklung von Insulinresistenz beitragen.

Es ist wichtig, dass Frauen mit einem hohen Risiko für

Input your prompt

Select Models

Sync Chat Histories

Clear Histories

Reload backend

Render markdown

Jaccard Similarity

hamming-deepseek-r1:70b	0.35	0.103	0.313	0.338
openai-gpt-4o-mini	0.441	0.134	0.35	1
gray-llama3.3:70b	0.388	0.116	1	0.35
gray-qwen:0.5b	0.145	1	0.116	0.134
gray-phi4:latest	1	0.145	0.388	0.441

Cosine Similarity (mxbai-embed-large)

hamming-deepseek-r1:70b	0.949	0.755	0.942	0.948	1
openai-gpt-4o-mini	0.958	0.791	0.95	1	0.948
gray-llama3.3:70b	0.951	0.769	1	0.95	0.942
gray-qwen:0.5b	0.773	1	0.769	0.791	0.755
gray-phi4:latest	1	0.773	0.951	0.958	0.949

Aktuelle & Nächste Schritte

• Aktuelle Schritte

- Bewertung der Leistungsfähigkeit von Encoder-Modellen wie BioBERT, ClinicalBERT und GerMedBERT anhand Precision, Recall, F1-Score und Loss mit aktuellen Datensätzen aus medizinischen Leitlinien [GGZ25].
- Bewertung der Leistungsfähigkeit von Decoder-Modellen wie DeepSeek-R1, Llama3.3, Gemma, ChatGPT anhand syntaktischer und semantischer Metriken mit QUEST.

• Nächste Schritte

- Weiterentwicklung von Ansätzen zur Bestimmung qualitativer, quantitativer und komplexer kausaler Relationen durch GPT-„Agenten“
- Erweiterung der Methodik und kontinuierliche Evaluation.
- Integration ins neue KiMeKo KI-med. Ökosystem zur Verbesserung von Diagnosesicherheit und Therapieentscheidungen aktueller und künftiger KI-med. Entwicklungen.



KiMeKo: KI-Med-Kollaborationsplattform

- **Projektziel:**

- Aufbau eines KI-Ökosystems zur Unterstützung der Entwicklung KI-gestützter Medizintechnologien in Norddeutschland.

- **Hintergrund:** KI SIGS

- Vorgängerprojekt KI SIGS: Konzeption eines „KI-Space“ für intelligente Gesundheitssysteme.
- Bisherige Partner: Norddeutsche KI-Institute in Kooperation mit führenden medizinischen Unternehmen und Universitätskliniken.

- **Unsere Rolle im KiMeKo-Projekt**

- Extraktion kausaler Zusammenhänge aus medizinischen Texten (Leitlinien, Berichte, ...)
- Umsetzung eines kausalen Alignments KI-basierter Expertensysteme an medizinischen Erkenntnisse
- Schließen einer Lücke zwischen LLM-basierter KI-Technologie und der praktischen Anwendung in der Medizin.



KiMeKo: KI-Med-Kollaborationsplattform

Große Sprachmodelle für die Extraktion von
kausalem Wissen aus medizinischen Texten

Robin Nicolay, Sebastian Bader, Thomas Kirste

Mobile Multimediale Informationssysteme
Institut für Visual & Analytic Computing
Fakultät für Informatik und Elektrotechnik