

Positionspapier der FMH: Die neue statistische WZW-Screening-Methode der Krankenversicherer – kurz erklärt.

Die Frage ob eine Ärztin oder ein Arzt seine Leistungen unwirtschaftlich erbringt, kann ausschliesslich im Rahmen einer dem Screening nachgelagerten Einzelfall-Analyse, welche die vergleichsweise hohen Kosten einer «auffälligen» Praxis im Detail untersucht, beantwortet werden!

1. Ausgangslage

Die Krankenversicherer sind nach Art. 56 des Bundesgesetzes über die Krankenversicherung (KVG) verpflichtet, zu überprüfen, ob die Betreuung der Patienten durch die medizinischen Leistungserbringer dem im KVG geforderten Gebot der Wirksamkeit, Zweckmässigkeit und Wirtschaftlichkeit (WZW-Kriterien gemäss Art. 32 KVG) folgt. Der erste Schritt der Wirtschaftlichkeitsüberprüfung besteht aus einem statistischen Screening-Verfahren. Dieses hat zum Ziel, mittels statistischer Methoden Leistungserbringer zu identifizieren, deren Kosten im Vergleich zu einem Vergleichskollektiv höher ausfallen (für eine Übersicht des ganzen Wirtschaftlichkeitsverfahrens siehe A. Häfeli & R. Schwarz. 2019. [Schweizerische Ärztezeitung; 31.07.2018; Umsetzung der TARMED Verträge tut langsam Not!]).

Seit 2004 kommt für dieses statistische Screening-Verfahren die sogenannte ANOVA-Methode (analysis of variance) zum Einsatz. Die **Kritik der Ärzteschaft** an der bisherigen statistischen Screening-Methode wurde bei der Überprüfung durch die Firma Polynomics AG bestätigt: die aktuelle Methode führt zu einem hohen Anteil an falsch positiv eingestuftem Leistungserbringern, d.h. «auffälligen» Ärzten, bei denen die höheren Kosten aber medizinisch begründbar waren. Die FMH, die santésuisse und Curafutura beauftragten deshalb die Polynomics AG, eine neue, verbesserte statistische Screening-Methode zur Identifikation von «auffälligen» Leistungserbringern zu entwickeln.

Dabei ist wichtig zu vermerken, dass nur die statistische Screening-Methode selbst überarbeitet wurde. Die nachgelagerten administrativen Prozesse der Wirtschaftlichkeitsüberprüfung durchgeführt von santésuisse waren nicht Gegenstand der Überarbeitung: die **nachgelagerte Einzelfall-Analyse** der erbrachten Leistungen der «auffälligen» Leistungserbringer wird wie bisher im Anschluss an das statistische Screening-Verfahren durchgeführt. Festzuhalten ist ausserdem, dass die Datenbasis der Wirtschaftlichkeitsverfahren bisher nicht Gegenstand der Untersuchung von Polynomics war, d.h. die Erfahrungen aus der Arbeit in den Kantonalen Paritätischen Vertrauenskommissionen PVK, dass die verschiedenen Praxisorganisationsformen und die Fragmentierung und Subspezialisierung der medizinischen Leistungserbringung einen negativen Einfluss auf die statistische Aussagekraft der verwendeten Methoden haben, müssen in einem anderen Kontext wissenschaftlich überprüft werden.

2. Grundsätzliche Anmerkungen zum statistischen Screening-Verfahren im Rahmen der Wirtschaftlichkeitsprüfung

Die statistische **Screening-Methode erbringt keinen abschliessenden Nachweis**, dass eine Ärztin oder ein Arzt unwirtschaftlich handelt und nicht alle Praxisbesonderheiten abschliessend korrigieren kann, deshalb kann die Screening-Methode immer nur ein Teil der Kontrolle der Wirtschaftlichkeit gemäss Art. 56 Abs. 6 KVG, und kein Ersatz für eine detaillierte Einzelfallbeurteilung sein. Analog zu einer Vorsorgeuntersuchung kann das statistische Screening-Verfahren lediglich Personen identifizieren, bei denen die untersuchten Daten «auffällige» Werte aufweisen. In einem medizinischen Screening muss bei auffälligen Werten durch nachfolgende diagnostische Untersuchungen das Vorliegen der Krankheit bestätigt werden. Analog dazu ist ein positives Screening-Resultat («Praxis ist auffällig») noch kein „Beweis“ für eine unwirtschaftliche Leistungserbringung durch den Arzt. Erst im Rahmen einer nachgelagerten Einzelfall-Analyse kann nach Erklärungen für die vergleichsweise hohen Kosten einer „auffälligen“ Praxis gesucht werden.

3. Eckwerte der neuen statistischen Screening-Methode

Die weiterentwickelte statistische Screening-Methode berücksichtigt zusätzlich Variablen zur Morbidität des Patientenguts einer Praxis. Dadurch soll eine verbesserte Screening-Genauigkeit erreicht werden. In einer von Polynomics AG durchgeführten Studie wurde gezeigt, dass dadurch Leistungserbringer mit überdurchschnittlichen Kosten nicht automatisch als „auffällig“ eingestuft werden (im Gegensatz zur bisherigen Methode), sondern nur dann, wenn diese überdurchschnittlichen Kosten nicht durch die Patienten-Morbidität der Praxis erklärt werden können. Der erweiterte Variablensatz bei der Berechnung der «Auffälligkeit» stellt einen wichtigen Fortschritt dar. Aber auch mit der verbesserten Datengrundlage lassen sich die spezifischen Kosten einer Praxis nicht vollständig herleiten. Es wird somit auch in Zukunft notwendig sein, für auffällige Praxen das Leistungsspektrum und die kostenrelevanten Praxisbesonderheiten zu berücksichtigen.

3.1. Datengrundlage der statistischen Screening-Methode: Was ist neu?

Wie bisher werden die Variablen pro Praxis nach Alters- und Geschlechtsgruppe der Patienten (als 5-Jahresgruppen) aggregiert. Die bei der weiterentwickelten statistischen Screening-Methode verwendeten Morbiditäts-Variablen sind nach aktuellem Stand:

- Wahlfranchisen
- Spitaleinweisung im Vorjahr
- Pharmazeutische Kostengruppen (PCG)

Die Verwendung der PCG ermöglicht es, Patienten, die hohe Kosten verursachen, aufgrund ihres Arzneimittelbedarfs zu ermitteln. Bei den verwendeten PCG handelt es sich um: *Asthma, COPD/schweres Asthma, Zystische Fibrose/Pankreasenzyme, Hoher Cholesterinspiegel, Morbus Crohn und Colitis ulcerosa, Depression, Diabetes Typ I, Diabetes Typ II, Epilepsie, Glaukom, Herzerkrankungen, HIV/AIDS, Hormonsensitive Tumore, Krebs, Nierenerkrankungen, Erkrankungen des Gehirns/Rückenmarks, Neuropathischer Schmerz, Parkinson, Psychose-Alzheimer-Sucht, Rheuma,*

Erkrankungen der Schilddrüse, Transplantationen, Bluthochdruck, ADHS (je nach Facharztgruppe finden nicht alle PCG Anwendung).

Eine bestimmte PCG wird nur dann für eine Facharztgruppe berücksichtigt, wenn mehr als 30 Ärzte innerhalb der Facharztgruppe eine Mindestmenge oder mehr Medikamente aus der entsprechenden PCG veranlasst haben. Die Mindestmenge wurde dabei als mindestens 180 definierten Tagesdosen (Defined Daily Dosis, DDD) pro 100 Patienten und Jahr definiert. Der Grund für diese Einschränkung liegt in der Stabilität der geschätzten Koeffizienten.

3.2. Berechnungsmethodik

Die statistischen **Berechnungen werden pro medizinische Fachspezialität** (Facharzttitel) separat durchgeführt, da davon auszugehen ist, dass die erklärenden Variablen wie beispielsweise die Altersstruktur des Patientenstamms bei den einzelnen Facharztgruppen nicht den gleichen Einfluss auf die Kosten haben.

Zur Anwendung kommt ein zweistufiges Verfahren zur Berechnung des Praxiseffekts (Praxisbesonderheiten).

1. Die initiale Berechnung des Praxiseffekts basiert auf der Kostenstruktur der Praxis und den Morbiditäts-Variablen mittels eines «Fixed-Effect-Modells» (Regressions-Modell) zur Berücksichtigung der Praxisbesonderheiten. Multivariate Regressionsmodelle sind darauf ausgerichtet, den durchschnittlichen Einfluss von erklärenden Faktoren (z.B. Morbiditäts-Variablen) auf eine Zielvariable (z. B. die Kosten) zu berechnen. Das Fixed-Effects-Modell erweitert die gewöhnliche lineare Regression, indem bei der Schätzung die nicht durch das Modell erklärten Unterschiede in zwei Komponenten aufgeteilt werden: erstens eine systematische Komponente pro Praxis, zweitens einen zufälligen Term. Für die Beurteilung der Praxis wird nur die systematische Komponente verwendet.
2. Anschliessend erfolgt die Berechnung des Einflusses des Praxisstandorts und eine dementsprechende Korrektur des Praxiseffekts um kantonale Faktoren, da es in einem Fixed-Effects-Modell nicht möglich ist Variablen zu berücksichtigen, welche pro Praxis nicht variieren.

In einem nächsten Schritt folgt die Berechnung des Indexwertes. Dieser Praxis-Index entsteht, wenn die Praxis mit den Praxis-Durchschnittskosten der jeweiligen Fachspezialität verglichen wird.

Vor dem Vergleich mit der Vergleichsgruppe wird ein Unsicherheitsfaktor¹ vom Praxiseffekt abgezogen, um das Risiko zu senken, dass eine Praxis fälschlicherweise als «auffällig» identifiziert wird, z.B. durch einige wenige Patienten mit hohen Medikamentenkosten. Dadurch wird die neue Screening-Methode weniger anfällig auf statistische Ausreisser. Der Unsicherheitsindikator erlaubt eine Aussage über die Präzision der Schätzung der praxisspezifischen Effekte (ähnlich einem Standardfehler). Der Unsicherheitsfaktor bezieht sich ausschliesslich auf den Praxiseffekt, und erlaubt keine Aussage zu den erklärenden Variablen.

¹ Der Unsicherheitsfaktor soll die statistische Unsicherheit bei der Berechnung des Praxiseffekts quantifizieren (ähnlich einer Standardabweichung).

Der Grenzwert für «Auffälligkeit» ist dabei unverändert: 30% höhere Kosten als der Durchschnitt des Vergleichskollektivs als Kriterium für «Auffälligkeit» (= Indexwert 130).

4. Beurteilung der verbesserten Screening-Methode

4.1. Fortschritte gegenüber der bisherigen Screening-Methode

Aus Sicht der FMH stellt die neue statistische Screening-Methode **eine deutliche Verbesserung** dar:

- Die neue statistische Screening-Methode ist **wissenschaftlich fundiert hergeleitet** und transparent. Die Methode soll gemäss vertraglicher Vereinbarung zwischen FMH, curafutura und Santésuisse dokumentiert und publiziert werden.
- Die Forderung der FMH nach einer **Reduktion der falsch positiven Leistungserbringer** wurde erfüllt. Mit der neuen statistischen Screening-Methode verringert sich der Einfluss statistischer Ausreisser deutlich und damit verringert sich der Anteil falsch positiver Leistungserbringer deutlich: Gemäss der Studie von Polynomics AG ist eine Reduktion um 50% gegenüber der alten ANOVA-Methode zu erwarten.
- Die **Morbidität des Patientenguts** auf Praxisebene wurde in die Berechnung integriert. Die verwendeten Morbiditäts-Variablen erfüllen im Rahmen des heute Machbaren die Forderungen der FMH.
- Die neue statistische Screening-Methode ist flexibler für die zukünftige Integration zusätzlicher Variablen zur Beschreibung von Praxisbesonderheiten, sie ist somit **einfacher weiterentwickelbar**.

4.2. Zukünftige Potentiale für Weiterentwicklungen

Die neue statistische Screening-Methode stellt **eine deutliche Verbesserung** dar. Trotzdem müssen die Screening-Methode und die Grundlagen der WZW-Verfahren kontinuierlich weiterentwickelt und verbessert werden. Zusätzlich ist es zwingend nötig einen quantitativen Vergleich der bisherigen und neuen Screening Methode bezüglich des Effekts auf die einzelne Praxis durchzuführen und transparent zu machen, d.h. pro Fachspezialität müssen die zur Anwendung kommenden Variablen und deren quantitativen Einfluss auf den Regressionsindex einzelner Praxen ausgewiesen werden, insbesondere auch, welche PCG überhaupt die Einschlusskriterien pro Fachspezialität erfüllen:

- **Krankheiten (Morbidität)**, welche hohe Behandlungskosten verursachen und nicht durch höhere Medikamenten- oder Hospitalisationskosten identifiziert werden, führen zu statistisch nicht erklärten hohen Kosten (z. Bsp.: Schmerzpatienten, palliative Betreuung, psychisch kranke Patienten).
- **Unterschiedliche Leistungsangebote** innerhalb derselben Facharztgruppe werden nicht erkannt (z. Bsp.: operativ tätig vs. Konservativ tätig).
- **Tätigkeit in zwei Fachgebieten** werden nicht erkannt (z. Bsp.: Kardiologie + hausärztliche Betreuung) – dies bereitet v.a. dann Probleme, wenn die Zusatzleistung mit den tarifarischen Grundleistungen abgerechnet werden muss.
- **Rein diagnostische Spezialisierungen** werden nicht erkannt, da die zugehörigen Medikamentenkosten bei anderen Leistungserbringern anfallen (z. Bsp. diagnostisch arbeitende Kardiologen, deren Patienten vom Hausarzt medikamentös betreut werden).

- **Spezielle Patientenkollektive** werden von den verwendeten Daten nicht korrekt abgebildet (sozial tiefe Schichten, Bevölkerung mit Migrationshintergrund, Fokussierung auf Krankheitsgruppe, welche fachgruppenspezifisch nicht durch PCG abgebildet wird (z. Bsp. HIV in der Fachgruppe Gynäkologie und Geburtshilfe).
- Die Integration von Fachgesellschaft-spezifischen Variablen würde es erlauben, homogenere und damit **genauere Vergleichsgruppen** zu definieren, z.B. indem nur Ärzte einer Fachgesellschaft, die bestimmte Behandlungen vornehmen, miteinander verglichen würden (z.B. invasive Eingriffe), und die Analyse der Daten auf Ebene des einzelnen Leistungserbringers statt der Praxis, könnte die Aussagekraft des Screenings bei verschiedenen Praxisorganisationsformen (Gruppenpraxen, Institute) verbessern.

4.3. Interpretation der Resultate des Screening-Verfahrens

Die Ergebnisse des statistischen Screening-Verfahrens sind nur dann korrekt interpretierbar, wenn auch die Schwächen des Verfahrens aus methodischer Sicht bekannt sind. In diesem Zusammenhang müssen folgende Einschränkungen berücksichtigt werden:

- a. Das **Fixed-Effects-Modell hat den Nachteil**, dass auch alle nicht beobachteten Praxisbesonderheiten in den praxisspezifischen Effekt eingehen. Falls z.B. ein Arzt ein bestimmtes Leistungsspektrum hat, das hohe Kosten verursacht, resultiert dies in einem hohen spezifischen Praxiseffekt. Dieser hohe Praxiseffekt ist aber nicht auf unwirtschaftliches Behandeln, sondern auf nicht beobachtete Praxisbesonderheiten zurückzuführen! Dies muss bei der Interpretation der Ergebnisse zwingend beachtet werden. Aus diesem Grund ist eine nachgelagerte Einzel Fallprüfung zwingend notwendig, um Praxisbesonderheiten wie beispielsweise ein spezifisches Leistungsspektrum, welche zu hohen Kosten führen, zu analysieren.
- b. Es existiert für „unwirtschaftliches Verhalten“ in der Medizin **kein Referenzstandard**, d.h. eine anerkannte und allgemeingültige Definition der (Un-)Wirtschaftlichkeit fehlt. Wirtschaftlichkeit wird oft definiert als das Verhältnis zwischen dem erreichten Resultat und dem dafür eingesetzten Aufwand. Es müsste also im Sinne einer «Kosten-Nutzen-Analyse» das klinische Resultat mit den Kosten der Behandlung korreliert werden. Das Fehlen eines solchen Referenzstandards führt dazu, dass nicht objektiv definiert werden kann, wo die Grenze der „unwirtschaftlichen Behandlungen“ liegt. Aus diesem Grund kann auch nicht beurteilt werden, welche Praxen im Rahmen des Screening-Verfahrens korrekt als «auffällig» identifiziert werden. Die verwendete Methodik geht vielmehr davon aus, dass die durchschnittlichen Behandlungskosten in einer Facharzttrichtung mit «wirtschaftlicher Behandlung» gleichgesetzt sind. In dieser Situation ist es zwingend notwendig im Anschluss an das Screening-Verfahren eine Einzelfallprüfung durchzuführen, um auffällige Kostenprofile zu analysieren.
- c. Ein zentraler Aspekt für die Beurteilung der Ergebnisse eines statistischen Screening-Verfahrens sind die Begriffe **Sensitivität** («alle Kranken werden korrekt erkannt») und **Spezifität** («alle Gesunden werden korrekt erkannt»). Aufgrund des Fehlens eines Referenzstandards ist es nicht möglich, die Sensitivität und Spezifität des verwendeten WZW-Verfahrens zu ermitteln.

5. Zusammenfassung

Das bisher verwendeten statistischen Screening-Verfahren der Versicherer, um bezüglich Kosten „auffällige“ Ärzte zu ermitteln, führten häufig zu falsch positiven Resultaten. Die FMH, die santésuisse und Curafutura beauftragten deshalb die Firma Polynomics AG, eine neue, verbesserte statistische Screening-Methode zu entwickeln.

Die weiterentwickelte statistische Screening-Methode ist wissenschaftlich fundiert hergeleitet und berücksichtigt neu zusätzlich Variablen zur Morbidität des Patientenguts einer Praxis.

Das zukünftige Entwicklungspotential der statistischen Screening-Methode besteht in der Integration weiterer Variablen zur Quantifizierung der **Patienten-Morbidität** (z.B. Diagnosedaten), und der **Verfeinerung der Vergleichsgruppen** durch Fachgesellschafts-spezifische Merkmale.

Es existiert für «(un)wirtschaftliches Verhalten» in der Medizin kein Referenzstandard, d.h. es kann nicht objektiv beurteilt werden, welche Praxen im Rahmen des statistischen Screening-Verfahren korrekt als «auffällig» identifiziert werden. In den aktuellen Methoden zur Wirtschaftlichkeitsprüfung werden die durchschnittlichen Behandlungskosten in einer Facharzttrichtung mit «wirtschaftlicher Behandlung» gleichgesetzt.

Die Frage ob eine Ärztin oder ein Arzt seine Leistungen unwirtschaftlich erbringt, kann ausschliesslich im Rahmen einer dem Screening nachgelagerten Einzelfall-Analyse, welche die vergleichsweise hohen Kosten einer «auffälligen» Praxis im Detail untersucht, beantwortet werden!