

Lukas Walter

## **Beweisrechtliche Anforderungen an verfügende Algorithmen**

**Aufgezeigt an der Wirtschaftlichkeitsprüfung nach Art. 56 KVG**

---

Durch den technologischen Fortschritt ist es möglich, dass Algorithmen automatisierte Entscheidungen treffen können. Soll ein Algorithmus über einen rechtlichen Tatbestand entscheiden bzw. eine Verfügung erlassen, ergeben sich eine Reihe rechtlicher Fragen. Dieser Beitrag widmet sich der Beziehung zwischen Algorithmus und Beweisrecht, und identifiziert beweisrechtliche Anforderungen an verfügende Algorithmen. Anhand eines Konzeptes aus der induktiven Statistik und am Beispiel der Wirtschaftlichkeitsprüfung nach Art. 56 KVG zeigt der Autor auf, wie ein verfügender Algorithmus auszugestaltet ist, damit dieser mit dem Beweisrecht vereinbar ist.

---

Beitragsart: LegalTech

Region: Schweiz

Rechtsgebiete: Artificial Intelligence & Recht; Datenschutz; LegalTech

Zitiervorschlag: Lukas Walter, Beweisrechtliche Anforderungen an verfügende Algorithmen, in: Jusletter IT 12. November 2020

## Inhaltsübersicht

1. Einleitung
2. Übersetzung des Beweisrechts in das Neyman-Pearson-Paradigma
3. Das Neyman-Pearson-Paradigma im Wirtschaftlichkeitsverfahren
  - 3.1. Nullhypothese: Der Leistungserbringer arbeitet effizient
  - 3.2. Erhöhtes Signifikanzniveau: Überwiegende Wahrscheinlichkeit
  - 3.3. Rechtspolitischer Zweck der Wirtschaftlichkeitsprüfung
4. Beweisrechtliche Anforderungen an einen Algorithmus
  - 4.1. Gestaltungsmöglichkeiten eines Algorithmus im Wirtschaftlichkeitsverfahren
  - 4.2. Anforderungen an die Nullhypothese
  - 4.3. Anforderungen an das Signifikanzniveau
5. Der umgekehrte Fall: Anforderungen eines Algorithmus an das Beweisrecht
6. Fazit

### 1. Einleitung

[1] Die Feststellung des rechtserheblichen Sachverhalts ist nicht nur richterliches Alltagsgeschäft, sondern auch äusserst schwierig. Der Zustand von Gewissheit ist dabei grundsätzlich unerreichbar, was für den Menschen, in dessen Natur es liegt nach Sicherheit zu streben, Unbehagen schafft.<sup>1</sup> Kann an dieser Stelle die Technik Abhilfe schaffen? Der technische Fortschritt ermöglicht es, dass Algorithmen selbstständig Entscheidungen treffen können (sog. automatisierte Einzelentscheidungen). Das Potential der Präzision und Geschwindigkeit von Algorithmen wird hoch gewertet, sie werden aber auch in verschiedener Hinsicht kritisiert.<sup>2</sup> Während das geltende Schweizer Recht automatisierte Einzelentscheidungen nicht konkret regelt, werden diesbezüglich spezifische Anforderungen *de lege ferenda* in Art. 19 E-DSG normiert.<sup>3</sup> Gleichzeitig werden mit der DSG Revision automatisierte Einzelentscheidungen für einzelne Rechtsbereiche ohne besondere Anforderungen zugelassen.<sup>4</sup>

[2] Wie in der Schweiz mit automatisierten Einzelentscheidungen rechtlich ganz grundsätzlich umgegangen werden soll, ist noch wenig erforscht.<sup>5</sup> Diesbezügliche Diskussionen adressieren unter anderem die Anwendung von Algorithmen im Verwaltungsrecht, insbesondere den automatisierten Erlass von Verfügungen.<sup>6</sup> Dabei wird jedoch nicht konkret behandelt, in welcher Beziehung die Ausgestaltung eines Algorithmus, der über einen rechtlichen Tatbestand entscheiden soll, zum zugehörigen Beweisrecht steht und welchen Einfluss dessen Ausgestaltung auf den mit der entsprechenden Norm verfolgten rechtspolitischen Zweck hat. Dieser Frage widmet sich der vorliegende Beitrag.

---

<sup>1</sup> SCHWEIZER MARK, Beweiswürdigung und Beweismass, Rationalität und Intuition, Tübingen 2015, 1.

<sup>2</sup> BOYD DANAH/CRAWFORD KATE, Critical Questions for Big Data, in: 15/Information, Communication & Society 2012, 662 ff., 662; THOUVENIN FLORENT/FRÜH ALFRED/GEORGE DAMIAN, Datenschutz und automatisierte Entscheidungen, in: Jusletter 26. November 2018, Rz. 4 ff.

<sup>3</sup> Bundesgesetz über den Datenschutz (Datenschutzgesetz), Entwurf vom 15. September 2017 (E-DSG), BBI 2017 7193, 7206 ff. Zur Tragweite von Art. 19 E-DSG, siehe ROSENTHAL DAVID, Der Entwurf für ein neues Datenschutzgesetz, in: Jusletter 27. November 2017, Rz. 100 ff.; THOUVENIN/FRÜH/GEORGE (Fn. 2), Rz. 27 ff.

<sup>4</sup> Vgl. Anhang zum E-DSG, 7235 ff. Für einen Überblick betroffener Tatbestände und kritische Würdigung, siehe RECHSTEINER DAVID, Der Algorithmus verfügt, in: Jusletter 26. November 2018, Rz. 16.

<sup>5</sup> THOUVENIN/FRÜH/GEORGE (Fn. 2), Rz. 2 f.

<sup>6</sup> Vgl. RECHSTEINER (Fn. 4); GLASER ANDREAS, Einflüsse der Digitalisierung auf das schweizerische Verwaltungsrecht, in: 114/SJZ 2018, 181 ff.; BRAUN BINDER NADJA, Weg frei für vollautomatisierte Verwaltungsverfahren in Deutschland, in: Jusletter IT 22. September 2016.

[3] In der schweizerischen Verwaltungsrechtspraxis ist der automatisierte Erlass von Verfügungen noch inexistent.<sup>7</sup> Die Beziehung Algorithmus und Beweisrecht soll aber nicht abstrakt, sondern an einer denkbaren Anwendung diskutiert werden. Als Beispiel dient das Wirtschaftlichkeitsverfahren des Krankenversicherungsrechts nach Art. 56 KVG.<sup>8</sup> Dabei geht es darum, dass Leistungserbringer, z.B. Arztpraxen, in der Krankenversicherung wirtschaftliche Leistungen erbringen sollen.<sup>9</sup> Diesem Rechtsgebiet kommt primär eine Beispielfunktion zu. Zudem ist die Wahl dadurch legitimiert, dass *de lege ferenda* für Bereiche des Versicherungsrechts – das KVG ist gemäss E-DSG zwar noch ausgenommen<sup>10</sup> – eine gesetzliche Grundlage vorgesehen wird, die automatisierte Einzelentscheidungen pauschal erlaubt.<sup>11</sup> Weiter werden im Beweisverfahren von Art. 56 KVG statistische Methoden berücksichtigt, was die Lehre kritisiert.<sup>12</sup> Ähnliche Bedenken sind, wie sich zeigen wird, auch für verfügende Algorithmen relevant.

[4] Die Beziehung zwischen der Ausgestaltung eines Algorithmus und dem Beweisrecht wird anhand des Neyman-Pearson-Paradigmas, einer bekannten Methode aus der induktiven Statistik, aufgezeigt. Diese Methode bietet einen einheitlichen analytischen Rahmen auf Stufe Beweisrecht und Algorithmus.<sup>13</sup> In einem ersten Schritt gilt es, die Logik des Beweisrechts in jene des Neyman-Pearson-Paradigmas zu übersetzen (Kapitel 2). Sodann wird das im Wirtschaftlichkeitsverfahren nach Art. 56 KVG geltende Beweisrecht mit der Terminologie des Neyman-Pearson-Paradigmas beschrieben (Kapitel 3). Folgend können die beweisrechtlichen Anforderungen an die Ausgestaltung eines entscheidenden Algorithmus im Anwendungsbereich des Wirtschaftlichkeitsverfahrens subsumiert und für weitere Rechtsbereiche generalisiert werden (Kapitel 4). Auch gilt es, die Sichtweise umzukehren und zu prüfen, welche Anforderungen denn ein Algorithmus an das bestehende Beweis- und Verfahrensrecht stellt (Kapitel 5). Abschliessend werden die Ergebnisse zusammengefasst und es wird auf weitere rechtliche Anforderungen an verfügende Algorithmen hingewiesen (Kapitel 6).

## 2. Übersetzung des Beweisrechts in das Neyman-Pearson-Paradigma<sup>14</sup>

[5] Eine rechtsanwendende Instanz, die einen Sachverhalt zu beurteilen hat, sieht sich mit dem Umstand konfrontiert, dass sie lediglich über lückenhafte Informationen der wirklichen Geschehnisse verfügt. Da ein Entscheid ergehen muss, ist es für die Behörde unausweichlich eine

---

<sup>7</sup> GLASER (Fn. 6), 184.

<sup>8</sup> Bundesgesetz über die Krankenversicherung vom 18. März 1999 (KVG; SR 832.10).

<sup>9</sup> KIESER UELI/GEHRING KASPAR/BOLLINGER SUSANNE, KVG/UVG Kommentar, Zürich 2018, Art. 56 KVG, Rz. 1 (zit. KIESER/GEHRING/BOLLINGER, Art. . . . KVG Rz. . . .).

<sup>10</sup> Vgl. aber die am 5. März 2020 eingereichte Motion Vasella 20.3013, welche die pauschale Ermächtigung zum Erlass automatisierter Einzelentscheidungen auch für das KVG vorsehen will.

<sup>11</sup> Vgl. Art. 96 Abs. 2 E-UVG, Art. 94a Abs. 2 E-MVG im Anhang zum E-DSG, 7271. Kritisch hierzu RECHSTEINER (Fn. 4), Rz. 16.

<sup>12</sup> Vgl. Kapitel 4.

<sup>13</sup> Als Analyseinstrument verwendet das Neyman-Pearson-Paradigma beispielweise HÄUSERMANN DANIEL M., Alkoholiker am Steuer: Wie eine falsch getestete Nullhypothese die Verkehrssicherheit gefährden kann, AJP 12/2014, 1636 ff. (zit. Alkoholiker am Steuer); HÄUSERMANN DANIEL M., Fehler erster und zweiter Art im Strafrecht – oder was das Nichtingangsetzen der Parkuhr mit induktiver Statistik zu tun hat, ZStrR 129/2011, 194 ff. (zit. Fehler erster und zweiter Art im Strafrecht).

<sup>14</sup> Die nachfolgenden Ausführungen orientieren sich an HÄUSERMANN, Alkoholiker am Steuer (Fn. 13); Fehler erster und zweiter Art im Strafrecht (Fn. 13). Auf einschlägige Literatur der Statistik wird dennoch verwiesen.

Hypothese über den tatsächlichen Sachverhalt zu bilden und diese anhand der verfügbaren, aber unvollständigen Informationen zu prüfen. Für diesen Konflikt bietet die induktive Statistik eine universell einsetzbare Methode, das Neyman-Pearson-Paradigma, welches ein Informations- in ein Entscheidungsproblem umwandelt. Diese Methode wird für statistische Testverfahren, u.a. für Algorithmen, die eine Aussage treffen sollen, angewendet. Mit dem Neyman-Pearson-Paradigma lässt sich aber auch das Beweisrecht beschreiben.<sup>15</sup> Dies erlaubt, wie sich zeigen wird, eine Untersuchung der Beziehung Algorithmus und Beweisrecht.

[6] Erstens wird beim Neyman-Pearson-Paradigma eine Ausgangshypothese, die *Nullhypothese*, aufgestellt. Sie wird als wahr angenommen, bis sie zugunsten der alternativen Hypothese verworfen werden kann.<sup>16</sup> Ein Beispiel für eine Nullhypothese wäre, dass ein Leistungserbringer i.S.v. Art. 56 KVG, beispielsweise eine Arztpraxis, effizient arbeitet und damit keine übermässigen Kosten verursacht. Beim Testen<sup>17</sup> einer Hypothese können zwei Fehler auftreten. Wird die Nullhypothese zu Unrecht verworfen, spricht man von einem Fehler erster Art; wird sie zu Unrecht akzeptiert, von einem Fehler zweiter Art. Dabei beschreibt das *Signifikanzniveau* die Wahrscheinlichkeit eines Fehlers erster Art. Obwohl es erstrebenswert erscheint beide Fehlerarten zu minimieren, ist dies bei gegebener Testmethodik gleichzeitig nicht möglich. Das Risiko von Fehlern der einen Art steigt, je zuverlässiger Fehler der anderen Art vermieden werden. Es besteht somit ein Zielkonflikt und es gilt normativ zu bestimmen, ob eher Fehler erster oder zweiter Art verhindert werden sollen. Die gleichzeitige Reduktion beider Fehlerarten kann lediglich durch eine verbesserte Testmethode und Datenqualität erreicht werden.<sup>18</sup>

[7] Im Beweisrecht folgt die Nullhypothese aus der objektiven Beweislast und entspricht dem Standpunkt jener Partei, die nicht beweisbelastend ist. Kann die Nullhypothese verworfen werden, ist der Beweis erfolgreich; ansonsten scheidet er. Das angewendete Signifikanzniveau ergibt sich aus dem Beweismass. Dieses bestimmt, ob das Gericht für das Vorhandensein einer bestimmten Tatsache einen strikten Beweis verlangt oder sich mit einem minderen Grad an Sicherheit begnügt, wobei Abstufungen von der Gewissheit bis zur überwiegenden Wahrscheinlichkeit bzw. Glaubhaftmachung möglich sind.<sup>19</sup> Gilt das Regelbeweismass der vollen Überzeugung, ist das Signifikanzniveau tief, höher ist es bei der überwiegenden Wahrscheinlichkeit, und noch höher bei der Glaubhaftmachung. Dies, weil mit einem abnehmenden Beweismass die Anforderungen an das Beweisen, also die Darlegung der Wahrheit, sinken, und damit die Wahrscheinlichkeit eines Fehlers erster Art – das Signifikanzniveau – zunimmt.<sup>20</sup>

---

<sup>15</sup> HÄUSERMANN, Alkoholiker am Steuer (Fn. 13), S. 1638; Fehler erster und zweiter Art im Strafrecht (Fn. 13), 194 ff.

<sup>16</sup> BORTZ JÜRGEN/SCHUSTER CHRISTOF, Statistik für Human- und Sozialwissenschaftler, 7. Auflage, Berlin 2010, 98 f.; HÄUSERMANN, Alkoholiker am Steuer (Fn. 13), 1638; Fehler erster und zweiter Art im Strafrecht (Fn. 13), 196.

<sup>17</sup> In der Statistik wird eine Nullhypothese mittels bestimmter statistischer Tests überprüft. BORTZ/SCHUSTER (Fn. 16), 99.

<sup>18</sup> BORTZ/SCHUSTER (Fn. 16), S. 100 f.; HÄUSERMANN, Alkoholiker am Steuer (Fn. 13) 1638; Fehler erster und zweiter Art im Strafrecht (Fn. 13) 194 ff.

<sup>19</sup> LARDELLI FLAVIO/VETTER MEINRAD, in: Geiser Thomas/Fountoulakis Christiana (Hrsg.), Basler Kommentar, Zivilgesetzbuch I, Art. 1–456 ZGB, 6. Auflage, Basel 2018, Art. 8 ZGB, Rz. 15 (zit. BSK-BEARBEITER, Art. . . . ZGB, Rz. . . .).

<sup>20</sup> BSK-LARDELLI/VETTER (Fn. 19), Art. 8 ZGB, Rz. 1; HÄUSERMANN, Alkoholiker am Steuer (Fn. 13), 1638.

### 3. Das Neyman-Pearson-Paradigma im Wirtschaftlichkeitsverfahren

[8] Nachfolgend gilt es, das im Wirtschaftlichkeitsverfahren geltende Beweisrecht mit dem Neyman-Pearson-Paradigma zu beschreiben. Die Ergebnisse dienen als Grundlage, um in Kapitel 4 die konkreten beweisrechtlichen Anforderungen an einen Algorithmus zu ermitteln.

[9] Das Gebot der Wirtschaftlichkeit durchdringt das KVG in vielfältiger Art. Die *sedes materiae* im Leistungsrecht sind die Art. 32 und 56 Abs. 1 KVG.<sup>21</sup> Art. 32 Abs. 1 KVG sieht vor, dass Leistungen wirksam, zweckmässig und wirtschaftlich sein müssen.<sup>22</sup> Gemäss Art. 56 Abs. 1 KVG hat sich der Leistungserbringer in seinen Leistungen auf das Mass zu beschränken, das im Interesse der Versicherten liegt und für den Behandlungszweck erforderlich ist. Diese Norm nimmt primär auf die Quantität medizinischer Leistungen Einfluss und regelt daher als *lex specialis* zu Art. 32 Abs. 1 KVG einen Anwendungsfall des allgemeinen Wirtschaftlichkeitsgebots im konkreten Behandlungsfall.<sup>23</sup> Dabei obliegt die Wirtschaftlichkeitskontrolle den Krankenversicherungen.<sup>24</sup> Da im KVG eine Pflichtleistungsvermutung besteht, gilt die im Einzelfall widerlegbare Vermutung, dass Leistungserbringer wirksame und zweckmässige Behandlungen anwenden.<sup>25</sup> Stellt ein Versicherer im Rahmen der Wirtschaftlichkeitskontrolle fest, dass eine Leistung über dem Mass von Art. 56 Abs. 1 KVG liegt, kann er die Vergütung verweigern, oder falls sie bereits bezahlt wurde, zurückfordern (Art. 56 Abs. 2 KVG). In einem Rückforderungsverfahren wird sich der aktivlegitimierte Krankenversicherer, ggf. mit dem Versicherten, und der passivlegitimierte Leistungserbringer – Hauptadressaten sind Ärzte – gegenüberstehen.<sup>26</sup>

#### 3.1. Nullhypothese: Der Leistungserbringer arbeitet effizient

[10] Im Wirtschaftlichkeitsverfahren gilt der Untersuchungsgrundsatz. Das Gericht stellt unter Mitwirkung der Parteien die für den Entscheid erheblichen Tatsachen fest, erhebt die notwendigen Beweise und ist in der Beweiswürdigung frei (Art. 89 Abs. 5 KVG). Im Verwaltungsrecht und somit im Krankenversicherungsrecht gilt gemäss herrschender Lehre auch die in Art. 8 ZGB<sup>27</sup> aufgeführte Beweislastregel.<sup>28</sup> Demnach tragen die Parteien nur dahingehend eine Beweislast, als der Entscheid zu Ungunsten desjenigen gefällt wird, der aus einem vorgebrachten Sachver-

---

<sup>21</sup> EUGSTER GEBHARD, Überarztung aus juristischer Sicht, Ein Konzert von Misstönen, in: Thomas Gächter/Myriam Schwendener (Hrsg.), Rechtsfragen zum Krankheitsbegriff – Entwicklung in der Praxis, Bern. 2009, Rz. 2 f. (zit. Überarztung, Rz. . . .); EUGSTER GEBHARD, Rechtsprechung des Bundesgerichts zum KVG, Zürich/Basel/Genf 2018, Art. 56 KVG, Rz. 1 (zit. Rechtsprechung des Bundesgerichts zum KVG, Art. . . . KVG, Rz. . . .).

<sup>22</sup> Es geht um die sog. WZW-Kriterien, das heisst um die Voraussetzung, dass die Leistung wirksam, zweckmässig und wirtschaftlich sein muss. Art. 32 KVG umschreibt die prinzipiell zu erfüllenden Voraussetzungen, damit die infrage stehende Leistung durch die Krankenversicherung vergütet wird. KIESER/GEHRING/BOLLINGER (Fn. 9), Art. 32 KVG, Rz. 1.

<sup>23</sup> EUGSTER, Überarztung (Fn. 21), Rz. 2 f.; Rechtsprechung des Bundesgerichts zum KVG (Fn. 21), Art. 56 KVG, Rz. 1.

<sup>24</sup> EUGSTER, Rechtsprechung des Bundesgerichts zum KVG (Fn. 21), Art. 56 KVG, Rz. 4.; BGE 127 V 43 E. 2b. und 51; 133 V 359 E. 6.1.

<sup>25</sup> KIESER UELI, Wirtschaftlichkeitsverfahren in der Krankenversicherung, Neuste Entwicklungen und eine Würdigung, Gutachten 2011, <https://www.physicianprofiling.ch/kieser2011.pdf> (besucht am 9. Mai 2020), 19 (zit. Wirtschaftlichkeitsverfahren).

<sup>26</sup> Für Ausführungen zum Rechtsweg und der Legitimation, siehe HÄFELI SIMON, Ruinöse Unrechtsprechung, in: Jusletter 18. August 2008, Rz. 7 ff.

<sup>27</sup> Schweizerisches Zivilgesetzbuch vom 10. Dezember 1907 (ZGB; SR 210).

<sup>28</sup> HÄFELIN ULRICH/MÜLLER GEORG/UHLMANN FELIX, Allgemeines Verwaltungsrecht, 5. Auflage, Zürich/Basel/Genf 2006, Rz. 1623 ff.; BSK-LARDELLI/VETTER (Fn. 19), Art. 8 ZGB, Rz. 27; EUGSTER, Überarztung (Fn. 21), Rz. 66.

halt Rechte ableiten will.<sup>29</sup> Mit Bezug auf geltend gemachte Verstösse gegen Art. 32 und 56 Abs. 1 KVG tragen die Krankenversicherer die Beweislast, wenn die von ihnen behauptete Unwirtschaftlichkeit der Behandlung trotz behördlicher Sachverhaltsabklärung unbewiesen bleibt, denn sie wollen aus dem unbewiesen gebliebenen Sachverhalt Rechte ableiten.<sup>30</sup>

[11] Da die Nullhypothese dem Standpunkt der nicht beweisbelasteten Partei, hier des Leistungserbringers, entspricht,<sup>31</sup> lautet diese im Rückforderungsverfahren: *Die Behandlung des Leistungserbringers ist wirtschaftlich*. Demnach bedeutet ein Fehler erster Art, dass der Unwirtschaftlichkeitsbeweis des Versicherers akzeptiert wird, obschon der Leistungserbringer tatsächlich eine wirtschaftliche Behandlung erbringt. Ein Fehler zweiter Art bedeutet hingegen, dass der Unwirtschaftlichkeitsbeweis des Versicherers nicht akzeptiert wird, obwohl der Leistungserbringer tatsächlich eine unwirtschaftliche Behandlung erbringt. Abbildung 1 fasst die Fehler erster und zweiter Art für die gegebene Nullhypothese zusammen.

	Die Behandlung des Leistungserbringers ist tatsächlich ...	
	wirtschaftlich.	unwirtschaftlich.
Unwirtschaftlichkeitsbeweis erbracht	Rückforderung trotz wirtschaftlicher Leistung (Fehler erster Art)	✓
Unwirtschaftlichkeitsbeweis nicht erbracht	✓	Keine Rückforderung trotz unwirtschaftlicher Leistung (Fehler zweiter Art)

Abbildung 1: Das Entscheidungsproblem nach Neyman-Pearson im Wirtschaftlichkeitsverfahren nach Art. 56 KVG; Quelle: Eigene Darstellung.

[12] Die hier gezeigte Beweislastverteilung erfasst den Grundsatz qua materiellem Recht (Art. 89 Abs. 5 KVG; Art. 8 ZGB), ist jedoch mit Bezug auf die Rechtsprechung des Bundesgerichts über die Würdigung von erbrachten Beweisen mittels statistischer Wirtschaftlichkeitsprüfung seitens Krankenversicherer zu relativieren. Die Lehre kritisiert diese Praxis dahingehend, dass faktisch eine Beweislastumkehr zulasten der Leistungserbringer stattfindet.<sup>32</sup> Dies wird in Kapitel 5 genauer erörtert.

### 3.2. Erhöhtes Signifikanzniveau: Überwiegende Wahrscheinlichkeit

[13] Wie allgemein im Sozialversicherungsrecht ist auch im Rückforderungsverfahren das Beweismass der überwiegenden Wahrscheinlichkeit anwendbar.<sup>33</sup> Das Gericht muss eine begründete Überzeugung erhalten, dass der erstellte Sachverhalt wahrscheinlich der Wirklichkeit ent-

<sup>29</sup> BGE 115 V 113 E. 8a.; HÄFELIN/MÜLLER/UHLMANN (Fn. 28), Rz. 351; BSK-LARDELLI/VETTER (Fn. 19), Art. 8 ZGB, Rz. 27; EUGSTER, Überarztung (Fn. 21), Rz. 66; HÄFELI (Fn. 26), Rz. 19.

<sup>30</sup> EUGSTER, Überarztung (Fn. 21), Rz. 66.; Rechtsprechung des Bundesgerichts zum KVG (Fn. 21), Art. 56 KVG, Rz. 6.

<sup>31</sup> Vgl. Kapitel 2.

<sup>32</sup> HÄFELI (Fn. 26), Rz. 20.; EUGSTER, Überarztung (Fn. 21), Rz. 64.

<sup>33</sup> EUGSTER, Rechtsprechung des Bundesgerichts zum KVG (Fn. 21), Art. 56 KVG, Rz. 19.; BSK-LARDELLI/VETTER (Fn. 19), Art. 8 ZGB, Rz. 19.

spricht, ihm keine konkreten Einwände entgegengehalten werden können und somit jeder erhebliche Zweifel ausgeschlossen ist.<sup>34</sup>

[14] In der Terminologie des Neyman-Pearson-Paradigmas beschreibt das Signifikanzniveau die Wahrscheinlichkeit eines Fehlers erster Art, hier die Rückforderung einer Leistung, die tatsächlich wirtschaftlich ist. Das Beweismass bestimmt das Signifikanzniveau.<sup>35</sup> Dieses ist beim Beweismass der überwiegenden Wahrscheinlichkeit relativ zum Regelbeweis der vollen Überzeugung höher, denn die Anforderungen an einen Unwirtschaftlichkeitsbeweis sind tiefer. Mit dem gesenkten Beweismass erhöht sich relativ zum strikten Beweis das Signifikanzniveau, also die Wahrscheinlichkeit, dass der Rückforderung einer Leistung stattgegeben wird, diese jedoch tatsächlich wirtschaftlich ist (Fehler erster Art).

### 3.3. Rechtspolitischer Zweck der Wirtschaftlichkeitsprüfung

[15] Die Wirtschaftlichkeitskontrolle ist ein belastender Eingriff in die Therapiewahlfreiheit; einerseits in die Wahlfreiheit der Versicherten, welche ein Teilgehalt der persönlichen Freiheit (Art. 10 Abs. 2 BV<sup>36</sup>) darstellt, und andererseits in die medizinische Berufsausübungsfreiheit der Leistungserbringer als Ausfluss der Wirtschaftsfreiheit (Art. 27 Abs. 2 BV). Der Eingriff lässt sich hingegen mit öffentlichen Interessen (Art. 36 Abs. 2 BV), namentlich mit jenen an der Erhaltung der Leistungsfähigkeit der obligatorischen Krankenpflegeversicherung sowie einer qualitativ hochstehenden medizinischen Versorgung, rechtfertigen. An sich ist die Wirtschaftlichkeitskontrolle verhältnismässig, jedoch müssen die Anforderungen an die Wirtschaftlichkeit der Leistung sowie die Kontrollmethoden auch im Einzelfall verhältnismässig erscheinen (Art. 36 Abs. 3 BV).<sup>37</sup>

[16] Dem allgemeinen Ziel des Wirtschaftlichkeitsgebots, die Krankenversicherung und indirekt die ganze Versicherungsgemeinschaft vor einer ungebührlichen Inanspruchnahme zu schützen,<sup>38</sup> trägt auch das Bundesgericht Rechnung. Dies wird am Beweismass der überwiegenden Wahrscheinlichkeit und mit Bezug auf das Neyman-Pearson-Paradigma ersichtlich.<sup>39</sup> Wie bereits erörtert, können Fehler erster und zweiter Art nicht gleichzeitig minimiert werden. Es besteht naturgemäss ein Zielkonflikt, der normativ zugunsten des einen oder des anderen Fehlers zu lösen ist.<sup>40</sup> Durch das gesenkte Beweismass bei gegebener Nullhypothese nimmt das Bundesgericht ein relativ höheres Signifikanzniveau und daher relativ mehr Fehler erster Art – Rückforderungen von Leistungen, die jedoch wirtschaftlich wären – in Kauf. Dafür sinkt die Wahrscheinlichkeit von Fehlern zweiter Art, also jene Fälle, in denen der Unwirtschaftlichkeitsbeweis nicht gelingt und damit Leistungen akzeptiert werden, die tatsächlich unwirtschaftlich sind. Das Bundesge-

---

<sup>34</sup> KIESER, Wirtschaftlichkeitsverfahren (Fn. 25), S. 16; HAEFELI (Fn. 26), Rz. 16. Die Definition der überwiegenden Wahrscheinlichkeit erscheint in sozialversicherungsrechtlichen Entscheiden weniger streng als in zivilrechtlichen Entscheiden: Mit Hinweisen auf die Rechtsprechung, siehe HÄUSERMANN, Alkoholiker am Steuer (Fn. 13), 1639, Fn. 21.

<sup>35</sup> Vgl. Kapitel 2.

<sup>36</sup> Bundesverfassung der Schweizerischen Eidgenossenschaft vom 18. April 1999 (BV; SR 101).

<sup>37</sup> EUGSTER, Überarztung (Fn. 21), Rz. 8.

<sup>38</sup> BGE 126 V 340 E. 2c.; KIESER/GEHRING/BOLLINGER (Fn. 9), Art. 56 KVG, Rz. 4.

<sup>39</sup> Vgl. Kapitel 3.2.

<sup>40</sup> Vgl. Kapitel 2.

richt gewichtet damit das öffentliche Interesse an tiefen Prämien in den Sozialversicherungen höher als jenes der Therapiewahlfreiheit.<sup>41</sup>

[17] Ob dieser Praxis, insbesondere mit Bezug auf die vom Bundesgericht vorgenommene Beweiswürdigungspraxis im Wirtschaftlichkeitsverfahren, die in Kapitel 5 angesprochen wird, zu folgen ist, kann an dieser Stelle offengelassen werden.

#### 4. Beweisrechtliche Anforderungen an einen Algorithmus

[18] Automatisierte Einzelentscheidungen können auf einem regelbasierten oder mit künstlicher Intelligenz ausgestatteten Algorithmus beruhen. Erstere sind insofern einfach ausgestattet, als sie lediglich gewisse Voraussetzungen für die Bestimmung eines Prozessausgangs prüfen, letztere stützen sich hingegen auf Korrelationen, das heisst Erfolgsaussichten basierend auf vergangenen vergleichbaren Fällen.<sup>42</sup> Bei der Ausgestaltung eines Algorithmus zur automatisierten Einzelentscheidung werden verschiedene Eigenschaften normativ bestimmt. Diese Aspekte können ebenfalls wie das Beweisrecht mit dem Neyman-Pearson-Paradigma beschrieben werden, denn Algorithmen machen von statistischen Modellen Gebrauch<sup>43</sup> und auch ihnen liegt aufgrund ihrer vorliegenden Funktion – den Erlass einer Verfügung – ein Entscheidungs- und Informationsproblem zugrunde. Aspekte, die bei der Ausgestaltung eines Algorithmus normativ bestimmt werden, sind auch die Nullhypothese sowie das Signifikanzniveau.<sup>44</sup>

[19] In Kapitel 3 wurde dargelegt, dass die im Wirtschaftlichkeitsverfahren geltende Nullhypothese lautet, dass *die Behandlung eines Leistungserbringers wirtschaftlich ist*, dass ein relativ zum strikten Beweis erhöhtes Signifikanzniveau besteht und der Zweck der Kontrolle darin liegt, dass Versicherer sowie die Versicherungsgemeinschaft vor einer ungebührlichen Inanspruchnahme geschützt werden sollen. Es stellt sich die Frage, wie die Nullhypothese und das Signifikanzniveau eines Algorithmus auszugestalten sind, damit dieser mit dem zugehörigen Beweisrecht und dem angestrebten rechtspolitischen Zweck von Art. 56 KVG konform ist.

[20] Im Wirtschaftlichkeitsverfahren nach Art. 56 KVG dürfen die Gerichte im Einklang mit dem Untersuchungsgrundsatz sämtliche Methoden zur Wirtschaftlichkeitsprüfung verwenden. In der Praxis geniessen statistische Methoden den Vorrang.<sup>45</sup> Diesbezüglich wird in der Lehre die vom Verband der Krankenkassen, santésuisse, eingesetzte ANOVA (Analysis of Variance)-Methode<sup>46</sup>

---

<sup>41</sup> HAEFELI (Fn. 26), Rz. 17.

<sup>42</sup> RECHSTEINER (Fn. 4), Rz. 2, Fn. 6; THOUVENIN/FRÜH/GEORGE (Fn. 2), Rz. 7.

<sup>43</sup> KELLEHER JOHN D./TIERNEY BRENDAN, *Data Science*, Cambridge, MA 2018, 23 f. Damit lässt sich das primär für die Statistik entwickelte Neyman-Pearson-Paradigma direkt anwenden.

<sup>44</sup> Deren Festlegung erfolgt in der Statistik im Rahmen des Hypothesentests. Für dessen Logik, siehe Kapitel 2; BORTZ/SCHUSTER (Fn. 16), 97.

<sup>45</sup> Im Gegensatz zur analytischen Methode. Für einen Überblick über die Methoden zur Wirtschaftlichkeitsprüfung, siehe EUGSTER, *Rechtsprechung des Bundesgerichts zum KVG* (Fn. 21), Art. 56 KVG, Rz. 4 ff; KIESER/GEHRING/BOLLINGER (Fn. 9), Art. 56 KVG, Rz. 8 ff.

<sup>46</sup> Im deutschen Varianzanalyse. Für eine statistische Auseinandersetzung mit der ANOVA-Methode zur Wirtschaftlichkeitsprüfung nach Art. 56 KVG siehe ROTH HANS-RUDOLF/STAHEL WERNER, *Die ANOVA-Methode zur Prüfung der Wirtschaftlichkeit von Leistungserbringern nach Art. 56 KVG*, Gutachten 2005, <https://www.docfind.ch/WZWAanovaGutachtenStahelRoth.pdf> (besucht am 9. Mai 2020); Polynomics AG, *Weiterentwicklung der statistischen Methode zur Prüfung der Wirtschaftlichkeit*, Schlussbericht, [https://www.tarifsuisse.ch/fileadmin/user\\_upload/tarifsuisse/Documents/Polynomics\\_Wirtschaftlichkeitspruefungen\\_Schlussbericht\\_2017-09-29\\_HP.pdf](https://www.tarifsuisse.ch/fileadmin/user_upload/tarifsuisse/Documents/Polynomics_Wirtschaftlichkeitspruefungen_Schlussbericht_2017-09-29_HP.pdf) (besucht am 23. Mai 2020).

zur Wirtschaftlichkeitsprüfung kontrovers diskutiert.<sup>47</sup> Die Methode nimmt eine automatische Vorselektion derjenigen Leistungserbringer vor, die auffallend hohe Kosten pro Erkrankten zu Lasten der Grundversicherung abrechnen.<sup>48</sup> Das Bundesgericht erachtet die ANOVA-Methode als geeignetes Beweismittel zur Feststellung der Unwirtschaftlichkeit sowie zur Bemessung des unwirtschaftlichen Mehraufwandes bzw. der Rückerstattungsforderung.<sup>49</sup> Die von der Lehre geschilderte Kritik bezüglich der Berücksichtigung dieser Methoden im Beweisverfahren wird nachfolgend integriert, um mögliche Konflikte von Algorithmen und Beweisrecht zu veranschaulichen.

#### 4.1. Gestaltungsmöglichkeiten eines Algorithmus im Wirtschaftlichkeitsverfahren

[21] Grundsätzlich bestehen sechs Möglichkeiten, einen Algorithmus mit Bezug auf die ihm zugrundeliegende Nullhypothese und dem vorgesehenen Signifikanzniveau auszugestalten. Die sechs Fälle sind in Abbildung 2 illustriert (Nr. 1–6).<sup>50</sup> Angewendet auf das Wirtschaftlichkeitsverfahren nach Art. 56 KVG wird sich zeigen, dass nur einer dieser sechs möglichen Algorithmen (Nr. 3) mit dem Beweisrecht gemäss bundesgerichtlicher Rechtsprechung und materiellem Recht vereinbar ist.

Formulierung Nullhypothese		Signifikanzniveau Konstellation (I) <i>H<sub>0(I)</sub>: Behandlung wirtschaftlich</i> Beweislast beim Krankenversicherer			Signifikanzniveau Konstellation (II) <i>H<sub>0(II)</sub>: Behandlung unwirtschaftlich</i> Beweislast beim Leistungserbringer		
		<i>tief</i>	<i>erhöht</i>	<i>hoch</i>	<i>tief</i>	<i>erhöht</i>	<i>hoch</i>
		$H_{0(I)} = H_{0(\text{Algorithmus})}$ bzw. $H_{0(II)} \neq H_{0(\text{Algorithmus})}$	Nr. 1	Nr. 3	Nr. 5	Nr. 5	Nr. 3
$H_{0(II)} = H_{0(\text{Algorithmus})}$ bzw. $H_{0(I)} \neq H_{0(\text{Algorithmus})}$	Nr. 2	Nr. 4	Nr. 6	Nr. 6	Nr. 4	Nr. 2	

**Legende**

■ Konform mit dem Beweisrecht. ■ Nicht konform mit dem Beweisrecht.

Abbildung 2: Gestaltungsmöglichkeiten eines Algorithmus im Falle der Beweislastregelung gemäss dem Wirtschaftlichkeitsverfahren nach Art. 56 KVG; Quelle: Eigene Darstellung.

<sup>47</sup> Die Varianzanalyse wurde im Rahmen von Art. 56 Abs. 6 KVG durch eine Vereinbarung von Versicherern und der Ärzteschaft als statistische Methode zur Wirtschaftlichkeitskontrolle festgelegt. EUGSTER, Rechtsprechung des Bundesgerichts zum KVG (Fn. 21), Art. 56 KVG, Rz. 44.

<sup>48</sup> ROTH/STAHEL (Fn. 46), 5.

<sup>49</sup> Ohne sich allerdings mit der Beweistauglichkeit auseinanderzusetzen. BGE 144 V 79; EUGSTER, Rechtsprechung des Bundesgerichts zum KVG (Fn. 21), Art. 56 KVG, Rz. 45.

<sup>50</sup> Lesehilfe zu Abbildung 2: Die Nullhypothese ist mit  $H_0$  abgekürzt. In der Konstellation (I) entspricht die Nullhypothese derjenigen des Beweisrechts im Wirtschaftlichkeitsverfahren, sie lautet also, dass die Behandlung des Leistungserbringers wirtschaftlich ist. Konstellation (II) erfasst den Fall der gegenteiligen Nullhypothese.  $H_{0(I)}$ ,  $H_{0(II)}$  und  $H_{0(\text{Algorithmus})}$  beschreibt die der jeweiligen Konstellation bzw. die dem Algorithmus zugrundeliegende Nullhypothese.

[22] Ein Algorithmus kann auf der gleichen Nullhypothese wie das Beweisrecht – *der Leistungserbringer erbringt eine wirtschaftliche Leistung* – basieren (Nr. 1, 3 & 5), oder auf der Umgekehrten, die lautet, dass der Leistungserbringer eine unwirtschaftliche Leistung erbringt (Nr. 2, 4 & 6). Weiter kann ein Algorithmus das gleiche Signifikanzniveau wie das Beweisrecht – *erhöhtes Signifikanzniveau; überwiegende Wahrscheinlichkeit* – vorsehen (Nr. 3 & 4) oder ein Abweichendes: Etwa ein tiefes Signifikanzniveau, wenn eine strikte Beweisregel vorgesehen wird (Nr. 1 & 2) oder ein Hohes im Falle der Glaubhaftmachung (Nr. 5 & 6) bei gegebener Nullhypothese von Konstellation (I).<sup>51</sup> In den Ausführungen der folgenden Kapiteln 4.2 und 4.3 wird jeweils auf Abbildung 2 Bezug genommen.

## 4.2. Anforderungen an die Nullhypothese

[23] Nicht konform mit dem anwendbaren Beweisrecht im Wirtschaftlichkeitsverfahren sind Algorithmen, die auf der gegenteiligen Nullhypothese beruhen (Nr. 2, 4 & 6). Die zum geltenden Beweisrecht gegenteilige Nullhypothese, dass ein Leistungserbringer eine Leistung grundsätzlich unwirtschaftlich erbringt, wäre mit dem Pflichtleistungsgrundsatz sowie der Beweislastregel von Art. 8 ZGB, wonach in einem Rückforderungsverfahren dem Krankenversicherer die Beweislast obliegt,<sup>52</sup> unvereinbar. In diesem Zusammenhang sind jedoch folgende Überlegungen zu berücksichtigen:

[24] Wird im Falle der vom Beweisrecht vorgesehenen Nullhypothese der Unwirtschaftlichkeitsbeweis erbracht, wird der Schluss gezogen, dass der Leistungserbringer unwirtschaftlich arbeitet. Hier wird die Nullhypothese, der Leistungserbringer arbeitet wirtschaftlich, verworfen. In der Statistik sowie auch bei der Anwendung von Algorithmen wird eine Nullhypothese erst dann verworfen, wenn ein signifikantes Ergebnis erzielt wird bzw. der Algorithmus eine gewünschte Präzision erreicht.<sup>53</sup>

[25] Wird im Falle der gegenteiligen Nullhypothese der Wirtschaftlichkeitsbeweis nicht erbracht, konnte ein Algorithmus kein genügend präziser Nachweis über die Wirtschaftlichkeit erbringen und die Nullhypothese damit nicht verwerfen. Hier stellt sich die Frage, ob die Nullhypothese, die nicht verworfen werden konnte, nun akzeptiert werden darf, das heisst ob auf die Unwirtschaftlichkeit der Leistung, die vermutet wird, geschlossen werden darf. Diese Schlussfolgerung wäre falsch. Korrekt wäre, dass ein nicht signifikantes Ergebnis kein Beleg dafür ist, dass die aufgestellte Nullhypothese richtig ist.<sup>54</sup> Indes lässt sich eine Testmethode, die für eine bestimmte Nullhypothese entwickelt wird, nicht ohne weiteres zur Überprüfung der gegenteiligen Nullhypothese einsetzen.<sup>55</sup> Wiederum kann nun gesagt werden, dass wenn der Nachweis der Unwirtschaftlichkeit im Falle der vom Beweisrecht vorgesehenen Nullhypothese nicht gelingt (diese somit nicht verworfen werden kann), nicht auf Wirtschaftlichkeit der Leistung geschlossen werden

---

<sup>51</sup> In Konstellation (II) ist die Nullhypothese sowie das Signifikanzniveau im Vergleich zur Konstellation (I) entsprechend der inversen Beweislastverteilung umgekehrt.

<sup>52</sup> Vgl. Kapitel 3. und 3.1.

<sup>53</sup> BORTZ/SCHUSTER (Fn. 16), 106.

<sup>54</sup> BORTZ/SCHUSTER (Fn. 16), 106.

<sup>55</sup> HÄUSERMANN, Alkoholiker am Steuer (Fn. 13), 1640. HÄUSERMANN legt dies mit folgendem Beispiel dar: Finden sich an einem Tatort Fingerabdrücke, beweist dies, dass die betreffende Person am Tatort war. Finden sich umgekehrt keine Fingerabdrücke an einem Tatort, beweist dies hingegen nicht, dass die Person nicht am Tatort war.

darf. Dies ist aus statistischer Sicht korrekt, jedoch nehmen das Beweisrecht bzw. die im Wirtschaftlichkeitsverfahren geltende Beweislast (Art. 8 ZGB) sowie der Pflichtleistungsgrundsatz eine Beweislosigkeit zugunsten der Leistungserbringer bewusst in Kauf.

[26] *Generell muss somit ein Algorithmus, der über einen rechtlichen Tatbestand entscheiden soll, auf der identischen Nullhypothese (der Position der nicht beweisbelastenden Partei) wie diejenige des zugehörigen Beweisrechts beruhen.*

[27] Mit Bezug auf die ANOVA-Methode wird diesem Umstand Rechnung getragen. Die Nullhypothese lautet nämlich, dass ein Leistungserbringer effizient arbeitet<sup>56</sup> und ist damit identisch mit jener des Beweisrechts im Wirtschaftlichkeitsverfahren.

### 4.3. Anforderungen an das Signifikanzniveau

[28] Weiter nicht konform mit dem anwendbarem Beweisrecht im Wirtschaftlichkeitsverfahren sind Algorithmen, die auf einem anderen Signifikanzniveau basieren (Nr. 1, 2, 5 & 6). Das Signifikanzniveau beschreibt die Wahrscheinlichkeit eines Fehlers erster Art.<sup>57</sup> Durch das relativ zum strikten Beweis gesenkte Beweismass der überwiegenden Wahrscheinlichkeit nimmt das Bundesgericht im Wirtschaftlichkeitsverfahren relativ mehr Fehler erster Art – Rückforderungen von Leistungen, die jedoch wirtschaftlich wären – in Kauf und trägt damit dem rechtspolitischen Zweck von Art. 56 KVG Rechnung.

[29] Würde ein verfügender Algorithmus im Anwendungsbereich von Art. 56 KVG ein aus Sicht der vom Beweisrecht vorgesehenen Nullhypothese tiefes Signifikanzniveau ausweisen (Nr. 1 & 2), würden relativ mehr Fehler zweiter Art – keine Rückforderung der Leistung trotz unwirtschaftlicher Behandlung – auftreten. Dies würde im Konflikt mit dem vom Bundesgericht vorgesehenen Beweismass bzw. Signifikanzniveau stehen.

[30] Würde der Algorithmus ein aus Sicht der vom Beweisrecht vorgesehenen Nullhypothese sehr hohes Signifikanzniveau ausweisen (Nr. 5 & 6), das heisst geringere Anforderungen an den Beweis stellen, könnte dies insofern zu einem überschüssenden Ergebnis führen, als in unverhältnismässiger Weise mehr Fehler erster Art in Kauf genommen würden; also Leistungserbringer als unwirtschaftlich deklariert werden, obwohl sie tatsächlich wirtschaftliche Leistungen erbringen. Das Bundesgericht sieht mit dem Beweismass der überwiegenden Wahrscheinlichkeit zwar relativ zum strikten Beweis ein erhöhtes Signifikanzniveau vor und gewichtet damit das Interesse an tieferen Prämien höher als jenes der Therapiewahlfreiheit.<sup>58</sup> Lediglich die Glaubhaftmachung von Unwirtschaftlichkeit bzw. das Vorsehen eines noch höheren Signifikanzniveaus wäre mit dem geltenden Beweisrecht jedoch wohl unvereinbar.

[31] Das Signifikanzniveau eines Algorithmus wird jedoch nicht bloss durch dessen abstrakte Festsetzung bestimmt, sondern massgeblich durch das zugrundeliegende Testverfahren sowie die Datenqualität beeinflusst.<sup>59</sup> Statistische Modelle können voreingenommen («biased») sein, wobei die Voreingenommenheit bereits in der Datengrundlage veranlagt sein kann.<sup>60</sup> Durch ein Testver-

---

<sup>56</sup> Polynomics AG (Fn. 46), 57.

<sup>57</sup> Vgl. Kapitel 2.

<sup>58</sup> Vgl. Kapitel 3.3.

<sup>59</sup> Vgl. Kapitel 2.

<sup>60</sup> BOYD/CRAWFORD (Fn. 2), S. 668; THOUVENIN/FRÜH/GEORGE (Fn. 2), Rz. 4 ff.

fahren sowie Daten, welche den tatsächlichen Sachverhalt bestmöglich erfassen, könnten sowohl Fehler erster als auch zweiter Art reduziert werden.<sup>61</sup>

[32] *Generell muss somit das Signifikanzniveau (die Wahrscheinlichkeit eines Fehlers erster Art) eines Algorithmus, der über einen rechtlichen Tatbestand entscheiden soll, mit demjenigen des zugehörigen Beweisrechts übereinstimmen. Diesbezüglich ist insbesondere die Verwendung einer adäquaten Testmethode und angemessenen Datengrundlage erforderlich.*

[33] Die Diskussion um Fehler erster und zweiter Art findet auch bezüglich der ANOVA-Methode statt. Die Ärzteschaft kritisierte, dass die bis zur beschlossenen Weiterentwicklung von 2016 geltende ANOVA-Methode zu einem hohen Anteil Fehler erster Art, das heisst zu einer relativ hohen Zahl als unwirtschaftlich deklariertes Ärzte, die jedoch die höheren Kosten medizinisch begründen konnten, neigte.<sup>62</sup> Am Testverfahren sowie der Datengrundlage der ANOVA-Methode wurde rege Kritik geäussert.<sup>63</sup>

[34] Mit der Weiterentwicklung der ANOVA-Methode wurde versucht, dieser Kritik Rechnung zu tragen.<sup>64</sup> Jedoch bleibt die Diskussion um Fehler erster und zweiter Art auch mit der weiterentwickelten Methode aktuell. Die Polynomics AG, welche mit der Entwicklung der Methode beauftragt wurde,<sup>65</sup> konstatiert: «*Es ist also eine Abwägung zu treffen zwischen dem Risiko, Praxen zu identifizieren, die korrekterweise nicht unwirtschaftlich sind (falsch positive) und dem Risiko, Praxen nicht zu erkennen, die eigentlich unwirtschaftlich wären (falsch negative).*»<sup>66</sup>

## 5. Der umgekehrte Fall: Anforderungen eines Algorithmus an das Beweisrecht

[35] In Kapitel 4.2 und 4.3 wurden die beweisrechtlichen Anforderungen an die Ausgestaltung von verfügenden Algorithmen erörtert. Es drängt sich jedoch auf, die Betrachtungsweise umzukehren. Denn die Einsetzung eines Algorithmus als gänzlich entscheidende Instanz oder zu berücksichtigendes Beweismittel im Verwaltungsrecht stellt auch von sich aus Anforderungen an das entsprechende Beweis- und Verfahrensrecht.

[36] Die Eigenschaft eines vollständig automatisierten Verwaltungsverfahrens besteht im Vergleich zum analogen Verwaltungsverfahren in der Automatisierung der Sammlung, Auswertung und Verifizierung der Sachverhaltsdaten. Somit erübrigt sich grundsätzlich eine menschlich gesteuerte Auslösung, um den Erlass eines Verwaltungsaktes anzustossen. Es wird ersichtlich, dass der für das Verwaltungsverfahren prägende Untersuchungsgrundsatz modifiziert werden müsste. Beim Einsatz zum Erlass von Verwaltungsakten hat die Behörde die für den Einzelfall bedeut-

---

<sup>61</sup> Vgl. Kapitel 2.

<sup>62</sup> FMH – Verbindung der Schweizerischen Ärztinnen und Ärzte, Positionspapier der FMH: Neue statistische WZW-Screening-Methode der Krankenversicherer – kurz erklärt, Stand 12.12.2019, [https://www.fmh.ch/files/pdf7/positionspapier\\_ag\\_wzw\\_fmh\\_20191212\\_d-v1.pdf](https://www.fmh.ch/files/pdf7/positionspapier_ag_wzw_fmh_20191212_d-v1.pdf) (besucht am 23. Mai 2019), 1.

<sup>63</sup> Für eine Übersicht mit Hinweisen auf die entsprechende Literatur, siehe ROMANENS MICHEL, Wirtschaftlichkeitsverfahren, die keine sind: rechtliches Gutachten im Licht unabhängiger Statistik, SÄZ 92/2011, 338 ff.

<sup>64</sup> Polynomics AG (Fn. 46), 26.

<sup>65</sup> KESSLER THOMAS/D'ANGELO MIRJAM/TRITTIN ANKE, Wirtschaftlichkeitsprüfung – Weiterentwicklung und Überprüfung der statistischen Methode, Neue Wege bei der Wirtschaftlichkeitsprüfung, SÄZ 98/2017, S. 208 f., 209.

<sup>66</sup> Polynomics AG (Fn. 46), S. 70. «Falsch positiv» und «falsch negativ» bedeuten Fehler erster bzw. zweiter Art.

samen tatsächlichen Angaben des Beteiligten zu berücksichtigen, die in einem automatisierten Verfahren grundsätzlich nicht ermittelt würden.<sup>67</sup>

[37] Dem Untersuchungsgrundsatz könnte bei einem vollständig automatisierten Verwaltungsverfahren Rechnung getragen werden, indem die spezifischen Eigenschaften der Betroffenen in die algorithmenbasierte Entscheidung einfliessen. In diesem Zusammenhang sollten in die weiterentwickelte ANOVA-Methode spezifische Charakteristika des betroffenen Leistungserbringers, sog. «Praxisbesonderheiten», berücksichtigt werden.<sup>68</sup>

[38] *Generell muss somit im Fall eines vollständig automatisierten Verwaltungsverfahrens der Untersuchungsgrundsatz modifiziert werden oder mittels der Methodik, die dem Algorithmus zugrunde liegt, Angaben von den Beteiligten berücksichtigt werden.*

[39] Die ANOVA-Methode bestimmt aber im Vergleich zu einem vollständig automatisierten Verwaltungsverfahren nicht alleine den Prozessausgang. Das Screening Verfahren mit der ANOVA-Methode soll Ärzte detektieren, die im Vergleich zu ihrem Vergleichskollektiv Kosten aufweisen, die über einem Toleranzbereich liegen. Dies bedeutet noch nicht, dass die statistisch auffälligen Ärzte unbedingt unwirtschaftliche Leistungen erbringen. Die Ärzte sind im Rahmen einer umfassenden Einzelfallbeurteilung eingehend zu prüfen, bevor ein Entscheid gefällt werden kann.<sup>69</sup> Das Bundesgericht berücksichtigt die statistischen Indexwerte folgendermassen. Nach der Rechtsprechung stellen überhöhte Indexziffern zunächst lediglich ein Indiz für eine Unwirtschaftlichkeit dar. Im Rahmen des Beweisverfahrens sei abzuklären, ob sich die mit dem Index dargelegte Unwirtschaftlichkeit rechtfertige bzw. wegen Besonderheiten der fraglichen Arztpraxis die überhöhten Indizes ganz oder teilweise begründet seien.<sup>70</sup>

[40] Die Lehre kritisiert, dass praxisgemäss ein überhöhter Indexwert des evaluierten Leistungserbringers nicht bloss ein Indiz für Unwirtschaftlichkeit, sondern eine Tatsachenvermutung ist, die einen vollen Beweis unwirtschaftlicher Behandlung begründet, wenn sie sich nicht durch spezifische Charakteristika des Leistungserbringers, den Praxisbesonderheiten, widerlegen lässt. Dabei ist der geprüfte Leistungserbringer mit Bezug auf die Unwirtschaftlichkeitsvermutung widerlegenden Praxisbesonderheiten beweisbelastend. Die Überschreitung eines gewissen Indexwertes hat daher faktisch eine Beweislastumkehr zulasten des Leistungserbringers zur Folge.<sup>71</sup>

[41] Die Vereinbarkeit mit der Beweisregel von Art. 8 ZGB ist hierbei fraglich, denn eine Umkehr der Beweislast ist grundsätzlich nur zulässig, wo sie durch eine gesetzliche Sondernorm angeordnet wird oder wo die bestehende Beweislastregelung rechtsmissbräuchlich ist.<sup>72</sup> Weder das KVG noch ein anderes Bundesgesetz enthalten eine Sondernorm, die eine solche Umkehr der Beweislast gestattet. Damit erscheint die Rechtsprechung widersprüchlich zur geltenden Beweisregel von Art. 8 ZGB.<sup>73</sup>

[42] Hiermit stellen algorithmenbasierte Einschätzungen über einen rechtlichen Tatbestand, die von einem Gericht im Beweisverfahren berücksichtigt werden, von sich aus Anforderungen an die

---

<sup>67</sup> GLASER (Fn. 6), 184.

<sup>68</sup> Polynomics AG (Fn. 46), 69.

<sup>69</sup> KESSLER THOMAS/BRUNNER LUKAS/TRITTIN ANKE, Neue Screening-Methode im Rahmen der Wirtschaftlichkeitskontrolle, SÄZ 99/2018, S. 1390 f., 1390.

<sup>70</sup> Mit Hinweisen auf die Rechtsprechung, siehe EUGSTER, Überarztung (Fn. 21), Rz. 61.

<sup>71</sup> EUGSTER, Überarztung (Fn. 21), Rz. 64; HAEPFELI (Fn. 26), Rz. 20.

<sup>72</sup> Mit Hinweisen auf die Lehre und Rechtsprechung, siehe EUGSTER, Überarztung (Fn. 21), Rz. 65.

<sup>73</sup> EUGSTER, Überarztung (Fn. 21), Rz. 67.

Rechtsprechung. *Generell gilt es zu beachten, dass ein Gericht bei der Berücksichtigung von algorithmenbasierten Einschätzungen im Beweisverfahren deren Präzision berücksichtigt und nicht bestehende beweis- und verfahrensrechtliche Grundsätze aushebelt.*

## 6. Fazit

[43] Im vorliegenden Beitrag wurden gestützt auf das Neyman-Pearson-Paradigma die beweisrechtlichen Anforderungen an Algorithmen zur automatisierten Einzelentscheidung ermittelt. Die Überlegungen wurden am Wirtschaftlichkeitsverfahren nach Art. 56 KVG aufgezeigt.

[44] Es wurden zwei zentrale Gestaltungsaspekte eines Algorithmus identifiziert, welche bestimmen, ob dieser mit dem zugehörigen Beweisrecht, spezifisch mit der Beweislast und dem Beweismass, vereinbar ist. Dies sind die *Nullhypothese* sowie das *Signifikanzniveau*. Beide müssen für einen Algorithmus gleich ausgestaltet sein wie es das Beweisrecht für den entsprechenden Tatbestand, über den der Algorithmus entscheiden soll, vorsieht. Werden diese Anforderungen nicht beachtet, kann der rechtspolitische Zweck, der mit der entsprechenden Norm angestrebt wird, verfehlt werden. Weiter wurde aufgezeigt, dass ein Algorithmus von sich aus Anforderungen an das Verfahrens- und Beweisrecht stellt. In einem vollautomatisierten Verwaltungsverfahren wäre grundsätzlich der Untersuchungsgrundsatz zu modifizieren. Berücksichtigt eine rechtsanwendende Behörde algorithmenbasierte Einschätzungen nur konsultativ, hat diese insbesondere die Präzision des Algorithmus in der Rechtsprechung zu beachten.

[45] *«Bei der Würdigung von Tatsachen irren sich viele, auch die Klügsten».*<sup>74</sup> Fehler verursachen auch die angeblich als die *«Klügsten»* bezeichneten Algorithmen. Dieser Umstand ist bekannt und soll auch vom Recht, insbesondere mit Art. 19 E-DSG, adressiert werden. Diese Bestimmung sieht unter anderem vor, dass eine betroffene Person darüber zu informieren ist, wenn eine Entscheidung ausschliesslich auf einer automatisierten Datenbearbeitung beruht. Weiter räumt sie ihr unter gewissen Umständen das Recht ein, ihren Standpunkt vorzutragen und eine Überprüfung durch eine natürliche Person zu verlangen. Wie RECHSTEINER darlegt, wird mit der DSG Revision jedoch nicht allen verfassungs- und verwaltungsrechtlichen Vorgaben für algorithmenbasierte Entscheidungen, insbesondere mit Blick auf die Grundrechte der Betroffenen und die allgemeinen Verfassungsgrundsätze, Rechnung getragen.<sup>75</sup> Indes bleiben beweisrechtliche Anforderungen, wie der vorliegende Beitrag aufzeigte, unberücksichtigt.

[46] In welchen Bereichen und wann automatisierte Einzelentscheidungen in der schweizerischen Verwaltungsrechtspraxis überhaupt Anwendung finden werden, ist fraglich. Berücksichtigt man das E-DSG, so sind auf Bundesebene Anwendungen vor allem im Steuer- und Abgabebereich sowie in der Unfall- und Militärversicherung vorgesehen.<sup>76</sup> In zeitlicher Hinsicht verrät ein Blick nach Deutschland, wo der Gesetzgeber den Grundsatzentscheid zugunsten vollautomatisierter Verwaltungsverfahren bereits 2016 fällte, dass bis zur Umsetzung eines vollautomatisierten Besteuerungsverfahrens mit bis zu sieben Jahren gerechnet wird.<sup>77</sup> Man darf dabei gespannt sein,

---

<sup>74</sup> Im Original: *«Facti interpretatio plerumque etiam prudentissimos fallat»*. Digesta Iustiniani, Liber XXII, 6, 2, in SCHWEIZER (Fn. 1), 1.

<sup>75</sup> RECHSTEINER (Fn. 4), Rz. 34 ff.

<sup>76</sup> RECHSTEINER (Fn. 4), Rz. 31.

<sup>77</sup> BRAUN BINDER (Fn. 6), Rz. 31.

welche weiteren Fragen sich durch die praktischen Erfahrungen mit automatisierten Einzelentscheidungen stellen werden.

---

LUKAS WALTER, Student der Rechtswissenschaft mit Wirtschaftswissenschaften an der Universität St. Gallen (HSG).