

Bundesgericht
Tribunal fédéral
Tribunale federale
Tribunal federal

9C 135/2022

Urteil vom 12. Dezember 2023

III. öffentlich-rechtliche Abteilung

Besetzung

Bundesrichter Parrino, Präsident,
Bundesrichter Stadelmann,
Bundesrichterin Moser-Szeless, Bundesrichter Beusch, Bundesrichterin Scherrer Reber,
Gerichtsschreiber Traub.

Verfahrensbeteiligte

A. _____,
vertreten durch Rechtsanwalt Marc Tomaschett,
Beschwerdeführer,

gegen

1. CSS Kranken-Versicherung AG, Recht & Compliance, Tribschenstrasse 21, 6005 Luzern,
2. Aquilana Versicherungen, Bruggerstrasse 46, 5400 Baden,
3. Supra-1846 SA, Avenue de la Rasude 8, 1006 Lausanne,
4. Sumiswalder Kranken- und Unfallkasse, Spitalstrasse 47, 3454 Sumiswald,
5. Genossenschaft Krankenkasse Steffisburg, Unterdorfstrasse 37, 3612 Steffisburg,
6. CONCORDIA Schweizerische Kranken- und Unfallversicherung AG, Rechtsdienst, Bundesplatz 15, 6002 Luzern,
7. Atupri Gesundheitsversicherung, Zieglerstrasse 29, 3007 Bern,
8. Avenir Assurance Maladie SA, Rue des Cèdres 5, 1920 Martigny,
9. KPT Krankenkasse AG, Wankdorffallee 3, 3014 Bern,

10. ÖKK Kranken- und Unfallversicherungen AG, Bahnhofstrasse 13, 7302 Landquart,
11. Vivao Sympany AG, Rechtsdienst, Peter Merian-Weg 4, 4052 Basel,
12. Kolping Krankenkasse AG, c/o Sympany Services AG, Peter Merian-Weg 4, 4052 Basel,
13. KLuG Krankenversicherung, Gubelstrasse 22, 6300 Zug,
14. EGK Grundversicherungen AG, Birspark 1, 4242 Laufen,

15. sodalis gesundheitsgruppe,
Balfrinstrasse 15, 3930 Visp,
16. SWICA Krankenversicherung AG,
Römerstrasse 38, 8401 Winterthur,
17. Galenos AG, Weltpoststrasse 19, 3015 Bern,
18. Mutuel Krankenversicherung AG,
Rue des Cèdres 5, 1920 Martigny,
19. Sanitas Grundversicherungen AG, Rechtsdienst,
Jänergasse 3, 8004 Zürich,
20. Philos Krankenversicherung AG, Rechtsdienst,
Rue des Cèdres 5, 1920 Martigny,
21. Assura-Basis SA,
Avenue Charles-Ferdinand-Ramuz 70, 1009 Pully,
22. Visana AG, Weltpoststrasse 19, 3015 Bern,
23. Agrisano Krankenkasse AG,
Laurstrasse 10, 5200 Brugg,
24. Helsana Versicherungen AG,
Zürichstrasse 130, 8600 Dübendorf,
25. sana24 AG, Weltpoststrasse 19, 3015 Bern,
26. vivacare AG, Weltpoststrasse 19, 3015 Bern,
alle handelnd durch Verein santésuisse,
Römerstrasse 20, 4500 Solothurn, und dieser
vertreten durch Advokat Dr. Vincent Augustin,
Beschwerdegegner.

Gegenstand
Krankenversicherung,

Beschwerde gegen das Urteil des Schiedsgerichts in Sozialversicherungsstreitigkeiten des Kantons Bern vom 18. Januar 2022 (200 19 551 SCHG).

Sachverhalt:

A.

Dr. med. A._____ ist Facharzt für Allgemeine Innere Medizin FMH in U._____ (BE). Er verfügt über eine Bewilligung zur Führung einer Praxisapotheke (sog. Selbstdispensation). Am 8. Juli 2019 führten verschiedene Krankenversicherer, vertreten durch den Branchenverband Santésuisse, beim Schiedsgericht in Sozialversicherungsstreitigkeiten des Kantons Bern Klage gegen Dr. med. A._____. Sie beantragten, der Beklagte sei wegen unwirtschaftlicher Behandlungsweise zu verpflichten, für das Jahr 2017 eine Rückzahlung von Fr. 336'368.- (gemäss Regressions-Index) oder von Fr. 790'761.- (gemäss ANOVA-Index), eventuell eine solche nach richterlichem Ermessen, zu leisten. Der Beklagte schloss auf Abweisung der Klage; eventuell sei eine alternative Prüfungsmethode anzuwenden.

Das Schiedsgericht vereinigte das Klageverfahren betreffend das Jahr 2017 mit einem hängigen Rückforderungsverfahren betreffend die Jahre 2013 bis 2015. Am 2. Dezember 2017 hatte es Dr. med. A._____ betreffend diesen Zeitraum zur Rückerstattung von Fr. 2'872'696.55 verpflichtet. Auf Beschwerde des Betroffenen hin hatte das Bundesgericht die Sache am 20. Dezember 2018 zur neuen Beurteilung an das Schiedsgericht zurückgewiesen (Urteil 9C 67/2018). Anlässlich einer Einigungsverhandlung vom 18. Juni 2020 verständigten sich die Parteien darauf, dass der Beklagte den Krankenversicherern für die Jahre 2013 bis 2015 per saldo aller Ansprüche einen Betrag von Fr. 500'000.- zahlt. Hinsichtlich des Jahres 2017 wurde keine Einigung erzielt. Das Schiedsgericht schrieb das Klageverfahren betreffend die Jahre 2013 bis 2015 als erledigt ab und führte es betreffend das Jahr 2017 weiter.

B.

Das kantonale Schiedsgericht hiess die Klage gut und verurteilte den Beklagten, den Klägern für das Jahr 2017 den Betrag von Fr. 266'998.40 zurückzuerstatten (Urteil vom 18. Januar 2022).

C.

Dr. med. A. _____ führt Beschwerde in öffentlich-rechtlichen Angelegenheiten mit den Anträgen, der schiedsgerichtliche Entscheid vom 18. Januar 2022 sei aufzuheben und die Klage abzuweisen. Eventuell sei die Sache zur Neuurteilung an die Vorinstanz zurückzuweisen und diese anzuweisen, zur Überprüfung der behaupteten Überarztung die analytische Methode anzuwenden, d.h. eine systematische Einzelfallprüfung oder repräsentative Einzelfallprüfung mit Hochrechnung resp. eine Einzelfallanalyse im Rahmen der Screening-Methode durchzuführen. Dem Rechtsmittel sei die aufschiebende Wirkung zuzuerkennen.

Die Krankenversicherer beantragen, die Beschwerde sei abzuweisen, soweit darauf eingetreten werden könne; eventuell sei die Sache an die Vorinstanz zurückzuweisen. Dr. med. A. _____ repliziert. Das Bundesamt für Gesundheit verzichtet auf eine Vernehmlassung.

Erwägungen:

1.

1.1. Bei Fusionen von Krankenversicherern tritt die Rechtsnachfolgerin in die materiellen Ansprüche und in die Prozessstellung der Vorgängerin ein. Das Rubrum dieses Entscheids ist von Amtes wegen anzupassen (vgl. Urteil 9C 508/2022 vom 15. Mai 2023 E. 1 mit Hinweisen).

1.2. Das Bundesgericht legt seinem Urteil den Sachverhalt zugrunde, den die Vorinstanz festgestellt hat (Art. 105 Abs. 1 BGG). Es kann die vorinstanzliche Sachverhaltsfeststellung auf entsprechende Rüge hin oder von Amtes wegen berichtigen oder ergänzen, wenn sie offensichtlich unrichtig (d.h. willkürlich) ist oder auf einer Rechtsverletzung im Sinn von Art. 95 BGG beruht (Art. 97 Abs. 1 und Art. 105 Abs. 2 BGG; BGE 145 V 57 E. 4). Recht wendet das Bundesgericht von Amtes wegen an (Art. 106 Abs. 1 BGG).

2.

Der Beschwerdeführer bringt vor, die Beschwerdegegner hätten die relative Verwirkungsfrist nicht eingehalten.

2.1. Nach Art. 25 Abs. 2 ATSG (in der bis Ende 2020 geltenden und hier anwendbaren Fassung; vgl. BGE 144 V 210 E. 4.3.1; 129 V 354 E. 1) erlischt der Rückforderungsanspruch mit Ablauf eines Jahres, nachdem die Versicherungseinrichtung davon Kenntnis erhalten hat, spätestens aber mit Ablauf von fünf Jahren nach der Entrichtung der einzelnen Leistung. Die Vorinstanz führt aus, die relative Verwirkungsfrist sei eingehalten, wenn die Rückforderung innerhalb eines Jahres nach Kenntnis der zur Beurteilung der Überarztung dienenden Statistik beim Schiedsgericht eingereicht wird (vgl. BGE 133 V 579 E. 4.3.4). Sie stellt fest, die Datenaufbereitung sei am 17. Juli 2018 erfolgt. Die am 8. Juli 2019 der Post übergebene Klageschrift wahre die Frist.

Der Beschwerdeführer wendet ein, nachdem er schon seit Jahren "überwacht" werde, wäre es der Santésuisse zumutbar gewesen, die Zahlen laufend zu erfassen und zu dokumentieren. Dies hätte schon vor dem 17. Juli 2018 geschehen müssen.

2.2. Nach der Rechtsprechung ist das Datum der Erstellung der Statistik massgeblich (Urteil K 9/07 vom 25. März 2008 E. 7.2). Der Beschwerdeführer macht zwar geltend, das Datum der aufbereiteten Statistik erscheine willkürlich, zumal auch in anderen Fällen stets der 17. Juli 2018 als Aufbereitungsdatum angegeben werde und es sehr unwahrscheinlich sei, dass die für die Santésuisse handelnde Dienstleisterin (SASIS AG) die dazugehörigen Statistiken alle am gleichen Tag erstellt habe. Es sind aber keine tatsächlichen Umstände ersichtlich, die es als willkürlich erscheinen lassen würden (vgl. Art. 105 Abs. 2 BGG; oben E. 1.2), auf das von der Vorinstanz festgestellte Datum abzustellen. Eine frühere Kenntnisnahme ist der Santésuisse auch mit Blick auf die Vorgeschichte (Verfahren betreffend die Statistikjahre 2013-2015) nicht anzurechnen. Denn erst die vollständigen, aggregierten Daten eines Jahres erlauben eine Wirtschaftlichkeitsprüfung für den betreffenden Zeitraum. Vom vorinstanzlich festgestellten Sachverhalt ausgehend ist der Verwirkungseinwand unbegründet.

3.

Strittig ist die Rückforderung der Krankenversicherer über Fr. 266'998.40. Grund der Forderung ist der Vorwurf an den Beschwerdeführer, seine im Jahr 2017 erbrachten ärztlichen Leistungen seien im Sinn von Art. 56 Abs. 1 KVG unwirtschaftlich gewesen (Art. 59 Abs. 1 lit. b KVG; BGE 141 V 25).

4.

4.1. Die zu Lasten der obligatorischen Krankenpflegeversicherung abgerechneten Leistungen müssen wirksam, zweckmässig und wirtschaftlich sein (Art. 32 Abs. 1 KVG). Der Leistungserbringer muss seine Leistungen auf das Mass beschränken, das im Interesse der Versicherten liegt und für den Behandlungszweck erforderlich ist (Art. 56 Abs. 1 KVG). Gegen Leistungserbringer, die gegen die im Gesetz vorgesehenen Wirtschaftlichkeits- und Qualitätsanforderungen verstossen, werden Sanktionen ergriffen. Dazu gehört die gänzliche oder teilweise Rückerstattung der Honorare, welche für nicht angemessene Leistungen bezogen wurden (Art. 59 Abs. 1 lit. b und Abs. 3 lit. a KVG).

Eine "Überarztung" (frz. polypragmasie, it. polipragmasia) liegt vor, wenn ein Arzt im Vergleich mit anderen Ärzten im gleichen Einzugsgebiet und mit etwa gleichem Krankengut im Durchschnitt erheblich mehr verrechnet, ohne dass er Besonderheiten geltend machen könnte, die den Durchschnitt beeinflussen (BGE 137 V 43 E. 2.2; 136 V 415 E. 6.2). Voraussetzung für die Beurteilung der Wirtschaftlichkeit nach der statistischen Methode (dazu E. 4.2) ist zunächst, dass die Referenzgruppe hinreichend ähnlich zusammengesetzt ist (BGE 137 V 43 E. 2.2). Nach der bisherigen, vor Einführung der (ab dem Rechnungsjahr 2017 anwendbaren) Screening-Methode (unten E. 4.3.2) geltenden Rechtsprechung ist eine unwirtschaftliche Tätigkeit nicht schon dann zu vermuten, wenn der Gesamtkostenindex der geprüften Arztpraxis den statistischen Mittelwert der Vergleichsgruppe (100 Indexpunkte) überschreitet. Vielmehr besteht eine Toleranz im Bereich von 120 bis 130 Indexpunkten. Darüber hinaus ist einem erhöhten Fallkostendurchschnitt infolge praxisspezifischer Kosteneffekte nach der bisherigen Praxis mit einem Zuschlag (zum Toleranzwert) Rechnung zu tragen, sofern die betreffenden Praxisbesonderheiten nicht bereits im Rahmen der anwendbaren Methode berücksichtigt sind (BGE 137 V 43 E. 2.2; Urteile 9C 259/2023 vom 18. September 2023 E. 5.2; 9C 656/2020 vom 22. September 2021 E. 4.2; GEBHARD EUGSTER, Krankenversicherung, in: Soziale Sicherheit, SBVR Bd. XIV, 3. Aufl. 2016, S. 680 Rz. 890). Einen Zuschlag rechtfertigen demnach Merkmale wie etwa sehr viele langjährige und/oder ältere Patientinnen und Patienten, eine überdurchschnittliche Zahl von Hausbesuchen, ein sehr hoher Anteil an ausländischen Patienten oder der Umstand, dass keine Notfallpatienten behandelt werden (vgl. Urteil 9C 558/2018 vom 12. April 2019 E. 8.2).

Die Frage, ob eine unwirtschaftliche ambulante ärztliche Praxistätigkeit gegeben ist, und - gegebenenfalls - die Rückforderung beurteilen sich nach unterschiedlichen Voraussetzungen. Für die Wirtschaftlichkeit ist eine Gesamtbetrachtung aller Kosten massgebend, die vom Verhalten des Arztes abhängig sind. Der Gesamtkostenindex umfasst direkte und veranlasste Kosten pro Erkrankten, also auch solche für verordnete Medikamente, Laborleistungen, Mittel und Gegenstände sowie Physiotherapie. Liegt der Gesamtkostenindex ausserhalb des Toleranzbereichs, ist das Wirtschaftlichkeitsgebot verletzt. In diesem Fall ist die Rückerstattungspflicht zu prüfen. Eine Rückforderung ist begründet, soweit auch der Index der (totalen) direkten Kosten pro Erkrankten den Toleranzwert übertrifft. Dementsprechend sind nur direkte Kosten (einschliesslich der vom betreffenden Arzt abgegebenen Medikamente) rückforderbar (BGE 137 V 43 E. 2.5; 133 V 37 E. 5.3; Urteile 9C 535/2014 vom 15. Januar 2015 E. 5.4 und 9C 732/2010 vom 7. April 2011 E. 3.3 und 3.4).

4.2. Die Rechtsprechung geht - auf der Grundlage der Situation vor Einführung des Screening-Modells (vgl. E. 4.3.2) - davon aus, die Wirtschaftlichkeit sei sowohl anhand der statistischen Methode (Durchschnittskostenvergleich) als auch der analytischen Methode (Einzelfallprüfung; dazu E. 4.4.2) - oder mittels einer Kombination beider Methoden - überprüfbar. Die statistische Methode beruht auf Statistiken, welche die Versicherer aufgrund der in Rechnung gestellten Vergütungen erstellen (Art. 76 lit. b KVV); dabei werden die einzelnen in Rechnung gestellten Leistungen dem Inhaber der auf der betreffenden Abrechnung vermerkten Zahlstellenregister (ZSR) -Nummer zugerechnet (BGE 135 V 237 E. 4.6.1; Urteil 9C 420/2022 vom 24. November 2022 E. 5.2).

Der Durchschnittskostenvergleich stand lange Zeit in der Kritik (dazu PATRIZIA GRATWOHL, Gilt die Screening-Methode als Beweismethode?: Rechtliche Einordnung der in Wirtschaftlichkeitsprüfungen nach Art. 56 KVG

verwendeten Regressionsanalyse, in: Jusletter 14. März 2022, Rz. 18 ff. mit Hinweisen; GEBHARD EUGSTER, KVG: Baustelle statistische Wirtschaftlichkeitsprüfung, in: Jusletter 27. August 2012, Rz. 13 f.). Diese Kritik wurde in parlamentarischen Initiativen aufgegriffen (GRATWOHL, a.a.O., Rz. 19 ff. mit Hinweisen). Die Diskussion führte schliesslich zur Einfügung von Art. 56 Abs. 6 in das KVG (Gesetzesänderung vom 23. Dezember 2011; in Kraft seit 1. Januar 2013). Nach dieser Bestimmung legen die Leistungserbringer und Krankenversicherer vertraglich eine Methode zur Kontrolle der Wirtschaftlichkeit fest.

4.3.

4.3.1. Dem gesetzlichen Auftrag von Art. 56 Abs. 6 KVG folgend verständigten sich der Berufsverband der Schweizer Ärztinnen und Ärzte (FMH) und die Verbände der Krankenversicherer (Santésuisse und Curafutura) in einem am 27. Dezember 2013/16. Januar 2014 abgeschlossenen Vertrag zunächst auf die Varianzanalyse ("An alysis o f Va riance", ANOVA) als statistische Methode zur Kontrolle der Wirtschaftlichkeit (Ziff. 1 des Vertrags; abrufbar, wie auch die späteren Vereinbarungen, unter <<https://tarifsuisse.ch/unsere-dienstleistungen/leistungsmanagement/wirtschaftlichkeitspruefung/>>). Zur Ermittlung des als Referenz dienenden Gesamtkostenindex resp. Fallgruppenwertes werden nach der ANOVA-Methode auf einer ersten Stufe die logarithmierten mittleren Kosten pro Arzt um den Effekt der Zusammensetzung seiner Patienten nach Alter und Geschlecht bereinigt; auf einer zweiten Stufe erfolgt eine Standardisierung nach Facharztgruppe (FMH, bundesweit) und Praxisstandort (Kanton; TROTTMANN/FISCHER/VON RECHENBERG/TELSER, Weiterentwicklung der statistischen Methode zur Überprüfung der Wirtschaftlichkeit: Schlussbericht der Polynomics AG, Studie im Auftrag von FMH, Santésuisse und Curafutura, 2017, S. 7).

Die Leistungserbringer und Versicherer kamen überein, das Varianzanalyse-Modell gemeinsam weiterzuentwickeln und unter anderem durch Morbiditätsvariablen zu ergänzen (Ziff. 2 des Vertrags). Inskünftig sollten möglichst viele äussere, durch das Verhalten des Arztes nicht direkt beeinflussbare kostenwirksame Faktoren bereits im statistischen Modell berücksichtigt werden. Die durchschnittlichen Behandlungskosten werden insbesondere durch die Morbidität der Patienten massgeblich beeinflusst. Mit dem Ausbau der Methode werden einerseits weniger Ärzte mit überdurchschnittlich morbidem Patientenkollektiv ungerechtfertigt auffällig; andererseits kann eine unwirtschaftliche Behandlungsweise von Leistungserbringern mit unterdurchschnittlich morbidem Patienten statistisch sichtbar werden. Da eine direkte Messung der Morbidität - etwa über diagnostische Codes - als schwierig gilt, bleibt nur eine indirekte Abbildung, so über Indikatoren wie Alter oder Medikamentenkonsum (KESSLER/D'ANGELO/TRITTIN, Neue Wege bei der Wirtschaftlichkeitsprüfung, in: SAeZ 2017 S. 208 f.).

4.3.2. Am 20. März 2018 vereinbarten die Verbände die Screening-Methode. Diese gilt für sämtliche Wirtschaftlichkeitskontrollverfahren ab dem Statistikjahr 2017. Es handelt sich um eine Weiterentwicklung der ANOVA-Methode. Die Screening-Methode besteht in einer zweistufigen Regressionsanalyse. Auf der ersten Stufe werden die - die Patienten betreffenden - bisherigen Morbiditätsfaktoren "Alter" und "Geschlecht" mit den Indikatoren "pharmazeutische Kostengruppen (Pharmaceutical Cost Groups, PCG)", "Franchisen" und "Spital- oder Pflegeheimaufenthalt im Vorjahr" ergänzt. Die den Leistungserbringer betreffenden Indikatoren der zweiten Stufe (Facharztgruppe und Standortkanton) bleiben unverändert. Die Vereinbarung verweist auf die Beschreibung der Screening-Methode resp. zweistufigen Regressionsanalyse im Schlussbericht der Polynomics AG von September 2017 (TROTTMANN und andere, a.a.O.).

Nicht zu verwechseln mit der zweistufigen Regressionsanalyse ist die Zweiteiligkeit der Wirtschaftlichkeitskontrolle insgesamt: Die Tarifpartner betrachten die vereinbarte Screening-Methode als bloss "ersten Schritt"; sie soll für die "Detektion von Ärzten mit auffälligen Kosten gemäss KVG zur Anwendung kommen" (dazu unten E. 4.4.1 a.E.).

4.3.3. Der Vollständigkeit halber ist auf eine nach dem angefochtenen Entscheid erfolgte aktualisierte Fassung des Vertrags über die Screening-Methode vom 1. Februar 2023 (in Kraft seit 1. Januar 2023) hinzuweisen. Diese erläutert das Modell zusätzlich: Bei Leistungserbringern mit erhöhten Indexwerten werde unter anderem auf der Basis des Regressionsberichts eine erste interne Einzelfallanalyse durch Santésuisse und/oder Versicherer durchgeführt (Ziff. 2 Abs. 4). Der Leistungserbringer erhalte Gelegenheit, seine Sicht darzulegen, d.h.

”allfällige im Rahmen der Screening-Methode nicht berücksichtigte Praxisbesonderheiten objektiv und nachvollziehbar aufzuzeigen, welche seine Praxis wesentlich von den Leistungserbringern seines Vergleichskollektivs unterscheiden und darum zu einem erhöhten Regressionsindexwert führen”. Die betreffenden Sachverhalte seien durch Santésuisse und/oder Versicherer zu prüfen und, wenn relevant, zu berücksichtigen. Mache der Leistungserbringer Praxisbesonderheiten geltend, sei er ”beweispflichtig” (Ziff. 2 Abs. 5). Es stehe ihm jederzeit frei, ”Rechtsvertreter oder weitere Experten zu bevollmächtigen und für die Verhandlungen bzw. das Verfahren beizuziehen” (Abs. 6). Der aktualisierte Vertrag beschränke nicht ”die Kompetenz der Versicherer zur Geltendmachung und Durchsetzung allfälliger weiterer Ansprüche, die sich auf eine andere Grundlage als den vorliegenden Screening-Vertrag stützen (z.B. Einzelfallkontrollen, analytische Methode, Tarifcontrolling, Rechnungskontrollen respektive retrospektive systematische Rechnungskontrollen, Durchschnittskostenvergleiche) ” (Ziff. 7 Abs. 1).

4.4. Die Screening-Methode als solche ist statistischer Natur.

4.4.1. Die ANOVA- und in deren Folge die Screening-Methode stellen eine Weiterentwicklung des Durchschnittskostenvergleichs dar. Der individuelle Fallwert (d.h. die vom überprüften Leistungserbringer während einer bestimmten Periode im Durchschnitt pro Patient durch Eigen- oder Drittleistungen ausgelösten Kosten des Krankenversicherers) wird dem Fallwert einer passenden Referenzgruppe (Gruppenfallwert) gegenübergestellt (GEBHARD EUGSTER, Wirtschaftlichkeitskontrolle ambulanter ärztlicher Leistungen mit statistischen Methoden, 2003, Rz. 498). Ein Durchschnittskostenvergleich weist auf eine Überarztung hin, wenn eine ins Gewicht fallende Zahl von Rechnungen desselben Arztes an eine Krankenkasse im Vergleich zur Zahl von Rechnungen von Ärzten in geographisch gleichem Tätigkeitsbereich und mit etwa gleichem Krankengut im Durchschnitt erheblich höher ist, ohne dass den Durchschnitt beeinflussende Besonderheiten geltend gemacht werden können (BGE 137 V 43 E. 2.2; JUANA VASELLA, in: Basler Kommentar zum KVG, 2020, N. 139 zu Art. 56 KVG). Die Berechnung der Indizes erfolgt aufgrund von Daten der Versicherer (Daten- und Tarifpool der SASIS AG, einem Unternehmen der Santésuisse-Gruppe).

Das Ziel, über die Definition der Referenzgruppe Vergleichbarkeit herzustellen, wird mittels einer sog. Regressionsanalyse verfolgt. Im Sprachgebrauch der Statistik meint ”Regression” den Einfluss einer (unabhängigen) Variablen (hier z.B. der Morbidität des Patientenkollektivs) auf eine (davon abhängige) Zielvariable (hier die Praxiskosten; vgl. DARIO PICECCHI, Das Wirtschaftlichkeitsgebot im Krankenversicherungsrecht, 2022, Rz. 566 mit Hinweisen). Anhand von gewichtenden Regressionsparametern resp. -koeffizienten wird bestimmt, in welchem Mass beispielsweise die Morbidität im konkreten Fall kostenwirksam ist. Dies erlaubt es, anerkannte praxisspezifische Effekte zu standardisieren, d.h. die betreffenden Abweichungen auf den Durchschnittswert der betreffenden Vergleichsgruppe zu normieren. Durch eine solche Neutralisierung verhaltensunabhängiger kostenrelevanter Faktoren kann letztlich der Kosteneffekt einer unwirtschaftlichen Behandlung isoliert werden (vgl. BGE 144 V 79 E. 5.3.1; KOHLER/LEBER/LANGHART/MOLINE/REUST/KESSLER, santésuisse Wirtschaftlichkeitsprüfung: Dokumentation der Umsetzung des Regressionsmodells, 2023, passim; PICECCHI, a.a.O., Rz. 567).

Im Unterschied zum ursprünglichen Durchschnittskostenvergleich wurden mit der Varianzanalyse (ANOVA-Methode) die Kosten der einzelnen Leistungserbringer erst einmal in Bezug auf die statistisch signifikanten Merkmale Alter und Geschlecht der Patienten sowie Kanton, in dem die ambulante ärztliche Tätigkeit ausgeübt wird, standardisiert: Die Kosten werden so ausgewiesen, als wäre die Alters- und Geschlechterverteilung in der untersuchten Praxis gleich wie in der Vergleichsgruppe als Ganze; ebenso wird simuliert, alle Leistungserbringer wären im selben Kanton tätig (Urteil 9C 558/2018 vom 12. April 2019 E. 7.1). Letzteres erlaubt es, den Vorteil einer möglichst grossen - gesamtschweizerischen - Vergleichsgruppe zu nutzen (vgl. BGE 144 V 79 E. 6.1) und gleichzeitig kantonale unterschiedliche Kostenniveaus (kantonale Unterschiede bezüglich TARMED-Taxpunktswerten, unterschiedliche Infrastruktur- und Personalkosten etc.) gerecht zu werden (PICECCHI, a.a.O., Rz. 555). Zur Verfügung stehen ANOVA-Indizes betreffend die direkten Kosten (ohne Medikamente), die Medikamentenkosten (direkt und veranlasst) sowie die totalen Kosten (jeweils pro erkrankte Person; Urteil 9C 656/2020 vom 22. September 2021 E. 3.3).

Die ab dem Statistikjahr 2017 anwendbare Screening-Methode ergänzt die im ANOVA-Modell enthaltenen Morbiditätsfaktoren ”Alter” und ”Geschlecht” mit den Indikatoren ”pharmazeutische Kostengruppen (PCG)”, ”Franchisen” und ”Spital- oder Pflegeheimaufenthalt im Vorjahr”. Diese Erweiterung um zusätzliche Morbidi-

tätsfaktoren bezweckt, Abweichungen von den Durchschnittskosten der betreffenden Facharztgruppe, die auf im Patientenkollektiv überdurchschnittlich häufig vorkommende, kostenintensive chronische Erkrankungen zurückzuführen sind, zu erfassen und von Auffälligkeiten abzugrenzen, die einer ineffizienten Behandlungsweise geschuldet sind.

Das Varianzanalysemodell ANOVA von 2014 wurde nur ansatzweise über seinen Zweck "zur Beurteilung von beobachteten Unterschieden zwischen Leistungserbringern innerhalb ihrer Vergleichsgruppe" beschrieben; im Übrigen wurde die Absicht zur Weiterentwicklung der Methodik (Ergänzung der "demographischen erklärenden Variablen" u.a. durch Morbiditätsvariablen, Erprobung und Validierung der Variablen auf der Grundlage "einer stetig sich verbessernden Datengrundlage in neuen statistischen Modellen") betont. Abgesehen von der erwähnten Integrierung von Morbiditätsfaktoren unterscheidet sich die Screening-Methode vom Vorgängermodell durch eine explizite Zweiteilung der Wirtschaftlichkeitsprüfung: Gegenstand der Vereinbarung ist allein die zweistufige Regressionsanalyse, die "als Screening-Methode zur Detektion von Ärzten mit auffälligen Kosten" dient. Auffällige Kosten sind nach diesem Konzept ausdrücklich nicht mit unwirtschaftlichem Arbeiten gleichzusetzen. Eine entsprechende Feststellung ist aufgrund einer Einzelfallanalyse ("zweiter Schritt") zu treffen, die nicht Gegenstand der Vereinbarung bildet.

4.4.2. Gegenstück der statistischen Vergleichsmethode ist die analytische Methode der Einzelfallprüfung, auch systematische Methode genannt. Sie besteht in einer konkreten Beurteilung der einzelnen Rechnungen eines Leistungserbringers auf ihre Wirtschaftlichkeit hin. Erhoben wird, ob der diagnostische und therapeutische Aufwand unter Berücksichtigung des Krankheitsbildes und des angestrebten Heilungserfolgs vertretbar war (VASELLA, a.a.O., N. 135 zu Art. 56 KVG; Urteil K 90/01 vom 27. November 2001 E. 3b; BGE 99 V 193 E. 3).

5.

Die Parteien und die Vorinstanz vertreten unterschiedliche Auffassungen darüber, ob die Vorgaben des tarifpartnerschaftlichen Vertrags betreffend die Screening-Methode nach formellen und materiellen Gesichtspunkten eingehalten sind (nachfolgend E. 5.1). Demgemäss ist auf die Ausgestaltung der Wirtschaftlichkeitsprüfung, namentlich die Zweiteilung in ein Screening zur "Detektion von Ärzten mit auffälligen Kosten" und eine anschliessende Einzelfallprüfung einzugehen; zu behandeln sind insbesondere die Auswirkungen auf die Methodenwahl (E. 5.2) und die beweisrechtlichen Folgen (E. 5.3). Strittig sind auch die Bemessung der Toleranzmarge im Rahmen des Screenings (E. 5.4) und die Handhabung kostenwirksamer Faktoren auf Stufe Einzelfallprüfung (E. 5.5). Schliesslich stellen sich Fragen im Verhältnis Wirtschaftlichkeitsprüfung und Schiedsgerichtsverfahren (E. 5.6-5.8).

5.1. Das kantonale Schiedsgericht erwägt, es bestehe kein Grund, dem Beklagten folgend anstelle der statistischen Methode eine analytische Einzelfallprüfung vorzunehmen. Es bestünden keine Praxisbesonderheiten, die von der angewendeten Screening-Methode ausgespart würden. Der Beschwerdeführer rügt, die Vorinstanz habe Praxisbesonderheiten nicht angemessen berücksichtigt. Er betont, das Screening diene nur der Feststellung von statistischen Auffälligkeiten, die gegebenenfalls eine vollständige Einzelfallprüfung analytischer Art auslösten. Die Beschwerdegegner vertreten die Ansicht, die dem Screening nachgelagerte Einzelfallprüfung komme nur zum Tragen, wenn zusätzliche, d.h. mit dem Screening-Modell noch nicht berücksichtigte Praxisbesonderheiten geltend gemacht würden.

5.2. Die seit dem Rechnungsjahr 2017 geltenden methodischen Vorgaben betonen die Zweiteilung der Wirtschaftlichkeitsprüfung.

5.2.1. Die Tarifpartner halten im Vertrag vom 20. März 2018 eingangs fest, die Screening-Methode finde "Anwendung als erster Schritt der Wirtschaftlichkeitskontrolle". Gegenstand des Vertrags sei nur die Screening-Methode. Diese solle "künftig für die Detektion von Ärzten mit auffälligen Kosten gemäss KVG zur Anwendung kommen". Das Screening beinhalte eine zweistufige Regressionsanalyse (vgl. oben E. 4.3.2). Weise ein Arzt auffällige Kosten auf, so heisse dies nicht per se, dass er unwirtschaftlich arbeite. Mittels "Einzelfallanalyse" sei näher abzuklären, ob dies zutrefte (Ziff. 1 und 2 des Vertrags).

5.2.2. In das Screening integrierte Faktoren verbessern die Treffgenauigkeit bei der Identifikation von Verdachtsfällen einer unwirtschaftlichen Behandlungsweise. Gleichzeitig bleibt aber intransparent, ob die einbezogenen Faktoren die tatsächlichen Verhältnisse effektiv abbilden. Je mehr Faktoren berücksichtigt werden, desto grösser wird diese Intransparenz. Umso wichtiger ist es, das provisorische Ergebnis des Screenings, d.h. den betreffenden Indexwert, anschliessend aufgrund einer vollständigen Einzelfallbetrachtung zu validieren.

5.2.3. Vor Einführung des Screening-Modells war ein Vorgehen über die analytische (statt statistische) Methode alternativ möglich (BGE 135 V 237 E. 4.6.1). Es ist davon auszugehen, dass die Tarifpartner die (vorläufige) Identifizierung von Auffälligkeiten, die auf eine unwirtschaftliche Behandlungsweise hindeuten könnten, ausschliesslich über das vereinbarte Screening bewerkstelligen wollen. Insofern besteht auch kein Raum mehr für eine Kombination der statistischen mit der systematischen Kontrolle, wie sie unter der ANOVA-Methode zulässig war, etwa um ein statistisches Ergebnis zu bestätigen, dessen Zuverlässigkeit zweifelhaft erschien (BGE 119 V 448 E. 4d; Urteile 9C 67/2018 vom 20. Dezember 2018 E. 12.3 und 9C 28/2017 vom 15. Mai 2017 E. 4.4; PICECCHI, a.a.O., Rz. 494). Einer solchen Verifizierung bedarf es nun nicht mehr, weil das Ergebnis des Screenings als solches nicht eine unwirtschaftliche Behandlungsweise aufzeigt, sondern erst einmal nur eine auffällige Kostenstruktur und damit die Notwendigkeit einer näheren Einzelfallprüfung.

5.2.4. Die statistische Auffälligkeit löst eine "Einzelfallanalyse" aus ("zweiter Schritt" der Wirtschaftlichkeitsprüfung). Dabei werden die in der zweistufigen Regressionsanalyse ermittelten individuellen Indexwerte auf die tatsächlichen Verhältnisse abgestimmt (vgl. HAEFELI/ SCHWARZ, Umsetzung der TARMED-Verträge - aber richtig!, SAeZ 2019 S. 1011 ff.; KESSLER/BRUNNER/TRITTIN, Neue Screening-Methode im Rahmen der Wirtschaftlichkeitskontrolle, SAeZ 2018 S. 1390 f.). Die Behandlungsweise des statistisch auffälligen Arztes wird individuell beurteilt; er erhält Gelegenheit, sein Kostenbild zu begründen. Bei der Einzelfallprüfung soll aber nicht eine Auswertung der Patientendossiers im Vordergrund stehen (vgl. KOHLER und andere, a.a.O., S. 3). Entgegen der Auffassung des Beschwerdeführers ist die "Einzelfallanalyse", mit welcher der im Screening erkannten Kostenauffälligkeit weiter nachgegangen wird, also nicht im Sinn der herkömmlichen analytischen Prüfmethode (alternativ zur statistischen Methode) zu verstehen.

Situativ wird freilich weiterhin je nach Bedarf auf ausgewählte Patientendossiers zurückzugreifen sein; so ist es etwa nicht ausgeschlossen, dass Praxisbesonderheiten, deren Kostenwirksamkeit nicht auf statistischem Weg bezifferbar ist, unter Beizug eines vertrauensärztlichen Dienstes anhand der stichprobeweisen Auswertung einer repräsentativen Anzahl von konkreten Rechnungen analysiert werden.

Anzumerken bleibt, dass das methodische Auswahlermessen nicht so weit geht, dass anstelle der Screening-Methode und der allfällig folgenden Einzelfallbeurteilung gleichsam voraussetzungslos auf eine andere Art der Prüfung ausgewichen werden dürfte (vgl. Vertrag betreffend die Screening-Methode in der Fassung vom 1. Februar 2023 Ziff. 7 Abs. 1; oben E. 4.3.3 a.E.). Eine solche methodische Beliebigkeit würde die Garantien einer gleichmässigen und angemessenen Kontrolle der Wirtschaftlichkeit unterlaufen, zu der der Gesetzgeber die beteiligten Akteure mit Art. 56 Abs. 6 KVG verpflichtet hat (vgl. E. 4.2).

5.3. Der Umstand, dass das Ergebnis des Screenings resp. der Regressionsanalyse noch keine Unwirtschaftlichkeit feststellt, zeitigt beweisrechtliche Auswirkungen.

5.3.1. Der Beschwerdeführer hat recht, wenn er geltend macht, dass mit dem Screening als solchem allein keine unwirtschaftliche Behandlung festgestellt wird und es demnach auch nicht Grundlage einer Rückforderung sein kann. Diese erfolgt immer erst aufgrund einer Einzelfallprüfung im Sinn der tarifpartnerschaftlichen Vereinbarung. Die Differenz zwischen dem - anhand der Screening-Methode, d.h. der zweistufigen Regressionsanalyse modifizierten - Durchschnittskostenwert des beurteilten Leistungserbringers und dem (auf 100 Punkte gestellten) Referenzwert der entsprechenden Vergleichsgruppe (zuzüglich Toleranzmarge) zeigt zunächst einmal nur, ob eine Auffälligkeit besteht; sie begründet einen entsprechenden Verdacht. Der im Screeningverfahren ermittelte "auffällige" Indexwert entspricht sozusagen einer red flag, einem Warnsignal hinsichtlich einer möglicherweise unwirtschaftlichen Behandlungsweise. Die diesfalls indizierte Einzelfallprüfung, bei der namentlich die individuellen Besonderheiten der Arztpraxis betrachtet werden, fördert anschliessend zutage, ob die zuvor infrage gestellte Wirtschaftlichkeit der Leistungserbringung auch tatsächlich im Sinn von Art. 59 KVG verletzt ist. Erst diese vollständige Feststellung kann Grundlage

einer Rückerstattungsklage beim kantonalen Schiedsgericht sein (vgl. Art. 59 Abs. 2 KVG).

5.3.2. Eine im Screening konstatierte Auffälligkeit erbringt somit keinen Nachweis für eine Überarztung, sondern leitet ein entsprechendes Feststellungsverfahren ein. Daraus folgt, dass die bislang herrschende Ansicht, die statistische Methode sei eine "Beweismethode" (BGE 136 V 415 E. 6.2), seit Einführung des Screening-Modells überholt ist (GRATWOHL, a.a.O., Rz. 32 f., 44 f.). Aus dem gleichen Grund führt eine im Screening angezeigte Auffälligkeit nicht zu einer Beweislastumkehr. Der allgemeinen Regel nach Art. 8 ZGB entsprechend bleiben nach Feststellung einer auffälligen Kostenstruktur grundsätzlich die Krankenversicherer beweisbelastet, zumal was die Folgen einer allfälligen Beweislosigkeit betrifft (sog. materielle Beweislast); dies indessen mit der Einschränkung, dass der betroffene Leistungserbringer im Zusammenhang mit geltend gemachten Praxisbesonderheiten zur Mitwirkung verpflichtet ist, soweit er über die Daten verfügt, die zur Interpretation der statistischen Daten erforderlich sind (vgl. Urteil 9C 259/2023 vom 18. September 2023 E. 5.6.2; VALÉRIE JUNOD, Polypragmasie: Analyse d'une procédure controversée, Cahiers Genevois et romands de sécurité sociale 2008, Nr. 40, S. 148). Soweit Praxisbesonderheiten nicht augenfällig sind, muss der Leistungserbringer jedenfalls glaubhaft machen, unter welchen Gesichtspunkten eine Einzelfallanalyse vorzunehmen ist und inwiefern die im Rahmen des Screenings erkannte Auffälligkeit einer Praxisbesonderheit zuzuschreiben sei.

5.4. Die Toleranzmarge ist praxisgemäss als Zuschlag von 20 bis 30 Punkten zum Referenzindex zu veranschlagen (oben E. 4.1). Die Vorinstanz fixiert die Toleranz bei 20 Punkten. Sie begründet dies damit, die neue Screening-Methode erhöhe die statistische Präzision wesentlich, indem sie Praxisbesonderheiten einschliesse, die unter der ANOVA-Methode noch nicht berücksichtigt gewesen seien.

Die Toleranzmarge ist Teil des Screenings von auffälligen Kostenstrukturen. Ihr Zweck ist es, den individuellen "Praxisstil" des Leistungserbringers zu respektieren. Sie wahrt den Grundsatz der ärztlichen Behandlungsfreiheit (BGE 103 V 145 E. 6; EUGSTER, in: SBVR Rz. 902; derselbe, Wirtschaftlichkeitskontrolle, Rz. 408, 806 ff.; derselbe, Überarztung aus juristischer Sicht, in: Rechtsfragen zum Krankheitsbegriff, Gächter/Schwendener [Hrsg.], 2009, Rz. 57). An dieser Stelle des - ansonsten der Kosteneffizienz verpflichteten - Verfahrens wird dem Umstand Rechnung getragen, dass der angestrebte Heilungserfolg (Behandlungsausgang) auf unterschiedlichen therapeutischen Wegen erreicht werden kann (vgl. PICECCHI, a.a.O., Rz. 524, 548; JUNOD, a.a.O., S. 154 f.). Soweit die Toleranz nicht dazu bestimmt ist, einer "technischen", methodenbedingten Standardabweichung Rechnung zu tragen, ist nicht ersichtlich, inwiefern die mit der Screening-Methode verbundene bessere Berücksichtigung der Morbidität - also einer Eigenschaft des Patientenkollektivs - den richtig verstandenen Toleranzbereich tangieren sollte. Auch quantifizierbare Praxisbesonderheiten sind nicht über die pauschale Toleranzmarge zu veranschlagen (EUGSTER, Wirtschaftlichkeitskontrolle, Rz. 807), sondern je für sich auf Grundlage einer statistischen Feststellung oder einer Schätzung (unten E. 5.5).

Die Vorinstanz überschreitet somit ihr methodisches Auswahlermessen (vgl. Urteil 9C 721/2020 vom 19. Oktober 2021 E. 4.3.2), wenn sie den Toleranzbereich mit Hinweis auf die verbesserte methodische Spezifikation in einen unveränderlichen Zuschlag von 20 Punkten verwandelt. Die Toleranzmarge bewegt sich unter Geltung der Screening-Methode weiterhin in einem Bereich von mindestens 20 bis höchstens 30 Punkten. Die Bestimmung des Zuschlags im Einzelfall steht im Ermessen der Krankenversicherer resp. des Schiedsgerichts. Für die Ermessensausübung wird beispielsweise begleitend sein, ob der betreffende Leistungserbringer auf bestimmte Krankheiten oder besondere Therapieformen spezialisiert ist, sofern diese nicht als Praxisbesonderheiten zu berücksichtigen sind.

5.5. Strittig ist des Weiteren, wie Faktoren, die die Kostenstruktur der Praxis des Beschwerdeführers beeinflussen, zu handhaben sind.

5.5.1. Trotz eines auffälligen Regressionsindexes können überdurchschnittliche Kosten gegebenenfalls durch praxisspezifische Effekte erklärt und damit der Verdacht einer unwirtschaftlichen Behandlungsweise widerlegt werden. Art und Umfang der Einzelfallprüfung richten sich nach augenscheinlichen oder glaubhaft gemachten Praxismerkmalen. Dies gilt ohne Weiteres, soweit die Variablen des zweistufigen Screenings die Merkmale nicht abschliessend erfassen. Es kann aber auch nötig erscheinen, bereits in der Regressionsanalyse berück-

sichtige Faktoren zu validieren (vgl. unten E. 5.5.3).

5.5.2. Das zu prüfende Merkmal kann praxistypologischer, "kategorialer" Art sein, namentlich wenn der Leistungserbringer über besondere Befugnisse (z.B. Führung einer Praxisapotheke) oder über eine fachliche Spezialisierung verfügt, die sich auf die Zusammensetzung des Patientenkollektivs auswirkt (vgl. EUGSTER, Wirtschaftlichkeitskontrolle, Rz. 415 ff. und 422 ff.).

Ein auf die Praxistypologie abzielendes Merkmal ist vorzugsweise schon im Rahmen des Screenings zu berücksichtigen, um falsch positive (auffällige) Treffer (oder auch falsch negative Befunde) von vornherein zu vermeiden. Dazu müsste indes eine entsprechende Referenzgruppe geschaffen werden (EUGSTER, Wirtschaftlichkeitskontrolle, Rz. 423; derselbe, Baustelle, N 28). Sofern keine spezifizierte Vergleichsgruppe gebildet werden kann (vgl. unten E. 6.2.3 und 6.6.1), bedarf es wenigstens eines entsprechenden Korrektivs auf der Stufe der Einzelfallprüfung, indem die Differenz zwischen Praxisindexwert und Vergleichsindex um den anhand statistischer Daten ausgewiesenen Praxiseffekt vermindert wird. Alternativ kann das Korrektiv auch durch eine ergänzende analytische (systematische) Einzelfallkontrolle eruiert werden, so mit einer stichprobeweisen Auswertung einer Anzahl repräsentativer Fälle der betreffenden Praxis und anschliessender Hochrechnung (vgl. Urteil 9C 67/2018 vom 20. Dezember 2018 E. 12.3; PETROV/HACK-LEONI, Methoden der Wirtschaftlichkeitskontrolle - ein Überblick, in: Pflgerecht 2021 S. 166 und 168).

5.5.3. Von den vorerwähnten "kategorialen" Merkmalen des Leistungserbringers zu unterscheiden sind Praxisbesonderheiten, die sich auf veränderliche Eigenschaften des Patientenkollektivs beziehen (vgl. EUGSTER, Wirtschaftlichkeitskontrolle, Rz. 409 ff.). Unter diesem Titel relevant sein können z.B. ein überdurchschnittlicher Ausländeranteil, häufige Hausbesuche in einem grossen Einzugsgebiet oder das Fehlen von Notfallpatienten (vgl. Urteil 9C 259/2023 vom 18. September 2023 E. 5.2). Doch auch Gesichtspunkte der Morbidität bedürfen einer näheren Betrachtung, soweit sie nicht bereits zuvor im Rahmen der Regressionsanalyse standardisiert worden sind. Entgegen der Auffassung von Beschwerdegegnern und Vorinstanz schliesst der Umstand, dass ein kostenwirksamer Faktor in das Screening-Modell integriert ist, eine weitere Berücksichtigung als Praxisbesonderheit nicht von vornherein aus. Zu bedenken ist, dass ein Morbiditätsfaktor nicht immer ohne Weiteres valid, d.h. im konkreten Fall tatsächlich geeignet ist, die erwartete Erkenntnis zu liefern. So kann es beispielsweise vorkommen, dass in einer Praxis aufgrund einer bestimmten fachlichen Spezialisierung überdurchschnittlich viele Chronischkranke mit hohem Betreuungsbedarf behandelt werden und aus krankheitsspezifischen Gründen trotzdem vergleichsweise wenig Medikamente verordnet oder abgegeben werden. In einer solchen Situation zeigt der Morbiditätsfaktor "Pharmaceutical Cost Groups" (PCG; vgl. E. 6.5.2) eine tiefe Morbidität an, die den tatsächlichen Verhältnissen nicht gerecht wird. Hier kann der Leistungserbringer auch unter der Screening-Methode geltend machen, dass ein entsprechendes Korrektiv im Einzelfall geprüft wird. Aus der (dossier- und/oder statistikgestützten) Einzelfallanalyse resultiert gegebenenfalls eine Korrektur, indem die Differenz zwischen Praxisindexwert und Vergleichsindex um den betreffenden Effekt vermindert wird.

5.5.4. Im Übrigen empfiehlt der (von den Tarifpartnern zum integrierten Bestandteil der Vereinbarung über die Screening-Methode erklärte) Schlussbericht der Polynomics AG, eine "Standardabweichung" resp. einen "Unsicherheitsfaktor für den praxisspezifischen Effekt" in die Betrachtung des Einzelfalls einzubeziehen: Danach könnten ausgeprägte "Ausreisser" einen praxisspezifischen Effekt verändern, obwohl sie - im Gegensatz zu Auffälligkeiten, die systematisch über den erwarteten Kosten liegen - wahrscheinlich nichts mit der Wirtschaftlichkeit der Behandlungsweise zu tun haben (TROTTMANN und andere, Schlussbericht der Polynomics AG, S. 8, 22 ff.; vgl. KOHLER und andere, a.a.O., S. 8). EUGSTER legt nahe, entsprechende Behandlungsfälle wenn möglich auszusondern anstatt sie als Praxisbesonderheit zu behandeln (in: SBVR, Rz. 890).

5.6. Der bisherige Verlauf des Verfahrens wirft sodann Fragen zum Verhältnis zwischen Wirtschaftlichkeitsprüfung und Schiedsgerichtsverfahren auf.

Angesichts des zweiteiligen Prüfverfahrens mit Screening und allfälliger Analyse des Einzelfalls darf die Auseinandersetzung mit kostenwirksamen Praxismerkmalen nicht auf das schiedsgerichtliche Verfahren verschoben werden, selbst wenn die Schiedsgerichte die für den Entscheid erheblichen Tatsachen unter Mitwirkung der Parteien - ergänzend - feststellen können, sie die notwendigen Beweise erheben und in der Beweiswürdi-

gung frei sind (Art. 89 Abs. 5 KVG). Bereits in der Phase der Einzelfallprüfung hat der Leistungserbringer die Möglichkeit, "allfällige im Rahmen der Screening-Methode nicht berücksichtigte Praxisbesonderheiten objektiv und nachvollziehbar aufzuzeigen, welche seine Praxis wesentlich von den Leistungserbringern seines Vergleichskollektivs unterscheiden und darum zu einem erhöhten Regressionsindexwert führen" (so nunmehr ausdrücklich festgehalten in Ziff. 2 Abs. 5 des Vertrags betreffend die Screening-Methode in der ab 1. Januar 2023 geltenden Fassung). Die Einzelfallprüfung durch die Krankenkassen resp. durch den von ihnen ermächtigten Verband ist somit partizipativ angelegt, dies auch mit dem Ziel, sich gegebenenfalls im Dialog mit dem überprüften Leistungserbringer gütlich zu einigen. Das kontradiktorische und förmliche Klageverfahren, in dem die prüfende Instanz zur Partei wird, ist weniger geeignet, die von hoher Technizität geprägte Aufarbeitung der Grundlagen des klageweise geltend gemachten Rückerstattungsanspruchs wirksam nachzuholen. Der Umstand, dass die Schiedsgerichte paritätisch mit einer Vertretung der Versicherer und der betroffenen Leistungserbringer besetzt sind (Art. 89 Abs. 4 KVG) und damit bis zu einem gewissen Grad auch über entsprechendes Fachwissen verfügen, ändert daran nichts. Die beim kantonalen Schiedsgericht einzubringende Klage muss daher auf den Ergebnissen einer kompletten Einzelfallprüfung beruhen. Aufgabe des Schiedsgerichts ist es, weiterhin bestrittene Erkenntnisse aus der Einzelfallprüfung zu überprüfen und den Rechtsstreit wenn möglich in einem Vermittlungsverfahren beizulegen oder andernfalls autoritativ zu entscheiden.

5.7. Vorliegend ergibt sich aus den Akten, dass keine Einzelfallprüfung stattgefunden hat.

5.7.1. In einem Schreiben vom 4. Juni 2019 informierte die Santésuisse den Beschwerdeführer über die Statistikdaten des Jahres 2017 und gab ihm die entsprechenden Kennzahlen bekannt. Sie stellte fest, der Regressionsindex von 144 Punkten überschreite die Toleranzschwelle und sei somit statistisch auffällig. Dem Schreiben lagen die Rechnungssteller-Statistiken 2013-2017 sowie die Wirtschaftlichkeitsprüfung für das Statistikjahr 2017 bei. Zum weiteren Vorgehen eröffnete die Santésuisse dem Beschwerdeführer, aufgrund der "bisherigen Abklärungen und Einschätzungen sowie den laufenden Gerichtsverfahren" werde eine "Eingabe aufgrund der Rechnungssteller-Statistik 2017 beim zuständigen kantonalen Schiedsgericht" folgen. Diese Eingabe diene in erster Linie der Fristwahrung. Die Berechnung der Rückforderung von Fr. 336'368.- berücksichtige noch keine weiteren Praxismerkmale. Zur Klärung der Kostenabweichung und gemeinsamen Lösungsfindung sei man nach wie vor an einer gütlichen Einigung interessiert. Das Schreiben schliesst mit den Sätzen: "Gerne erwarten wir Ihre Rückmeldung bis am 21. Juni 2019, ob Sie an bilateralen Gesprächen interessiert sind oder nicht. Ohne Ihren Gegenbericht gehen wir davon aus, dass Sie die weiteren Verhandlungen auf dem gerichtlichen Wege führen möchten. In diesem Fall sehen wir von einem Sistierungsantrag der Eingabe ab".

5.7.2. Mit Rückforderungsklage vom 8. Juli 2019 machten die Krankenversicherer für das Statistikjahr 2017 einen Regressionsindex totale Kosten von 144 Punkten geltend. Der Beklagte sei in der Regressionsanalyse mit der Facharztgruppe Allgemeine Innere Medizin zu vergleichen, die schweizweit 5'497 Praxen umfasse. Die Differenz zwischen dem Index totale Kosten des Arztes und dem Referenzindex 120 (d.h. einschliesslich einer Toleranzmarge von 20 Prozent) entspreche einer Kostenüberschreitung von 24 Punkten. Bei totalen Kosten (TK) von Fr. 2'332'413.- (einschliesslich veranlasster Kosten) und totalen direkten Kosten (TDK) von Fr. 2'018'212, einem Regressionsindex (RI) von 144 und Toleranzbereich (TB) von 120 errechne sich eine Rückforderungssumme von Fr. 336'368.- (Formel: $TK:RI \times (RI-TB) \times TDK:TK$). Betreffend den Ablauf des Wirtschaftlichkeitsverfahrens wiesen die Kläger auf das Schreiben vom 4. Juni 2019 hin. Der Beklagte habe nicht innert gesetzter Frist Stellung genommen.

5.7.3. Während des vorinstanzlichen Verfahrens bemühten sich die Parteien um einen Vergleich. Ein solcher kam aber nicht zustande. In einer Stellungnahme vom 18. März 2021 an das Schiedsgericht führten die Kläger aus, die Einzelfallanalyse sei in diesem Fall gemacht worden. Der Beklagte habe alle zusätzlichen, d.h. im Regressionsindex nicht eingeschlossenen, "auf Diagnosen beruhenden" Kostenfaktoren vorbringen können. Diese seien auch berücksichtigt worden. Man sei zum Ergebnis gekommen, dass die hohen Kosten aufgrund der vorliegenden Informationen, auch aus früheren Statistikjahren, nicht mit Praxisbesonderheiten erklärt werden könnten. Um auch bis dahin nicht bekannte Faktoren berücksichtigen zu können, sei der Beklagte mit dem

Schreiben vom 4. Juni 2019 mit allen Details dokumentiert und um Stellungnahme ersucht worden. Durch Stillschweigen habe er den klägerischen Ausführungen implizit zugestimmt.

5.7.4. Das Schiedsgericht hält fest, die Anwendung der statistischen Methode eines Durchschnittskostenvergleichs in Form der zweistufigen Regressionsanalyse sei nicht zu beanstanden. Die Klage werde mit den eingereichten Regressionsberichten rechtsgenügend dokumentiert. Damit sei die vom Beklagten monierte fehlende Nachvollziehbarkeit der Regressionsindizes formal behoben und eine Beurteilung der Wirtschaftlichkeitsprüfung zuverlässig gewährleistet. Die zweistufige Regressionsanalyse nach dem Screening-Modell erlaube (im Vergleich mit der ANOVA-Methode) eine präzisere Erfassung der statistischen Auffälligkeiten resp. deren spezifischere Identifikation. Die neu entwickelte Methode sehe nicht vor, bei statistisch auffälligen Ärzten in einem zweiten Schritt systematisch die analytische Methode anzuwenden, d.h. eine repräsentative Anzahl von Rechnungen nach dem diagnostischen und therapeutischen Aufwand zu überprüfen. Die Zweistufigkeit beziehe sich allein auf das Vorgehen zur Festlegung des Regressionsindex. Bei statistisch auffälligen Ärzten systematisch die analytische Methode anzuwenden scheine auch nicht zielführend, weil dies die neu erarbeitete Regressionsanalyse infrage stellen würde: Es bleibe dabei, dass die Überarztung anhand einer statistischen Auswertung und eines Durchschnittskostenvergleichs festgestellt werde. Liege der Index über 120 Punkten, sei eine Überarztung im Rahmen des ersten Prüfschritts grundsätzlich plausibilisiert. Im zweiten Prüfschritt sei danach zu klären, ob weitere Praxisbesonderheiten vorlägen, die allenfalls eine Erhöhung des Referenzindex rechtfertigten. Entgegen der Auffassung des Beklagten, der mit der Gesamtheit der Fachärzte Allgemeine Innere Medizin zu vergleichen sei, bestehe kein Anlass, die Referenzgruppe aufgrund absolvierter Weiterbildungen (u.a. Akupunktur, chinesische Arzneitherapie, Sonographie) enger zu fassen. Dem Umstand, dass seine Patienten wegen dieser Weiterbildungen eine vom Durchschnitt abweichende Morbidität aufwiesen, werde mit dem erweiterten Variablenansatz der Screening-Methode hinreichend Rechnung getragen. Der Index sei bereits entsprechend bereinigt; die Morbidität könne nicht ein zweites Mal berücksichtigt werden. Des Weiteren werde die Morbidität des Patientenkollektivs u.a. durch Pharmaceutical Cost Groups (PCG) abgebildet. Damit würden Versicherte mit kostenintensiven, chronischen Krankheiten und entsprechend hohem Leistungsbedarf aufgrund ihres Arzneimittelbedarfs (Wirkstoff und Menge) identifiziert. Weil die Polymorbidität der Patienten bereits in den Ergebnissen der Regressionsanalyse enthalten sei, stelle sie von vornherein keine Praxisbesonderheit dar (E. 5 des angefochtenen Urteils).

Im Anschluss an diese Erwägungen korrigierte die Vorinstanz den eingeklagten Rückforderungsbetrag, indem sie in der von den Klägern verwendeten Berechnungsformel für die Rückforderung den Regressionsindex Gesamtkosten mit dem Regressionsindex "Arztkosten, Taxpunkte (direkt)" (179 Punkte) ersetzte (vgl. oben E. 4.1 a.E.) und so (ausgehend von direkten Arztkosten 2017 von Fr. 810'046.-) ein Rückforderungsbetrag von Fr. 266'998.40 ermittelte (E. 6).

5.8.

5.8.1. Das zweiteilige Wirtschaftlichkeitsverfahren setzt voraus, dass sich eine (allfällige) Klage auf das Ergebnis einer Einzelfallprüfung stützt. Nach der klaren Konzeption des Screening-Vertrags ist der Leistungserbringer aufzufordern, die Besonderheiten seiner Praxis zu benennen und zu substantiieren. Im vorliegenden Fall haben die Krankenversicherer vorab "Verhandlungen auf dem gerichtlichen Wege" in Aussicht gestellt. Es genügt indessen nicht, eine Klage anzukündigen und dem Adressaten Gelegenheit zur Stellungnahme zu geben, ohne ihm dabei vor Augen zu führen, dass die Klage auf davon abhängigen Abklärungen beruhen wird. Wenn aus Gründen der Fristwahrung mit der Einreichung einer Klage nicht mehr zugewartet werden kann, ist dem Schiedsgericht zu beantragen, das Klageverfahren bis zum Abschluss der Einzelfallprüfung zu sistieren.

5.8.2. Entgegen den Erwägungen der Vorinstanz dokumentieren Regressionsberichte allein keine vollständige Wirtschaftlichkeitsprüfung. Das Schiedsgericht schätzt die Tragweite der tarifvertraglich vorgeschriebenen Einzelfallprüfung unzutreffend ein, wenn es ausführt, die neue Methode sehe nicht mehr vor, dass im Fall auffälliger Kosten in einem zweiten Schritt systematisch die analytische Methode anzuwenden sei. Richtig ist zwar, dass die Einzelfallprüfung nicht auf diesem Weg erfolgen muss (oben E. 5.2.4). Die zweistufige Regressionsanalyse ist indessen nicht mit der Zweiteilung der Wirtschaftlichkeitsprüfung in eine (vorläufige) Regressionsanalyse und eine anschliessende Einzelfallanalyse zu verwechseln. Die Einzelfallprüfung ist der tarifver-

traglichen Konzeption nach eine umfassende; individuelle Praxismerkmale sind auf ihre effektive Tragweite hin abzuklären, auch wenn sie bereits im erweiterten Variablensatz der - bloss zur Aufdeckung von Verdachtsfällen bestimmten - Screening-Methode erfasst worden sind (E. 5.5.1 und 5.5.3).

5.8.3. Nach dem Gesagten sind die vorinstanzlichen Entscheidungsgrundlagen unvollständig. Der Beschwerdeführer rügt daher zu Recht, dass die Beschwerdegegner resp. die Vorinstanz die Einzelfallanalyse als zweiten Teil der Wirtschaftlichkeitsprüfung nicht vollständig umgesetzt haben. Dies verletzt Art. 56 KVG. Unbegründet ist hingegen die Auffassung des Beschwerdeführers, dass diese Einzelfallprüfung in jedem Fall mit einer Anwendung der analytischen Methode im herkömmlichen Sinn einhergehe.

6.

Der Beschwerdeführer verfügt über eine kantonale Bewilligung zur Führung einer Praxisapotheke (sog. Selbstdispensation). Er rügt, die Vorinstanz habe diese Besonderheit zu Unrecht als bedeutungslos angesehen.

6.1. Die Kantone bestimmen, unter welchen Voraussetzungen Ärzten eine Bewilligung zur Führung einer Apotheke erteilt wird (vgl. Art. 37 Abs. 3 KVG [in der bis Ende 2021 gültigen Fassung; BBl 2018 3155] sowie Art. 4 Abs. 1 lit. k und Art. 24 Abs. 1 lit. b des Bundesgesetzes vom 15. Dezember 2000 über Arzneimittel und Medizinprodukte [HMG, SR 812.21]; Urteil 2C 53/2009 vom 23. September 2011 E. 3). Nach dem seit Anfang 2022 geltenden Art. 40 Abs. 2 KVV berücksichtigen die Kantone bei der Festlegung der Voraussetzungen, unter denen Ärzte mit einer Bewilligung zur Führung einer Apotheke den zugelassenen Apothekern gleichgestellt sind, insbesondere die Zugangsmöglichkeiten der Patienten zu einer Apotheke.

Während die Deutschschweizer Kantone, mit Ausnahme von Basel-Stadt und Aargau, die Selbstdispensation kennen, ist diese in der Romandie (einschliesslich der zweisprachigen Kantone Freiburg und Wallis) und im Tessin nur in Ausnahmefällen zugelassen. In Bern und Graubünden bestehen Mischsysteme. Der über eine Bewilligung zur Führung einer Praxisapotheke verfügende Beschwerdeführer praktiziert im Kanton Bern. Hier ist die ärztliche Abgabe von Medikamenten grundsätzlich auf Notfälle, Hausbesuche und die Erstversorgung beschränkt. Indessen ist die wiederholte Abgabe (Dauermedikation) durch Ärzte für Ortschaften vorgesehen, in denen die Notfallversorgung mit Arzneimitteln nicht durch mindestens zwei öffentliche Apotheken gewährleistet ist (Art. 31 Abs. 2 und Art. 32 Abs. 1 lit. a des bernischen Gesundheitsgesetzes vom 2. Dezember 1984; BSG 811.01).

6.2. Selbstdispensierende Mediziner werden die Medikamente wohl zu einem überwiegenden Teil, soweit im eigenen Lagerbestand vorhanden, selbst abgeben. Falls dieser Umstand für die Wirtschaftlichkeitsprüfung bedeutsam ist, kann er grundsätzlich auf verschiedene Arten berücksichtigt werden.

6.2.1. Im Fall des Beschwerdeführers erwog das Bundesgericht mit Urteil 9C 67/2018 vom 20. Dezember 2018 betreffend die Statistikjahre 2013 bis 2015, die ANOVA-Methode enthalte Indizes betreffend die direkten Behandlungskosten (ohne Medikamente), die Medikamentenkosten (direkte und veranlasste) sowie die totalen Kosten pro erkrankter Person (E. 10). Es gebe keinen ANOVA-Index, der die direkten Kosten einschliesslich der vom Arzt selber abgegebenen Medikamente umfasse (E. 12.2.1). Nach der Rechtsprechung zur statistischen Methode des Durchschnittskostenvergleichs (und damit auch zur ANOVA-Methode) stelle die kantonale Bewilligung zur Selbstdispensation ein Praxismerkmal dar, das bei der Bildung der Vergleichsgruppe zu berücksichtigen sei, dem also nicht erst bei der Bemessung der (allenfalls) rückzuerstattenden Vergütungen Rechnung getragen werde. Ärzte ohne Selbstdispensation gehörten nicht zur massgeblichen Vergleichsgruppe. Die Spezialität "Internist mit Bewilligung zur Führung einer Apotheke" lasse die bestehenden Facharztgruppen "Allgemeine Innere Medizin" (Kanton Bern oder ganze Schweiz) nicht als genügend homogen erscheinen, da sie Internisten mit und solche ohne Privatapotheke umfassten (E. 12.2.3 mit Hinweisen).

In diesem Sinn sei das kantonale Schiedsgericht bundesrechtswidrig vorgegangen. Es werde auf entsprechend angepasster Grundlage neu über die streitige Rückerstattungspflicht zu entscheiden haben. Sollte sich für die enger gefasste Facharztgruppe "Allgemeine Innere Medizin mit Bewilligung zur Führung einer Praxisapotheke" (ganze Schweiz) kein ANOVA-Index berechnen lassen oder wäre dies unverhältnismässig (was die Versicherer zumindest glaubhaft zu machen hätten), müsse ein (reiner) Durchschnittskostenvergleich durchgeführt werden, allenfalls kombiniert mit der analytischen Methode (Überprüfung einer repräsentativen Anzahl von

Rechnungen in Bezug auf den diagnostischen und therapeutischen Aufwand), sofern weitere geltend gemachte Praxisbesonderheiten dazu Anlass gäben (E. 12.3).

6.2.2. Nachdem das Bundesgericht im Verfahren 9C 67/2018 betreffend die Statistikjahre 2013 bis 2015 die Sache an das kantonale Schiedsgericht zurückgewiesen hatte, führte dieses Beweismassnahmen durch; die Parteien nahmen Stellung. Am 18. Juni 2020 einigten sich die Parteien auf eine Regelung betreffend die Jahre 2013 bis 2015. Das im Rückweisungsentscheid 9C 67/2018 E. 12.3 skizzierte Verfahren wurde daher nicht weiterverfolgt.

6.2.3. In einem anderen Fall hielt das Bundesgericht fest, es sei nicht bekannt, ob sich "für eine wie auch immer modifizierte, d.h. enger gefasste Facharztgruppe 'Allgemeine Innere Medizin' ganze Schweiz ein ANOVA-Index berechnen" lasse (Urteil 9C 558/2018 vom 12. April 2019 E. 8.3). Aus einem Verfahren betreffend die Statistikjahre 2015 und 2016 ergibt sich, dass der im Urteil 9C 67/2018 E. 12.3 geforderte engere ANOVA-Index (Vergleichsgruppe "Ärzte Allgemeine Innere Medizin mit Bewilligung zur Führung einer Praxisapothek") jedenfalls bis Oktober 2020 nicht vorlag; das kantonale Schiedsgericht hatte dort ausgeführt, die Klägerschaft habe sich bis anhin nicht in der Lage gesehen, eine entsprechend angepasste ANOVA-Auswertung beizubringen (Urteil 9C 721/2020 vom 19. Oktober 2021 E. 4.1).

6.3. Die Vorinstanz geht davon aus, nach der (ab dem Statistikjahr 2017 anwendbaren) neuen Screening-Methode sei eine Differenzierung der Vergleichsgruppen nach Ärzten mit und solchen ohne Selbstdispensation überholt. Die Frage der Medikamentenkosten und damit der Selbstdispensation sei in die neue Methode integriert worden. Diese enthalte unter anderem den Morbiditätsfaktor "Pharmazeutische Kostengruppe (PCG)" und ordne die darin erfassten - direkt abgegebenen oder veranlassten - Medikamente resp. die entsprechenden Kosten der ZSR-Nummer zu. Im Hinblick auf die Kontrolle, ob ein Arzt seinen Patienten nur effektiv indizierte, der Morbidität entsprechende Medikamente abgebe resp. verschreibe, würden alle durch den Arzt selbst oder über Apotheken abgegebenen Medikamente erfasst (angefochtenes Urteil E. 5.3.4). Die Beschwerdegegner schliessen sich dieser Auffassung an.

Der Beschwerdeführer rügt, die Vorinstanz habe diesbezügliche Auflagen im Urteil 9C 67/2018 (betreffend die Rückerstattungsjahre 2013 bis 2015) nicht beachtet, obwohl sie auch hinsichtlich des hier strittigen Jahres 2017 einschlägig seien.

Ob das im Urteil 9C 67/2018 Gesagte sinngemäss auf den vorliegenden Fall zu übertragen oder ob die betreffenden Vorgaben angesichts der veränderten methodischen Grundlagen überholt sind, ist im Folgenden unter zwei Aspekten zu beurteilen: Zunächst ist die Frage überhaupt nur relevant, wenn die Selbstdispensation erhebliche Kostenunterschiede bewirkt (sogleich E. 6.4). Sofern dies zu bejahen ist, bleibt zu prüfen, ob die Erweiterung der tarifvertraglich vorgesehenen Morbiditätsfaktoren durch "Pharmazeutische Kostengruppen" (PCG) die Selbstdispensation erfasst und kostenmässig neutralisiert (E. 6.5).

6.4. Die Frage, ob von einem signifikanten Kostenunterschied zwischen Arztpraxen mit und ohne Praxisapothekenzuzugehen ist, stellt sich bei den Medikamentenkosten selbst und bei den mit der Medikation zusammenhängenden Behandlungskosten.

6.4.1. Die Wirtschaftlichkeit (Art. 56 und 59 KVG) beurteilt sich nach dem Gesamtkostenindex der direkten und veranlassten Kosten (BGE 137 V 43 E. 2.5.6; oben E. 4.1). Insoweit spielt an sich keine Rolle, ob die Arzneimittel - deren Preise behördlich administriert sind (vgl. Art. 65b ff. KVV) - direkt durch eine praxiseigene Apotheke oder auf ärztliches Rezept durch eine externe Apotheke abgegeben werden. Die geforderte Berücksichtigung der Selbstdispensation ist hingegen dann jedenfalls relevant, wenn die beiden Formen der Medikamentenabgabe unterschiedlich kostenwirksam sein sollten.

6.4.2. In einem früheren Entscheid nahm das Bundesgericht an, die Selbstdispensation sei kostenneutral. Es sei unerheblich, ob die in den Gesamtkosten der ärztlichen Tätigkeit enthaltenen Medikamente direkt vom Arzt oder über einen Apotheker abgegeben würden. Die Selbstdispensation könne gar kostensenkend wirken (Urteil 9C 570/2015 vom 6. Juni 2016 E. 8.1). Verhielte es sich tatsächlich so, wäre die im Urteil 9C 67/2018 formulierte Vorgabe, die Wirtschaftlichkeitsprüfung nach Art. 56 Abs. 6 und Art. 59 Abs. 1 lit. b KVG müsse dem

Praxismerkmal "Bewilligung zur Führung einer Praxisapotheke" Rechnung tragen, gegenstandslos. Die Frage ist daher zu vertiefen.

6.4.3. Eine gesundheitsökonomische Studie der Universität Bern aus dem Jahr 2016, in welcher der Kosteneffekt der Selbstdispensation nach der Regressionsmethode von anderen Einflussfaktoren (z.B. Zusammensetzung der Patientenschaft) getrennt wurde, zeigt anhand der Daten von 3400 Arztpraxen in der Deutschschweiz, dass eine selbstdispensierende Arztpraxis pro Patient und Jahr bis über ein Drittel höhere Medikamentenkosten zulasten der obligatorischen Krankenpflegeversicherung verursacht als eine vergleichbare Praxis ohne Patientenapotheke. Bei Medikamentenkosten von rund 280 Franken pro Patient und Jahr betrage dieser Mehraufwand etwa 90 bis 100 Franken, also 32 bis 35 Prozent (BORIS KAISER/CHRISTIAN SCHMID, Does Physician Dispensing Increase Drug Expenditures?: Empirical Evidence from Switzerland, in: Health Economics 2016, Vol. 25 S. 71 ff.; zusammenfassend: KAISER/SCHMID, Kostenanreize im Gesundheitswesen: Das Beispiel der Medikamentenabgabe, in: Die Volkswirtschaft 8-9/2017 S. 66 f.). Weitere Studien kamen zu qualitativ übereinstimmenden Ergebnissen (KAISER/SCHMID, Kostenanreize, S. 67 mit Hinweisen). Einzig im Rahmen einer Studie fand sich kein statistisch signifikanter Einfluss der Selbstdispensation auf die Gesamtausgaben für Leistungen der obligatorischen Krankenpflegeversicherung (TROTSMANN/FRÜH/REICH/TELSER, Auswirkungen der Medikamentenabgabe durch die Ärzteschaft [Selbstdispensation] auf den Arzneimittelkonsum und die Kosten zu Lasten der OKP, Schlussbericht, Studie im Auftrag des Bundesamts für Gesundheit [BAG], 2014, S. 55 f.).

Dass die Medikamentenkosten selbstdispensierender Arztpraxen im Vergleich zu solchen ohne diese Besonderheit in der Grössenordnung von einem Drittel höher ausfallen können, wird mit wirtschaftlichen Anreizen im Zusammenhang mit der Marge beim Medikamentenverkauf erklärt. Dieser Effekt soll vor allem auf einer Mengenausweitung beruhen; hingegen scheine der durchschnittliche Preis selbstdispensierter Medikamente gar etwas tiefer als derjenige verschriebener Arzneien (KAISER/SCHMID, Kostenanreize, S. 67). Eine weitere Untersuchung von Forschern der Universität Bern aus dem Jahr 2017 kommt ebenfalls zum Schluss, dass selbstdispensierende Ärzte mehr (aber nicht teurere) Arzneimittel abgeben (bei Grundversorgern Mengenausweitung von 56 Prozent; BURKHARD/SCHMID/WÜTHRICH, Financial incentives and physician prescription behavior: Evidence from dispensing regulations, <<https://boris.unibe.ch/145820/>>; publ. auch in: Health Economics 2019, Vol. 28 S. 1114 ff.).

6.4.4. Bei Selbstdispensation fallen neben höheren Medikamentenkosten möglicherweise - in einem geringeren Umfang - auch höhere Behandlungskosten an, dies wohl wegen des zeitlichen Aufwands (KAISER/SCHMID, Kostenanreize, S. 66 f.). Denkbar ist auch, dass die Selbstdispensation durch Grundversorger zu einer höheren Anzahl Konsultationen führt (TROTSMANN/FRÜH/REICH/TELSER, a.a.O., S. 11).

Es erscheint plausibel, dass die Selbstdispensation nicht nur mit höheren Medikamentenkosten, sondern darüber hinaus auch mit höheren Begleitkosten der Medikamentenabgabe verbunden sein kann. Leistungen wie die medikamentenbezogene Beratung und Aufklärung des Patienten durch den selbstdispensierenden Arzt werden unter dem Titel der Behandlung nach dem allgemeinen TARMED-Zeittarif (Art. 43 Abs. 2 lit. a KVG) verrechnet. Dagegen werden die pharmazeutischen Grundleistungen der Apotheken bei einem Bezug von Arzneien auf Rezept bloss pauschal entschädigt ("Leistungsorientierte Abgeltung" [LOA]; vgl. Art. 25 Abs. 2 lit. h und Art. 43 Abs. 2 lit. c KVG, Art. 4a Abs. 1 lit. a und Abs. 2 KLV). Bei der Verschreibung eines Arzneimittels zuhanden einer externen Apotheke entstehen zwar etliche Aufwände, die wegen der Trennung von Behandlung und Medikation nötig werden (z.B. Kontrolle des ärztlichen Rezepts und Führung eines Patientendossiers durch den Apotheker). Diese fallen kostenmässig aber vergleichsweise wenig ins Gewicht, da sie den Apotheken nur im Umfang eines "Medikamenten-Checks" pro verordnetem Medikament resp. eines "Bezugs-Checks" pro Patient, Tag und Leistungserbringer pauschal vergütet werden.

6.4.5. Zusammengefasst zeigen empirische Erhebungen an, dass bei Selbstdispensation signifikant höhere Medikamentenkosten entstehen dürften. Zudem gibt es Anhaltspunkte, dass auch die Behandlungskosten etwas höher ausfallen. Unter diesen Voraussetzungen ist die Prämisse des Urteils 9C 67/2018 - die Kostenstruktur von selbstdispensierenden Internisten unterscheide sich erheblich von derjenigen entsprechender Fachärzte ohne diese Praxisbesonderheit - für die Feststellung von Verstössen gegen das Wirtschaftlichkeitsgebot nach Art. 56 Abs. 1 und Art. 59 KVG weiterhin relevant (vgl. auch GEBHARD EUGSTER, KVG: Statistische Wirt-

schaftlichkeitsprüfung im Wandel, in: Jusletter 25. Juni 2012, Rz. 79, wonach in Kantonen mit einem Mischsystem Ärzte mit resp. ohne Selbstdispensation bezüglich der direkten Kosten "grundlegend unterschiedliche Leistungsspektren" aufwiesen, weshalb notwendigerweise Vergleichsgruppen mit und ohne Selbstdispensation zu bilden seien; vgl. auch derselbe, Baustelle, Rz. 28).

6.5. Zu prüfen bleibt, ob die Problematik der ungleich zusammengesetzten Vergleichsgruppen infolge der Weiterentwicklung der ANOVA-Methode zum Screening-Modell gegenstandslos geworden ist.

6.5.1. Die Vorinstanz geht davon aus, die Abbildung der Morbidität durch pharmazeutische Kostengruppen (Pharmaceutical Cost Groups, PCG) mache die im Urteil 9C 67/2018 getroffenen Festlegungen obsolet.

6.5.2. Die tarifvertraglich definierten Morbiditätsfaktoren sollen den Einfluss einer spezifischen gesundheitlichen Charakteristik des Patientenkollektivs auf die Praxiskosten erfassen. "Morbidität" ist ein Überbegriff für die Prävalenz (bestehende Fälle) und die Inzidenz (Auftreten neuer Fälle) von Krankheiten; im vorliegenden Zusammenhang ist sie als Gradmesser der Behandlungskostenintensität pro Behandlungsfall oder Patientenkollektiv zu verstehen (EUGSTER, Baustelle, Rz. 19). Die Morbiditätsfaktoren haben bei der Mehrzahl der Facharztgruppen einen statistisch signifikanten Einfluss auf die Kosten (TROTSMANN und andere, Schlussbericht der Polynomics AG, S. 69 f.).

Bei den Pharmaceutical Cost Groups handelt es sich um einen von vier Morbiditätsfaktoren gemäss Screening-Methode (neben Alter und Geschlecht, Franchisen sowie Spital- resp. Pflegeheimaufenthalt im Vorjahr). Das Instrument wurde ursprünglich im Hinblick auf den Risikoausgleich unter den sozialen Krankenversicherern geschaffen (Art. 16 ff. KVG). Dort dienen die PCG als Indikatoren zur Feststellung von morbiditätsbezogenen Unterschieden in den Versichertenkollektiven (TROTSMANN und andere, Schlussbericht der Polynomics AG, S. 41 Ziff. 6.2.4). Eine "pharmazeutische Kostengruppe" umfasst Arzneimittel mit bestimmten Wirkstoffen, die für die Behandlung eines Krankheitsbildes eingesetzt werden, das (für sich allein oder zusammen mit einer anderen Erkrankung) besonders kostenintensiv ist (vgl. Art. 4 Abs. 1bis der Verordnung vom 19. Oktober 2016 über den Risikoausgleich in der Krankenversicherung [VORA; SR 832.112.1]). Beispiele dafür sind Arzneimittel gegen Asthma, die chronisch obstruktive Lungenerkrankung COPD, Diabetes (Typ I und Typ II), Epilepsie, Herzerkrankungen, Krebs und Nierenerkrankungen. Anhand der Menge abgerechneter Arzneimittel mit entsprechend spezifizierten Wirkstoffen kann die Häufigkeit bestimmter chronischer, kostenintensiver Krankheiten in einem Patientenkollektiv indirekt abgebildet werden (BILL/MEYER/TELSER, Aktualisierung der PCG-Liste für den Schweizer Risikoausgleich: Studie im Auftrag des BAG, 2019, S. 5 und 8).

6.5.3. Die Morbiditätsfaktoren verbinden die auffällige Kostenstruktur einer Arztpraxis mit einer allenfalls überdurchschnittlichen Anzahl von Chronischkranken im Patientenbestand. Die Zuordnung der Patienten in pharmazeutische Kostengruppen hilft bei der Beurteilung, ob die Kosten für die (eigene und/oder veranlasste) Abgabe von Medikamenten als Teil der gesamten Praxiskosten angemessen sind. Eine andere Frage ist, ob die PCG, die zur Identifizierung und Quantifizierung von Morbiditätsunterschieden bestimmt sind, den im Urteil 9C 67/2018 geforderten Vergleich ausschliesslich mit selbstdispensierenden Arztpraxen ersetzen können.

Die in E. 6.4.3 zitierten Studien weisen mehrheitlich darauf hin, dass die ärztliche Medikamentenabgabe insgesamt höhere Kosten verursacht als eine ärztliche Verschreibung mit anschliessendem Bezug des Medikaments in einer externen Apotheke. Es gibt einen unabhängig von der Morbidität des Patientenkollektivs bestehenden kostenstrukturellen Unterschied zwischen Arztpraxen mit und ohne Selbstdispensation. Die Integration der pharmazeutischen Kostengruppen (und weiterer Morbiditätsfaktoren) in die statistische Prüfmethode ebnet den fraglichen Unterschied somit höchstens teilweise ein. Die PCG standardisieren die entsprechenden Mehrkosten soweit, wie sie auf eine überdurchschnittliche Morbidität des Patientenkollektivs zurückzuführen sind. Das grundsätzliche Mehr an Arztkosten, bedingt durch die Praxisbesonderheit "Selbstdispensation" als solche, gleichen sie aber nicht aus. Sie erfassen auch keine anderen Kostenunterschiede, die sich aus der Führung einer Praxisapotheke ergeben. Entgegen der Auffassung von Vorinstanz und Beschwerdegegnern begründet somit die Screening-Methode, namentlich die Aufnahme der PCG als zusätzlichen Morbiditätsfaktor, keine Veränderung gegenüber der Vorgängermethode, die die im Urteil 9C 67/2018 gemachten Überlegungen zur Definition der Vergleichsgruppe gegenstandslos machen würde.

6.6. Insgesamt verletzt die Ausserachtlassung der Praxisbesonderheit "Selbstdispensation" die bundesrechtlichen Anforderungen an die Feststellung einer unwirtschaftlichen Behandlungsweise. Da die Gründe, die für die Rückweisung durch das Urteil 9C 67/2018 massgebend gewesen sind, grundsätzlich auch unter dem neuen Prüfsystem beachtlich bleiben, sind die damaligen Vorgaben den aktuellen Verhältnissen anzupassen.

6.6.1. Die Wirtschaftlichkeit der Behandlungsweise im Jahr 2017 muss auf der Grundlage einer ergänzten Klage neu geprüft werden (oben E. 5.6). Dabei ist sinngemäss auf die Vorgaben im Urteil 9C 67/2018 E. 12.3 zurückzugreifen (vgl. E. 6.2.1). Bei der Selbstdispensation handelt es sich um eine Frage des Praxistyps. An sich wäre danach in erster Linie die Berechnung eines Indexes gefordert, der nur selbstdispensierende Internisten enthält (E. 5.5.2). Dem Urteil 9C 721/2020 vom 19. Oktober 2021 E. 4.1 allerdings ist zu entnehmen, dass bis dahin kein entsprechend angepasster Regressionsindex vorlag (E. 6.2.3). Wegen den kantonal unterschiedlichen Systemen der Medikamentenabgabe (E. 6.1) wäre wohl schon die Bildung einer gesamtschweizerischen Vergleichsgruppe (vgl. E. 4.4.1) schwierig. Im vorliegenden Verfahren haben die Beschwerdegegner bekräftigt, das mit der FMH vereinbarte zweistufige Regressionsmodell zur Berechnung der Indizes ermögliche keine Anpassung des Vergleichskollektivs auf Stufe Regressionsindex. Auf die Frage des vorinstanzlichen Instruktionsrichters, ob die nicht während des ganzen Statistikjahrs über eine Berufsausübungsbewilligung verfügenden Ärzte aus der Referenzgruppe entfernt werden könnten, führte die

Santésuisse namens der Krankenversicherer aus, das System weise eine Interdependenz zwischen allen Facharztgruppen auf. Würden Vergleichskollektive willkürlich verändert, beeinflusste dies auch alle anderen Indizes (Stellungnahme vom 18. März 2021). Entsprechend können wohl auch nicht selbstdispensierende Ärzte nicht ausgeklammert werden. Demnach wäre nach dem Rückweisungsentscheid 9C 67/2018 an sich alternativ auf einen (reinen) Durchschnittskostenvergleich auszuweichen. Ein solcher berücksichtigt indessen weder die im Screening-Modell vorgesehenen standardisierenden Faktoren noch das im Bereich der Einzelfallanalyse u.a. beachtliche Praxismerkmal "Selbstdispensation". Nachdem die Morbidität nicht standardisiert worden ist, dürfte sich ein "reiner" Durchschnittskostenvergleich ebenfalls nur auf nicht selbstdispensierende Praxen beziehen. Ansonsten könnte er vor allem durch den Umstand verfälscht werden, dass das bernische Gesundheitsgesetz die dauerhafte Selbstdispensation nur für Ortschaften vorsieht, in denen die Notfallversorgung mit Arzneimitteln nicht durch mindestens zwei öffentliche Apotheken gewährleistet ist (oben E. 6.1). Die über eigene Patientenapotheken verfügenden Praxen befinden sich mithin vorwiegend im ländlichen Raum. Das typische Patientenkollektiv einer Praxis auf dem Land ist seiner Morbidität nach erfahrungsgemäss anders zusammengesetzt als dasjenige einer städtischen Praxis.

6.6.2. Im Rahmen der Screening-Methode könnte dem Merkmal der Selbstdispensation beispielsweise Rechnung getragen werden, indem der Referenzindex "Gesamtkosten", der u.a. einen Teilindex "direkte und veranlasste Medikamentenkosten" enthält, anhand empirischer Erkenntnisse soweit erhöht wird, wie wenn Letzterer ausschliesslich auf den von selbstdispensierenden Ärzten generierten Medikamentenkosten beruhen würde. Die Anpassung des Referenzindex könnte datenbasiert erfolgen, etwa auf der Grundlage eines Vergleichs der Kostenstruktur von Praxen mit und ohne Selbstdispensation im Kanton Bern, oder, falls dies mit unverhältnismässigem Aufwand verbunden sein sollte, anhand von Studien, wonach bei Selbstdispensation Mehrkosten von rund einem Drittel anfallen (E. 6.4.3). Bei der Umlegung der gesamtschweizerischen Daten auf den Standortkanton des Beschwerdeführers (zweite Stufe der Regressionsanalyse) wird das im Kanton Bern bestehende gemischte System der Medikamentenabgabe (E. 6.1) wohl nicht berücksichtigt. Die Mehrkosten wären diesfalls soweit zu veranschlagen, wie es dem gesamtschweizerischen Anteil selbstdispensierender Praxen entspricht. Beträgt dieser Anteil beispielsweise 50 Prozent, so wäre der komplementäre Teil-Referenzindex "direkte und veranlasste Medikamentenkosten" um die Hälfte der Mehrkosten zu erhöhen, also um 15 Prozent, sofern auf die erwähnten Studien abgestellt wird. Da allerdings kein Index der direkten Behandlungs- und übrigen Kosten (exklusive Medikamente) existiert, wäre zur Ermittlung des dem Durchschnittskostenwert des Beschwerdeführers gegenüberzustellenden Gesamt-Referenzwerts der Anteil eigener Medikamentenabgabe zu schätzen und abzuziehen.

7.1. Zusammenfassend bleibt festzuhalten, dass sich die Wirtschaftlichkeitsprüfung nach der tarifpartnerschaftlich vereinbarten Screening-Methode stets aus einer Regressionsanalyse (Screening) und, bei auffälligem Resultat, einer anschliessenden Einzelfallprüfung zusammensetzt. Die Einzelfallprüfung deckt sich nicht mit der herkömmlichen analytischen Methode (E. 5.2). Ein auffälliges Ergebnis der Regressionsanalyse bedeutet keine Feststellung von Unwirtschaftlichkeit; insoweit handelt es sich nicht um eine "Beweismethode" (E. 5.3). Die im Rahmen des Screenings zu veranschlagende Toleranzmarge trägt vor allem der ärztlichen Behandlungsfreiheit Rechnung; die neue Methode wirkt sich nicht darauf aus (E. 5.4). Praxistypologische Merkmale (z.B. Selbstdispensation) sind auf Stufe der Einzelfallprüfung zu berücksichtigen, wenn sie nicht als Screening-Faktor implementiert werden können (E. 5.5.2). Praxisbesonderheiten, die sich auf Eigenschaften des Patientenkollektivs beziehen, kann gestützt auf Patientendossiers oder Statistiken Rechnung getragen werden, soweit sie nicht schon im Rahmen der Regressionsanalyse standardisiert worden sind (E. 5.5.3). Da im vorliegenden Fall eine Einzelfallprüfung unterblieben ist, sind die vorinstanzlichen Entscheidungsgrundlagen unvollständig (E. 5.6-5.8). Die Führung einer Praxisapotheke (Selbstdispensation) ist grundsätzlich kostenrelevant, so dass sie als praxistypologisches Merkmal in die Wirtschaftlichkeitsprüfung einfließen muss (E. 6.4). Der in die Screening-Methode integrierte Morbiditätsfaktor Pharmaceutical Cost Group (PCG) macht die Berücksichtigung der Selbstdispensation nicht entbehrlich (E. 6.5).

7.2. Die Beschwerde ist im Eventualstandpunkt begründet und die Sache im Sinn der Erwägungen (vgl. E. 5.7-5.8; E. 6.6) an die Vorinstanz zurückzuweisen, damit sie auf der Grundlage einer ergänzten Klage neu entscheide (E. 5.6). Die vom Beschwerdeführer geltend gemachten weiteren Praxismerkmale (wie etwa Weiterbildungen im Bereich Alternativmedizin) sind gegebenenfalls in die Einzelfallprüfung einzubeziehen.

8.

Mit dem Entscheid in der Sache ist das Gesuch um aufschiebende Wirkung der Beschwerde gegenstandslos.

9.

Die Rückweisung an die Vorinstanz zur neuen Beurteilung (mit offenem Ausgang) gilt für die Frage der Auf-erlegung der Gerichtskosten wie auch der Parteientschädigung als vollständiges Obsiegen, unabhängig davon, ob die Rückweisung beantragt oder ob das Begehren im Haupt- oder Eventualantrag gestellt wird (BGE 137 V 210 E. 7.1; Urteil 8C 136/2021 vom 7. April 2022 E. 7). Ausgangsgemäss haben die Beschwerdegegner die Gerichtskosten zu tragen (Art. 66 Abs. 1 BGG) und dem Beschwerdeführer eine Parteientschädigung zu bezahlen (Art. 68 Abs. 2 BGG).

Demnach erkennt das Bundesgericht:

1.

Die Beschwerde wird teilweise gutgeheissen. Das Urteil des Schiedsgerichts in Sozialversicherungsstreitigkeiten des Kantons Bern vom 18. Januar 2022 wird aufgehoben und die Sache zur neuen Entscheidung an dieses zurückgewiesen.

2.

Die Gerichtskosten von Fr. 8'000.- werden den Beschwerdegegnern auferlegt.

3.

Die Beschwerdegegner haben den Beschwerdeführer für das bundesgerichtliche Verfahren mit Fr. 6'000.- zu entschädigen.

4.

Dieses Urteil wird den Parteien, dem Schiedsgericht in Sozialversicherungsstreitigkeiten des Kantons Bern und dem Bundesamt für Gesundheit schriftlich mitgeteilt.

Luzern, 12. Dezember 2023

Im Namen der III. öffentlich-rechtlichen Abteilung
des Schweizerischen Bundesgerichts

Der Präsident: Parrino

Der Gerichtsschreiber: Traub