

Le recourant, médecin spécialiste en médecine interne générale, a été recherché par plusieurs assureurs-maladie, représentés par l'association Santésuisse, pour traitement non économique. Ceux-ci lui réclamaient la restitution de montants compris entre CHF 148'024 et 219'375 pour l'année 2019, au motif que ses coûts de traitement étaient excessifs. Le 9 mars 2022, le Tribunal arbitral a condamné le recourant à rembourser un total de CHF 156'455.75. Le recourant a saisi le TF en demandant l'annulation de cette décision.

Concernant le **contrôle de l'économicité**, le TF a rappelé que les traitements médicaux doivent respecter les prescriptions d'économicité prévues aux art. 32 et 56 LAMal, et que les traitements inutiles sont considérés comme non économiques. En cas de violation, le prestataire peut être sanctionné, notamment par la **restitution des montants perçus** (art. 59 LAMal) (c. 4.1). Une prestation est jugée non économique si un prestataire facture de manière significativement plus élevée que ses pairs de la même spécialité, sans justification valable (type de patientèle, offre de prestations, etc.). Les coûts sont comparés via un **indice de régression**, qui met en relation la valeur moyenne des coûts par cas du prestataire contrôlé avec celle d'un groupe de référence comparable (c. 4.2).

Depuis 2017, une **méthode standardisée de dépistage** (« screening ») fondée sur une analyse de régression est utilisée pour détecter les prestataires aux coûts anormalement élevés. Cette méthode est un outil préliminaire et non une preuve directe d'absence d'économicité, qui doit être confirmée par une **analyse individuelle au cas par cas** (c. 4.3). Une **marge de tolérance, comprise entre 20 et 30 points d'indice**, est déduite de l'indice de régression afin de prendre en compte le style individuel du médecin. Cette marge, appliquée lors du screening, est forfaitaire et provisoire, tandis que **la marge définitive est individualisée en fonction des particularités du prestataire lors de l'examen approfondi** (c. 4.4.2).

Dans le cas d'espèce, l'analyse a révélé un indice de coûts de 158 points pour le prestataire, dépassant la valeur seuil de tolérance fixée à 120 points (c. 6.3.1). Le TF a jugé que **l'indice seul ne suffit pas à prouver l'absence d'économicité** dès lors que des particularités du cabinet peuvent influencer les résultats et qu'aucune analyse individuelle spécifique n'a été effectuée, notamment concernant les prestations telles que la psychothérapie déléguée ou la médecine manuelle. **L'absence d'examen individualisé constitue une violation de l'art. 56 LAMal** (c. 6.4.2).

Le litige portait également sur la marge de tolérance à appliquer. Le TF a confirmé que celle-ci doit se situer entre 20 et 30 points d'indice pour respecter la liberté thérapeutique et le style individuel de pratique médicale. Il a précisé que, **sauf circonstances particulières, une déduction de 20 points reste conforme à la jurisprudence constante, sans qu'une marge fixe de 30 points soit obligatoire** (c. 8.2.2 et 8.2.3).

La demande de remboursement doit être calculée en fonction de la différence entre les coûts facturés et la valeur attendue selon le groupe de référence, en tenant compte de la marge de tolérance. Ce calcul doit être **précis et justifié**, notamment en tenant compte toutes les particularités validées lors de l'examen individuel. En l'occurrence, l'instance précédente n'avait pas suffisamment appliqué ces critères, ne considérant que les frais médicaux directs pour le remboursement (c. 10).

Le TF a donc annulé la décision du tribunal arbitral et renvoyé l'affaire pour réexamen, soulignant la nécessité d'un examen approfondi des conditions individuelles et des structures de coûts pour une évaluation équitable des demandes de remboursement.

Auteur : Alexandre Lehmann, avocat à Lausanne

Sachverhalt:

A.

Dr. med. A. ist Facharzt für Allgemeine Innere Medizin. Am 30. Juni 2021 reichten verschiedene Krankenversicherer vertreten durch den Branchenverband Santésuisse beim Schiedsgericht in Sozialversicherungsstreitigkeiten des Kantons Bern gegen ihn eine Rückforderungsklage wegen unwirtschaftlicher Behandlungsweise ein. Sie beantragten, der Beklagte sei zu verpflichten, für das Jahr 2019 Vergütungen aus obligatorischer Krankenpflegeversicherung im Betrag von Fr. 148'024.- (gemäss Regressions-Index), eventuell Fr. 219'375.- (gemäss ANOVA-Index), zurückzuerstatten. Es sei vorzumerken, dass sich die Klägerinnen das Recht vorbehielten, den Rückforderungsbetrag nach Abschluss des Beweisverfahrens anzupassen.

B.

Mit Urteil vom 9. März 2022 hiess das Schiedsgericht in Sozialversicherungsstreitigkeiten des Kantons Bern die Klage gut und verurteilte Dr. med. A., für das Jahr 2019 den Betrag von insgesamt Fr. 156'455.75 zurückzuerstatten.

C.

C.a. Dr. med. A. führt Beschwerde in öffentlich-rechtlichen Angelegenheiten mit den Rechtsbegehren, das angefochtene Urteil sei aufzuheben und die Klage abzuweisen; eventuell sei die Sache zur Neuurteilung an die Vorinstanz zurückzuweisen. Gleichzeitig beantragt er, der Beschwerde sei die aufschiebende Wirkung zuzuerkennen.

C.b. Am 9. September 2022 weist das kantonale Schiedsgericht das am 5. April 2022 gestellte Gesuch des Beschwerdeführers ab, das schiedsgerichtliche Urteil vom 9. März 2022 sei aufzuheben und die Frist zur Einreichung einer Klageantwort wiederherzustellen. Das Bundesgericht hebt darauf die zwischenzeitlich verfügte Sistierung des Verfahrens auf.

Erwägungen:

1.

Bei Namensänderungen und Fusionen von Krankenversicherern gehen die Ansprüche der Versicherungsträger auf ihre Rechtsnachfolger über. Die Parteibezeichnung wird im Rubrum von Amtes wegen angepasst (Urteil 9C_115/2023 vom 29. Mai 2024 E. 1.1).

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).

3.

Strittig ist, ob das kantonale Schiedsgericht Bundesrecht verletzt, indem es den Beschwerdeführer in Anwendung von Art. 59 Abs. 1 lit. b KVG für das Jahr 2019 wegen unwirtschaftlicher Behandlungsweise (sog. "Überarztung") zu einer Rückerstattung entgoltenen Kosten im Betrag von Fr. 156'455.75 verpflichtet.

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). Unter verschiedenen zweckmässigen Vorkehrungen gilt die bei vergleichbarem medizinischem Nutzen kostengünstigste Alternative als wirtschaftlich (BGE 139 V 135 E. 4.4.3). Unnötige therapeutische oder diagnostische Massnahmen sind unzweckmässig und damit ohne Weiteres auch unwirtschaftlich (zur Publ. vorgesehene Urteil 9C_125/2022 vom 10. September 2024 E. 2.2.1.1). Gegenüber Leistungserbringern, die gegen die im Gesetz vorgesehenen Wirtschaftlichkeits- und Qualitätsanforderungen verstossen, werden Sanktionen ergriffen. Dazu gehört die gänzliche oder teilweise Rückerstattung der Honorare, die für nicht angemessene Leistungen bezogen wurden (Art. 59 Abs. 1 lit. b und Abs. 3 lit. a KVG).

4.2. Gegenstand der Wirtschaftlichkeitsprüfung ist die Gesamtheit der unter einer Zahlstellenregister (ZSR) -Nummer in Rechnung gestellten und vergüteten Leistungen nach Art. 24 ff. KVG. Das Gebot einer wirtschaftlichen Leistungserbringung ist verletzt, wenn ein Leistungserbringer im Vergleich mit den anderen Leistungserbringern gleicher Fachrichtung (Referenzgruppe) pro Patient erheblich mehr verrechnet, ohne dass dieser Mehraufwand von der ärztlichen Behandlungsfreiheit gedeckt ist oder durch kostenwirksame Besonderheiten (hinsichtlich Leistungsangebot, Zusammensetzung Patientenkollektiv etc.) erklärt wird (vgl. BGE 150 V 129 E. 4.1; 137 V 43 E. 2.2; 136 V 415 E. 6.2). Als Grundlage für die Identifikation der einzelnen Leistungserbringer, ihrer Patientenzahl sowie der verursachten Kosten dient die ZSR-Nummer in Verbindung mit dem Datenpool der SASIS AG, einer Tochtergesellschaft des Krankenkassenverbands Santésuisse (<https://www.sasis.ch/kategorie/dp/>; vgl. DARIO PICECCHI, Das Wirtschaftlichkeitsgebot im Krankenversicherungsrecht, 2022, Rz. 527). Dieser Datenbestand führt die Angaben der Versicherer über die unter einer ZSR-Nummer vergüteten Leistungen zusammen (LARISA PETROV, Weiterentwicklung der Wirtschaftlichkeitskontrolle nach KVG, 2024, Rz. 131). Die vom überprüften Leistungserbringer während einer bestimmten Periode im Durchschnitt pro Patient durch Eigen- oder Dritteleistungen ausgelösten Kosten des Krankenversicherers (individueller Fallwert) werden den entsprechenden Durchschnittskosten der Referenzgruppe (Gruppenfallwert, entspricht einem Indexwert von 100 [Prozent-]Punkten) gegenübergestellt (BGE 150 V 129 E. 4.4.1; 135 V 237 E. 4.6.1; PETROV, a.a.O., Rz. 261). Die Referenzgruppe muss ihrer Zusammensetzung nach dem geprüften Leistungserbringer hinreichend ähnlich sein (BGE 150 V 129 E. 4.1; 137 V 43 E. 2.2; dazu unten E. 8.1.4 f.).

4.3. Nach Art. 56 Abs. 6 KVG legen die Leistungserbringer und Krankenversicherer vertraglich eine Methode zur Kontrolle der Wirtschaftlichkeit fest.

Dem gesetzlichen Auftrag folgend vereinbarten die beteiligten Verbände (Tarifpartner) am 20. März 2018 die Screening-Methode. Diese gilt für sämtliche Wirtschaftlichkeitskontrollverfahren ab dem Statistikjahr 2017 (vgl. nunmehr auch die aktualisierte Fassung des Vertrags vom 1. Februar 2023, in Kraft mit Wirkung seit 1. Januar 2023). Das Screeningverfahren dient dazu, Leistungserbringer zu identifizieren, deren auffällige Kostenstruktur eine möglicherweise unwirtschaftliche Behandlungsweise nahelegt (BGE 150 V 129 E. 5.3.1). Auffällige Kosten sind indessen nicht mit unwirtschaftlicher Behandlungsweise gleichzusetzen: Die Tarifpartner halten im Ingress des Vertrags über die Screening-Methode fest, diese finde "Anwendung als erster Schritt der Wirtschaftlichkeitskontrolle". Unmittelbarer Gegenstand der Vereinbarung sei allein die Regressionsanalyse, die "als Screening-Methode zur Detektion von Ärzten mit auffälligen Kosten" diene. Ob ein Arzt mit auffälligen Kosten unwirtschaftlich arbeite, sei mittels Einzelfallanalyse näher abzuklären (Ziff. 1 und 2 des Vertrags).

Das standardisierte Screening erlaubt es, kostenauffällige Leistungserbringer frühzeitig auf die Möglichkeit hinzuweisen, dass sie gegen die gesetzlichen Wirtschaftlichkeitsanforderungen verstossen

könnten (vgl. Urteil K 57/95 vom 5. Juli 1996 E. 4). Das ist nach Treu und Glauben schon deswegen geboten, weil eine Rückforderung nach Art. 59 Abs. 1 lit. b KVG kein Verschulden des Leistungserbringers voraussetzt (BGE 141 V 25 E. 8.4; vgl. GEBHARD EUGSTER, Rechtsprechung des Bundesgerichts zum KVG, 2. Aufl. 2018, N. 2 zu Art. 59 KVG; ISABELLE HÄNER, in: Basler Kommentar zum KVG/KVAG, 2020, N. 12 und 28 zu Art. 59 KVG; PICECCHI, a.a.O., Rz. 579).

4.4. Die Screening-Methode als "erster Schritt" der Wirtschaftlichkeitsprüfung besteht in einer Regressionsanalyse, mit welcher der individuelle Fallwert (Regressionsindex) ermittelt wird, und in einem Vergleich zwischen diesem und dem Gruppenfallwert. Je nach Ergebnis erfolgt in einem "zweiten Schritt" eine nähere Prüfung der Wirtschaftlichkeit (dazu unten E. 4.5).

4.4.1. Der (vorläufige) Fallwert des geprüften Leistungserbringers wird mittels einer zweistufigen Regressionsanalyse bestimmt. Die tarifpartnerschaftliche Vereinbarung installiert auf der ersten Stufe - die Patienten betreffend - die Morbiditätsindikatoren "Alter", "Geschlecht", "pharmazeutische Kostengruppen (Pharmaceutical Cost Groups, PCG)", "Franchisen" und "Spital- oder Pflegeheimaufenthalt im Vorjahr". Auf der zweiten Stufe kommen die den Leistungserbringer betreffenden Faktoren "Facharztgruppe" und "Standortkanton" hinzu (Näheres bei PETROV, a.a.O., Rz. 308 ff.).

Die Regressionsanalyse erlaubt es, kostenwirksame Effekte gewisser Praxiseigenschaften zu standardisieren, das heisst die betreffenden Abweichungen auf den Durchschnittswert der jeweiligen Facharztgruppe zu normieren. Anhand von gewichtenden Regressionsparametern resp. -koeffizienten bestimmt sich, in welchem Mass beispielsweise die Morbidität des konkreten Patientenkollektivs kostenwirksam ist. Hierbei werden etwa Abweichungen von den Durchschnittskosten, die auf im Patientenkollektiv überdurchschnittlich häufig vorkommende, kostenintensive chronische Erkrankungen zurückzuführen sind, erfasst und von Auffälligkeiten abgegrenzt, die einer ineffizienten Behandlungsweise geschuldet sind. Durch eine solche Neutralisierung verhaltensunabhängiger kostenrelevanter Faktoren kann letztlich der Kosteneffekt einer unwirtschaftlichen Behandlung isoliert werden (BGE 150 V 129 E. 4.4.1 mit Hinweisen).

4.4.2. Eine Arztpraxis ist nicht schon dann kostenauffällig, wenn der Kostenindex der geprüften Praxis den Durchschnittskostenindex der Vergleichsgruppe (= 100 Indexpunkte) überschreitet. Vielmehr ist beim Fallkostenvergleich eine Toleranzmarge von 20 bis 30 Indexpunkten vom Fallwert resp. Kostenindex des geprüften Leistungserbringers abzuziehen (BGE 137 V 43 E. 2.2 mit Hinweisen). Zweck dieser Marge ist es, im Rahmen des Grundsatzes der ärztlichen Behandlungsfreiheit dem individuellen Praxisstil des Leistungserbringers Rechnung zu tragen (vgl. GEBHARD EUGSTER, Krankenversicherung, in: Soziale Sicherheit, SBVR Bd. XIV, 3. Aufl. 2016, S. 682 Rz. 902; MIRINA GROSZ, Einzelfallprüfung vs. statistikbasierte Tatsachenermittlung im Verwaltungsrecht: Die Wirtschaftlichkeitsprüfung gemäss Art. 56 KVG, in: BJM 2024 S. 265).

Die Vorinstanz geht davon aus, die Neuerungen des Screening-Modells rechtfertigten es, den Toleranzabzug auf stets 20 Punkte zu beschränken. Die methodische Umstellung reduziert indessen nicht die Notwendigkeit, dem individuellen Praxisstil des Leistungserbringers angemessen viel Raum zu belassen (BGE 150 V 129 E. 5.4). Einem Einbezug von statistischen Streuungen, Unschärfen u.Ä. in diesem Rahmen fehlte es ausserdem an empirischer Validität (vgl. GEBHARD EUGSTER, KVG: Baustelle statistische Wirtschaftlichkeitsprüfung, in: Jusletter 27. August 2012, Rz. 37).

Der Beschwerdeführer wiederum postuliert eine Marge von fixen 30 Punkten; dies entspreche ständiger Praxis. So wurde es denn auch in einer Vielzahl von Fällen gehandhabt, die vom Bundesgericht beurteilt wurden (z.B. BGE 144 V 79 E. 3 und 6; Urteile 9C_67/2018 vom 20. Dezember 2018 E. 7, 9C_267/2017 vom 1. März 2018 E. 4 und 9C_558/2018 vom 12. April 2019 E. 6); allerdings ging es dort jeweils nicht um die Frage nach dem angemessenen Umfang der Marge. Zudem beruft sich der Beschwerdeführer auf Äusserungen der Tarifpartner, aus denen die generelle Gebotenheit einer 30

Punkte-Marge hervorgehe. Nachdem keine neuen Erkenntnisse vorlagen, die eine abweichende Betrachtung nahelegten, hat das Bundesgericht inzwischen aber an der bisherigen Rechtsprechung festgehalten, soweit diese eine - nach den individuellen Umständen differenzierende - Spanne von 20 bis 30 Indexpunkten vorsieht (Urteil 9C_166/2022 vom 9. Dezember 2024 E. 8.3.2; vgl. auch unten E. 8.2).

Die bisherige Rechtsprechung ist freilich auf dem Boden der früheren Methodik entstanden. In der nunmehr geltenden zweiteiligen Wirtschaftlichkeitsprüfung erfolgt das Screening allein anhand eines mit statistischen Daten unterlegten Regressionsberichts zum jeweiligen Leistungserbringer, während die Umstände, die letztlich für die individuelle Festlegung der Toleranzmarge massgebend sind, gutenteils erst bei einer etwaigen Einzelfallprüfung (E. 4.5 hernach) bekannt werden. Aufgrund dieser neuen Ausgangslage ist die Handhabung der Toleranzmarge anpassungsbedürftig. Es ist angezeigt, sie schon im ersten, grundsätzlich alle Leistungserbringer betreffenden Teil der Wirtschaftlichkeitsprüfung (Screening) provisorisch zu veranschlagen. Andernfalls würden sehr viele Leistungserbringer auffällig, die tatsächlich gar nicht gegen das Gebot einer wirtschaftlichen Behandlungsweise verstossen. Die vorläufige Marge ist insoweit pauschaler Natur. Es wäre in erster Linie Sache der Tarifpartner, zu Screeningzwecken einen allgemeingültigen (vorläufigen) Toleranzwert festzulegen. Die definitive, individualisierte Marge von mindestens 20 und höchstens 30 Punkten kann erst im zweiten Teil der Wirtschaftlichkeitsprüfung resp. im schiedsgerichtlichen Verfahren bestimmt werden. Die Festlegung ihres Ausmasses im Einzelfall steht im pflichtgemässen Ermessen der Krankenversicherer resp. des Schiedsgerichts. Dabei wird beispielsweise massgebend sein, ob der Leistungserbringer in kostenwirksamer Weise auf besondere Krankheiten oder Therapieformen spezialisiert ist, die indessen nicht als Praxisbesonderheiten berücksichtigt werden können (vgl. BGE 150 V 129 E. 5.4; unten E. 9.2 a.E.).

4.4.3. Die vertragliche Festlegung der Methodik und der in der Regressionsanalyse verwendeten Faktoren dient einer gleichmässigen Identifizierung von Leistungserbringern mit erhöhten Kosten ("Verdachtsfälle"). Da das Ergebnis der Regressionsanalyse allein noch keine abschliessende Aussage über die Wirtschaftlichkeit beinhaltet, schliesst die vertraglich getroffene Auswahl bestimmter Faktoren den späteren Einbezug weiterer Beweismittel und -themen nicht aus; der Grundsatz der freien Beweiswürdigung ist gewahrt (vgl. BGE 150 V 129 E. 5.3.2; kritisch UELI KIESER, Interessenabwägung - ein sichtender Gang durch das Sozialversicherungsrecht, in: ZSR 2024 II S. 200 ff. und GROSZ, a.a.O., S. 261; vgl. auch LARA DE HELLER, La polypragmasie et ses conséquences dans l'assurance-maladie, in: SZS 2024 S. 321).

4.5. Indiziert das Screening eine unwirtschaftliche Behandlungsweise, indem der Regressionsbericht zum betreffenden Leistungserbringer einen über den (vorläufigen) Toleranzbereich hinaus erhöhten Indexwert offenbart, so führen die Versicherer resp. die Santésuisse eine Einzelfallprüfung durch.

4.5.1. In diesem "zweiten Schritt" der Wirtschaftlichkeitsprüfung gilt es, die definitive Toleranzmarge festzulegen (oben E. 4.4.2). Zudem werden Praxisbesonderheiten ermittelt, die sich auf die Kostenstruktur des geprüften Leistungserbringers auswirken. Erst die Einzelfallprüfung führt gegebenenfalls zur Feststellung eines Verstosses gegen das Gebot einer wirtschaftlichen Behandlungsweise. Die Tarifpartner sehen vor, dass der Leistungserbringer Gelegenheit erhält, 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" (vgl. aktualisierte Fassung des Screening-Vertrags vom 1. Februar 2023, Ziff. 2 Abs. 4 und 5; BGE 150 V 129 E. 4.3.3).

Die Tarifpartner gehen davon aus, dass der aufgrund eines auffälligen Regressionsindexes entstandene Verdacht unwirtschaftlicher Behandlungsweise insoweit ausgeräumt wird, wie die überdurch-

schnittlichen Kosten durch praxisspezifische Effekte bedingt sind. Der individuelle Indexwert reduziert sich um die (definitive) Toleranzmarge (oben E. 4.4.2) und im Umfang des Kosteneffekts von signifikanten Eigenschaften des Leistungserbringers und des Patientenkollektivs (BGE 150 V 129 E.5.5). Ein solcher Effekt ist, wo möglich, anhand von individuellen Daten des Leistungserbringers und statistischen Vergleichswerten auszuweisen. Besonderheiten, deren Kostenwirksamkeit nicht auf statistischem Weg bezifferbar sind, können allenfalls auch auf analytischem Weg (anhand ausgewählter Patientendossiers und Rechnungen) festgestellt werden (BGE 150 V 129 E. 5.2.4 mit Hinweisen). Möglicherweise ist es aus Gründen der rechtsgleichen Behandlung und auch der Praktikabilität angezeigt, auf dem Weg der tarifpartnerschaftlichen Vereinbarung für typische Konstellationen von Praxisbesonderheiten Richtwerte festzulegen.

4.5.2. Als kostenwirksame Praxisbesonderheiten anerkannt werden seit jeher Merkmale wie beispielsweise sehr viele langjährige und/oder ältere Patienten, eine überdurchschnittliche Zahl von Hausbesuchen, ein sehr hoher Anteil an ausländischen Patienten oder der Umstand, dass keine Notfallpatienten behandelt werden (vgl. etwa Urteil 9C_558/2018 vom 12. April 2019 E. 8.2.1). Angebotsseitig ergeben sich Praxisbesonderheiten vor allem bei klar kostenwirksamen Spezialisierungen im internistischen Behandlungsspektrum (vgl. unten E. 9.3) oder beispielsweise auch, wenn eine Praxis aufgrund grosser zeitlicher Verfügbarkeit die Funktion einer "Permanence" hat, was die Breite der medizinischen Versorgung von bestehenden Patienten tendenziell erweitert (zur Publ. vorgesehenes Urteil 9C_166/2022 vom 9. Dezember 2024 E. 9.1). Praxisschwerpunkte, die ein typisches Merkmal der ärztlichen Grundversorgung betonen (z.B. eine psychosoziale und sozialmedizinische Ausrichtung), begründen indessen nur dann eine Praxisbesonderheit, wenn sie den Anteil von (z.B. morbiditätsbedingt) kostenintensiven Patienten erheblich vergrössern; im Übrigen können sie bei der Bemessung der Toleranzmarge bedeutsam sein (erwähntes Urteil 9C_166/2022 E. 9.2.3 sowie unten E. 9.2). Spezialisierte Praxisangebote, die sich ausserhalb der üblichen ärztlichen Grundversorgung bewegen, bilden womöglich Gegenstand einer eigenen Wirtschaftlichkeitsprüfung (erwähntes Urteil 9C_166/2022 E. 9.2.2 sowie unten E. 9.1.2).

Nachfrageseitig definierte Praxisbesonderheiten entstehen namentlich aus der Art, Schwere und Häufigkeit von Erkrankungen im Patientenkollektiv (Morbidität). Im Screening-Verfahren wird die Morbidität über Indikatoren ("Proxies") wie z.B. die pharmazeutischen Kostengruppen PCG bloss indirekt erschlossen (vgl. 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 [= Anhang des Screening-Vertrags vom 20. März 2018], S. 26 Ziff. 4.6 und S. 70; zu den PCG vgl. Anhang zur Verordnung des EDI vom 14. Oktober 2019 über die Umsetzung des Risikoausgleichs in der Krankenversicherung [VORA-EDI; SR 832.112.11] und BGE 150 V 129 E. 6.5.2). Es ist daher möglich, dass die Regressionsanalyse kostenwirksame Effekte der Morbidität nicht (abschliessend) neutralisiert (vgl. oben E. 4.4.1) und sie deswegen (zusätzlich) unter dem Titel einer Praxisbesonderheit berücksichtigt werden müssen (dazu BGE 150 V 129 E. 5.5.3); so etwa im Fall von Krankheiten, bei denen sich der spezifische Behandlungsaufwand nicht in einer PCG-normierten Medikation niederschlägt. Zur Erfassung der Morbidität ist unter anderem auf Angaben des Leistungserbringers über das Vorkommen und die Verteilung von Diagnosen im Patientenkollektiv zurückzugreifen (Urteil 9C_166/2022 vom 9. Dezember 2024 E. 5.5.2; vgl. TRÜMPLER/WERDER, Klarstellung des Bundesgerichts: Wirtschaftlichkeitsverfahren brauchen Einzelfallprüfungen, in: Jusletter 29. April 2024, Rz. 17); dies, zumal entgegen Art. 42 Abs. 3bis KVG in den Abrechnungsdaten diagnostische Informationen regelmässig fehlen (vgl. BGE 151 V 30 E. 5.7.4).

4.5.3. Der Leistungserbringer ist zur Mitwirkung verpflichtet. Soweit Praxisbesonderheiten nicht augenscheinlich sind und die Krankenversicherer die Verhältnisse von sich aus abklären, muss er glaubhaft machen, unter welchen Gesichtspunkten eine Einzelfallbetrachtung angezeigt und inwiefern die im Rahmen des Screenings erkannte Kostenauffälligkeit auf Praxisbesonderheiten zurückzuführen

ist (BGE 150 V 129 E. 5.3.2, 5.5.1 und 5.6; EUGSTER, SBVR, S. 683 f. Rz. 908 f.). Die Substantiierungsobliegenheit des Leistungserbringers umfasst die Besonderheit als solche und deren grundsätzliche Kostenrelevanz. Er muss den allfälligen Kosteneffekt aber nicht exakt quantifizieren; dies ist dem Prüfverfahren der Krankenversicherer vorbehalten (Urteil 9C_166/2022 vom 9. Dezember 2024 E. 5.5.3 [nicht zur Publ. vorgesehen]).

4.6. Die Wirtschaftlichkeit einer ambulanten ärztlichen Praxistätigkeit im Sinn von Art. 56 KVG wird auf der Grundlage des Gesamtkostenindexes beurteilt, der direkte und veranlasste Kosten umfasst. Derzeit wird der Gegenstand einer allfälligen Rückforderung durch den Index der direkten Kosten bezeichnet; veranlasste Kosten, die zur Feststellung eines unwirtschaftlichen Verhaltens beigetragen haben, sind nicht zurückzuerstatten (BGE 137 V 43 E. 2.5.5).

4.6.1. Zur Beantwortung der Frage, ob die Behandlungsweise im betreffenden Statistikjahr wirtschaftlich war, sind alle Kosten massgebend, die vom Verhalten des Leistungserbringers abhängig und in seinem Verantwortlichkeitsbereich entstanden sind (vgl. PICECCHI, a.a.O., Rz. 142 und 528). Eine Gesamtbetrachtung der direkten und veranlassten Kosten ist erforderlich, weil ein Arzt, der Behandlungen selber durchführt, die andere Ärzte an Dritte auslagern würden, dem Wirtschaftlichkeitsgebot entsprechend handeln kann, indem er zwar überdurchschnittlich hohe direkte, dafür aber unterdurchschnittliche veranlasste und gesamthaft durchschnittliche oder möglicherweise sogar unterdurchschnittliche Gesamtkosten verursacht (BGE 137 V 43 E. 2.5.6; PETROV, a.a.O., Rz. 279).

Veranlasste Kosten sind solche für verordnete Medikamente sowie Vergütungen für "Leistungen anderer Leistungserbringer" (BGE 137 V 43 E. 2.3), im Einzelnen: Laborleistungen (Art. 35 Abs. 2 lit. f KVG), der Untersuchung oder Behandlung dienende Mittel und Gegenstände (Art. 35 Abs. 2 lit. g KVG) sowie Leistungen auf ärztliche Anordnung oder im ärztlichen Auftrag (Art. 35 Abs. 2 lit. e KVG) durch freiberufliche (nichtärztliche) Leistungserbringer nach Art. 47 ff. KVV, wie z.B. Physiotherapie, Ergotherapie, Logopädie, Ernährungsberatung oder Neuropsychologie (BGE 133 V 37 E. 5.3.3; JUANA VASELLA, in: Basler Kommentar KVG/KVAG, 2020, N. 36 zu Art. 35 KVG).

Keine veranlassten Kosten sind Leistungen anderer Ärzte, an die der Leistungserbringer seine Patienten zur selbständigen spezialisierten Weiterbehandlung überwiesen hat (PETROV, a.a.O., Rz. 143 und 327). Dies gilt sinngemäss auch für verordnete nichtärztliche Leistungen nach Art. 35 Abs. 2 lit. e KVG ohne direkten sachlichen Zusammenhang mit der einschlägigen ärztlichen Therapie. Nicht in die Wirtschaftlichkeitsprüfung von Grundversorgern einbezogen werden daher Leistungen wie die psychologische Psychotherapie (Art. 50c KVV); für diese ist selbst dann eine separate Wirtschaftlichkeitsprüfung durchzuführen, wenn sie unter der ZSR-Nummer des Grundversorgers erbracht wird (E. 9.1.2 f.). Hingegen werden die Kosten für ausgelagerte ärztliche Leistungen (z.B. radiologische Untersuchungen) dem Gesamtkostenindex des auftraggebenden Arzt zugerechnet, wenn sie als Bestandteil der von diesem durchzuführenden Therapie erscheinen (E. 9.3.2.4).

4.6.2. Gegenstand der Rückforderung sind nur die direkten Kosten (einschliesslich etwa der vom betreffenden Arzt selbst abgegebenen Medikamente), weil nur die Rückforderung von "dem Leistungserbringer zu Unrecht bezahlte[n]" Vergütungen gesetzlich vorgesehen ist (vgl. Art. 56 Abs. 2 zweiter Satz KVG; BGE 137 V 43 E. 2.5). Hingegen erfolgt die Berechnung der Rückforderung wiederum anhand des (bereinigten) Gesamtkostenindexes (unten E. 10.2-10.4).

5.

Der Beschwerdeführer rügt vorab, die Vorinstanz habe sein rechtliches Gehör (Art. 29 Abs. 2 BV) verletzt, indem sie ihn nicht zu ihrer Absicht angehört habe, einen höheren Rückzahlungsbetrag festzulegen als im Klagebegehren (Berechnung gemäss Regressionsindex) beantragt. Nach dem kantonalen Prozessrecht ist das Schiedsgericht nicht an die Begehren der Parteien gebunden; es muss

aber das rechtliche Gehör gewähren, wenn es etwa der klagenden Partei mehr zuspricht, als sie verlangt hat (vgl. Art. 46 Abs. 2 des bernischen Gesetzes betreffend die Einführung der Bundesgesetze über die Kranken-, die Unfall- und die Militärversicherung [EG KUMV; Systematische Sammlung [BSG] 842.11] und Art. 92 Abs. 3 des Gesetzes über die Verwaltungsrechtspflege [VRPG; BSG 155.21]; RUTH HERZOG, in: Kommentar zum Gesetz vom 23. Mai 1989 über die Verwaltungsrechtspflege des Kantons Bern, 2. Aufl. 2020, N. 9 zu Art. 92 VRPG). Wie es sich damit verhält, kann hier mit Blick auf den Ausgang des Verfahrens offenbleiben.

6.

Zunächst gilt es zu prüfen, ob die klagenden Krankenversicherer und das Schiedsgericht methodisch rechtskonform vorgegangen sind.

6.1. Der Beschwerdeführer wendet sich gegen die vorinstanzliche Auffassung, bereits die Regressionsanalyse erbringe den Beweis einer unwirtschaftlichen Behandlung. Das Schiedsgericht verletze Art. 56 Abs. 6 KVG, indem es ausser Acht lasse, dass die Tarifparteien mit dem Screening-Vertrag ausdrücklich eine andere Vorgehensweise gewählt hätten. Die erforderliche Einzelfallanalyse habe im vorliegenden Fall nicht stattgefunden. In der Klage fehle jegliche Auseinandersetzung mit den Praxisbesonderheiten. Die Unwirtschaftlichkeit als Grundlage für die eingeklagte Rückerstattung sei nicht nachgewiesen und die Klage schon deswegen abzuweisen.

6.2. Die seit dem Rechnungsjahr 2017 geltenden methodischen Vorgaben betonen die Zweiteilung der Wirtschaftlichkeitsprüfung (oben E. 4.3-4.5). Eine im Screening konstatierte Auffälligkeit leitet eine entsprechende Einzelfallprüfung ein. Die Santésuisse als mit der Wirtschaftlichkeitskontrolle betrauter Verband der Versicherer ist gemäss Screening-Vertrag für beide Teile der Wirtschaftlichkeitsprüfung zuständig; sie nimmt also nach der Regressionsanalyse auch die Einzelfallbeurteilung vor. Im Hinblick auf diese ist der Leistungserbringer zur Mitwirkung verpflichtet, d.h. angehalten, die Besonderheiten seiner Praxis zu benennen und zu substantiieren (oben E. 4.5.3). Das partizipativ angelegte Verfahren gibt den Parteien Gelegenheit, sich untereinander gütlich zu einigen (BGE 150 V 129 E. 5.6). Hat der überprüfte Leistungserbringer im Rahmen der Einzelfallprüfung keine Praxismerkmale geltend gemacht, die nahelegen, dass die im Screening-Verfahren ermittelte Überschreitung wegen Praxiseffekten tatsächlich geringer ist oder wegfällt, so ist dies in der Rückforderungsklage festzustellen (Urteil 9C_166/2022 vom 9. Dezember 2024 E. 6.3).

Die beim kantonalen Schiedsgericht einzubringende Klage muss auf den Ergebnissen einer kompletten Einzelfallbeurteilung beruhen. Diese kann nicht in das schiedsgerichtliche Verfahren verlagert werden. Das kontradiktorische und förmliche Klageverfahren, in dem die vormals mit der objektiven Wirtschaftlichkeitskontrolle betraute (Verwaltungs-) Instanz zur Partei wird (vgl. GEBHARD EUGSTER, Wirtschaftlichkeitskontrolle ambulanter ärztlicher Leistungen mit statistischen Methoden, 2003, Rz. 362), ist wenig geeignet, die von hoher Technizität geprägte Aufarbeitung der Grundlagen der eingeklagten Rückforderung wirksam nachzuholen. Aufgabe des Schiedsgerichts ist es, weiterhin bestrittene Feststellungen zu überprüfen und den Rechtsstreit wenn möglich in einem Vermittlungsverfahren beizulegen oder andernfalls autoritativ zu entscheiden (BGE 150 V 129 E. 5.6). Soweit das anwendbare Verfahrensrecht es vorsieht, kann auch schon vor Anhebung einer Klage ein Vermittlungsverfahren beansprucht werden (für den Kanton Bern: Art. 44 f. EG KUMV).

6.3. Der Hergang des bisherigen Verfahrens zeigt, dass die Einzelfallprüfung - im Wesentlichen der Abgleich des Screening-Ergebnisses mit konkreten Praxismerkmalen - vorliegend weitgehend dem (hierfür wie erwähnt wenig geeigneten) schiedsgerichtlichen Prozess überlassen wurde.

6.3.1. In der beim Schiedsgericht eingereichten Klageschrift vom 30. Juni 2021 erklärten die Beschwerdegegnerinnen, die Regressionsanalyse weise für das Statistikjahr 2019 einen Gesamtkostenindex von 158 Punkten aus. Der Beschwerdeführer werde mit dem Kollektiv der Facharztgruppe

"Allgemeine Innere Medizin" verglichen (im Statistikjahr 2019 schweizweit 5'250 Praxen mit mindestens 50 Erkrankten oder insgesamt mindestens Fr. 100'000.- Bruttoleistungen). Damit sei der berechnete Regressionsindex des Beklagten statistisch aussagekräftig. Die Kostenüberschreitung um 38 Indexpunkte (Differenz zwischen dem Gesamtkostenindex des Arztes und dem Index der Vergleichsgruppe, abzüglich einer Toleranzmarge von 20 Punkten) qualifiziere die Tätigkeit des Beklagten als unwirtschaftlich. Die Berechnung der Rückforderung erfolge auf der Grundlage nur der direkten, nicht auch der veranlassten Kosten. Bei Gesamtkosten (GK) von Fr. 1'037'609.- (einschliesslich veranlasster Kosten) und totalen direkten Kosten (TDK) von Fr. 615'469.-, einem Regressionsindex Totale Kosten (RI) von 158 und einem Toleranzwert (TB) von 120 Punkten errechne sich eine Rückforderungssumme von Fr. 148'024.- (Formel: $GK:RI \times (RI-TB) \times TDK:GK$).

Mit der Klage reichte die Santésuisse u.a. die vom Beklagten abgerechneten Kosten (Jahresdaten 2019 aus dem Datenpool der SASIS AG), den Regressionsbericht (Indexzahlen betreffend das Statistikjahr 2019, Morbiditätsindikatoren) sowie die Namensliste und eine Kostenaufstellung des Referenzkollektivs "Facharztgruppe Allgemeine Innere Medizin" ein.

Am 14. Januar 2022 teilte die Santésuisse dem Schiedsgericht namens der Klägerinnen mit, Vergleichsverhandlungen seien gescheitert. Der Beklagte reichte innert gesetzter Frist keine Klageantwort ein. Daraufhin schloss das kantonale Schiedsgericht den Schriftenwechsel und das Beweisverfahren.

6.3.2. Die Vorinstanz hält im angefochtenen Urteil fest, der arithmetische Mittelwertvergleich erbringe den vollen Beweis für eine unwirtschaftliche Behandlung. Es bestehe kein Anlass, die Vergleichsgruppe (Fachärzte "Allgemeine Innere Medizin") mit Blick auf vom Beklagten erworbene Fähigkeitsausweise (Delegierte Psychotherapie, Praxislabor) und Schwerpunkte (Manuelle Medizin, Sportmedizin) enger zu fassen, zumal die entsprechenden Weiterbildungen auch im Vergleichskollektiv enthalten seien. Die für die Regressionsanalyse massgebenden Daten und die Werte der zugrundeliegenden Morbiditätsindikatoren ergäben sich aus dem Regressionsbericht. Durch die Morbiditätsindikatoren würden auch kostensteigernde Faktoren, die aus den Weiterbildungen folgten, erkannt und ausgeglichen. Praxisbesonderheiten, die nicht bereits durch die in der Regressionsanalyse verwendeten Morbiditätsvariablen abgebildet würden, seien nicht erkennbar. Der Beklagte mache dergleichen auch nicht geltend. Demnach sei die Wirtschaftlichkeitsprüfung auf der Grundlage der vertraglich vereinbarten und (grundsätzlich wie konkret) beweistauglichen Regressionsanalyse durchzuführen. Der Beklagte weise gemäss Klage betreffend die Gesamtkosten im Statistikjahr 2019 158 Indexpunkte aus. Bei der Veranschlagung des Toleranzbereichs sei zu berücksichtigen, dass die Regressionsanalyse im Vergleich zur Vorgängermethode (ANOVA) zusätzliche Faktoren einbeziehe, so dass sich die Aussagekraft des Modells und dadurch die Güte der Wirtschaftlichkeitsüberprüfung wesentlich verbessere. Der Toleranzbereich sei daher auf 20 Indexpunkte zu beschränken. Im Umfang der verbleibenden Überschreitung werde der Beklagte rückerstattungspflichtig.

6.3.3. Nachdem das angefochtene Urteil am 9. März 2022 ergangen war, liess der Beschwerdeführer vor Schiedsgericht beantragen, der Entscheid sei aufzuheben und die Frist zur Klageantwort sei wiederherzustellen; er sei unverschuldet - krankheitshalber - verhindert gewesen. Mit Urteil vom 9. September 2022 wies die Vorinstanz das Gesuch ab. Dieser Entscheid wurde rechtskräftig.

6.4.

6.4.1. Nach der Konzeption des Screening-Vertrags (oben E. 6.2) setzt die Klage voraus, dass: der im Screening als kostenauffällig erkannte Leistungserbringer Gelegenheit hatte, (nicht offensichtliche) Praxisbesonderheiten zu benennen und zu substantiiieren resp. andere Einwendungen (z.B. hinsichtlich des Umfangs der Toleranzmarge) zu erheben; der Verband der Versicherer die nötigen Abklärungen und Auswertungen statistischer (oder analytischer) Art getätigt hat; die Parteien wenn möglich auf

eine gütliche Einigung hingearbeitet haben; dies bevor die prüfende Instanz schliesslich feststellt, inwieweit individuelle Praxiseffekte die im Screening angezeigte Indexdifferenz verändern, und gegebenenfalls eine entsprechende Rückforderung klageweise geltend macht (Urteil 9C_166/2022 vom 9. Dezember 2024 E. 6.4.4).

6.4.2. Gemessen an diesen Erfordernissen sind die vorinstanzlichen Entscheidungsgrundlagen unvollständig. Die Wirtschaftlichkeitsprüfung kann nur dann unmittelbar auf Grundlage der Regressionsanalyse abgeschlossen werden, wenn feststeht, dass keine besonderen Faktoren die Kostenstruktur des Leistungserbringers beeinflussen (vgl. BGE 150 V 129 E. 5.5), oder wenn der Leistungserbringer nicht Gesichtspunkte glaubhaft machen kann, die als Praxisbesonderheiten eine Einzelfallprüfung bedingen, resp. wenn er nicht plausibilisiert, inwiefern die im Rahmen des Screenings erkannte Auffälligkeit einer solchen Praxisbesonderheit zuzuschreiben sei; ebenso, wenn er die Mitwirkung bei der Erhebung von Daten verweigert, die zur Interpretation der statistischen Daten erforderlich sind (vgl. oben E. 4.5.3 mit Hinweisen). Der beim Schiedsgericht eingereichten Klage der Beschwerdegegnerinnen ist hierzu einzig der allgemeine Hinweis zu entnehmen, es obliege "dem statistisch auffälligen Arzt, mittels Praxisbesonderheiten den Vermutungsbeweis der Unwirtschaftlichkeit zu widerlegen". Vorliegend bedarf insbesondere näherer Betrachtung, inwiefern die augenscheinlichen Praxisbesonderheiten - namentlich das in Teilen einer anderen medizinischen Disziplin zuzuordnende (Delegierte Psychotherapie) oder hinsichtlich einer Spezialisierung (z.B. Manuelle Medizin) deutlich vom üblichen Tätigkeitsfeld internistischer Praxen abweichende Behandlungsspektrum (dazu unten E. 9) - für die Wirtschaftlichkeitsprüfung relevant sind.

Der Beschwerdeführer rügt daher zu Recht, dass die Beschwerdegegnerinnen die Einzelfallbeurteilung nicht umgesetzt haben. Dies verletzt Art. 56 KVG. Daran ändert der Umstand nichts, dass der Beschwerdeführer im schiedsgerichtlichen Verfahren keine Klageantwort eingereicht und die Vorinstanz sein Gesuch um Wiederherstellung der Frist abgelehnt hat (oben E. 6.3.3). Nach dem Gesagten wäre es erforderlich gewesen, den Leistungserbringer schon vor Einreichung der Klage aufzufordern, kostenwirksame Besonderheiten seiner Praxis zu benennen und zu substantiieren (BGE 150 V 129 E. 5.8.1). Die betreffenden Ergebnisse hätten im vorinstanzlichen Verfahren von Amtes wegen berücksichtigt werden müssen. Die unterbliebene Einzelfallbeurteilung kann, wie erwähnt (E. 6.2), nicht im schiedsgerichtlichen Verfahren nachgeholt werden. Die fehlende Klageantwort im schiedsgerichtlichen Verfahren gereicht dem Beschwerdeführer daher nicht zum Nachteil.

7.

Im Hinblick auf die Fortsetzung des Verfahrens (Neubeurteilung auf der Grundlage einer ergänzten Klage; unten E. 11.1) bleibt auf die in der Sache erhobenen Rügen des Beschwerdeführers einzugehen. Die Vorbringen betreffen die Ermittlung der im Screening-Verfahren gegenüberzustellenden Fallkosten der Vergleichsgruppe (richtige Zusammensetzung des Referenzkollektivs; nachfolgend E. 8.1) und des Leistungserbringers (Toleranzmarge; E. 8.2), die im Rahmen einer Einzelfallprüfung zu veranschlagenden Korrekturen (Praxisbesonderheiten; E. 9) sowie die Berechnung der Rückforderung (E. 10).

8.

8.1. Der Beschwerdeführer beanstandet, dass die Zusammensetzung der Vergleichsgruppe (Referenzkollektiv), an deren durchschnittlichem Fallwert sein im Screening-Verfahren ermittelter Regressionsindex gemessen wird, allein anhand des Facharztstitels erfolgt.

8.1.1. Er macht geltend, so werde das Erfordernis, Gleiches mit Gleichem zu vergleichen, verfehlt. Es müsse über die jeweilige Facharztgruppe hinaus differenziert werden. Vor allem bei besonderen Leistungsangeboten (Praxisschwerpunkten) mit entsprechender Ausstattung der Praxis und/oder bei einem erhöhten Anteil von Patienten mit besonders kostspieligem Therapiebedarf müsse das Profil

der in der Vergleichsgruppe enthaltenen Leistungserbringer weitgehend mit dem Profil der geprüften Praxis übereinstimmen. Dies bedinge eine Anpassung des Referenzkollektivs. Sein Patientenstamm sei nicht mit jenem eines gewöhnlichen Grundversorgers mit internistischem Facharztstitel vergleichbar. Die Fähigkeitsausweise "Praxislabor", "Röntgen", "Sportmedizin (SGSM) " und "Delegierte Psychotherapie (FMPP) " sowie der interdisziplinäre Schwerpunkt "Manuelle Medizin (SAMM) " wirkten sich auf seine Tätigkeit aus; das gelte es im Rahmen des Vergleichskollektivs zu berücksichtigen.

Der Beschwerdeführer fordert mithin eine Beschränkung der Vergleichsgruppe auf Leistungserbringer, die nicht nur hinsichtlich der fachärztlichen Disziplin (hier: Allgemeine Innere Medizin), sondern auch im Leistungsprofil - namentlich was Behandlungsschwerpunkte angeht - mit seiner Praxis übereinstimmen (vgl. GEBHARD EUGSTER, KVG: Statistische Wirtschaftlichkeitsprüfung im Wandel, in: Jusletter 25. Juni 2012, Rz. 8; DERS., a.a.O., Jusletter 27. August 2012, Rz. 21). Dies führt zur Frage, wie weit das Referenzkollektiv schon im Screening-Verfahren auf (der prüfenden Instanz bekannte) kostenrelevante Spezialisierungen des Leistungserbringers abzustimmen ist, resp. wie weit jene stattdessen bei der anschliessenden Einzelfallprüfung zu einer Korrektur des Fallkostenwerts führen.

8.1.2. Ansatzpunkt der Zusammensetzung einer Vergleichsgruppe ist der Weiterbildungstitel gemäss Gesetzgebung über die universitären Medizinalberufe. Anstelle der früheren privatrechtlichen Facharztstitel der FMH bestehen seit 1. Juni 2002 eidgenössische Weiterbildungstitel (vgl. Art. 5 Abs. 2 und Art. 20 des Bundesgesetzes vom 23. Juni 2006 über die universitären Medizinalberufe [Medizinalberufegesetz, MedBG; SR 811.11]; Art. 2 Abs. 1 und Anhang 1 der Verordnung vom 27. Juni 2007 über Diplome, Ausbildung, Weiterbildung und Berufsausübung in den universitären Medizinalberufen [Medizinalberufeverordnung, MedBV; SR 811.112.0]; Botschaft vom 3. Dezember 2004 zum Bundesgesetz über die universitären Medizinalberufe, BBl 2005 184 Ziff. 1.2). Die vormaligen Disziplinen "Innere Medizin" und "Allgemeinmedizin" sind 2011 im hier interessierenden fachärztlichen Weiterbildungstitel "Allgemeine Innere Medizin" aufgegangen. Unter diesem Titel praktizieren Mediziner, die sich der Grundversorgung, also der hausärztlichen Tätigkeit, widmen (zum "Curriculum Hausärztin/Hausarzt" vgl. Schweizerisches Institut für ärztliche Weiter- und Fortbildung [SIWF], Weiterbildungsprogramm Fachärztin oder Facharzt für Allgemeine Innere Medizin, Stand August 2023[<https://www.siwf.ch/weiterbildung/facharztstitel-und-schwerpunkte/allgemeine-innere-medizin.cfm>]). Daneben gibt es im Bereich der Inneren Medizin selbständige Weiterbildungstitel, die sich auf die Diagnose und Behandlung von Erkrankungen bestimmter innerer Organe konzentrieren (z.B. Gastroenterologie, Hämatologie, Kardiologie). Die hier relevante Vergleichsgruppe der Allgemeinen Inneren Medizin umfasst nur Grundversorger. Ärztliche Leistungserbringer, die (zusätzlich) über einen spezialisierten Weiterbildungstitel verfügen und gleichzeitig haus- und spezialärztlich tätig sind, dürfen, sofern möglich, nur im Umfang ihrer allgemeininternistischen Tätigkeit einbezogen werden (vgl. PETROV, a.a.O., Rz. 151 und 338 f.; GEBHARD EUGSTER, Überarztung aus juristischer Sicht, in: Rechtsfragen zum Krankheitsbegriff, Gächter/Schwendener [Hrsg.], 2009, S. 105).

8.1.3. Der Beschwerdeführer begründet seine Forderung nach einer differenzierteren Zusammensetzung der Vergleichsgruppe unter anderem mit dem Hinweis, praxismässig müssten ja auch Merkmale wie etwa eine kantonale Bewilligung zur Führung einer Praxisapotheke (sog. Selbstdispensation) berücksichtigt werden. Vor Einführung der Screening-Methode mahnte das Bundesgericht in der Tat zuweilen eine grössere Homogenität des Referenzkollektivs an: So erblickte es in der kantonalen Bewilligung zur Selbstdispensation ein Praxismerkmal, das bei der Bildung der Vergleichsgruppe grundsätzlich berücksichtigt werden müsse. Bei der Wirtschaftlichkeitskontrolle einer internistischen Praxis mit Bewilligung zur Führung einer Apotheke seien nicht selbstdispensierende Ärzte aus dem Referenzkollektiv zu entfernen (Urteil 9C_67/2018 vom 20. Dezember 2018 E. 12.2.3).

Unter der Geltung der aktuellen Methode knüpfte das Bundesgericht bei der bisherigen Rechtsprechung an und erwog, ein Praxismerkmal solle vorzugsweise schon bei der Zusammensetzung des Referenzkollektivs berücksichtigt werden, damit beim Screening eine akzeptable Trefferqualität erzielt werde. Für den Fall, dass keine entsprechend spezifizierte Vergleichsgruppe geschaffen werden kann, sah das Gericht vor, den Kosteneffekt eines solchen Merkmals auf Stufe Einzelfallprüfung zu veranschlagen, dies anhand statistischer Daten oder alternativ auch auf analytischem Weg durch Auswertung von Falldossiers (BGE 150 V 129 E. 5.5.2). In jenem Verfahren war es aus (näher bezeichneten) methodischen Gründen denn auch kaum möglich, das Vergleichskollektiv zur Ermittlung des Regressionsindex dahin anzupassen, dass es ausschliesslich selbstdispensierende Arztpraxen enthielt (dazu BGE 150 V 129 E. 6.6.1).

8.1.4. Das Erfordernis, wonach die Vergleichsgruppe dem geprüften Leistungserbringer "hinreichend ähnlich" sein muss (vgl. oben E. 4.2 a.E.), ist mit Blick auf die Vorgaben des Screening-Modells systemkonform umzusetzen.

Seit dem Methodenwechsel per Statistikjahr 2017 dient das Screening erst einmal der Identifizierung von möglicherweise unwirtschaftlich behandelnden Leistungserbringern ("red flag"). Die Regressionsvariablen (oben E. 4.4.1) schaffen die Voraussetzungen dafür, dass potentiell unwirtschaftlich arbeitende Leistungserbringer hinreichend treffgenau identifiziert werden können. Ein Vergleich zwischen dem Regressionsindex des geprüften Leistungserbringers und dem Durchschnittswert der Vergleichsgruppe wird umso aussagekräftiger, je spezifischer das Referenzkollektiv zusammengesetzt ist (EUGSTER, SBVR, S. 679 f. Rz. 889). Falsch-positive - oder falsch-negative - Screening-Ergebnisse sollen möglichst minimiert und (unnötige) Einzelfallprüfungen verhindert werden.

Andererseits erfolgt die Feststellung eines tatsächlichen Verstosses gegen das Wirtschaftlichkeitsgebot immer erst bei der Einzelfallprüfung. Kostenwirksame Besonderheiten, die in der Referenzgruppe nicht ausreichend repräsentiert sind (vgl. die zu publ. E. 9.2.3 des Urteils 9C_166/2022 vom 9. Dezember 2024), können sowieso nur durch eine individuell-konkrete Beurteilung abschliessend erfasst und gewichtet werden. Die Aufgabenteilung zwischen Screening und Einzelfallprüfung reduziert daher die Notwendigkeit, "praxistypologische" Merkmale (dazu BGE 150 V 129 E. 5.5.2) möglichst schon im Referenzkollektiv abzubilden (vgl. dazu EUGSTER, a.a.O., Jusletter 27. August 2012, Rz. 48). Ohnehin käme dies nur für gemeinsame Merkmale einer genügend grossen, klar abgrenzbaren Untergruppe des betreffenden Facharztkollektivs (z.B. selbstdispensierende Ärzte) infrage.

Hinzu kommt Folgendes: Versuche, die Vergleichsgruppe auf Leistungserbringer mit möglichst vielen gemeinsamen Merkmalen zu beschränken, führen zu Abgrenzungsproblemen, weil die betreffenden Folgen für das Kostenprofil des geprüften Leistungserbringers im Rahmen der Regressionsanalyse mitunter nicht abschliessend quantifizierbar sind und sie demnach in der Einzelfallprüfung ohnehin nochmals aufgegriffen werden müssen (vgl. BGE 150 V 129 E. 5.5.3; oben E. 4.5.2 zweiter Abs.). Eine standardisierte Zusammensetzung der Vergleichsgruppe scheint zudem notwendig, um das Screening auf ökonomische Art und Weise umzusetzen; der Aufwand, den die Krankenversicherer bei der Wirtschaftlichkeitskontrolle betreiben, muss seinerseits verhältnismässig und damit wirtschaftlich sein (Art. 19 Abs. 1 KVAG; BGE 151 V 30 E. 5.6.3).

8.1.5. Insgesamt soll die Individualisierung nach konkreten Eigenschaften des Leistungserbringers (Praxisschwerpunkte etc.) in erster Linie im Rahmen der Einzelfallprüfung gewährleistet und entsprechend nicht mehr - wie vor Einführung der Screening-Methode gebräuchlich (vgl. z.B. Urteil 9C_535/2014 vom 15. Januar 2015 E. 6.2.2 [in BGE 141 V 25 nicht publ.]) - vorrangig als Frage der richtigen Zusammensetzung der Vergleichsgruppe behandelt werden. Demnach ist das nach dem Facharztstitel "Allgemeine Innere Medizin" zusammengesetzte Referenzkollektiv dem Status des Beschwerdeführers "hinreichend ähnlich", sofern sich das Kollektiv auf internistische Praxen beschränkt, die als Grundversorger auftreten (E. 8.1.2).

8.2. Weiter ist strittig, ob die Toleranzmarge bei der Bestimmung des individuellen Fallwerts im Vorhinein auf 20 oder aber auf 30 Punkte zu beschränken ist.

8.2.1. Die Vorinstanz vertritt die Auffassung, verglichen mit der Vorgängermethode ANOVA beziehe die Regressionsanalyse zusätzliche (Morbiditäts-) Faktoren mit ein, so dass sich die Aussagekraft des Modells und die Güte des Verfahrens dadurch wesentlich verbesserten (vgl. oben E. 4.4.2). Anstelle des nach bisheriger Rechtsprechung bestehenden Toleranzbereichs von 120 bis 130 Indexpunkten werde daher ein Wert von 120 Indexpunkten massgebend.

Der Beschwerdeführer rügt, damit stelle die Vorinstanz auf eine unbelegte Behauptung der Beschwerdegegnerinnen ab, die insbesondere in seinem konkreten Fall nicht zuzutreffen brauche. Die Morbiditätsvariablen, namentlich die PCG, korrigierten das Kostenbild ohnehin nur unzureichend. Der dem tarifpartnerschaftlichen Methodenvertrag zugrundeliegende Schlussbericht der Polynomics AG halte zudem fest, dass erst Praxen, die einen Indexwert von über 130 Punkten aufweisen, also mehr als 30 Prozent über dem erwarteten Wert ihrer Facharztgruppe liegen, als auffällig gälten. Wenn die Vorinstanz die Grenze im Vorhinein bei bloss 120 Punkten ansetze, so verletze sie Art. 56 Abs. 6 KVG.

8.2.2. Die mit der Screening-Methode verbundene Berücksichtigung zusätzlicher Morbiditätsvariablen beeinflusst den Umfang des Toleranzbereichs nicht (E. 4.4.2). Zweck der Toleranzmarge ist es, im Rahmen des Grundsatzes der ärztlichen Behandlungsfreiheit dem individuellen Praxisstil des Leistungserbringers Rechnung zu tragen, d.h. dem Umstand, dass der angestrebte Heilungserfolg (Behandlungsergebnis) auf unterschiedlichen therapeutischen Wegen erreicht werden kann. Die Toleranz ist nicht dazu bestimmt, einer methodenbedingten Standardabweichung Rechnung zu tragen (BGE 150 V 129 E. 5.4 mit Hinweisen).

8.2.3. Der Beschwerdeführer bezieht sich unter anderem auf eine Formulierung im Schlussbericht der Polynomics AG von September 2017, den die Tarifpartner zum integrierten Bestandteil der Vereinbarung über die Screening-Methode erklärt haben. Darin ist zwar ausschliesslich von einer Toleranz von 30 Prozentpunkten über dem erwarteten Wert der betreffenden Facharztgruppe die Rede (a.a.O. S. 11 Ziff. 1.5). Dies bedeutet indessen nicht, dass der Toleranzbereich damit vertraglich fixiert worden wäre. Weder dem Vertrag selbst noch dessen Anhang (dem erwähnten Schlussbericht) ist zu entnehmen, dass der Wert von 130 Indexpunkten, wie er in der Praxis offenbar meist verwendet wird (oben E. 4.4.2), in jedem Fall massgebend sein soll. Die Vertragsparteien treffen keine generelle Festlegung, sondern verweisen auf die Rechtsprechung, die immer von einem Toleranzbereich von 120 bis 130 Indexpunkten ausgegangen ist. Insoweit beträgt die vom individuellen Indexwert des Beschwerdeführers abzuziehende Toleranzmarge weiterhin mindestens 20 und höchstens 30 Prozentpunkte (so schon die Urteile 9C_166/2022 vom 9. Dezember 2024 E. 8.3.2 und 9C_126/2023 vom 4. März 2024 E. 7.3; zur differenzierten Handhabung des Abzugs im Verlauf einer zweiteiligen Wirtschaftlichkeitsprüfung vgl. oben E. 4.4.2). Die Vorinstanz wird den Toleranzabzug nach pflichtgemässen Ermessen festlegen (vgl. E. 11.1).

9.

Weiter ist auf die Rolle verschiedener geltend gemachter Praxisbesonderheiten im Rahmen der Einzelfallprüfung einzugehen.

9.1. Der Beschwerdeführer rügt zunächst, die überdurchschnittlichen Kosten, die aus der Integrierung von delegierter Psychotherapie in seine internistische Praxis resultierten, seien zu Unrecht nicht berücksichtigt worden.

9.1.1. Er bringt vor, er delegiere Patienten mit psychischen Problemen, die er selbst hausärztlich betreue, für die psychotherapeutische Behandlung an einen von ihm angestellten Psychologen. Dessen

Leistungen würden ebenfalls über seine ZSR-Nummer abgerechnet und flössen so in die Wirtschaftlichkeitsprüfung ein. Die Behandlung der betreffenden Patienten sei deutlich (zeit-) aufwendiger und entsprechend kostenintensiver als der Durchschnitt des Patientenkollektivs. Er habe die Tarifposition "Delegierte psychotherapeutische Behandlung in der Arztpraxis, Einzelsetting, pro 5 Min" im Statistikjahr 2019 9'318-mal abgerechnet, während dies in der Vergleichsgruppe durchschnittlich nur 315-mal der Fall gewesen sei. Seine direkten Kosten im Bereich delegierte Psychotherapie betrügen Fr. 100'353.-, diejenigen der Vergleichsgruppe Fr. 3'502.-. Von den 3'672 Praxen im Vergleichskollektiv führten gerade einmal deren 145 überhaupt die delegierte Psychotherapie durch. Soweit die psychisch kranken Patienten erheblich überdurchschnittliche Kosten verursachten, müsse dies als Praxisbesonderheit berücksichtigt werden.

9.1.2. Das Modell der "delegierten Psychotherapie" existierte im hier massgebenden Statistikjahr 2019 noch. Auf den 1. Juli 2022 wurde es durch ein Anordnungsmodell ersetzt; seit dem 1. Januar 2023 gibt es keine delegierte Psychotherapie mehr (zur Entstehung, Definition, KVG-rechtlichen Bedeutung und Ablösung der delegierten Psychotherapie vgl. das zur Publ. bestimmte Urteil 9C_166/2022 vom 9. Dezember 2024 E. 9.2.2.2).

Eine Wirtschaftlichkeitskontrolle von Leistungen der delegierten Psychotherapie war von vornherein nur erforderlich und zulässig, soweit der Krankenversicherer nicht gestützt auf einen vertrauensärztlichen Bericht Kostengutsprache für eine fortgesetzte (über 40 Sitzungen hinausreichende) Therapie erteilt hat (vgl. Art. 3b Verordnung des EDI vom 29. September 1995 über Leistungen in der obligatorischen Krankenpflegeversicherung [Krankenpflege-Leistungsverordnung, KLV; SR 832.112.31]). Soweit der zuständige Krankenversicherer sie genehmigt hat, gilt eine psychotherapeutische Behandlung als in wirtschaftlicher Weise erbracht (Urteile 9C_570/2015 vom 6. Juni 2016 E. 7.2 und 7.3, K 172/97 vom 23. April 1999 E. 5). Im Übrigen steht die mit einem hohen Zeitaufwand verbundene Psychotherapie - die in delegierter Form als ärztliche Leistung gilt - ausserhalb des Leistungsspektrums einer internistischen Praxis und kann daher nicht als Praxisbesonderheit behandelt werden. Ebensowenig ist es angezeigt, dieses Zusatzangebot über eine spiegelbildlich zusammengesetzte Vergleichsgruppe zu erfassen (vgl. dazu auch E. 8.1). Vielmehr braucht es hier für die unter den Titeln der allgemeinmedizinischen Behandlung einerseits und der delegierten Psychotherapie andererseits entstandenen Kosten je eine separate Wirtschaftlichkeitsprüfung (oben E. 4.6.1; erwähntes Urteil 9C_570/2015 E. 7.3; vgl. auch Urteil K 50/00 vom 30. Juli 2001 E. 6; EUGSTER, Wirtschaftlichkeitskontrolle, Rz. 424). Patienten, die sowohl ärztliche Grundversorgung wie auch Psychotherapie beanspruchen, werden zwei verschiedenen Patientenkollektiven zugeordnet. Für psychotherapeutische Leistungen, die noch nicht qua Kostengutsprache genehmigt sind, werden eigene Prokopf-Fallwerte berechnet (erwähntes Urteil 9C_166/2022 E. 9.2.2.3).

9.1.3. Nach dem Gesagten ist die nachträgliche Wirtschaftlichkeitskontrolle im Bereich der Psychotherapie erheblich eingeschränkt. Angesichts dessen hat das Bundesgericht mit Urteil 9C_166/2022 vom 9. Dezember 2024 festgehalten, es sei den fachkundigen Tarifpartnern resp. dem Schiedsgericht zu überlassen, inwiefern sie das vorstehend skizzierte Vorgehen (Aufteilung der Wirtschaftlichkeitsprüfung) als sinnvoll und praktikabel beurteilen (a.a.O. E. 9.2.2.3 a.E.). Unter diesem Vorbehalt ist die Frage im Rahmen der neuen schiedsgerichtlichen Beurteilung (unten E. 11.1) aufzugreifen. Sofern es mit Blick auf den Umstand, dass die psychotherapeutischen Leistungen zu einem erheblichen Teil einer (neuen) Wirtschaftlichkeitsprüfung entzogen sind (dazu EUGSTER, a.a.O., Jusletter 25. Juni 2012, Rz. 32 f.), überhaupt noch angezeigt und verhältnismässig erscheint, sind die im streitgegenständlichen Statistikjahr 2019 erbrachten Leistungen des Beschwerdeführers in die zwei Gruppen "ärztliche Grundversorgung" und "delegierte Psychotherapie" aufzuteilen; die Tarifpool-Daten der SASIS AG (vgl. oben E. 6.3.1) weisen auch die Fakturabeträge (nach TARMED-Positionen) für delegierte psychotherapeutische Behandlungen aus. Anhand der jeweiligen Anzahl entsprechend behandelter Patienten ist gegebenenfalls für beide Bereiche je ein Fallwert zu ermitteln und demjenigen eines entsprechenden Referenzkollektivs gegenüberzustellen.

9.2. Der Beschwerdeführer bringt weiter vor, im Patientenkollektiv seiner Praxis finde sich ein erheblich überdurchschnittlicher Anteil von Patienten mit psychischen Problemen, deren Behandlung auch im Rahmen der Grundversorgung deutlich (zeit-) aufwendiger und entsprechend kostenintensiver sei. Dabei handelte es sich allenfalls dann um eine Praxisbesonderheit, wenn sich mithilfe einer signifikanten Zahl entsprechender Diagnosen belegen liesse (vgl. erwähntes Urteil 9C_166/2022 E. 5.5.2 zweiter Abs. und [zu publ.] E. 9.3.3 a.E.), dass sich die Morbidität vom durchschnittlichen Patientengut internistischer Praxen deutlich unterscheidet. Diesfalls wäre zu prüfen, wie weit sich die hausärztliche Grundversorgung dadurch erheblich aufwendiger gestaltet, und inwiefern die abweichende Morbidität nicht schon vorweg durch die Regressionsanalyse (anhand einschlägiger PCG [oben E. 4.5.2] wie z.B. Depression) neutralisiert wurde (E. 4.4.1).

Hingegen anerkennt die Rechtsprechung den Schwerpunkt Psychosomatik und Sozialmedizin als solchen nicht als Praxisbesonderheit für internistische Leistungserbringer, weil und soweit diese Disziplinen als typische Bestandteile der ärztlichen Grundversorgung erscheinen (Urteil 9C_570/2015 vom 6. Juni 2016 E. 7.4). Wenn dies zu einem grösseren Anteil kostenintensiver Patienten führen sollte, wäre das aber zu berücksichtigen. Die Kostenrelevanz eines nicht als Praxisbesonderheit anerkannten Schwerpunkts ist allenfalls im Zusammenhang mit der Bestimmung des Umfangs der Toleranzmarge beachtlich (erwähntes Urteil 9C_166/2022 E. 9.2.3).

9.3. Der Beschwerdeführer macht weitere Praxisbesonderheiten in Form von Leistungsangeboten und Spezialisierungen geltend. Seine internistische Arztpraxis decke ein breites Leistungsspektrum ab, was sich in einer besonderen, kostenwirksamen Zusammensetzung der Patienten niederschlage.

9.3.1. Zu Mehrkosten führe einmal das Anfertigen von Röntgenaufnahmen in der eigenen Praxis. Nur gut zwei Drittel der Vergleichspraxen verfügten überhaupt über ein Röntgengerät. Über TARMED-Kapitel 39 (bildgebende Verfahren) seien bei ihm rund zweieinhalbmal mehr Leistungen abgerechnet worden als im Schnitt der Vergleichsgruppe. Ebenfalls durch vergleichsweise hohen Aufwand gekennzeichnet seien Wundversorgungen und Kleinchirurgie, etwa bei Diabetikern; dabei handle es sich um Therapien, die er - anders als viele seiner Berufskollegen - selber durchführe, anstatt die betreffenden Patienten an (wesentlich teurere) Chirurgen zu überweisen.

Sodann weist der Beschwerdeführer auf zwei Spezialisierungen seiner Praxis hin: Ein angestellter Arztkollege erbringe Akupunkturleistungen, die über seine ZSR-Nummer abgerechnet würden. Da diese komplementärmedizinische Behandlungsmethode nicht mit dem Einsatz von Medikamenten verbunden sei, gleiche der Morbiditätsindikator "Pharmazeutische Kostengruppen PCG" den damit verbundenen erhöhten Aufwand nicht aus (vgl. dazu oben E. 4.4.1; BGE 150 V 129 E. 5.5.3 und 6.5). Ausserdem verfüge er über einen Schwerpunkt "Manuelle Medizin". Da er diese Behandlungen selber durchführe, erübrige sich die Konsultation eines Spezialisten resp. verhindere dies kostspielige anderweitige Abklärungen und Behandlungen. Auf sein eigenes Kostenbild wirke sich das aber nachteilig aus. Insgesamt folgten aus den genannten Praxisbesonderheiten jährliche Mehrkosten von Fr. 131'030.-, was einen erheblichen Teil der totalen direkten Arztkosten von Fr. 406'785.- ausmache.

9.3.2. Nachfolgend sind die geltend gemachten Praxisbesonderheiten mit Blick auf die Frage, ob sie vom Leistungsangebot des Referenzkollektivs rechtserheblich abweichen, annäherungsweise einzuordnen; abschliessend zu beurteilen sind sie aufgrund der Feststellungen der Parteien resp. des kantonalen Schiedsgerichts im weiteren Verlauf des Verfahrens (unten E. 11.1).

9.3.2.1. Der Indexwert des geprüften Leistungserbringers muss gegebenenfalls korrigiert werden, wenn der Leistungserbringer aufgrund besonderer fachlicher Qualifikationen (Fähigkeitsausweise) regelmässig Behandlungen anbietet, die sich kostenmässig vom durchschnittlichen Leistungsangebot der Referenzgruppe deutlich abheben (vgl. erwähntes Urteil 9C_166/2022 E. 9.2.3 [zur Publ.

vorgesehen]). Sollte die Praxis des Beschwerdeführers infolge der geltend gemachten Spezialisierung in Wundversorgung und Kleinchirurgie überproportional häufig entsprechend kostspielige Leistungen erbringen, wäre also zu prüfen, ob eine wesentliche Praxisbesonderheit gegeben ist.

9.3.2.2. Zusätzliche Praxisangebote können zudem dazu führen, dass sich Behandlungen beim geprüften Arzt konzentrieren, die ansonsten durch andere Leistungserbringer erbracht und als Fremdleistungen nicht in dessen Kostenstatistik einfließen würden (dazu E. 4.6.1). Dies erhöht die pro Patient erbrachte Leistungsmenge und damit den Praxisfallwert.

Bei der Akupunktur handelt es sich unter Umständen um eine wesentliche Erweiterung des Behandlungsspektrums, wenn sie durch eine ärztliche Fachkraft (vgl. Art. 4b lit. a KLV) unter der ZSR-Nummer des Beschwerdeführers appliziert wird. Würde der Beschwerdeführer die betreffenden Patienten zur komplementärmedizinischen Behandlung an einen anderen ärztlichen Leistungserbringer überweisen, so fielen der kassenpflichtige Aufwand wohl nicht unter die ihm anzurechnenden veranlassten Kosten.

9.3.2.3. Derweil stellt die manualmedizinische Versorgung von akuten und chronischen Funktionsstörungen des Bewegungsapparats zwar ebenfalls eine Erweiterung des Behandlungsspektrums dar. Der Sache nach ist sie aber eher nicht als Praxisbesonderheit zu begreifen, zumal wenn die als "schonende Therapie von Rückenschmerzen oder Bewegungseinschränkungen der Gelenke mittels gezielter Handgriffe" (<https://www.samm.ch/de/weiterbildung/>) beschriebene Behandlung an die Stelle von Physiotherapie tritt, die dem überweisenden ärztlichen Leistungserbringer als veranlasste Leistung anzurechnen wäre. Auch dürfte das Anbieten von Manueller Medizin die Zusammensetzung des Patientenkollektivs kostenstrukturell nicht bedeutsam beeinflussen, anders als es etwa bei einem dezidiert psychosomatisch-schmerzmedizinischen Praxisprofil denkbar ist (oben E. 9.2).

9.3.2.4. Nicht als Praxisbesonderheit eines internistischen Leistungserbringers gelten wesentliche diagnostische oder therapeutische Bestandteile einer im Rahmen der medizinischen Grundversorgung erfolgenden Behandlung, die - erbrächte der kontrollierte Erbringer sie nicht selbst - ausnahmsweise auch als veranlasste ärztliche Leistungen in die Wirtschaftlichkeitsprüfung einfließen würden (oben E. 4.6.1 a.E.).

Der Umstand, dass der Beschwerdeführer Röntgendiagnostik in Eigenleistung erbringt und entsprechende direkte Kosten anfallen, verhält sich hinsichtlich des Gesamtkostenindex grundsätzlich neutral, weil im Gegenzug die Bilanz der veranlassten Kosten entlastet wird. Lagerte der Beschwerdeführer die diagnostische Leistung an radiologische Institute aus, so müssten die betreffenden Kosten Teil seines Gesamtkostenindex bleiben. Insoweit könnte radiologische Diagnostik höchstens dann als Praxisbesonderheit gelten, wenn dargetan wäre, dass die entsprechende apparative Ausstattung das Leistungsspektrum der Praxis entscheidend prägt (EUGSTER, Wirtschaftlichkeitskontrolle, Rz. 420); ebenso, wenn im Vergleich mit einer externen Versorgung unvermeidbare Mehrkosten entstünden, die bei der Mehrzahl der Vergleichspraxen nicht zu Buche schlagen.

10.

Der Beschwerdeführer rügt schliesslich, das kantonale Schiedsgericht habe den Betrag der Rückforderung bundesrechtswidrig berechnet.

10.1. Im angefochtenen Urteil schickt die Vorinstanz voraus, die Frage der Wirtschaftlichkeit sei anhand einer Gesamtbetrachtung der direkten und veranlassten Kosten (Gesamtkostenindex) zu beurteilen. Im Jahr 2019 habe der beklagte Leistungserbringer 158 Indexpunkte erreicht. Der Toleranzwert von 120 Indexpunkten sei deutlich überschritten, weshalb er - hinsichtlich der direkten

Kosten - grundsätzlich rückerstattungspflichtig werde. Die Regressionsanalyse weise Indizes betreffend direkte Arztkosten, direkte und veranlasste Medikamenten-, Labor-, Mittel- und Gegenstände- und Physiotherapiekosten sowie direkte und veranlasste Kosten aus. Es fehle demnach ein Index, der (ergänzend zum Index direkte Arztkosten) nur die direkten Medikamentenkosten etc. erfasse. Bezüglich dieses Teils der direkten Kosten könne keine Rückforderung erfolgen, weil bei Verwendung der vorhandenen Indizes auch veranlasste Kosten in die Rückforderung einfließen würden, was nicht zulässig sei. Die Rückerstattung sei damit allein anhand des Regressionsindex für die direkten Arztkosten zu bemessen. Dieser betrage für das Jahr 2019 195 Punkte, was weit über der Toleranzgrenze von 120 Punkten liege. Ausgehend von direkten Arztkosten von Fr. 406'785.- resultiere ein Rückerstattungsbetrag von Fr. 156'455.75 (Fr. 406'785.- x [75/195]).

10.2.

10.2.1. Die Frage, ob die Behandlungsweise des geprüften Leistungserbringers insgesamt als wirtschaftlich im Sinn von Art. 56 Abs. 1 und Art. 59 Abs. 3 lit. a KVG erscheint, beantwortet sich nach dem individuellen Gesamtkostenindex der direkten und veranlassten Kosten pro erkranktem Patient, der um die definitive Toleranzmarge sowie um den Effekt allfälliger Praxisbesonderheiten vermindert worden ist (E. 4.4.2 und 4.5). Die Gesamtbetrachtung der Kosten erlaubt es, überdurchschnittliche direkte Kosten allenfalls durch unterdurchschnittliche veranlasste Kosten auszugleichen (BGE 137 V 43 E. 2.5.6; dazu und zum Umfang der veranlassten Kosten oben E. 4.6.1). Rückerstattungspflichtig sind hingegen nur die direkten Kosten (BGE 137 V 43 E. 2.5.5; oben E. 4.6.2). Die Vorinstanz hat den Index der direkten Kosten darüber hinaus auch zur Bestimmung der Rückerstattungsquote verwendet.

10.2.2. Indessen wirkt sich die Gesamtbetrachtung der verursachten Kosten notwendigerweise auch bei der Berechnung der Rückerstattung aus. Nach der seit BGE 137 V 43 geltenden Rechtsprechung erfasst die Rückerstattungspflicht zwar ausschliesslich die (totalen) direkten Kosten, bestehend aus den Arztkosten und den Kosten für selbst abgegebene Medikamente sowie für weitere unter der ZSR-Nummer des Leistungserbringers erbrachte Leistungen (namentlich Labor, Mittel und Gegenstände, Physiotherapie; vgl. E. 4.6.1). Aus diesem Grund kommt eine Rückforderung nur infrage, wenn der Index der direkten Kosten pro erkranktem Patient den um den Effekt von Praxisbesonderheiten erweiterten Toleranzwert übertrifft (BGE 150 V 129 E. 4.1 a.E. mit Hinweisen; Urteil 9C_535/2014 vom 15. Januar 2015 E. 5.4). Ansonsten würde es schon am Rückforderungssubstrat fehlen: Rückforderbar sind wie erwähnt allein die direkten Kosten. Das heisst aber nicht, dass sich die Berechnungsformel auch hinsichtlich der Bestimmung der Rückerstattungsquote auf den Index der direkten Kosten stützt. Die Motive für die Änderung der diesbezüglichen Rechtsprechung (BGE 137 V 43 E. 2.5) stellen klar, dass sich die Massgeblichkeit der (totalen) direkten Kosten auf den Gegenstand (das betragliche Substrat) der Rückerstattung bezieht. Sollen unterdurchschnittliche veranlasste Kosten im Sinn der zitierten Rechtsprechung kompensatorisch berücksichtigt werden, so muss bei der Bestimmung der mit unwirtschaftlichem Verhalten übereinstimmenden Rückerstattungsquote auf den Gesamtkostenindex abgestellt werden.

10.3. Zusammengefasst ist ein Verstoß gegen das Gebot der wirtschaftlichen Behandlungsweise gegeben, wenn der Gesamtkostenindex der direkten und veranlassten Kosten pro abgerechnetem Patient bezogen auf die Referenzgrösse nach Abzug der Toleranzmarge und der auf Praxisbesonderheiten entfallenden Quote grösser als 100 ist. Eine Rückforderung kommt infrage, wenn dies zusätzlich auch hinsichtlich des (in gleicher Weise korrigierten) Indexes der direkten Kosten zutrifft. Das Rückforderungssubstrat beschränkt sich auf die totalen direkten Kosten. Wegen der gebotenen integralen Betrachtung der direkten und der veranlassten Kosten erfolgt die Berechnung der Rückerstattungsquote und des rückforderbaren Betrags indessen auf der Grundlage des Gesamtkostenindexes.

Die Berechnungsformel lautet mithin: Totale direkte Kosten multipliziert mit dem Quotienten aus dem durch Toleranzbereich und Quote der Praxisbesonderheiten bereinigten Gesamtkostenindex (minus 100) und dem (unbereinigten) Gesamtkostenindex.

10.4.

10.4.1. Für das Jahr 2019 beträgt der Gesamtkostenindex des Beschwerdeführers 158 Punkte. Der Toleranzbereich und die Quote der Praxisbesonderheiten, um die der Gesamtkostenindex korrigiert wird, sind im weiteren Verfahren noch festzulegen (E. 11.1). Im Hinblick auf die Feststellung, ob auch bezüglich des Rückforderungssubstrats (totale direkte Kosten) allein eine unwirtschaftliche Behandlungsweise feststellbar ist, kommen beim einschlägigen Index die gleichen Abzüge zum Tragen.

10.4.2. Der Beschwerdeführer beanstandet, dass im Regressionsbericht 2019 kein Index bezüglich der totalen direkten Kosten (d.h. einschliesslich der vom Arzt selber abgegebenen Medikamente) aufgeführt ist (vgl. hinsichtlich der Vorgängermethode ANOVA Urteil 9C_558/2018 vom 12. April 2019 E. 8.2.5) und die Vorinstanz infolgedessen für die Berechnung der Rückforderung lediglich den Index betreffend die direkten Arztkosten heranzieht. Damit stelle die Vorinstanz den rechtserheblichen Sachverhalt unvollständig fest (Art. 97 Abs. 1 BGG). Die Rückforderung dürfe den aus unwirtschaftlicher Behandlung erlangten Vorteil nicht überschreiten. Die Anwendung des richtigen Indexes wirke sich erheblich auf eine allfällige Rückerstattung aus. Das vorinstanzliche Vorgehen bestrafe ihn, den Beschwerdeführer, letztlich dafür, dass er tiefe direkte Medikamentenkosten ausweise. Bei der Bemessung der Rückforderung sei daher auf die totalen - d.h. auch selbstdispensierte Medikamente umfassenden - direkten Kosten resp. deren Index abzustellen. Die Vorinstanz übersehe, dass sich ein Index der totalen direkten Kosten (einschliesslich Medikamente) anhand der aufbereiteten Daten durchaus berechnen lasse. Bezeichnenderweise enthalte der Regressionsbericht für das vorangegangene Statistikjahr 2018 einen einschlägigen Index. Dass dies für das Folgejahr 2019 nicht der Fall sei, dürfe ihm nicht zum Nachteil gereichen.

In der Tat wäre es nicht systemkonform, allein auf die direkten Arztkosten abzustellen (so schon das erwähnte Urteil 9C_558/2018 E. 8.2.5); dies zumal beispielsweise eine intensiviertere Therapie mit einem geringeren Medikationsbedarf einhergehen kann und in einem solchen Fall (bei Selbstdispensation) der Index der totalen direkten Kosten tiefer ausfällt als derjenige der Arztkosten allein. Zutreffend weist der Beschwerdeführer darauf hin, dass der Statistikbericht für das Vorjahr (2018) einen Regressionsindex "Totale direkte Kosten" enthält; dieser liegt im Übrigen mit 167 Punkten effektiv tiefer als derjenige für direkte Arztkosten (186 Punkte). Es ist davon auszugehen, dass aus den statistischen Grundlagen auch für das streitgegenständliche Jahr 2019 ein Index der totalen direkten Kosten abgeleitet werden kann.

10.4.3. Sofern sowohl nach dem bereinigten Gesamtkostenindex wie auch nach dem Index der totalen direkten Kosten Unwirtschaftlichkeit ausgewiesen ist, gilt es die allfällige Rückerstattung nach der Formel gemäss E. 10.3 zu berechnen.

11.

11.1. Aus den dargelegten Gründen ist das angefochtene Urteil aufzuheben und die Sache an die Vorinstanz zum weiteren Verfahren (oben E. 6.4.2 und 8-10) zurückzuweisen (Art. 107 Abs. 2 BGG). Das kantonale Schiedsgericht wird den Parteien Gelegenheit zur Vervollständigung der Wirtschaftlichkeitsprüfung geben und anschliessend auf der Grundlage einer ergänzten Klage über die Frage der wirtschaftlichen Behandlungsweise im Statistikjahr 2019 sowie gegebenenfalls über die Rückforderung entscheiden.

11.2. Die Gründe für die teilweise Gutheissung der Beschwerde resp. die Rückweisung der Sache zur neuen Entscheidung (vgl. E. 6 und 7) ergeben sich im Wesentlichen aus BGE 150 V 129 sowie aus dem (teilweise zur amtlichen Publikation bestimmten) Urteil 9C_166/2022 vom 9. Dezember 2024. Die Erwägungen betreffend Handhabung von Praxisbesonderheiten (E. 9) und Berechnung der Rückerstattung (E. 10) konkretisieren ebenfalls bestehende Rechtsprechungen. Der vorliegende Rückweisungsentscheid präjudiziert die Beurteilung im Einzelfall nicht. In der Sache werden sich die Parteien im Verfahren vor Schiedsgericht zu den strittigen Punkten äussern können. Weitere Rügen des Beschwerdeführers sind unbegründet (vgl. E. 8). Daher ist kein Schriftenwechsel durchzuführen (Art. 102 Abs. 1 BGG; so auch Urteile 9C_126/2023 vom 4. März 2024 E. 10 und 9C_38/2023 vom 21. April 2023 E. 7).

11.3. Mit dieser Entscheidung ist das Gesuch um aufschiebende Wirkung der Beschwerde gegenstandslos.

12.

Im Hinblick auf die Kosten- und Entschädigungsfolgen gilt die Rückweisung der Sache zur Ergänzung der Entscheidungsgrundlagen als vollständiges Obsiegen der beschwerdeführenden Partei, unabhängig davon, ob sie die Rückweisung beantragt resp. ob sie das Begehren im Haupt- oder Eventualantrag stellt (BGE 137 V 210 E. 7.1).

Ausgangsgemäss tragen die Beschwerdegegnerinnen die Gerichtskosten (Art. 66 Abs. 1 BGG). Sie bezahlen dem Beschwerdeführer eine Parteientschädigung (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 9. März 2022 wird aufgehoben. Die Sache wird zur neuen Entscheidung an die Vorinstanz zurückgewiesen.

2.

Die Gerichtskosten von Fr. 6'000.- werden den Beschwerdegegnerinnen auferlegt.

3.

Die Beschwerdegegnerinnen entschädigen den Beschwerdeführer für das bundesgerichtliche Verfahren mit Fr. 8'000.-.

4.

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