

Kurzfassung

Fragestellung

Hauptziel der vorliegenden Untersuchung war die Bewertung des Patientennutzens aus einer Knochendichtemessung zur Prävention von osteoporosetypischen Frakturen bei Personen ohne vorbestehende Frakturen und der damit gegebenenfalls verbundenen Folgen. Dieses Ziel impliziert verschiedene Teilziele: Erstens kann die Nutzenbewertung anhand unterschiedlicher Arten von Evidenz vorgenommen werden (Prüfung der Anwendung der Knochendichtemessung als Intervention [Teilziel 1a] oder Prüfung der aus der Anwendung resultierenden Konsequenz [Teilziel 1b]). Zweitens können 2 prinzipiell unterschiedliche Anwendungssituationen unterschieden werden (Messung zur Feststellung einer Osteoporose oder Messung im Rahmen eines Therapiemonitorings [Teilziel 2]). Drittens sind neben der als Goldstandard angesehenen Dual-Energy X-Ray Absorptiometry (DXA) (Messort Hüfte und / oder Lendenwirbelsäule) andere apparative Methoden oder klinische Scores zur Bestimmung einer erniedrigten Knochendichte als mögliche Alternativen zu bewerten (Teilziel 3). Und viertens sollen neben der primären Osteoporose der Frau, für die der Erkenntnisstand vergleichsweise hoch ist, auch die primäre Osteoporose des Mannes und sekundäre Osteoporoseformen betrachtet werden (Teilziel 4).

Explizit von der Bewertung ausgenommen war die Nutzenbewertung der Knochendichtemessung beim Vorliegen bzw. bei einer Anamnese von osteoporosetypischen Frakturen. Eine Knochendichtemessung bei diesen Patienten ist bereits im Leistungskatalog der gesetzlichen Krankenversicherung enthalten.

Methoden

Teilziel 1 (Knochendichtemessung zur Feststellung einer Osteoporose)

Für Teilziel 1a sollten Studien an erwachsenen Personen beiderlei Geschlechts ohne vorbestehende osteoporosetypische Frakturen herangezogen werden. Dabei sollte die Effektivität von Versorgungsstrategien mit Knochendichtemessung (Knochendichtemessung + anschließende befundabhängige Behandlung) mit der Effektivität von Versorgungsstrategien ohne Knochendichtemessung hinsichtlich der Inzidenz von klinisch manifesten Frakturen und ihrer krankheitswertigen Folgen verglichen werden. Hierfür sollten randomisierte klinische Studien (RCT) herangezogen werden. Wären keine RCT auffindbar gewesen, hätten gegebenenfalls eindeutig prospektiv geplante nicht randomisierte kontrollierte (Interventions-)Studien (CCT) mit adäquater Berücksichtigung des Problems einer möglichen Strukturungleichheit (unfairer Vergleich) betrachtet werden sollen.

Für Teilziel 1b wurden Studien mit derselben Population wie für Teilziel 1a eingeschlossen. Bei den Personen sollte die Indikationsstellung zur Behandlung im Wesentlichen auf dem Nachweis einer verminderten Knochendichte beruhen – die Population musste u. a. über einen

erniedrigten Knochendichtemesswert definiert sein. Für den Nutznachweis anhand der aus der Anwendung einer diagnostischen Messmethode resultierenden Konsequenz wurden Studien herangezogen, die einen Wirksamkeitsnachweis kausaler Therapieprinzipien (= auf die Beeinflussung der Knochendichte ausgerichteter Behandlungen) hinsichtlich der Inzidenz von klinisch manifesten Frakturen und ihrer krankheitswertigen Folgen erbringen. Hierfür wurden RCT herangezogen, die als Vergleichsbehandlung Placebo bzw. eine Sham-Behandlung oder keine Behandlung einsetzen.

Für den Einschluss der (Interventions-)Studien war eine Beobachtungsdauer von mindestens 2 Jahren Voraussetzung. Dieser Zeitraum wird von der Leitlinie der European Medicines Agency (EMA) 2006 [1] zur Untersuchung medikamentöser Interventionen bei primärer Osteoporose bei postmenopausalen Frauen als angebracht beschrieben, um Daten hinsichtlich Frakturen und Sicherheit zu erheben.

Teilziel 2 (Therapiemonitoring)

Für den Nutznachweis eines Therapiemonitorings mittels Knochendichtemessung sollte wiederum die bereits beschriebene Population betrachtet werden. Es sollten Studien herangezogen werden, die folgende Strategien verglichen: eine Strategie, in der die Fortführung, die Anpassung oder das Absetzen einer knochenwirksamen Behandlung (vs. Placebo oder keine Behandlung bzw. Sham-Behandlung) von den Befunden des Knochendichtemonitorings abhängig gemacht wird und eine Strategie, in der kein Monitoring durchgeführt wurde oder in der die Befunde keinen Einfluss auf die Indikationsstellung zur Weiterbehandlung hatten. Es sollten sowohl RCT als auch CCT (Voraussetzung vgl. Teilziel 1a) sowie die unter Teilziel 1a und 1b bereits beschriebenen Zielgrößen betrachtet werden.

Auch hier galt die Voraussetzung einer mindestens 2 Jahre langen Studiendauer (vgl. Teilziel 1).

Teilziel 3 (Diagnostische Güte)

Hierzu wurden Studien an erwachsenen Personen mit klinischem Verdacht auf Osteoporose eingeschlossen. Der Frakturstatus zu Studienbeginn war nicht relevant. Der Goldstandard DXA galt als Referenz zur Feststellung einer Osteoporose. Gegen ihre diagnostischen Eigenschaften (Stellung der Diagnosen „Osteopenie“ oder „Osteoporose“) wurden die diagnostischen Eigenschaften der übrigen apparativen Verfahren (QCT, periphere DXA-Verfahren, DXL, DXR und Ultraschallverfahren) und klinischen Scoresysteme verglichen. Hätte in Teilziel 1 für ein anderes Verfahren als den Goldstandard ein Nutzen abgeleitet werden können, wäre auch dieses Verfahren als Referenzmethode eingesetzt worden.

Die Frage nach der diagnostischen Güte von alternativen apparativen Knochendichtemessverfahren und klinischen Scores im Vergleich zur DXA war anhand von Querschnitts-

studien zu beantworten, in denen alle Probanden sowohl mit der zu bewertenden Methode als auch mit der Referenzmethode untersucht wurden (Phase-3-Studien nach [2]).

Teilziel 4 (Zusammenhang zwischen Knochendichte und osteoporosetypischen Frakturen und ihren krankheitswertigen Folgen bei Männern mit Verdacht auf primäre Osteoporose bzw. bei gemischtgeschlechtlichen Populationen mit Verdacht auf sekundäre Osteoporose)

Die Frage in Teilziel 4 nach dem Zusammenhang zwischen Knochendichte und patientenrelevanten Zielgrößen (vgl. hierzu Teilziele 1 und 2) wurde hinsichtlich der primären Osteoporose ausschließlich für Männer untersucht. Hinsichtlich der sekundären Osteoporose wurden sowohl Frauen als auch Männer betrachtet. Primär wurden Ergebnisse aus prospektiven Kohortenstudien und von nested case control studies (= eingebetteten Fall-Kontroll-Studien) herangezogen. Dabei musste die Messung der Knochendichte zu Studienbeginn mit dem Goldstandard DXA erfolgt sein. Falls in Teilziel 1 für eine weitere Messmethode gezeigt worden wäre, dass damit identifizierte Personen von einer Behandlung profitieren können, wären auch solche Studien eingeschlossen worden, die diese Messmethode als Referenz eingesetzt haben.

Eine Beobachtungsdauer von mindestens einem Jahr war Voraussetzung für den Einschluss von Studien. Die EMEA-Leitlinie (siehe oben) empfiehlt diesen Beobachtungszeitraum für Männer mit Osteoporose zur Untersuchung medikamentöser Interventionen, die zur Behandlung postmenopausaler Frauen bereits zugelassen sind.

Für alle zu untersuchenden Fragestellungen wurden Studien herangezogen, die kaukasische Populationen einschlossen.

Außerdem galt für alle Teilziele, dass radiologisch nachgewiesene Wirbelkörperfrakturen ausschließlich deskriptiv dargestellt wurden.

Linked Evidence

Lagen für ein zu prüfendes Verfahren (Indextest) keine Interventionsstudien vor, sollte im vorliegenden Bericht versucht werden, über die Verknüpfung von Ergebnissen aus Therapie und Diagnosestudien mögliche Aussagen zum Nutzen für den Indextest abzuleiten (= „Linked Evidence“). Voraussetzungen dafür waren erstens, dass ergebnissichere Therapiestudien (Teilziel 1b) für mittels der Referenzmethode identifizierte Personen einen Nutzen hinsichtlich patientenrelevanter Zielgrößen zeigen (unter der Annahme, dass testnegative Personen keinen Nutzen haben und der Schaden einer Nichtbehandlung den potenziellen Nutzen der Knochendichtemessung für testpositive Personen nicht überwiegt). Und zweitens, dass diagnostische Querschnittsstudien zeigen, dass der Indextest weitgehend dieselben Personen wie der Referenztest identifiziert (Teilziel 3). Sind in diesem Fall die in den Querschnittsstudien als testpositiv identifizierten Personen mit der Population der Thera-

piestudien vergleichbar, kann davon ausgegangen werden, dass auch die über den Indextest identifizierten Personen höchstwahrscheinlich von einer Behandlung profitieren.

Literaturrecherche

Die systematische Literaturrecherche erfolgte in den Datenbanken MEDLINE, EMBASE und CENTRAL. Aufgrund der umfassenden Fragestellung wurden darüber hinaus je nach Teilziel weitere Datenbanken durchsucht. Erfasst wurde der Zeitraum von 1990 bis Juni 2008. Zusätzlich wurden Literaturverzeichnisse relevanter Sekundärpublikationen (systematische Übersichten, Health Technology Assessment [HTA]-Berichte) durchsucht.

Ergebnisse

Teilziel 1 (Knochendichtemessung zur Feststellung einer Osteoporose)

Mangels Studien zur Untersuchung der Fragestellung in Teilziel 1a kann keine direkte Aussage für den Nutzen oder Schaden einer Versorgungsstrategie mit Knochendichtemessung abgeleitet werden.

Für Teilziel 1b wurden 10 Interventionsstudien identifiziert. Bis auf eine dieser Studien, die ein Nahrungsergänzungsmittel vs. Placebo untersuchte, verglichen die Studien jeweils ein Arzneimittel mit Placebo (Alendronat, Risedronat, Strontium Ranelat, Fluorid, Kalzitinin, Raloxifen). Aus diesen 9 Studien wurden insgesamt 19 732 Frauen und Männer in der vorliegenden Bewertung einbezogen, wobei es deutlich mehr Frauen als Männer waren. In 7 der 9 Studien wurde zur Diagnosestellung der Goldstandard DXA eingesetzt, in jeweils einer Studie die Dual-Photon Absorptiometry (DPA) bzw. die Single-Photon-Absorptiometry (SPA).

- Zusammenfassend liegt ein Hinweis auf einen Nutzen einer Behandlung vor, für Frauen ab 65 Jahren und ohne Vorfrakturen, die eine mit einer zentralen DXA gemessene Knochendichte von $T < -2,5$ aufwiesen. Dies gilt hinsichtlich klinisch manifester Frakturen (ausgenommen Schädel-, Gesichts- und Wirbelkörperfrakturen), Hüftfrakturen und klinisch manifeste Wirbelkörperfrakturen. Mit Ausnahme der klinisch manifesten Wirbelkörperfrakturen zeigte sich zudem eine Wechselwirkung zwischen Knochendichte und Therapieeffekt.
- Für Handgelenkfrakturen liegt weder ein Hinweis auf noch ein Beleg für einen Nutzen einer Behandlung für Frauen mit erniedrigter Knochendichte und ohne Vorfrakturen vor.
- Zu frakturassoziierten Funktionseinschränkungen und Schmerzen kann weder ein Hinweis auf noch ein Beleg für einen Nutzen einer Behandlung abgeleitet werden, da für das für den vorliegenden Bericht zu untersuchende Kollektiv keine aussagekräftigen Daten gefunden wurden.

- Zur frakturassoziierten Mortalität und zur gesundheitsbezogenen Lebensqualität wurden in den Studien keine Daten für das für den vorliegenden Bericht zu untersuchende Kollektiv beschrieben.
- Bei der Bewertung der unerwünschten Ereignisse zeigte sich zum Nachteil von Raloxifen ein Hinweis darauf, dass thromboembolische Ereignisse häufiger auftreten.
- Es liegt ein Hinweis darauf vor, dass die Einnahme von Strontium Ranelat häufiger zu Übelkeit, Durchfall, Kopfschmerzen sowie Hautentzündungen oder Ekzemen als Placebo führt. Es liegt auch ein Hinweis darauf vor, dass Frauen unter Strontium Ranelat häufiger die Therapie aufgrund unerwünschter Ereignisse abbrechen als unter Placebo.

Neben dem Goldstandard zeigte sich für kein anderes Verfahren, dass hierüber identifizierte Personen von einer auf die Knochendichte ausgerichteten Behandlung profitieren könnten.

Teilziel 2 (Therapiemonitoring)

Zur Untersuchung der Fragestellung in Teilziel 2 lagen keine Studien vor. Eine Aussage hinsichtlich des Einsatzes einer Knochendichtemessung zur Fortführung, Anpassung oder zum Absetzen einer knochenwirksamen Behandlung ist daher nicht möglich und der Nutzen und Schaden eines Therapiemonitorings bleiben unklar.

Teilziel 3 (Diagnostische Güte)

In Teilziel 1a oder Teilziel 1b wurde kein anderes Verfahren als die zentrale DXA identifiziert, für das ein Nutzen hinsichtlich patientenrelevanter Zielgrößen abgeleitet werden konnte. Aus diesem Grund wurden für Teilziel 3 ausschließlich Studien herangezogen, die die zentrale DXA als Referenz einsetzten.

Für die Untersuchung der Testgüte wurden insgesamt 67 Studien identifiziert. Diese Studien untersuchten im Vergleich zum Goldstandard quantitative Ultraschallverfahren, Risikoscores und vereinzelt Verfahren, die im Folgenden unter „sonstige Verfahren“ gefasst wurden (siehe unten). Die Studien wurden im vorliegenden Bericht nach den untersuchten (Index-)Verfahren gruppiert und entsprechend dieser Gruppierung beschrieben. Einige der Studien untersuchten mehr als einen Indextest, sodass im Folgenden Mehrfachnennungen von Studien vorkommen.

In 33 Studien wurden quantitative Ultraschallverfahren, in 28 Studien wurden 19 verschiedene Risikoscores und in 16 Studien „sonstige Verfahren“ untersucht. Unter „sonstige Verfahren“ wurden solche Verfahren subsumiert, für die weniger als 5 Studien pro Indexverfahren identifiziert wurden. Hierzu gehören periphere DXA-Verfahren, die radiografische Absorptiometrie (RA), die Single-Energy X-Ray Absorptiometry (SXA), die Dual-X-Ray Laser Absorptiometry (DXL), die Digital X-Ray Radiogrammetry (DXR) sowie Kombinationen von apparativen Verfahren und klinischem Score (Ultraschall + Simple

Calculated of Osteoporosis Risk Estimation [SCORE], RA des Kiefers + Osteoporosis Index of Risk [OSIRIS]). Die Studien waren für jede der Indexgruppen sehr heterogen. Diese Heterogenität lag jeweils zumeist an unterschiedlichen Studienpopulationen, unterschiedlich angelegten Selektionskriterien, der unterschiedlichen Ausführlichkeit der Beschreibung von Basisdaten – insbesondere von Knochendichteausgangswerten, der Verwendung verschiedener Referenzkollektive zur Ableitung von T-Scores und auch am Einsatz unterschiedlicher Messorte bzw. -parameter und Trennwerte sowohl für den Index- als auch für den Referenztest. Für die meisten Studien wurde ein hohes Verzerrungspotenzial der Ergebnisse abgeleitet. Keines der untersuchten Indexverfahren erfüllte die für den vorliegenden Bericht definierte Mindestanforderung an einen Indextest. Diese bezog auf die unteren Grenzen der einseitigen Konfidenzintervalle von Sensitivität und Spezifität, die jeweils einen Wert von mindestens 85 % haben mussten. Da die Mindestanforderung nicht erfüllt ist, kann nicht davon ausgegangen werden, dass eines der untersuchten Indextestverfahren ausreichend gut in der Lage ist, dieselben Personen wie der Goldstandard DXA zu identifizieren.

Teilziel 4 (Zusammenhang zwischen Knochendichte und osteoporosetypischer Frakturen und ihren krankheitswertigen Folgen)

Auch für Teilziel 4 wurden ausschließlich Studien herangezogen, die den Goldstandard DXA als Referenz einsetzten (vgl. oben Teilziel 3).

Insgesamt wurden 11 Studien zur Untersuchung des Zusammenhangs zwischen Knochendichte und Frakturen identifiziert. Sieben dieser Studien untersuchten patientenrelevante Zielgrößen und 4 Studien ausschließlich radiologisch nachweisbare Wirbelkörperfrakturen. Drei der Studien zu patientenrelevanten Zielgrößen untersuchten Männer mit Verdacht auf primäre Osteoporose untersuchten und 4 Studien bezogen sich auf gemischtgeschlechtliche Populationen mit Verdacht auf sekundäre Osteoporose.

Drei Studien, die Männer mit Verdacht auf primäre Osteoporose einschlossen, untersuchten den Zusammenhang zwischen Knochendichte und klinisch manifesten Frakturen sowie Mortalität. In keiner der Studien erfolgte eine Verblindung der Männer gegenüber ihrem Testergebnis. Die 3 Studien untersuchten zwischen 709 und 5995 selbstständig lebende Männer (mittleres Alter zwischen ca. 68 und 74 Jahren). Eine weitere Studie dokumentierte ausschließlich Daten zum Risiko von radiologisch nachgewiesenen Wirbelkörperfrakturen. Zusammenfassend zeigten sich folgende Ergebnisse für die patienterlevanten Zielgrößen:

- Es liegt ein Beleg für einen statistischen Zusammenhang zwischen erniedrigter Knochendichte und einem erhöhten Hüftfrakturrisiko vor, wenn die Knochendichte per Goldstandard DXA am Schenkelhals erhoben wurde. Ein Hinweis auf einen solchen Zusammenhang liegt vor, wenn die Knochendichte an der Lendenwirbelsäule oder Gesamthüfte erhoben wurde.

- Es liegt ein Hinweis auf einen statistischen Zusammenhang zwischen niedriger Knochendichte und einem erhöhten Risiko für klinisch manifeste Frakturen vor, wenn die Knochendichte per DXA an einem der zentralen Messorte gemessen wurde.

Vier Studien schlossen Frauen und Männer nach Lungen- bzw. Herztransplantation sowie Hämodialysepatientinnen und -patienten ein und untersuchten patientenrelevante Zielgrößen. Die Fallzahlen waren recht klein und lagen zwischen 14 und 106 Frauen und Männer. Alle Studien wiesen ein hohes Verzerrungspotenzial der Ergebnisse auf. Drei weitere Studien untersuchten lediglich den Zusammenhang zwischen Knochendichte und radiologisch nachgewiesenen Wirbelkörperfrakturen.

Für Risikogruppen für eine sekundäre Osteoporose liegt weder ein Beleg für noch ein Hinweis auf einen statistischen Zusammenhang zwischen niedriger Knochendichte und einem erhöhten Risiko für klinisch manifeste Frakturen bzw. Gesamtmortalität vor.

Weder für Männer mit Verdacht auf primäre Osteoporose noch für Personen mit Verdacht auf sekundäre Osteoporose fanden sich in den eingeschlossenen Studien Daten für die Endpunkte frakturassoziierte Funktionseinschränkungen, frakturassoziierte Schmerzen, frakturassoziierte Mortalität sowie Veränderungen der gesundheitsbezogenen Lebensqualität bei und nach Frakturen.

Linked Evidence

Aus den Therapiestudien konnte für Personen ohne Vorfrakturen zwar ein Hinweis auf einen Nutzen durch eine auf die Knochendichte ausgerichtete medikamentöse Behandlung abgeleitet werden, allerdings zeigte sich für keinen Indextest, dass er ausreichend gut ist, um dieselben Patienten wie der Goldstandard DXA zu identifizieren. Eine Verknüpfung der Ergebnisse aus Therapie- und diagnostischen Querschnittsstudien und eine Ableitung einer Aussage für eines der zu untersuchenden (Index-)Verfahren war daher nicht möglich.

Fazit

Die Kernfrage des vorliegenden Berichts war, ob Frauen oder Männer mit primärer oder sekundärer Osteoporose, aber ohne vorbestehende osteoporosetypische Frakturen von einer Versorgungsstrategie mit Diagnostik einer erniedrigten Knochendichte bzw. eines erhöhten Frakturrisikos und Behandlung im Vergleich zu einer Versorgungsstrategie ohne Knochendichtemessung bzw. Erhebung des Frakturrisikos profitieren. Mangels Studien kann keine direkte Aussage zum Nutzen oder zum Schaden einer Versorgungsstrategie mit Knochendichtemessung abgeleitet werden.

Es liegt jedoch ein Hinweis auf einen Nutzen der Knochendichtemessung für Frauen ab 65 Jahren ohne vorbestehende osteoporosetypische Frakturen vor, bei denen die Knochendichte mittels zentraler DXA-Messung bestimmt wird. Dieser Hinweis leitet sich aus dem Nachweis

einer Wechselwirkung zwischen Knochendichte und Therapieeffekten ab. Dieser Nachweis beruht maßgeblich auf den Ergebnissen einer Studie.

Für einzelne Therapiemöglichkeiten zur Behandlung einer Osteoporose ohne Frakturmanifestation bestehen Hinweise auf einen Schaden durch unerwünschte Ereignisse. Die Ergebnisse zu unerwünschten Ereignissen beruhen jedoch auf Auswertungen, in die auch Frauen mit Vorfrakturen eingeschlossen wurden.

Der Nutzen sowie der Schaden eines Therapiemonitorings mithilfe einer Knochendichtemessung bleiben mangels Studien unklar und sind somit nicht belegt.

Es fand sich keine gleichwertige Alternative zur Erhebung der Knochendichte bzw. des Frakturrisikos, die als Ersatz zur zentralen DXA-Messung eingesetzt werden kann.

Für Männer mit Verdacht auf eine primäre Osteoporose existiert ein Beleg für einen statistischen Zusammenhang zwischen niedriger Knochendichte und erhöhtem Frakturrisiko. Für Risikogruppen für sekundäre Osteoporoseformen liegt weder ein Beleg für noch ein Hinweis auf einen solchen Zusammenhang vor. Daraus kann nicht geschlossen werden, dass solch ein Zusammenhang nicht besteht.

Schlagwörter: Osteoporose, Osteodensitometrie, DXA, QCT, QUS, Frakturen, systematische Übersicht, Nutzenbewertung