

## **Kurzfassung**

### **Hintergrund**

Der Gemeinsame Bundesausschuss (G-BA) hat mit Schreiben vom 21.12.2006 das Institut für Qualität und Wirtschaftlichkeit im Gesundheitswesen (IQWiG) mit der Recherche, Darstellung und Bewertung des aktuellen medizinischen Wissensstandes zur Positronenemissionstomographie (PET), auch als Integrationsgerät von PET und Computertomographie (PET/CT), bei 14 verschiedenen Krankheitsbildern beauftragt. Der vorliegende Vorbericht bezieht sich auf den Teilauftrag zu Ovarialkarzinomen.

### **Fragestellung**

Der vorliegende Bericht verfolgte 2 Ziele:

- 1) Ermittlung des patientenrelevanten Nutzens der PET bzw. PET/CT

Das primäre Ziel des Berichts war die Beschreibung des patientenrelevanten Nutzens, den Ärzte und Patienten bei der Primärdiagnostik, beim Primärstaging, beim Restaging und bei der Rezidivdiagnostik von Ovarialkarzinomen von dem bildgebenden Verfahren PET bzw. PET/CT erwarten können. Unter „Nutzen“ wurden die Veränderungen verstanden, die kausal auf den Einsatz der PET bzw. PET/CT zurückzuführen sind und für den Patienten fassbare Konsequenzen haben.

- 2) Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT.

Sollten zu wenige aussagekräftige Studien zur Ermittlung des patientenrelevanten Nutzens identifiziert werden (Ziel 1), so sollte zusätzlich eine systematische Bewertung der diagnostischen und prognostischen Güte der PET bzw. PET/CT vorgenommen werden (Ziel 2). Dabei sollte geprüft werden, inwieweit die PET bzw. PET/CT den diagnostischen Standardverfahren ohne PET überlegen ist. Das bedeutet: Verbessert sich durch den Einsatz der PET bzw. PET/CT die Primärdiagnostik, das Primärstaging, das Restaging oder der korrekte Ausschluss von Rezidiven? Ebenso sollte geprüft werden, ob sich mittels PET und PET/CT zuverlässigere prognostische Aussagen im Rahmen der eben genannten Indikationen treffen lassen, als es mit den bisherigen diagnostischen Standardverfahren möglich ist.

### **Methoden**

Für die Nutzenbewertung sollten im Rahmen einer systematischen Übersicht (randomisierte) kontrollierte Studien (RCTs) – z. B. Strategie mit vs. ohne PET – mit patientenrelevanten Endpunkten (z. B. verringerte Mortalität / Morbidität) berücksichtigt werden.

Für die Beurteilung der diagnostischen und prognostischen Güte sollte ein „Review of Reviews“, also eine Bewertung auf der Basis publizierter Evidenzsynthesen, dienen. Für den Zeitraum und Fragestellungen, die die aktuellste Evidenzsynthese mit ihrer Literatursuche nicht abdeckt, sollte die relevante Primärliteratur (prospektive Kohorten- und Querschnittsstudien) durch eigene Ergänzungsrecherchen zusätzlich identifiziert werden.

Im Rahmen der Ergänzungsrecherche wurde eine systematische Literaturrecherche nach RCTs und nach Studien zur diagnostischen und prognostischen Güte in den folgenden Datenbanken durchgeführt: EMBASE, MEDLINE und Cochrane Central Register of Controlled Trials (Clinical Trials). Zur Ermittlung von Evidenzsynthesen wurde zusätzlich in folgenden Datenbanken recherchiert: Cochrane Database of Systematic Reviews (Cochrane Reviews), Database of Abstracts of Reviews of Effects (Other Reviews) und Health Technology Assessment Database (Technology Assessments). Die Literaturrecherche umfasste den Zeitraum bis zum 16.02.2010. Eine Aktualisierung der Recherche ist für den Zeitraum nach Publikation des Vorberichts geplant. Darüber hinaus wurden die vom G-BA übermittelten Unterlagen, öffentlich zugängliche Studienregister, die im Rahmen der Anhörung zum vorläufigen Berichtsplan übermittelten Unterlagen sowie die Datenbanken von Leitlinienerstellern durchsucht. Daneben wurde in Kongressbänden recherchiert. Die Literaturverzeichnisse potenziell relevanter Evidenzsynthesen wurden ebenfalls durchsucht.

Das Literaturscreening wurde von 2 Reviewern unabhängig durchgeführt. Nach der Bewertung der Studienqualität wurden die Ergebnisse der einzelnen Studien nach Fragestellungen geordnet und beschrieben. Darüber hinaus wurden die Studien, die im Rahmen der Ergänzungsrecherche eingeschlossen wurden, im Hinblick auf ihre Übertragbarkeit auf den deutschen Versorgungskontext bewertet.

## **Ergebnisse**

### ***Patientenrelevanter Nutzen***

Durch die systematische Recherche nach publizierter Literatur konnte keine vergleichende Studie identifiziert werden, die eine Aussage zum patientenrelevanten (Zusatz-)Nutzen der PET bzw. PET/CT im Krankheitsbild der Ovarialkarzinome ermöglicht hätte. Auch die Suche in Kongressbänden und Studienregistern lieferte keine Hinweise auf noch laufende vergleichende Studien.

### ***Diagnostische und prognostische Güte***

Die Suche nach diagnostischen und prognostischen Studien erfolgte zweischrittig. Zunächst wurde systematisch nach methodisch hochwertigen Evidenzsynthesen gesucht, die vorhandene Studien in den jeweiligen Indikationen zusammenfassten. Diese Suche wurde durch eine weitere Recherche ergänzt, um Primärstudien zu identifizieren, die nach dem Einschlusszeitraum der eingeschlossenen Evidenzsynthesen publiziert wurden (Ergänzungsrecherche).

6 Evidenzsynthesen und 7 Primärstudien entsprachen den Einschlusskriterien des vorliegenden Berichts. Die Evidenzsynthesen enthielten insgesamt 31 Primärstudien, sodass dieser Bericht auf insgesamt 38 Studien basiert (Primärdiagnostik [n = 4], Primärstaging [n = 2], Restaging [n = 4], Rezidiverkennung [n = 28], indikationsübergreifend [n = 2]; eine Studie liefert Daten zu 2 Indikationen).

### *Primärdiagnostik*

Für die Indikation der Primärdiagnostik von Ovarialkarzinomen konnten 2 Primärstudien aus den Evidenzsynthesen und weitere 2 aus der Ergänzungsrecherche identifiziert werden. Der bivariat gepoolte Schätzer für die Sensitivität der FDG-PET/CT (**3 Studien**) lag bei 92,6 %, jener für die Spezifität bei 89,0 %. Als Referenztest dienten jeweils die Operationsergebnisse und die Histologie. Die Sensitivität der (nicht-integrierten) PET betrug in der Studie von **Kawahara 2004** 78,3 %, die Spezifität 86,7 %. Die Konfidenzintervalle dieser Schätzer waren aufgrund der geringen Fallzahlen relativ weit. Nur 1 der 4 Studien (**Nam 2010**) stellte Vergleiche zwischen der PET/CT, der Kombination aus CT und MRT sowie dem Doppler-Ultraschall an. Die PET/CT erzielte dabei eine Sensitivität von 91,7 % bei einer Spezifität von 92,0 %. Die Ergebnisse zur Kombination aus CT und MRT können aufgrund der hohen Anzahl fehlender Werte nicht interpretiert werden, weshalb keine Bewertung und Darstellung erfolgt. Der Doppler-Ultraschall hatte eine geringere Sensitivität (76,4 %) aber eine höhere Spezifität (94,4 %) als die PET/CT.

### *Primärstaging*

2 Evidenzsynthesen enthielten für die Indikation Primärstaging insgesamt 1 Primärstudie. Diese entsprach allerdings nicht den Einschlusskriterien des vorliegenden Berichts, weil sie nur läsionsbasierte Angaben enthielt. Die Ergänzungsrecherche identifizierte 2 Primärstudien, die die diagnostische Güte der PET/CT beim Primärstaging untersuchten.

In der Studie von **Kitajima 2008b** wurden durch die PET/CT 34 von 40 Patientinnen in die korrekten Stadien eingeteilt (85 %). 3 Patientinnen (7,5 %) wurden aufgrund der PET zu hoch und 3 zu niedrig eingestuft. Im Vergleich dazu wurden durch die CT 27 (67,5 %) Patientinnen richtig, 5 (12,5 %) zu hoch und 8 (20,0 %) zu niedrig eingestuft.

Die Studie von **Nam 2010** untersuchte neben der Primärdiagnostik auch die diagnostische Güte der PET/CT beim primären Staging. Bei 71 von 91 Patientinnen (78,0 %) wurde durch die PET/CT das korrekte Stadium ermittelt. In 14 Fällen (15,4 %) wurde ein zu hohes Stadium, in 6 Fällen (6,6 %) ein zu niedriges Stadium diagnostiziert. Ein Vergleich zu anderen diagnostischen Verfahren fand in dieser Studie in der Indikation des Primärstagings nicht statt.

### *Restaging*

Aus den 3 Evidenzsynthesen entsprachen 2 Primärstudien den Einschlusskriterien des vorliegenden Berichts und der Indikation des Restaging. Die nicht integrierte PET wies in der Studie von **Piccio 2003** eine Sensitivität von 69,2 % und eine Spezifität von 91,7 % auf. Die Sensitivität der PET/CT erreichte in der Studie von **Sironi 2004** einen Wert von 52,9 %, die Spezifität 85,7 %. Auch hier dienten Operationsergebnisse und Histologie als Referenztest. Aufgrund der geringen Fallzahlen in beiden Studien waren die Präzisionen aller Schätzungen sehr gering und die Konfidenzintervalle entsprechend weit.

Die in diesen Studien angestellten Vergleiche zwischen der PET bzw. PET/CT (integriert oder Seite an Seite), der CT und dem CA 125 zeigten keine eindeutigen Unterschiede hinsichtlich der diagnostischen Güte.

In der Ergänzungsrecherche wurden weitere 2 Primärstudien identifiziert. Beide Studien waren als Prognosestudien angelegt, beide wiesen ein hohes Verzerrungspotenzial auf.

Anhand der Angaben der Studie von **Kurosaki 2006** („prognostic accuracy study“; Endpunkt 2-Jahres-Überleben) konnte die Sensitivität berechnet werden und erreichte einen Wert von 45,5 %. Die Spezifität lag bei 85,7 %. In der Prognosestudie von **Risum 2010** wurde die PET/CT eingesetzt, um das klinische Stadium IV vom Stadium III zu unterscheiden. In einem multivariaten Modell, welches neben dem Ergebnis der PET/CT (Stadium IV oder niedriger), die komplette chirurgische Tumorentfernung sowie eine Bewertung im Leistungsindex der Gynecologic Oncology Group (GOG) einschloss ( $>2$ ), war die prognostische Aussagekraft der PET/CT nicht statistisch signifikant (keine Angabe zur Effektgröße).

In den beiden Studien fanden sich weder Vergleiche der PET bzw. PET/CT mit anderen diagnostischen Verfahren noch Ergebnisse zu Subgruppen oder anderen Effektmodifikatoren.

#### *Rezidiverkennung*

Aus den 6 Evidenzsynthesen entsprachen 25 Primärstudien den Einschlusskriterien des vorliegenden Berichts in der Indikation der Rezidiverkennung. Darüber hinaus wurden über die Ergänzungsrecherche 2 weitere Primärstudien identifiziert. Von den 27 identifizierten Primärstudien wiesen 7 ein niedriges Verzerrungspotenzial auf.

Aus 7 Studien, die die PET/CT und 4 Studien, die die PET direkt mit der Kombination aus CT und MRT verglichen, ließ sich eine bessere diagnostische Güte der PET bzw. PET/CT erkennen. In drei Studien wurde die PET mit dem CA 125 direkt verglichen. Hier zeigte sich eine höhere Sensitivität der PET bei vergleichbarer Spezifität. Der direkte Vergleich zwischen PET/CT und CA 125 wurde nur in einer Studie angestellt und lieferte gleichwertige Ergebnisse zur diagnostischen Güte.

In den indirekten bivariaten Vergleichen erzielte die PET die höchste gepoolte Sensitivität (91,1 %) bei gleichzeitig hoher Spezifität (94,2 %). Die höchste gepoolte Spezifität erreichte das CA 125 (95,0 %) bei einer Sensitivität von 74,7 %. Die PET/CT erzielte geringfügig niedrigere Werte, als die nicht integrierte PET (Se = 90,3 %; Sp = 89,8 %). Die Kombination aus CT und MRT lieferte eine gepoolte Sensitivität von 66,1 % und eine Spezifität von 83,3 %.

Die indirekten bivariaten Analysen konnten keine Unterschiede in der diagnostischen Güte von PET und PET/CT finden. Im Vergleich zur CT in Kombination mit der MRT fiel die diagnostische Güte sowohl der PET als auch der PET/CT besser aus (p-Wert  $\leq 10$  %). Auf Grundlage des indirekten Vergleichs der PET mit dem CA 125 kann eine höhere

diagnostische Güte der PET weder belegt noch ausgeschlossen werden. Hier überlappten sich die Konfidenzregionen nur geringfügig. Der indirekte bivariate Vergleich zwischen der PET/CT und dem CA 125 konnte keinen Unterschied in der diagnostischen Güte zeigen.

Die univariaten Analysen deuten an, dass in Subgruppen von Patientinnen, in denen bereits ein Verdacht auf ein Rezidiv besteht, die Sensitivität der PET bzw. PET/CT höher sein könnte als in gemischten Kollektiven.

Die Robustheit dieser gepoolten Ergebnisse wurde in mehreren Sensitivitätsanalysen bestätigt.

### **Fazit**

Der Nutzen der PET bzw. PET/CT bei Ovarialkarzinomen ist nicht belegt.

Zu der 2. Fragestellung des Berichts, der diagnostischen und prognostischen Güte, konnten aus 6 eingeschlossenen Evidenzsynthesen und der Ergänzungsrecherche insgesamt 38 Primärstudien eingeschlossen werden.

Zu den Indikationen Primärdiagnostik, Primärstaging und Restaging liegen bisher zu wenige hochwertige Studien vor, um eine verlässliche Aussage zur diagnostischen und prognostischen Güte der PET oder PET/CT treffen zu können. Unklar bleibt insbesondere der Vorteil dieses metabolischen Verfahrens gegenüber den morphologischen bildgebenden Verfahren.

Aus 11 Studien, die die PET bzw. PET/CT direkt mit einer Kombination aus CT und MRT bei der Rezidiverkennung verglichen, zeigte sich eine bessere diagnostische Güte der PET bzw. PET/CT. Die meisten dieser Studien wurden bei Patientinnen mit Verdacht auf ein Rezidiv (steigender Tumormarker oder positive bzw. unklare konventionelle Bildgebung) durchgeführt. Zur Rezidiverkennung bei Patientinnen ohne Verdacht ist die Testgüte der PET bzw. PET/CT gegenüber der morphologischen Bildgebung bisher nicht ausreichend untersucht. Ob die bessere diagnostische Güte der PET bzw. PET/CT bei Patientinnen mit Verdacht auf ein Rezidiv auch zu einer Verbesserung patientenrelevanter Endpunkte führt, ist bisher nicht untersucht.

Weitere methodisch hochwertige Studien sind dringend erforderlich, um den patientenrelevanten Nutzen oder Schaden der PET bzw. PET/CT (speziell bei der Rezidivdiagnostik) von Ovarialkarzinomen verlässlich bewerten zu können. Bisher konnten weder laufende noch abgeschlossene vergleichende Studien zum patientenrelevanten Nutzen der PET bzw. PET/CT bei Ovarialkarzinomen identifiziert werden.

**Schlagwörter:** Positronenemissionstomographie, Computertomographie, Primärdiagnostik, Staging, Restaging, Rezidiv, systematische Übersicht, Ovarialkarzinom

**Keywords:** Positron-Emission Tomography; Tomography, X-Ray Computed; Diagnosis; Staging; Recurrence; Systematic Review; Ovarian Neoplasms