

Dünnschicht-Zytologie zur Früherkennung des Zervixkarzinoms

Ergebnisbericht – aktualisierte Fassung

Recherche

Datum der Erstrecherche: 24.11.2011

Datum der Aktualisierungsrecherche: 15.12.2014

PICO-Fragestellung:

Population: Frauen ab 20 Jahren (keine weiteren Einschränkungen)

Intervention: „Dünnschicht-Zytologie“, auch genannt „Flüssigkeitszytologie“ oder „flüssigkeitsbasierte Zytologie“; englisch „LBC = Liquid Based Cytology“ zur Früherkennung des Zervixkarzinoms und seinen Vorstufen

Kontrolle (Control): konventionelle Zytologie (PAP-Abstrich; conventional cytology: CC)

Zielgröße (Outcome):

- Sterblichkeit an Zervixkarzinom,
- Inzidenz an invasivem Zervixkarzinom,
- Inzidenz unerwünschter Ereignisse
- Anteil unbrauchbarer Proben

Erläuterung der PICO-Fragestellung: Ein Nutzen beider Tests (konventionelle und Dünnschicht-Zytologie) läge in der Senkung der Inzidenz an invasivem Zervixkarzinom und in der Folge einer Senkung der Sterblichkeit an einem Zervixkarzinom. Nach Möglichkeit sollten auch Aussagen zum Potential von Überdiagnosen in die Bewertung einfließen. Überdiagnosen sind Erkrankungen, die ohne Früherkennung zu Lebzeiten nie entdeckt (diagnostiziert) worden wären und somit keine Beschwerden verursacht hätten. Mit beiden Tests (CC und LBC) sollen vor allem Vorstufen des Zervixkarzinoms erkannt und in der Folge durch weitere Maßnahmen entfernt werden, so dass es erst gar nicht zu einer invasiven Krebserkrankung kommt. Bei der Erkennung früher Krebsformen wäre die Zielerkrankung zwar schon vorhanden, es könnte aber dennoch ein günstiger Einfluss zumindest bei der Zervixkarzinom-Mortalität durch die Intervention erzielt werden.

Die Dünnschicht-Zytologie ist technisch aufwändiger durchzuführen und in der Durchführung teurer.

Kommentar zur Aktualisierung-Recherche:

In der Erstbewertung (Ergebnisbericht 1.0) wurde beschrieben, dass die identifizierten Studien auf die Frage, ob die Dünnschichtzytologie (LBC) der konventionellen Zytologie (CC, PAP-Abstrich) in Bezug auf eine Senkung der Zervixkarzinom-Mortalität bzw. der Inzidenz des invasiven Zervixkarzinoms als Früherkennungsmethode überlegen ist, keine Antworten liefern konnten. Die 3 Reviews (mit zwei betrachteten RCTs, ASCUS-LSIL Triage Study und NETHCON-Studie) und der G-BA Bericht lieferten lediglich orientierende Angaben zur Testperformance mit Aussagen zur Sensitivität und prädiktivem Wert (PPV).

Im Rahmen der Aktualisierungs-Recherche (Suchzeitraum ab April 2011) in den unten definierten Datenbanken wurden 1 Review (Whitlock et al., 2011) und eine Originalarbeit (Klug et al., 2013) identifiziert. Bei dem Review handelt es sich um die Publikation zum AHRQ Bericht (Vesco 2011), der bereits in die Erstbewertung aufgenommen wurde. Das Review liefert daher keine neuen Erkenntnisse. Die Originalarbeit von Klug et al. (2013) liefert lediglich aktuelle Ergebnisse zur Testgüte von LBC vs. CC, trifft allerdings keine Aussagen zu der o.g. Fragestellung hinsichtlich Mortalität und Inzidenz.

Zusammenfassend wurden keine aktuellen Studien identifiziert, aus denen Aussagen zur Bewertung des Nutzens und Schadens der Dünnschicht-Zytologie zur Früherkennung des Zervixkarzinoms abgeleitet werden können. Nur zur orientierenden Bewertung der diagnostischen Güte konnten aus einer aktuellen deutschen Studie Ergänzungen aufgenommen werden.

Suchbegriffe:

deutsch: (Dünnschicht-Zytologie oder Flüssigzytologie) und Zervixkarzinom und (Früherkennung oder Screening)

englisch: (LBC or liquid) based cytology cervical cancer screening

Datenbank	gefundene Dokumente	verwendete Dokumente
IQWiG (Berichte)	9 Treffer: nicht relevant (Fragestellung HPV-Test und Zervixkarzinom)	0
Cochrane (Reviews)	1 Treffer: nicht relevant (kein Vergleich zur CC)	0

G-BA	1 Treffer: in Erstbewertung verwendet	0
AWMF (S2e und S3)	1 Treffer: nicht relevant (HPV-Infektion)	0
NICE (Guidance documents)	5 Treffer: 1x veraltet und ersetzt 3x nicht relevant, andere Themen 1x in Erstbewertung verwendet	0
CRD (DARE und HTA)	9 Treffer: 1x verwendet 8x nicht relevant (andere Publikationssprache bzw. andere Themen)	Whitlock EP et al. 2011 (siehe AHRQ)
AHRQ/USPSTF (Index-Suche)	1 Treffer: Whitlock 2011: Publikation zum AHRQ-Bericht (Vesco 2011), der in der Erstbewertung verwendet wurde	0
PubMed	28 Treffer: 1x verwendet 26x nicht relevant (Fragestellung HPV-Test oder andere Themen) 1x doppelt mit CRD Treffer (Whitlock 2011)	Klug SJ, et al.: A randomized trial comparing conventional cytology to liquid-based cytology and computer assistance. Int J Cancer. 2013 Jun 15;132(12):2849-57

Eingeschlossene Dokumente: Beschreibung, Qualitätsbewertung, Extraktion

Ausgewertete Originalarbeiten

Studie	Studientyp und PICO-Fragestellung	Evidenz
Klug SJ, et al.: A randomized trial comparing conventional cytology to liquid-based cytology and computer assistance. Int J Cancer. 2013 Jun 15;132(12):2849-57	Blockrandomisierte Studie zum Vergleich Dünnschichtzytologie (LBC) vs. konventionelle Zytologie (CC) hinsichtlich Testgüte von histologisch gesicherten Gebärmutterhalsläsionen	<p>Zur Früherkennung eines Zervixkarzinoms und seiner Vorstufen wurden in 20 Gynäkologiepraxen in der Rhein-Saar-Studie 20627 Frauen untersucht. Die zur routinemäßigen Vorsorgeuntersuchung gekommenen Frauen wurden zufällig den Untersuchungen zugeordnet. Die Zuordnung wechselte wöchentlich zwischen den Praxen. Die Ergebnisse der LBC-Tests wurden herkömmlich (manuell) und computerassistent (LBC CAS) ausgewertet.</p> <p>Für unsere Bewertung beziehen wir uns auf den Vergleich konventionelle Zytologie gegen manuelle Dünnschichtzytologie.</p> <p>Evidenz zur Treffsicherheit der Tests</p> <p>In der Gruppe mit konventioneller Zytologie wiesen insgesamt 93 Frauen einen positiven zytologischen Befund auf (mindestens „low-grade intraepitheliale Neoplasie“ (LSIL+)). Im LBC-Arm wiesen insgesamt 260 (2,3%) Frauen einen positiven Befund auf. Frauen, bei denen ein positiver Befund vorlag, wurden weiterführenden Untersuchungen bzw. Eingriffen wie Kolposkopien, Biopsien, Konisationen und Hysterektomien unterzogen. Differenziert betrachtet erhielten 67,7% (n=63) der Frauen im CC-Arm Kolposkopien oder andere weiterführende Untersuchungen im Vergleich zu 54,7% (n=205) im LBC-Arm. Frauen mit einem</p>

		<p>negativen Testergebnis wurden keiner weiteren Untersuchung zugeführt.</p> <p>Die relative Sensitivität (Verhältnis der Erkennungsraten der histologisch bestätigten Befunde (mind. CIN2) zwischen LBC und CC unter Berücksichtigung des Cut-offs der Zytologie (LSIL+) für LBC manuell vs. CC lag für LSIL+ und CIN2+ bei 2,74 [95% CI 1,66-4,53] und für LSIL+ und CIN3+ (mindestens hochgradige Dysplasie oder Carcinoma in situ) bei 2,87 [95% CI 1,55-5,32] (statistisch signifikant). Es fanden sich also mit manueller Dünnschichtzytologie 2,74-mal so viele Fälle mit CIN2+ als mit der konventionellen Zytologie und rund 2,9-mal so viele CIN3+Stufen.</p> <p>Die relative Sensitivität für LBC manuell vs. CC bei HSIL+ (high grade intraepitheliale Läsion) und CIN2+ wird mit 3,22 [95% CI 1,73-5,99] sowie für HSIL+ und CIN3+ mit 3,86 [95% CI 1,97-7,56] angegeben (beides statistisch signifikant).</p> <p>Beim relativen positiv prädiktiven Wert (Verhältnis der PPVs der einzelnen Tests) zeigten sich für alle CIN-Grade keine signifikanten Unterschiede zwischen der konventionellen Zytologie und der manuell durchgeführten LBC.</p> <p>Der Anteil unbrauchbarer Proben wurde mit 0,04% für CC und 0,31% für LBC ohne Assistenz angegeben.</p> <p><u>Evidenz zum Schaden</u> Es wurden keine Aussagen zu möglichen Schäden getroffen.</p> <p>Bei der Interpretation der Ergebnisse ist zu berücksichtigen, dass nur eine orientierende Bewertung der diagnostischen Güte möglich ist. Angaben zur genauen Sensitivität sowie zur Spezifität fehlen, da insbesondere keine Untersuchung der Test-Negativen mit dem Referenzverfahren (Kolposkopie) erfolgte. Auch ein nicht unwesentlicher Anteil der Test-Positiven wurde keiner weiterführenden Untersuchung mittels Referenzverfahren unterzogen. Die entsprechenden Werte wurden für die Analysen imputiert.</p>
--	--	--

Verwendete Reviews aus der Erstbewertung

Verwendete Reviews	Reviewqualität 1. Methodik der Recherche/ Literaturauswahl systematisch? 2. Ende des Suchzeitraums? 3. Ergebnispräsentation ausführlich?	Einschlusskriterien für Studienauswahl Design und ggf. PICO-Erläuterung	Evidenz
<p>Unterausschuss "Prävention" des Gemeinsamen Bundesausschusses: Teilbericht Früherkennung des Zervixkarzinoms, hier: -neue Technologien im Rahmen des Primärscreenings, - Qualitätssicherung 2007</p>	<p>1. ja 2.: Februar 2005 3. ja</p>	<p>7 per Handsuche identifizierte HTA Berichte. Suchstrategie eines dieser Berichte (Karnon et al. 2004) verwendet für anschließenden Zeitraum bis Februar 2005. Einschlusskriterien Karnon, diese wiederum gemäß NICE Kriterien. Frage: Ist die Einführung der Dünnschichtzytologie derzeit geeignet, das Programm zu verbessern? Auswahl der Studien gemäß G-BA-Verfahrensordnung</p>	<p>Studienlage: Bewertete HTA-Berichte (Broadstock 2000, Karnon et al. 2004, McCrory et al. 1999, Noorani et al 2003, Payne et al. 2000, Siebert et al. 2003, MSAC 2002). Die Update-Recherche ergab 142 Dokumente für 1/2002-2/2005; 7 Studien wurden extrahiert (Bergeron et al. 2003; Ferreccio et al. 2003; Harkness et al. 2003; Sass MA. 2004; Schorge et al. 2002; Mattosinho de Castro Ferraz Mda et al. 2004; Coste et al. 2003).</p> <p>Evidenz zur Treffsicherheit der Tests: Karnon et al. 2004 hätten per Metaanalyse eine relative Sensitivitätssteigerung von 12% bezüglich der Entdeckung von Dysplasien \geq LSIL berechnet, Noorani et al 2003 hätten anhand von 11 Studien die Sensitivitätssteigerung durch LBC gegenüber CC mit 11% angegeben. Auch der britische HTA-Bericht von Payne et al. 2000 habe für alle untersuchten 11 Studien mindestens eine gleichwertige oder auch bessere Sensitivität für LBC aufgezeigt. Fast alle HTA Berichte warnten vor Überinterpretation: die ermittelten Sensitivitätsunterschiede zwischen CC und LBC könnten Folge von Verzerrungen durch z.B. mangelhafte Verblindung sein: Meist fehlende Angaben zur Spezifität, da Gold-Standard-Untersuchung bei Test-Negativen häufig nicht erfolgt sei. Verweis auf den systematischen Review (Davey et. al. 2006) mit eingeschlossener Metaanalyse: dortige Aussage: „kaum Studien von hoher Qualität“. Zur Bewertung der LBC gäbe es systematische Übersichtsarbeiten, die jedoch gemäß § 18 Abs. 2 der VerfO nur Studien der Evidenzklasse IIb (z. B. Kohortenstudien) umfassten.</p> <p>Evidenz zum Nutzen: HTA-Berichte und Studien aus der Update-Recherche lieferten keine ausreichende Evidenz dafür, dass die klinische Effektivität von LBC im Vergleich zu CC besser sei. Die Autoren des G-BA-Berichts lehnen die Einführung von LBC anstelle des konventionellen Pap-Tests (CC) zur</p>

			<p>Verbesserung der Testgüte des Primärtest ab.</p> <p>Anteil unbrauchbarer Proben im Vergleich: Nur ein Zitat im Bericht: Harkness et al, 2003: Die Anzahl der unbrauchbaren Befunde war bei der PAP-Gruppe (CC; n = 177) war signifikant höher als beim ThinPrep (LBC; n = 81). Beim ThinPrep wurden deutlich häufiger ASCUS (nicht signifikant), LSIL und HISL (beide signifikant) Befunde berichtet.</p> <p>Evidenz zu potentiellen Schäden durch CC und LBC (beiden Verfahren gemeinsam, Autoren sehen keine Unterschiede): Keine Aussage</p>
<p>Vesco KK, et al. : Screening for Cervical Cancer: A Systematic Evidence Review for the U.S. Preventive Services Task Force. Evidence Synthesis No. 86 ; AHRQ Publication No. 11-05156-EF-1. 2011</p>	<p>1. ja 2. Januar 2000 - September 2010 (update des 2007 erschienenen Reports 3. ja</p>	<p>Auswahl gemäß -USPSTF -NICE (für RCTs) -QUADAS tool for quality assessment of diagnostic accuracy studies. Schwache Studien wurden hiernach eliminiert. 5 Fragestellungen: 1) In welchem Alter soll Screening beginnen? Welche Risikoadaptation soll für das jeweilige Alter erfolgen? 2) Verbessert LBC die Sensitivität /Spezifität /diagnostische Trennschärfe und mindert den Anteil unbrauchbarer Proben gegenüber CC? 3) Schäden durch LBC versus CC</p> <p>2 weitere Fragen: zu HPV-Tests, hier nicht relevant</p>	<p>Studienlage: 4.262 Abstracts und 641 Artikel zu 35 Studien wurden für alle 5 Fragestellungen ausgewertet. Zu keiner der hier referierten 3 Fragestellungen wurden RCTs identifiziert mit den Endpunkten „Inzidenz invasiver Zervixkarzinome“ oder „Zervixkarzinom-Mortalität“. Als Gold-Standard wurde die Kolposkopie und/oder Biopsie bei allen Test-Positiven gefordert und mindestens bei einer Stichprobe der Test-Negativen.</p> <p><u>Evidenz zur Treffsicherheit der Tests:</u> Zu Frage 1 (Alter) wurden keine die beiden Verfahren (CC, LBC) differenzierenden Aussagen gemacht. Zu Frage 2 (LBC besser als CC?) wurde 1 RCT mit 88.988 Frauen („NETHCON trial“ in 246 niederländischen Allgemeinpraxen, Frauen im Alter 30-60 J; Siebers et al 2009), eine Kohortenstudie und eine Querschnittsstudie identifiziert. Beide Beobachtungsstudien hatten zusammen n=7.404 Teilnehmerinnen. NETHCON fand keinen Unterschied für die Schwelle CIN2+ (cervical intraepithelial neoplasia: mindestens mäßiggradige Vorstufen eines Gebärmutterhalskrebses; adjustiert für Alter, Ort, Stadt/Land und Studienzeitraum: OR:1.00 [95% CI, 0.84 - 1.20]) oder für die Schwelle CIN3+ (mindestens hochgradige Dysplasie oder Carcinoma in situ, adjustiert s.o. OR: 1.05 [95% CI, 0.86 - 1.29]). Für die Schwelle “CIN3+” verfehlte der relative PPV für LBC knapp die statistische Signifikanz, vergleichen mit CC (PPV für ASC-US+ (atypische Plattenepithelzellen unklarer Bedeutung), 1.17 [95% CI, 0.99 - 1.39]; der PPV für LSIL+ war dagegen knapp signifikant: 1.17 [95% CI, 1.01 - 1.36]). Der relative falsch-positive Anteil (RFPP) für LBC betrug 0.90 (95% CI, 0.82 - 0.99) für die Schwelle von CIN2+ und 0.89 (95% CI, 0.82 - 0.98) bei einer Schwelle von CIN3+, verglichen mit CC.</p>

			<p>Die Studie von Ronco et al 2007 („NTCC trial“) wird zwar als RCT beschrieben, dieses Design wurde aber nur im Vergleich zu LBC mit kombiniertem HPV-Test versus CC gewählt, so dass diese Studie für uns unberücksichtigt bleibt. Auch waren die per Kolposkopie evaluierenden Ärzte nicht blind für den primär durchgeführten Test. In den beiden Beobachtungsstudien aus Südafrika (Taylor et al 2006; n=5.647, keine verblindete Zuordnung/Evaluation) und Frankreich (Coste et al 2003; n=1.757, split-sample Design) konnten absolute Testparameter vergleichend gemessen werden, da alle Frauen einer Goldstandard-Untersuchung unterzogen worden waren. Beim <i>split-sample</i>-Design werden beide Tests vom selben Untersucher nacheinander durchgeführt, die Reihenfolge der abgenommenen Tests kann zu Verzerrungen des Ergebnisses führen. Für LBC und CC gab es in beiden Beobachtungsstudien keine signifikanten Unterschiede bzgl. Sensitivität, Spezifität, falsch positiv Anteilen oder des PPV hinsichtlich der Schwellen CIN2+ und CIN3+, allerdings waren die Konfidenzintervalle breit (zu kleine Fallzahlen).</p> <p>Anteil unbrauchbarer Proben im Vergleich: Der Anteil der unbrauchbaren Proben wurde für LBC niedriger als bei CC angegeben (LBC: NETHCON: 0,37%, NTCC: 2,6%; für CC: NETHCON: 1,09% und NTCC:4,1%).</p> <p>Evidenz zu potentiellen Schäden durch CC und LBC (beiden Verfahren gemeinsam, Autoren sehen keine Unterschiede):</p> <ul style="list-style-type: none"> - durch die Entnahme der Probe - durch unnötige Abklärung falsch-positiver Proben - durch unnötige psychologischer Belastung durch falsch-positiven Befund - ökonomische Konsequenzen im Kontext von Wiedereinbestellungen bei unbrauchbaren Proben. <p>Fazit der Autoren: Ihre Ergebnisse seien im Einklang mit 2 vorausgegangenen Systematischen Evidenz Reviews mit liberaleren Einschlusskriterien: Arbyn et al 2008, Davey et al 2006. Jene hatten die RCTs (NETHCON, NTCC) noch nicht berücksichtigen können, es zeige sich, dass LBC wohl weniger unbrauchbare Proben liefere als CC. Schiffman et al 2009 führten die eher ökonomischen und logistischen Gründe an, die in den USA zu einer Bevorzugung der LBC geführt haben.</p>
Karnon J, et al. : Liquid-	1. ja	Frage: Nutzen und Kosten-Effektivität der	Studienlage:

<p>based cytology in cervical screening: an updated rapid and systematic review and economic analysis. Health Technol Assess, 2004; 8 (20)</p>	<p>2. 1999–2002 (Update von Payne et al. Health Technol Assess 2000;4(18):1–73) 3. ja</p>	<p>LBC für Zervix-Screening im Vergleich zu CC.</p>	<p>4 HTA Reports, 784 Primär-Studien, 15 ausgewählt (auch aus dem Zeitraum vor 1999), 5 Studien in Metaanalyse eingeschlossen. Von den 15 Studien bezogen sich 10 Studien auf Hoch-Risiko- Populationen, z.B. teilweise mit Frauen mit gesichertem Karzinom. Keine Studien zum Endpunkt: Inzidenz/Mortalität Zervixkarzinom.</p> <p>Evidenz zur Treffsicherheit der Tests: Die gepoolte Sensitivität für CC wurde mit 71.5% und mit 80.1% für LBC angegeben, damit sei die LBC um 12% besser als CC, wenn die Schwelle für einen positiven Befund als „low-grade squamous intraepithelial lesion (LSIL)“ definiert wird. In den Niedrig-Risiko- Populationen (eher dem Screening Setting entsprechend) war der Vorsprung für LBC noch größer. Die meisten Studien basierten auf dem Produkt „ThinPrep“, nur wenige arbeiteten mit „SurePath“. Für 6 Studien konnten Spezifitätsunterschiede per Metaanalyse untersucht werden, es zeigte sich kein Unterschied LBC/CC. Methodisch wurde kritisiert, dass die meisten Studien Split-Sample Studien seien (Reihenfolge der beiden Probenahmen kann beeinflusst werden, Untersucher nicht blind für Test) und vom Hersteller der Kits finanziert worden (Interessenkonflikt).</p> <p>Anteil unbrauchbarer Proben im Vergleich: Der Anteil lag bei 24 von 38 einbezogenen Studien höher bei der CC als bei der LBC, bei 3 Studien gab es keinen Unterschied und bei 9 Studien lag der Anteil bei LBC höher.</p> <p>Evidenz zu potentiellen Schäden durch CC und LBC: Keine Aussage</p> <p>Fazit der Autoren: Studienlage 2003 bietet eine höhere Sicherheit (als Payne et al 2000) zur besseren Kosteneffektivität der LBC gegenüber CC. Weitere Untersuchungen seien erforderlich, insbesondere RCTs.</p>
<p>Davey E, et al. : Effect of study design and quality on unsatisfactory rates, cytology classifications, and accuracy in liquid-based versus conventional</p>	<p>1. ja 2. 1966– bis 15. Januar 2004 3. ja</p>	<p>Frage: Diagnostische Performance in Abhängigkeit vom Studiendesign/ unbrauchbarer Proben, nach Anteilen der Klassifikationen: normal,</p>	<p>Studienlage: 56 Studien erfüllten Einschluss-Kriterien, darunter ein kleines RCT (Obwegeser et al 2001). Eine Metaanalyse mit LBC n=483.050 und CC n=662.401 wurde durchgeführt. Die 5 besten Studien waren solche mit paarweiser Probennahme (4 mit Split-Sample Design) ohne Kenntnis des anderen Befundes und verblindete Überprüfung mindestens aller diskordanten Befundpaare per Referenz-</p>

<p>cervical cytology: a systematic review. Lancet, 2006; 367: 122-132</p>		<p>ASCUS, LSIL HSIL und Vergleich mit Referenz-Standard.</p>	<p>Test. Nur 5 von 56 Studien waren im Screening-Kontext durchgeführt worden.</p> <p>Evidenz zur Treffsicherheit der Tests: In den 3 Studien mit hoher Qualität wurden mit CC mehr hochgradige Veränderungen als bei LBC befundet. Wurden Studien hoher und mittlerer Qualität gepoolt, klassifizierte LBC einen höheren Anteil an ASCUS-Befunden Nur 4 der 58 Datasets erlaubten die Berechnung von Sensitivität und Spezifität, ein Drittel der Studien wurde ohne Referenzuntersuchung durchgeführt und kaum eine der mit Referenzuntersuchung durchgeführten Studien hielt den validierenden Untersucher verblindet. Es gäbe keinen Unterschied in der diagnostischen Performance.</p> <p>Anteil unbrauchbarer Proben im Vergleich: Metaanalyse über 48 Studien: kein Unterschied hinsichtlich mangelhafter Proben LBC Anteil: 0.75%, CC: 0.81%: Die gewichtete Differenz LBC-CC betrug: -0.14% (95% CI -0.33% bis 0.06%).</p> <p>Evidenz zu potentiellen Schäden durch CC und LBC: Keine Aussage</p> <p>Fazit der Autoren: Kein Hinweis, dass LBC besser sei als CC. LBC wies keinen niedrigeren Anteil an unbrauchbaren Proben auf. Niedrige Studienqualität wird beklagt, RTCs gefordert.</p>
---	--	--	--