

Health Technology Assessment

Herausgeber:

Deutsche Agentur für Health Technology Assessment des
Deutschen Instituts für Medizinische Dokumentation und Information
(DAHTA@DIMDI)

In der elektronischen Zeitschrift gms Health Technology Assessment der Deutschen Agentur für Health Technology Assessment des Deutschen Instituts für Medizinische Dokumentation und Information (DAHTA@DIMDI) werden Forschungsbeiträge, Untersuchungen, Umfragen usw. als Diskussionsbeiträge im Auftrag des Bundesministeriums für Gesundheit und Soziale Sicherung veröffentlicht. Die Verantwortung für den Inhalt obliegt den jeweiligen Autoren bzw. der jeweiligen Autorin / Autor.



Deutsche Agentur für Health Technology Assessment des
Deutschen Instituts für Medizinische Dokumentation und Information
(DAHTA@DIMDI)

**Informationssystem
Health Technology Assessment (HTA)
in der Bundesrepublik Deutschland**

PAP-Test zum Screening auf Zervixkarzinom

W. Frank, B. Konta, C. Peters-Engl

1. Auflage 2005

© DAHTA@DIMDI. Alle Rechte, auch die des Nachdrucks von Auszügen, der photomechanischen Wiedergabe und der Übersetzung, vorbehalten.

DIMDI

Waisenhausgasse 36 - 38a

50676 Köln

Telefon: 0221 / 4724 - 1

Telefax: 0221 / 4724 - 444

Vorbemerkungen

Die Deutsche Agentur für Health Technology Assessment beim Deutschen Institut für Medizinische Dokumentation und Information (DAHTA@DIMDI) hat den Arbeitskreis für wissenschaftsbasierte Gesundheitsversorgung (ARWIG) mit dem HTA-Bericht „PAP-Test zum Screening auf Zervixkarzinom“ beauftragt.

Mit der vorliegenden Arbeit wird eine endgültige Version nach der Einarbeitung der Kommentare des Gutachterverfahrens eingereicht.

Das interne Gutachten, eine inhaltliche und eine formale Kontrolle der Arbeit, wurde von Mitarbeitern von DAHTA@DIMDI durchgeführt.

Die externe Begutachtung erfolgte durch Herrn PD Volker Schneider, Labor für Zytodiagnostik, Freiburg.

Die Basis der Finanzierung des Gesamtberichts bildet der gesetzliche Auftrag nach Artikel 19 des GKV-Gesundheitsreformgesetzes 2000 und erfolgte durch die Deutsche Agentur für Health Technology Assessment des Deutschen Instituts für Medizinische Dokumentation und Information (DAHTA@DIMDI) im Auftrag des Bundesministeriums für Gesundheit und Soziale Sicherung.

Inhaltsverzeichnis

1	Gesundheitspolitischer Hintergrund	1
2	Wissenschaftliche Kurzfassung	2
2.1	Einleitung	2
2.2	Fragestellung	2
2.3	Methodik	2
2.4	Ergebnisse	2
2.5	Diskussion.....	2
2.6	Schlussfolgerung.....	3
3	Hauptdokument	4
3.1	Einleitung	4
3.2	Fragestellung	4
3.3	Methodik	4
3.4	Medizinische Grundlagen.....	6
3.4.1	Das Zervixkarzinom (Gebärmutterhalskrebs).....	6
3.4.2	Epidemiologie des Zervixkarzinoms	7
3.4.3	Der Papanicolaou-Test (Pap-Test).....	9
3.4.4	Pap-Test - Neue Technologien.....	11
3.5	Screening gegen Zervixkarzinom.....	12
3.5.1	Screening im Allgemeinen	12
3.5.2	Kritik zum Screening gegen Zervixkarzinom	15
3.5.3	Das Zervixkarzinom-Screening in Deutschland.....	16
3.5.4	Europäische Übersicht über das Zervixkarzinom-Screening.....	18
3.6	Ethische Aspekte	23
3.6.1	Grundsätze der medizinischen Ethik	23
3.6.2	Evidenz ethischer Aspekte bei Screeninguntersuchungen	24
3.6.3	Diskussion der ethischen Aspekte.....	27
3.7	Juristische Grundlagen	27
3.8	Verfahrensbewertung des Papanicolaou-Tests	28
3.8.1	Testgütekriterien des konventionellen Pap-Tests sowie der neuen Technologien.....	28
3.8.2	Screeningintervall	30
3.9	Einfluss der Untersuchungsintervalle - Ergebnisse	30
3.9.1	Einfluss der Intervalle auf die Testkennwerte.....	31
3.9.2	Einfluss der Intervalle auf 100.000 Frauen	33
3.9.3	Einfluss der Intervalle für die Bundesrepublik Deutschland	34
3.9.4	Schlussfolgerungen	37
3.10	Ökonomische Bewertung	38
3.10.1	Einfluss der Intervalle auf die Kostenwirksamkeit.....	38
3.10.2	Kosten	41
3.10.3	Diskussion.....	42
3.11	Zusammenfassende Diskussion aller Ergebnisse.....	43

3.12	Schlussfolgerung	44
4	Anhang	45
4.1	Richtlinien der American Cancer Society zur Früherkennung von Krebs (englisch)	45
4.2	Empfehlungen der US Preventive Services Task Force (englisch).....	46
4.3	Abkürzungsverzeichnis.....	47
4.4	Tabellenverzeichnis	47
4.5	Abbildungsverzeichnis	48
5	Literaturverzeichnis	49

1 Gesundheitspolitischer Hintergrund

Weltweit ist das Zervixkarzinom die zweithäufigste Krebserkrankung bei Frauen, wobei 80 % der Diagnosen in Entwicklungsländern gestellt werden. Das Risiko schwankt im Ländervergleich um den Faktor 20. Bei geeigneten Früherkennungsmaßnahmen kann ein großer Teil der Erkrankungen im Frühstadium erkannt sowie einfach und erfolgreich behandelt werden.

In Deutschland verstarben laut Bundesgesundheitsberichterstattung 2.207 Frauen 1995 an einem Zervixkarzinom, die standardisierte Mortalitätsrate lag bei 4,2 je 100.000 Einwohner. Seit den 70-er Jahren geht die Mortalität deutlich zurück, in den letzten Jahren um rund 40 %. Diese Entwicklung wird vor allem mit erfolgreichen Screeningmaßnahmen und der Früherkennung der Karzinome in Zusammenhang gebracht. Die Inzidenzrate (Neuerkrankungsrate) wird in Deutschland aus den Angaben der Register im Saarland und im Osten geschätzt, da ein bundesweites Krebsregister fehlt. Danach erkranken jährlich in Deutschland etwa 6.200 Frauen an der invasiven Form des Gebärmutterhalskrebses.

Das Screening gegen Zervixkarzinom wurde in West-Deutschland 1971 eingeführt. 1991 wurde das Programm auf den Osten (neue Bundesländer) ausgeweitet. Davor wurde seit den 50-er Jahren in opportunistischer Form auf Zervixkarzinom untersucht. Als Screeningmethode der Wahl wurde von Anfang an der Papanicolaou-Abstrich-Test (Pap-Test) eingesetzt. Bei diesem Verfahren handelt es sich um einen Zellabstrich von Muttermund und Gebärmutterhalskanal; es ist also einfach in einer großen Zahl durchzuführen. Mit dem Pap-Test können entzündliche Veränderungen der Zellen oder mögliche Vorstufen des Karzinoms genauso wie Krebszellen erkannt werden, auch wenn noch keine Beschwerden bestehen. Nach dem Schema von Papanicolaou können die untersuchten Zellen in fünf Befundgruppen nach Schweregraden unterschieden werden. In Deutschland werden Zellen nach der Münchner Nomenklatur II^{52, 60} beurteilt.

Hier wird ein Screeningintervall von einem Jahr, bei einem Beginn ab dem 20. Lebensjahr empfohlen. Hinsichtlich der Intervalle und der Altersgruppen existieren in Europa (15 Staaten der europäischen Gemeinschaft (EU)) sehr große Unterschiede in den Screeningprogrammen, obwohl alle den Pap-Test als Screeningmethode der Wahl einsetzen. Die Screeningintervalle schwanken zwischen einem und fünf Jahren, wobei drei Jahre häufig sind. Die Altersgruppen werden oft erst ab 23, 25 oder 30 Jahren einbezogen, die obere Grenze ist oftmals 60 Jahre. Daher werden in Deutschland mit einem Screening ab 20 Jahren ohne obere Altersgrenze und mit der Screeninghäufigkeit von einem Jahr im Vergleich sehr viele Pap-Tests durchgeführt, was teuer und möglicherweise ineffizient ist, da einige neue wissenschaftliche Arbeiten zum Pap-Test dessen Güte anzweifeln und daher ein Screeningintervall von länger als einem Jahr für ausreichend erachten.

Mit vorliegender Arbeit soll festgestellt werden, welchen Einfluss das Untersuchungsintervall auf das System der Früherkennung von Zervixkarzinom hat. Die neuesten Erkenntnisse aus internationaler Sicht sollen zu diesem Thema herangezogen werden und damit eine Überprüfung der Testgüte des Pap-Tests, sowie seiner bestmöglichen Einsatzmöglichkeiten erarbeitet werden. Empfehlungen, die ermöglichen, das deutsche Screening auf Zervixkarzinom den neuesten Erkenntnissen anzupassen, sollen gegeben werden.

2 Wissenschaftliche Kurzfassung

2.1 Einleitung

Im Rahmen der EU unterscheiden sich bei Anwendung desselben Tests die Modalitäten des Screenings auf Zervixkarzinom beträchtlich. So schwankt das empfohlene Untersuchungsintervall zwischen einem (Deutschland) und fünf Jahren. Auch die Altersspanne, innerhalb der untersucht werden soll, ist teils sehr unterschiedlich.

2.2 Fragestellung

Im Rahmen dieser Arbeit soll auf der Basis der wissenschaftlichen Evidenz festgestellt werden, welchen Einfluss unterschiedliche Untersuchungsintervalle auf das Vermögen des Screenings zur Früherkennung der Erkrankung haben. Dabei soll insbesondere auf den Nutzen und den Aufwand der damit verbunden ist, eingegangen werden. Im Speziellen sollen die Berechnungen die Konsequenzen für die Bundesrepublik Deutschland erhellen.

2.3 Methodik

Nach der Recherche der besten verfügbaren Literatur im Sinne der evidenzbasierten Medizin (EBM) werden anhand einer statisch-epidemiologischen Modellierung die Konsequenzen der unterschiedlichen Untersuchungsintervalle geprüft. Dabei werden eine nichtgescreente Bevölkerung und eine mit unterschiedlichen Intervallen gescreente Bevölkerung gegenübergestellt. Die kombinierten Testkenngrößen des Pap-Tests, eine Umrechnung auf 100.000 Frauen, für die Bevölkerung Deutschlands und eine ökonomische Perspektive werden eingenommen.

2.4 Ergebnisse

Anhand des zentralen Kriteriums, der Anzahl der erkennbaren Fälle kann festgestellt werden, dass sich die Ergebnisse des Untersuchungsintervalls von einem bis drei Jahren fast überhaupt nicht unterscheiden. Es sind aber deutliche Unterschiede in der Zahl der zu erwartenden Fehldiagnosen (falsch-positive Resultate) als auch im für das Screening erforderlichen finanziellen Aufwand erkennbar. Entsprechend der wissenschaftlichen Literatur als auch der Modellrechnungen kann das Aufrechterhalten eines einjährigen Untersuchungsintervalls für das Massenscreening nicht empfohlen werden.

Die Teilnehmerate ist für den Ertrag des Screenings ein wesentlicher Einflussfaktor. Ihre Steigerung um 10 % hat einen viel deutlicheren positiven Effekt auf die Früherkennung als die Diskussion über das Untersuchungsintervall von einem bis zu zehn Jahren.

2.5 Diskussion

Ein einjähriges Screening erscheint vor dem Hintergrund der großen Gefahr falsch-positiver Resultate als auch des erforderlichen Aufwands suboptimal. Die generelle Festlegung eines Untersuchungsintervalls ist insofern problematisch, als es nicht spezielle Risikokonstellationen berücksichtigen kann. Das Screeningintervall zu optimieren ist aus wissenschaftlicher Sicht nur möglich, wenn es Bewertungen gibt, die auch psychologische oder ethische Konsequenzen der falsch-positiven oder falsch-negativen Ergebnisse mit einbeziehen. Neuere Technologien, wie die flüssigkeitsbasierte Zytologie (FBC) sind mit teils schwacher Evidenzlage belegt, sodass Unsicherheiten in der sinnvollen Berücksichtigung dieser Testverfahren gegeben sind. Gleiches gilt für die Evidenzlage des Tests auf einen anerkannten Risikofaktor, dem humanen Papillomavirus (HPV).

Der Steigerung der Teilnehmerate ist besonderes Augenmerk zu schenken. Dabei ist es bereits als ein großer Erfolg zu betrachten, wenn Frauen, die noch nie eine entsprechende Untersuchung haben vornehmen lassen, auch nur ein einziges Mal in ihrem Leben dazu bewegt werden könnten.

2.6 Schlussfolgerung

Das Screening auf Früherkennung des Zervixkarzinoms sollte nicht generell auf ein Untersuchungsintervall festgelegt werden, sondern die spezielle Risikokonstellation der Frau miteinbeziehen. Generell ist ein einjähriges Untersuchungsintervall im Vergleich zu einem zwei- oder dreijährigen Screening sehr ineffizient und mit größerem Aufwand verbunden. Dem steht die erhebliche Gefahr von falsch-positiven Testergebnissen bei kurzen Untersuchungsintervallen gegenüber.

Insgesamt soll die eventuelle Risikokonstellation der Frau bzw. der Patientin in evidenzbasierter Form in das angebotene Screening aufgenommen werden. Dies führt zu einer Optimierung des Screenings vor dem Hintergrund der Anpassung der Testauswahl sowie der Untersuchungsintervalle an die Risikokonstellation sowie die vorangegangenen Testergebnisse. In einigen Ländern sind diese ergebnisabhängigen Untersuchungsintervalle umgesetzt.

Die Optimierung des Screenings vor dem Hintergrund der Personenbiographie ermöglicht beträchtliche Ressourcenschonung sowohl für das Gesundheitssystem als auch für die betroffene Person. Die Optimierung muss daher die Erkennungswahrscheinlichkeit, als auch die Gefahr von falsch-positiven Resultaten, die zwangsläufig mit der Durchführung der Untersuchung verbunden ist – und zwar in einem beträchtlichem Ausmaß – miteinbeziehen. Dies ermöglicht auch, die persönlichen Präferenzen der betroffenen Frau (risikoaversives bis risikofreudiges Gesundheitsverhalten) mit einzubeziehen.

Anzustreben wäre auf Basis dieser Arbeit eine gemeinsame Aktion der Versicherungsverbände, die die Optimierung des Screenings zum Ziel hat. Diese Aufgabe wäre auf der Basis der gegenwärtigen Evidenzlage durchführbar und könnte die persönliche Situation der untersuchten Frau berücksichtigen. Eine solche Vorgangsweise wäre nicht nur innovativ im Sinne eines „Disease Managements“ oder eines modernen Gesundheitsmanagements, sondern würde zu Ressourceneinsparungen und zur gesamten Verkleinerung des seelischen Leids infolge von Fehldiagnosen führen. Darüber hinaus könnte durch die Berücksichtigung von individuellen Präferenzen die Teilnehmerate möglicherweise gesteigert werden, was als zentrales Kriterium für den Erfolg des Screenings betrachtet werden muss.

3 Hauptdokument

3.1 Einleitung

Der Pap-Test ist eine seit vielen Jahren verwendete Screeninguntersuchung zur Früherkennung des Zervixkarzinoms. Die Sinnhaftigkeit seiner Verwendung gilt als unbestritten. In der vorliegenden Arbeit soll die Evidenz für die Häufigkeit der Verwendung des Pap-Tests geprüft werden, um Erkenntnissen der letzten Jahre Rechnung tragen zu können.

Der Pap-Test (Synonym für zytologischer Abstrich, Abstrich oder Zervixabstrich) ist in (West-)Deutschland im Rahmen des Krebsvorsorgeprogramms bereits seit 1971 in Verwendung. Die Screeningintervalle werden mit einem Jahr empfohlen bzw. die krankenversicherte weibliche Bevölkerung hat ab dem 20. Lebensjahr Anspruch auf die kostenlose Inanspruchnahme einer entsprechenden Untersuchung. Da dieses Verfahren im gesamten europäischen Raum Verwendung findet, sollen die Evidenz hinter den stark unterschiedlichen Untersuchungsintervallen geprüft und eventuelle Empfehlungen zur Neustrukturierung des Screenings in Deutschland vorgeschlagen werden.

3.2 Fragestellung

Die zentralen Kenntnisse über die Erkennung von Zervixkarzinomen haben sich seit vielen Jahren nicht verändert. Seit Beginn der Verwendung des Pap-Tests wird dieses Verfahren in allen europäischen Staaten zur Früherkennung von Zervixkarzinom eingesetzt. Die Screeningmethoden unterscheiden sich jedoch in der Häufigkeit der Verwendung des Testverfahrens als auch in der Art des Screenings (opportunistisches, organisiertes Screening).

Es ist unklar, in welchen Untersuchungsintervallen der Pap-Test angewendet werden soll. Dazu müssen sowohl die Testkennwerte (Sensitivität, Spezifität) als auch ökonomische Gegebenheiten mit berücksichtigt werden. Wenn durch zu niedrige Testkennwerte zu hohe Fallzahlen an falsch-negativen sowie falsch-positiven Testergebnisse zu beobachten wären, so hätte dies nicht nur für die betroffenen Frauen unerwünschte Auswirkungen, sondern würde auch das medizinische Finanzierungssystem zum Teil unnötig belasten.

Nachdem die Forschung auf dem Gebiet der Screeningmethoden und der medizinisch-diagnostischen Verfahren sehr dynamisch ist, kann mit einer guten Evidenzlage zur Beurteilung auch älterer Testmethoden wie dem Pap-Test gerechnet werden.

3.3 Methodik

Die folgende Methodik wird im vorliegenden Bericht angewendet:

Zunächst wird für den europäischen Raum eine Recherche über die Häufigkeit der Anwendung einer Screeninguntersuchung durchgeführt. Dabei wird auch die eingesetzte Technik recherchiert.

Nach Klärung der angewendeten Methode (Pap-Test) wird seine Güte erforscht. Diese Kriterien sind die epidemiologischen Eigenschaften zur Beurteilung eines diagnostischen Tests.

Methoden der EBM: Die für die Beurteilung relevante Literatur wird in internationalen Literaturdatenbanken nach vordefinierten Schlagworten vorgenommen. Zu den einzelnen Bereichen (Medizin, Epidemiologie, Ökonomie, Ethik) werden unterschiedliche Suchmodi eingesetzt.

Berechnungsmodelle: Die Erkenntnisse über den wissenschaftlichen Sachstand werden für die Durchführung von Berechnungen anhand eines statistisch-epidemiologischen Berechnungsmodells gewonnen. Dieses Modell ist für die Beurteilung von Screeningverfahren im Arbeitskreis für wissenschaftsbasierte Gesundheitsversorgung (ARWIG) seit Jahren im Einsatz und zahlreiche Screeningverfahren wurden damit bereits im medizinischen oder ökonomischen Sinne evaluiert.

Literaturrecherche

Die Literaturrecherche wurde in Kooperation mit DAHTA@DIMDI durchgeführt, wobei Erkenntnisse aus einer Machbarkeitsstudie in die Recherchestrategie eingeflossen sind. Aus der Aufgabenstellung und dem Wunsch nach einer möglichst umfassenden Beurteilung im Sinne eines HTA-Gutachtens wurden vier voneinander unabhängige Themenkomplexe identifiziert.

Folgende vier Bereiche wurden getrennt recherchiert und analysiert:

1. Screening auf Zervixkarzinom mittels Pap-Test.
2. Epidemiologie des Zervixkarzinoms.
3. Ökonomische Analysen.
4. Ethische Bezüge.

Ad 1: Screening auf Zervixkarzinom mittels Pap-Test (komprimierte Darstellung)

Vier Detailrecherchen:

Pap-Test and # Screening and # cost-benefit and # analysis

Pap-Test and # screening and # interval and # benefit

Pap-Test and # Cervix and # screening and # policy and # interval

Cervix and # screening and # cancer and # cost-benefit and # outcome

In den Datenbanken: EMBASE (EM90), MEDLINE (ME90), CANCERLIT (falls durch MEDLINE nicht vollständig abgedeckt), CATFILEplus (CATLINE), COCHRANE – CDSR + DARE, BIOSIS (BA93), SCISEARCH (IS90)

Eingeschränkt auf: Clinical Trials, Meta-analysis and Randomised controlled trial and reviews

Regionale Einschränkung: Europa, Nordamerika und Australien.

sowie:

HTA-Datenbank zum Schlagwort: Cervix*(x, cal) Screening (last 5 years)

Ausschluss von Doppel- und Mehrfachzitationen.

Handsuche

Ad 2: Epidemiologie des Zervixkarzinoms

Cervix and # cancer and # epidemiology and # incidence and # prevalence

in den Datenbanken EMBASE sowie MEDLINE (ME90)

Ad 3: Ökonomische Analysen

Screening and # economics and # cancer

in den Datenbanken: EconLit, HSPROJ, HSTAT, DARE, HTA sowie NHS EED

Ad 4: Ethische Bezüge

Screening and # (false-positive or false-negative) and consequences

in den Datenbanken: ETHMED sowie eventuell EUROETHICS

sowie:

Screening and # treatment and # coping

in der Datenbank PSYCINFO

In der systematischen Literatursuche konnten 249 Dokumente identifiziert werden, die von zwei unabhängigen Gutachtern einer Validitäts- sowie einer Qualitätsprüfung unterzogen wurden. Die verbliebene Literatur wurde DAHTA@DIMDI mit der Bitte zur Bereitstellung übermittelt.

3.4 Medizinische Grundlagen

3.4.1 Das Zervixkarzinom (Gebärmutterhalskrebs)

Gebärmutterhalskrebs ist in Deutschland an der achten Stelle der Krebserkrankungen. Seit Einführung des Früherkennungsprogramms konnte die Gefahr, die diese Diagnose in sich birgt, durch die Erkennung und Behandlung von Vor- und Frühformen bereits erheblich reduziert werden.

Diese Krebsform, in der medizinischen Fachsprache Zervix- bzw. Portiokarzinom genannt, entsteht meist an der Übergangszone vom Zylinderepithel zum Plattenepithel der Portio: 95 %

3.4.1.1 Diagnose

Nach mikroskopischer Erkennung von auffälligen (dysplastischen oder tumorösen) werden Zellveränderungen mittels Pap-Test festgestellt (wird im Kapitel „Der Papanicolau-Test (Pap-Test)“ beschrieben), der eine Klassifizierung enthält, die eine Abstufung nach Schweregraden zulässt. Bei bereits makroskopisch sichtbaren Läsionen an der Portio erfolgt die bioptische Gewebeuntersuchung durch Knipsbiopsie, Elektroschlinge und Konisation, mit anschließender histopathologischer Aufarbeitung. Zunächst wird der Tumortyp klassifiziert, die Invasionstiefe und horizontale Ausbreitung, sowie Lymph- und Gefäßeinbruch bestimmt. Je nachdem wie stark sich die Tumorzellen mikroskopisch von normalen „ausgereiften“ Zellen unterscheiden, können Grade der „Bösartigkeit“ eines Tumors unterschieden werden. Dieses „Grading“ umfasst die Stufen G1, G2 und G3. Tumorgewebe vom Grad 1 (G1) ist der normalen Schleimhaut vergleichsweise ähnlich. Tumore vom Grad 3 (G3) unterscheiden sich dagegen stark von normalem Zervixgewebe, wachsen besonders schnell und sind aggressiver. G2 liegt dazwischen.

Für das definitive therapeutische Konzept sind jedoch Faktoren wie Tumorgröße, örtliche Ausdehnung, Lymphknotenbefall und Metastasen von Bedeutung. Diese können durch ein Klassifizierungssystem (TNM-System (TNM = Tumor-Node-Metastasen)) beurteilt werden. Die FIGO-Klassifikation (FIGO = International Federation of Gynecology and Obstetrics) ist ebenfalls ein System, das die genannten Faktoren berücksichtigt. Es findet im gynäkologischen Bereich Anwendung (siehe „Klassifikationssysteme des Zervixkarzinoms“). Eine exakte pathologische Stadieneinteilung ist jedoch nur nach radikaler operativer Entfernung eines Tumors möglich.

3.4.1.2 Risikofaktoren

Verschiedene Einflüsse müssen zusammenwirken, um schrittweise die Entartung des Gewebes durch Krebszellen zu fördern. Als wichtigste Auslöser gelten heute Infektionen mit Papillomaviren, besonders mit den Typen 16 und 18. Papillomaviren sind sexuell übertragbar und verursachen auch genitale Warzen, die so genannten Kondylome. Bei Männern sind Papillomaviren an der Entstehung von Peniskarzinomen beteiligt. Raucherinnen haben ebenfalls ein deutlich erhöhtes Risiko an Zervixkarzinom zu erkranken, denn die krebserregenden Bestandteile des Zigarettenrauchs sind auch im Sekret des Gebärmutterhalses nachweisbar und können dort ihre schädigende Wirkung entfalten. Als weitere Risikofaktoren sind eine hohe Promiskuität sowie ein niedriger sozialer Status von Bedeutung.

3.4.1.3 Behandlung des Zervixkarzinoms

Die prätherapeutische klinische Stadieneinteilung und die individuelle Situation der Patientin entscheiden über die Operabilität des Karzinoms. Die Wertheimsche Radikaloperation als Standardtherapie für operable Stadien größer als Stadium Ia1 wird heute durch eine individualisierte abgestufte Radikalität ersetzt. Bei der Behandlung der fortgeschrittenen Zervix-

karzinome (Ib2 - IIb) gibt es wesentliche Unterschiede zwischen den USA und Deutschland. Während in den USA prinzipiell eine Strahlentherapie durchgeführt wird, erfolgt in vielen deutschen Zentren die Indikationsstellung zur Operation. Eine Kombination beider Therapiemodalitäten ist mit einer erheblichen Morbiditätssteigerung verbunden und dann notwendig, wenn es zu einer klinischen Unterschätzung des Stadiums oder zu einer unvollständigen Tumorsektion kommt. Alle fortgeschrittenen Stadien werden mittels primärer Strahlentherapie bzw. kombinierter Chemo / Irradiatio behandelt.

3.4.2 Epidemiologie des Zervixkarzinoms

Werden Vorstufen und Frühformen mitberechnet, ist das Gebärmutterhalskarzinom in Deutschland die vierthäufigste Krebserkrankung von Frauen, etwa gleich häufig wie der Krebs des Gebärmutterkörpers. Pro Jahr erkranken etwa 5.800 Frauen (diese Zahl variiert je nach Quelle bis zu 6.200, Bundesgesundheitsberichterstattung 1998). Vorstufen und Frühformen des Krebses werden überwiegend bei Frauen im Alter von 20 bis 40 Jahren festgestellt. Sie sind gut behandelbar, und ein Fortschreiten der Erkrankung wird damit vermieden. Die bei der Diagnose bereits fortgeschrittenen Tumore zeigen eine Häufung ab dem 45. Lebensjahr (Information des deutschen Krebsforschungsinstituts Heidelberg).

3.4.2.1 Inzidenz und Mortalität

Nach Schätzungen der Arbeitsgemeinschaft bevölkerungsbezogener Krebsregister in Deutschland zitiert nach Fischer²⁰ steht das Zervixkarzinom bezüglich der Neuerkrankungen mit 3,3 % 1997 an achter Stelle mit jährlich 5.800 erkrankten Frauen. Der Verlust an Lebenserwartung für eine an einem Zervixkarzinom erkrankte Frau, bezogen auf alle Stadien, wird mit durchschnittlich neun Jahren angegeben. Schneider et al.⁵³ gehen von einer 100-fach höheren Inzidenz zervikaler präkanzeröser Läsionen aus, verglichen mit dem invasiven Karzinom.

Nach Einführung von Screeningprogrammen in zahlreichen Ländern sank die Inzidenz des invasiven Zervixkarzinoms³⁴. Genauso wird in unzähligen Arbeiten darauf hingewiesen, dass seit Einführung von Screeningmaßnahmen die Mortalität an Zervixkarzinom um 40 bis 90 % gesunken ist^{34, 35, 43}.

Dabei muss bedacht werden, dass in einem Gesundheitssystem sowohl die Inzidenz- als auch die Mortalitätsraten des Zervixkarzinoms stark vom Screeningverhalten der Frauen abhängen (Teilnahmerate). Ebenso wirkt sich das Screeningverhalten auf die Verteilung der Tumorstadien zum Diagnosezeitpunkt aus.

Bezüglich des Zusammenhangs von schwerer Dysplasie bzw. Oberflächenkarzinom und dem invasiven Zervixkarzinom wird anhand von Übergangsraten geschätzt, dass sich - bezogen auf alle diagnostizierten Dysplasien – einer von acht (entspricht 12,5 %) der diagnostizierten Fälle eines Oberflächenkarzinoms zu einem invasiven Karzinom weiterentwickeln würde, falls dieses unbehandelt bliebe⁴⁷.

3.4.2.2 Prävalenz

Entscheidend für die Beurteilung eines Screenings ist die Prävalenz der Zielerkrankung in der Bevölkerung. Durch das Pap-Test-Screening ist die Prävalenz des Zervixkarzinoms unzweifelhaft deutlich gesunken. Soll nun nachträglich das bereits eingeführte Screening beurteilt werden, ist es insofern unzulässig, eine durch das Screening abgesenkte Prävalenz zur Basis der Beurteilung zu machen, da diese aufgrund der gesunkenen Prävalenz deutlich negativer ausfiele als im Fall der Evaluation anhand einer Bevölkerung, in der kein Screening stattfindet. Die negativen Aspekte der Epidemiologie zum Zeitpunkt der Überprüfung (z.B. geringer positiver Voraussagewert (PPV)) würden dem etablierten Screening angelastet werden, ohne dass die positiven Aspekte (Vergleich gescreente gegenüber ungescreenter Bevölkerung) dem gegenüber gestellt werden könnten. Dies kann sogar so weit führen dass die Kenngrößen des Screenings so schlecht wären, dass von der weiteren Durchführung des Screenings Abstand

genommen werden müsste. Im Fall des Zervixkarzinoms würde dies in der Folge zu einem deutlichen Anstieg der Inzidenz, Prävalenz und Mortalität führen.

Daraus muss geschlossen werden, dass eine Prävalenz des Zervixkarzinoms in einer ungescreenten Bevölkerung für eine Beurteilung des Screenings herangezogen werden muss. Daraus lassen sich neutrale Ergebnisse ableiten, die eine Evaluation oder eine Studie vor Einführung eines Screenings in einer Bevölkerung ergeben hätte.

Die Schätzung der Prävalenz in einer ungescreenten Bevölkerung kann mit Hilfe des kumulativen Lebenszeitriskos approximativ vorgenommen werden. Das Lebenszeitrisiko eines Zervixkarzinoms im Fall keines Screenings wurde mit Hilfe von schwedischen Daten geschätzt. Es wird für die Altersgruppe von 30 bis 74 Jahren (das kumulative Risiko der Geburtsjahreskohorte 1915, die praktisch noch kein Screening erhalten hatte) auf 2 % vermutet⁴⁷. Demnach würden in einer Bevölkerung, in der es kein Screening auf Zervixkarzinom gibt, zwei von 100 Frauen (jede 50-ste Frau) im Laufe ihres Lebens an der invasiven Form des Zervixkarzinoms erkranken.

Die anhand einer Modellrechnung festgestellte erwartete jährliche Prävalenz in einer ungescreenten Bevölkerung über 18 Jahre wird mit 5,9 auf 10.000 pro Jahr angegeben⁵⁶. Umgerechnet anhand der Altersverteilung und der Prävalenzverteilung ergäbe dies ein Lebenszeitrisiko für eine Gesamtbevölkerung von rund 2,06 %⁵⁵.

Zu deutlich differierenden Resultaten kommen Cruickshank et al.¹¹ für Frauen zwischen 50 und 60 Jahren. Die Prävalenz des invasiven Karzinoms aus einer 11-jährigen Geburtsjahreskohorte (retrospektiv analysiert) von 23.440 Frauen betrug in der Gesamtkohorte für fünf Jahre 0,059 %, wohingegen für die gut gescreente Gruppe nur 0,011 % (82 % Senkung). Umgerechnet auf eine Gesamtbevölkerung würde dies ungefähr 0,71 % in der Gesamtgruppe sowie 0,13 % Prävalenz in der gescreenten Gruppe entsprechen (eigene Umrechnung).

3.4.2.3 Epidemiologie des humanen Papillomavirus (HPV)

In geschätzten 95 % ist HPV beim vorliegenden Zervixkarzinom koexistent. HPV wird als Mitursache für Zervixkarzinom betrachtet, jedoch entwickelt sich nicht aus jedem Fall von HPV ein Zervixkarzinom. HPV ist also eine notwendige, jedoch nicht hinreichende Bedingung für die Entwicklung eines Zervixkarzinoms. Dies wird durch die unterschiedlichen Inzidenzen der beiden Krankheiten verdeutlicht. Während die HPV-Infektion stark altersabhängig ist, wird bei 35-jährigen Frauen die Prävalenz für HPV zwischen 1 und 8 % angegeben. Dies entspricht für eine einzige Alterskohorte der 35-jährigen Frauen in Deutschland zwischen 7.000 und 56.000 Frauen, denen jedoch eine Erkrankung von rund 5.000 Fällen in der Gesamtbevölkerung an Zervixkarzinom gegenübersteht.

Bestimmte Typen von HPV weisen einen stark epidemiologischen Zusammenhang mit dem Zervixkarzinom auf. Die Typen 16 und 18 sind zwei von mehr als 70 onkogenetischen Typen, die die stärkste Beziehung mit dem Zervixkarzinom zeigen. Sowohl Inzidenz als auch Prävalenz von HPV sind bei Frauen unter 25 Jahren am größten³³. Von Ho et al.²⁸ wird eine kumulative Dreijahresinzidenz von 43 % berichtet (jährlich 14 %, aus einer Kohorte von College-Studentinnen).

Die Zeit, bis eine HPV-Infektion wieder völlig verschwindet, wird mit zwischen ein und zwei Jahren angegeben^{16, 27, 28, 68}. Die Prävalenz von HPV wird bei Mandelblatt et al.³⁶ in einer Modellrechnung bei 20 bis 24-jährigen Frauen mit 24,5 % beschrieben, die Wahrscheinlichkeit innerhalb von fünf Jahren zu HPV zu gelangen mit bis zu 28,4 %, die Regression von HPV zu ‚Gesund‘ ebenfalls mit 28,4 %. Daraus ist eine statistisch (nicht individuell) nahezu vollständige Ausheilung der Infektion abzuleiten.

Umgekehrt betrachtet kann jedoch davon ausgegangen werden, dass ein beträchtlicher Anteil der Frauen mit persistierender HPV-Infektion im Alter von 35 Jahren früher oder später ein Zervixkarzinom entwickeln wird. Der Anteil der Frauen mit persistierender HPV-Infektion im Ver-

hältnis zur einmaligen Diagnostizierung eines HPV-Virus ist nach gegenwärtiger wissenschaftlicher Sachlage unbekannt.

In einer 10-jährigen Kohortenanalyse^{8, 57} mit 23.810 Frauen wurde folgendes festgestellt:

Tabelle 1: Häufigkeit von atypischen Zellen (ASC) oder negative und positive Tests auf karzinogene humane Papillomavirus-Typen.

Pap-Test	Anzahl	Anzahl und Anteil negativer onkogenetischer HPV-DNA	Anzahl und Anteil positiver onkogenetischer HPV-DNA
Negativ	20.156 (96,9)	17.594 (87,3)	2.562 (12,7)
ASC oder schlechter	654 (3,1)	237 (36,2)	417 (63,8)
ASC allein	476 (2,3)	202 (42,4)	274 (57,6)
LSIL oder schlechter	178 (0,9)	35 (19,7)	143 (80,3)
Gesamt	20.810	17.831	2.979

LSIL = Low grade intraepithelial lesion. HPV-DNA = Humanes Papillomavirus – Desoxyribonuclein-acid.

Quelle: Sherman et al.⁵⁷

Aus dieser Darstellung ist eine Koexistenz von Zervixkarzinomanzeichen ableitbar, nicht jedoch eine Ursache-Wirkungsbeziehung zwischen Infektion und Krebsanzeichen, da diese Analyse keine dynamische Betrachtung erlaubt.

3.4.3 Der Papanicolaou-Test (Pap-Test)

Bei diesem Test handelt es sich um einen Zellabstrich von Muttermund und Gebärmutterhalskanal. Die Untersuchung kann von Ärzten für Allgemeinmedizin, von Gynäkologen oder von speziell geschultem Pflegepersonal durchgeführt werden. Dabei wird die Scheide mit einem eingeführten Metallspatel schmerzlos entfaltet, so dass sich der äußere Muttermund direkt einsehen lässt. Das Zellmaterial wird vom Arzt mit Hilfe von zwei Watteträgern oder eines Spatels entnommen sowie haltbar gemacht. Anschließend wird es an ein für die Beurteilung qualifiziertes Labor gesendet, wo es gefärbt und unter dem Mikroskop begutachtet wird. Diese spezielle Methode zur Evaluation von Zellen wurde von dem Anatomen George Papanicolaou in den USA entwickelt. Sie wird deshalb Papanicolaou-Abstrich-Test, kurz Pap-Test, genannt. Ein Nachweis der Infektion mit HPV ist heute möglich, aber noch nicht ausreichend bezüglich der Konsequenzen für die betroffenen Frauen geprüft, als dass der Nachweis zu den etablierten Diagnoseverfahren gerechnet werden könnte. Deshalb ist ein zusätzlicher Test auf HPV noch nicht Teil des gesetzlichen Früherkennungsprogramms. Im Rahmen der gesetzlichen Vorsorgeuntersuchungen kann der Pap-Test in Deutschland für jede Frau über 20 Jahre einmal jährlich von den Krankenversicherungen finanziert, in Anspruch genommen werden.

Die regelmäßige Abstrichuntersuchung bietet die Chance, Veränderungen so frühzeitig zu erkennen, dass die betroffenen Frauen mit großer Sicherheit nicht erst durch einen größeren chirurgischen Eingriff, sondern durch einen weniger belastenden Eingriff wie die Konisation dauerhaft geheilt werden können. Darunter versteht man das Herausschneiden eines Gewebekegels aus dem Muttermund und dem Gebärmutterhals mit dem Skalpell, der Hochfrequenzschlinge oder mit dem Laserstrahl. Die Exzisionsmethode ist für eine histologische Aufarbeitung unumgänglich und bietet höchstmögliche diagnostische Sicherheit. Die zahlreichen von diesem Gewebekegel hergestellten Gewebeschnitte werden im Labor unter dem Mikroskop begutachtet. Der Eingriff ist klein und relativ komplikationslos. Im Zusammenhang mit der Konisation wird immer auch eine Ausschabung der Gebärmutter Schleimhaut durchgeführt, um eine weitere Ausdehnung der Veränderungen auf den Innenraum der Gebärmutter auszuschließen.

Mit dem Pap-Test können entzündliche Modifikationen der Zellen oder mögliche Vorstufen von Krebs genauso wie Krebszellen erkannt werden, auch wenn noch keine Beschwerden bestehen. Die Zellen werden in Deutschland nach der Münchner Nomenklatur II^{52, 60} beurteilt.

Tabelle 2: Die Papanicolaou-Test-Gruppen.

Pap-I	Normales Zellbild. Unverdächtiger Zellbefund.
Pap-II	Einzelne atypische, aber nicht schwerwiegend veränderte (dysplastische) Zellen sind vorhanden, ausgelöst durch entzündliche Prozesse. Unverdächtiger Befund.
Pap-II-W	Entzündliche Zellveränderungen.
Pap-II-W(K)	Weist auf eine mögliche Infektion mit dem humanen Papillomavirus (HPV) hin.
Pap-III-D	Verdacht auf leichte bis mittelschwere Zelldysplasie.
Pap-III	Schwere entzündliche oder degenerative Veränderungen und / oder schlecht erhaltenes Material. Zelldysplasie (Zellveränderung) und Carcinoma in situ* nicht auszuschließen. Unklarer Befund.
Pap-IV-A	Verdacht auf schwere Dysplasie oder Carcinoma in situ*.
Pap-IV-B	Verdacht auf schwere Dysplasie oder Carcinoma in situ*. Krebserkrankung nicht sicher auszuschließen.
Pap-V	Zellen eines bösartigen Tumors sind nachweisbar.

* = Carcinoma in situ: Oberflächenkarzinom, Karzinom, dass die Basalmembran noch nicht durchbrochen hat⁴⁸.
 Quellen: ³, Information der Krebsforschung Heidelberg.

Die Befunde Pap-I und -II sind unbedenklich. Bei einem solchen Befund wird in Deutschland empfohlen, die nächste Kontrolle im Rahmen der Früherkennungsuntersuchung in einem Jahr durchführen zu lassen. Sind bei Pap-II die Veränderungen etwas ausgeprägter, kann eine Entzündungsbehandlung oder hormonelle Behandlung mit einer anschließenden Kontrolluntersuchung des Muttermundes durchgeführt werden. Pap-II-W(K) macht weitere Abstriche zur Kontrolle erforderlich. Der Befund Pap-III ist häufig noch rückbildungsfähig, bei Pap-IV liegen meist schon echte Krebsvorstufen oder gar frühe Krebsformen vor.

Bei Pap-III müssen kurzfristig ein weiterer Zellabstrich und eine Lupenuntersuchung (Kolposkopie) erfolgen, bei Verdacht auf eine Entzündung auch eine entsprechende Behandlung. Bilden sich die Zellveränderungen nicht zurück, kann zur Sicherung der Diagnose auch eine Gewebeentnahme (Biopsie) nötig sein. Bei Pap-III-D werden der Abstrich und die Kolposkopie in Abständen von drei Monaten wiederholt. Bleiben die verdächtigen Zellveränderungen längerfristig bestehen, muss dieser Befund durch Entnahme von Gewebe gesichert werden. Diese Gewebeentnahme (z.B. durch Knipsbiopsie oder durch Konisation) stellt meistens die einzige erforderliche Therapie dar, sofern die feingewebliche Untersuchung ergibt, dass die verdächtigen Bezirke im Gesunden entfernt wurden.

Bei Pap-IV-A und -B ist eine Konisation (wie oben beschrieben) erforderlich, da der Verdacht auf ein Oberflächenkarzinom oder auf ein schon etwas tiefer ins Gewebe vorgedrungenes Karzinom sonst nicht sicher ausgeschlossen werden kann.

Bei einem Befund Pap-V wurden im Abstrich Zellen gefunden, die von einem bösartigen Tumor stammen. Daraus ergibt sich der dringende Verdacht auf ein Karzinom, das bereits in tiefere Schichten eingedrungen ist. Die Diagnose wird durch eine Gewebeentnahme (Biopsie) gesichert. Bestätigt sich der Verdacht, muss eine definitive Therapie mittels Operation oder Bestrahlung erfolgen.

3.4.3.1 Klassifikationssysteme des Zervixkarzinoms

Im Folgenden sollen zwei gebräuchliche Klassifikationssysteme für verschiedene Stadien des Zervixkarzinoms aufgeführt werden. Sie werden angewandt, wenn bereits ein Tumor festgestellt werden konnte.

Tabelle 3: Stadieneinteilungen des Zervixkarzinoms nach FIGO (Fédération Internationale de Gynécologie et d'Obstétrique, 1994) und TNM-Klassifikation (1997).

	TNM-Klassifikation	Beschreibung
0	TIs	Oberflächenkarzinom.
I	T1	Invasives, auf die Zervix begrenztes Karzinom.
la	T1a	Präklinisches, ausschließlich mikroskopisch diagnostiziertes mikroinvasives Karzinom.
la1	T1a1	Minimale Stromainvasion von ≤ 3 mm Tiefe und ≤ 7 mm in horizontaler Ausdehnung.
la2	T1a2	Stromainvasion von 3 - 5 mm Tiefe und ≤ 7 mm in horizontaler Ausdehnung (Mikrokarzinom).
lb	T1b	Makroinvasives, klinisch sichtbares Karzinom.
lb1	T1b1	Tumordurchmesser ≤ 4 cm.
lb2	T1b2	Tumordurchmesser < 4 cm.
II	T2	Das Karzinom hat die Zervix überschritten, jedoch die Beckenwand noch nicht erreicht und / oder das untere Drittel der Vagina noch nicht befallen.
Ila	T2a	Befall der oberen 2 / 3 der Vagina, Parametrium frei.
Ilb	T2b	Befall der oberen 2 / 3 der Vagina, Parametrium infiltriert, Beckenwand nicht erreicht.
III	T3	Tumorausbreitung bis zur Beckenwand und / oder bis in das untere Scheidendrittel.
IIIa	T3a	Befall des distalen Drittels der Vagina, Parametrium frei.
IIIb	T3b (jedes N)	Tumorausbreitung bis zur Beckenwand und / oder Hydronephrose oder stumme Niere, auch Befall der Vagina, regionäre Lymphknotenmetastasen.
IVa	T4 (jedes N)	Infiltration der Schleimhaut von Rektum und / oder Blase und / oder Überschreitung der Grenzen des kleinen Beckens.
IVb	T1-4 (jedes N,M1)	Fernmetastasen außerhalb des kleinen Beckens.

TMM = Tumor-Node-Metastasen Klassifikationssystem.

Quelle: Pfeleiderer⁴⁶.

3.4.4 Pap-Test - Neue Technologien

In den letzten zehn Jahren wurden verschiedene neue Technologien zur Feststellung von Zervixkarzinomen entwickelt, um die unzureichenden Testkennwerte (vor allem Sensitivität) des Pap-Tests auszugleichen. Darunter ist vor allem die FBC, synonym: Dünnschichtpräparation, oder Monolayer-Zytologie, zu nennen. Die dafür am Markt angebotenen Geräte AutoCyte PREP (Tri Path Imaging) oder ThinPrep (Cytoc Corp) sollen geeignet sein, die durch Abnahme- bzw. Präparationsfehler beim Pap-Test verantwortlichen falschen Testergebnisse zu vermeiden.^{4,5}

Weitere am Markt befindliche Systeme können ebenfalls Steigerungen der Sensitivität erwarten lassen. Über die Spezifität ist praktisch mit Ausnahme einiger weniger Aussagen keine ausreichend gesicherte Evidenz verfügbar. Thinprep®, AutoPap®, AutoCyte® sind eingetragene Marken bzw. geschützte Namen.

3.5 Screening gegen Zervixkarzinom

3.5.1 Screening im Allgemeinen

3.5.1.1 Arten des Screenings

Unter Screening (Früherkennung, Vorfelddiagnostik) wird eine auf eine bestimmte Krankheit gerichtete diagnostische Maßnahme (Anwendung eines Testverfahrens) verstanden, die darauf abzielt, noch symptomlose Personen (potenzielle Krankheitsträger) möglichst im Frühstadium der Erkrankung zu identifizieren und sie frühzeitig einer effektiven Behandlung zuzuführen. Durch die Früherkennung können in den meisten Fällen weniger aggressive Behandlungsmethoden angewandt, der Krankheitsverlauf abgeschwächt oder das Sterberisiko vermindert werden. Hauptsächlicher Zweck von Früherkennungsmaßnahmen ist die Senkung der krankheitsspezifischen Morbidität und / oder Mortalität^{25, 64}.

Folgende verschiedene Typen von Screening lassen sich unterscheiden:

- *Das Massen-Screening (mass screening)* bezieht die Gesamtbevölkerung oder zumindest einen ganzen Bevölkerungsteil mit ein.
- *Das selektive Screening (selective screening)* zielt auf bestimmte Hochrisikogruppen in der Bevölkerung ab.

Hier wird zwischen

a) *zielgruppenspezifischem Screening*, das bei bestimmten Personen entweder nach Alter, Geschlecht oder Risikogruppe durchgeführt wird und

b) *altersgruppenspezifischem Screening*, das Personengruppen eines bestimmten Alters auswählt, (da die Erkrankung nur in diesem Alter auftreten kann (z.B. verschiedene Neoplasien, oder Phenylketonurie),

unterschieden.

- *Das opportunistische Screening* nutzt die Kontaktnahme der Zielpersonen, wenn sie wegen eines anderen Problems zum Arzt kommen, um sie zu einer weiteren Untersuchung (Screeningtest) zu überweisen. Diese Screeningform kommt beispielsweise bei Blutdruckmessungen im Zusammenhang mit kardiovaskulären Erkrankungen, bei Abstrichen im Zusammenhang mit Gebärmutterhalskrebs oder bei digitalen rektalen Untersuchungen im Zusammenhang mit Prostata- und Mastdarmkrebs zum Tragen.
- *Das offensive oder organisierte Screening* wird in Form eines Programms durchgeführt, das so organisiert ist, die gesamte ausgewählte Bevölkerungsgruppe in einem gewissen Zeitraum mit dem diagnostischen Früherkennungsprogramm zu untersuchen. Es ist populationsbezogen, wobei bestimmte Bevölkerungsteile oder Subpopulationen (z.B. Arbeitsplatzuntersuchungen, Immigranten), beispielsweise in Form von öffentlichen Kampagnen, zu einer Früherkennungsuntersuchung eingeladen bzw. aufgefordert werden. Die Initiative zur Früherkennung geht hier von den medizinisch Verantwortlichen aus.

Konzeptionen zu Früherkennungsmaßnahmen basieren auf einer von zwei impliziten Annahmen. Gemäß der ersten können die Entdeckung und die Elimination von Risikofaktoren oder Vorformen einer Krankheit, deren Ausbruch verhindern. Entsprechend der zweiten Annahme können die frühzeitige Entdeckung und Behandlung einer Krankheit in der präklinischen Phase (Symptomfreiheit) deren Prognose verbessern.

3.5.1.2 Effektivitätsgrundsätze

Die Weltgesundheitsorganisation (WHO)⁷³ hat zehn Kriterien definiert, die zur Effektivitätsbeurteilung eines Früherkennungsprogramms in Betracht gezogen werden sollten und an denen sich der bundesdeutsche Gesetzgeber bei der Definition von Voraussetzungen für Früherkennungsuntersuchungen (§25 SGB V) teilweise orientiert hat.

Dies sind die folgenden Kriterien:

- Die Erkrankung sollte ein bedeutendes Gesundheitsproblem darstellen.
- Die Erkrankung sollte ein diagnostisch erfassbares latentes oder präsymptomatisches Stadium aufweisen.
- Zum natürlichen Verlauf der Erkrankung einschließlich des Übergangs vom latenten zum manifesten Stadium sollten adäquate Erkenntnisse vorliegen.
- Für Patienten mit diagnostizierter Erkrankung sollte eine anerkannte Therapiemöglichkeit bestehen.
- Für die Früherkennung der Erkrankung sollte ein ausreichend eindeutiges Test- bzw. Untersuchungsverfahren vorhanden sein.
- Der Test bzw. die Untersuchung für die Früherkennung der Erkrankung sollte für die Zielpopulation annehmbar sein.
- Es sollte Einvernehmen darüber bestehen, wer als Patient zu behandeln ist.
- Die für die Früherkennung, Diagnose und Therapie benötigte Infrastruktur sollte vorhanden sein.
- Die Kosten pro identifizierten Erkrankungsfall (einschließlich Diagnose und Therapie) sollten in einem ökonomisch ausgewogenen Verhältnis zu den möglichen Gesamtausgaben im Gesundheitswesen stehen.
- Die Früherkennung von Erkrankungsfällen sollte ein kontinuierlicher Prozess (Inzidenz-Screening) und kein Einmal-Projekt (Prävalenz-Screening) sein.

Im Fall des Screenings auf Zervixkarzinom können alle Anforderungen als erfüllt betrachtet werden. Wie ein Screening im Detail zu organisieren ist, hängt darüber hinaus von lokalen, epidemiologischen, ökonomischen und versorgungstechnischen Gesichtspunkten ab.

3.5.1.3 Testgütekriterien

„Ideale“ Tests zeichnen sich durch eine hohe Sicherheit aus, Erkrankte als auch Gesunde zuverlässig zu klassifizieren²⁵. Sensitivität (Prozentsatz der Personen, die korrekt als krank erkannt wurden) und Spezifität (Prozentsatz der Personen, die korrekt als gesund erkannt wurden) sind Charakteristika, die für einen Test zur Beurteilung ermittelt werden müssen. Dies geschieht in der Regel durch den Vergleich der Testmethode mit dem besten verfügbaren diagnostischen Referenzverfahren (Goldstandard).

Die wichtigsten Testgütekriterien sind die Sensitivität und die Spezifität, sowie die Vorhersagewerte PPV und NPV (negative predicted value, dt.: negativ prädiktiver Vorhersagewert):

1. Sensitivität steht für das Vermögen eines Tests, tatsächlich Erkrankte als krank zu identifizieren.
2. Spezifität ist als die Fähigkeit definiert, Gesunde richtig als gesund zu identifizieren.
3. Der PPV stellt unter Realbedingungen die Wahrscheinlichkeit dar, dass eine Person mit positivem Testergebnis tatsächlich erkrankt ist. Der PPV ist abhängig von der Krankheitsprävalenz und der Spezifität des Tests.
4. Der NPV bildet die Wahrscheinlichkeit ab, dass eine Person mit negativem Testergebnis tatsächlich nicht erkrankt ist³⁷.

Besonderes Augenmerk ist im Bereich des Screenings auf die Spezifität zu richten. Da die Zielpopulation von Screeningmaßnahmen durch Definition aus symptomlosen und damit überwiegend gesunden Personen besteht, führt eine geringe Spezifität zwangsläufig zu einer großen Zahl falsch-positiver Testergebnisse. Dies bedeutet, dass die Sicherheit, mit der aus dem Auftreten eines positiven Testresultats auf das Vorliegen der Erkrankung geschlossen werden kann (prädiktiver Wert des positiven Tests oder PPV), gering ausfällt. Die negativen

Folgen von falsch-positiven Testergebnissen werden im Kapitel „Falsch-positive Resultate“ im Detail beschrieben.

Es sei darauf hingewiesen, dass Sensitivität und Spezifität als Funktion des gewählten Grenzwerts zur Trennung von Gesunden und Kranken (diagnostic cut-off) zu verstehen sind. Durch eine Verschiebung des Grenzwerts kann auf beide Charakteristika Einfluss genommen werden. Allerdings geht die Verbesserung einer Größe bei ein und demselben Test in der Regel mit der Verschlechterung der anderen Größe Hand in Hand. Bei der Festsetzung eines Grenzwerts muss daher entschieden werden, inwieweit der Sensitivität oder der Spezifität eines Tests Vorrang gegeben wird, was ein Abwägen zwischen den Folgen von Fehlzurechnungen, sowohl von falsch-positiven als auch falsch-negativen Testergebnissen erforderlich macht²⁵.

Durch Kombination von Tests kann zusätzlicher Einfluss auf das Trennverhalten (Diskriminationsfähigkeit) genommen werden. Hier bieten sich parallele oder sequentielle Teststrategien an.

Bei der parallelen Anwendung von Tests wird das Ergebnis als positiv gewertet, wenn entweder alle Tests positiv sind (die Sensitivität gegenüber den Einzeltests wird niedriger, die Spezifität höher) oder mindestens ein Test positiv ausfällt (die Sensitivität steigt gegenüber den Einzeltests, während die Spezifität abnimmt).

Bei der sequentiellen Anwendung von Tests wird der Einsatz weiterer diagnostischer Maßnahmen vom Ergebnis des jeweils vorangehenden Tests abhängig gemacht. Normalerweise wird bei negativem Ausfall des Tests auf eine weitere diagnostische Abklärung verzichtet, womit das sequentielle Testen zu einer Erhöhung der Spezifität führt^{25, 70}.

3.5.1.4 Systematische Fehler (Informationsverzerrungen)

In Evaluationen und Studien von Screeningprogrammen muss darauf geachtet werden, systematische Fehler, die methodisch möglich sind, zu vermeiden. Folgende systematischen Fehler sind häufig:

Verification bias

Diese Informationsverzerrung kommt zustande, wenn die Testergebnisse des zuerst durchgeführten Testverfahrens Einfluss auf die Durchführung des Referenzverfahrens (zweites, zur Absicherung eingesetztes Testverfahren) nimmt. In diesen Fällen werden Sensitivität und / oder Spezifität falsch eingeschätzt.

Selection bias (volunteer bias)

Dieser Fehler kommt vor, wenn Teilnehmer einer Früherkennungsmaßnahme aus Gesundheitsbewusstsein freiwillig teilnehmen, sich jedoch von der Zielpopulation unterscheiden. Auf diese Weise wird ein Ausschnitt aus der Bevölkerung untersucht, der nicht repräsentativ ist.

Length time bias

Erkrankungen mit langsamer Entwicklung (z.B. langsam wachsende Tumore) zeichnen sich durch relativ lange Phasen im Vorstadium aus. Deshalb können in einem Screening solche Stadien häufiger auffallen und damit überrepräsentiert sein. Es werden auf diese Weise mehr Fälle entdeckt. Jedoch hat dies keine Auswirkung auf die Mortalität.

Lead time bias

Mit „Lead time“ ist der screeningbedingte vorverlegte Zeitpunkt der Diagnosestellung gemeint. Im Vergleich dazu wird dieselbe Diagnose bei Patienten einer ungescreenten Bevölkerungsgruppe später gestellt. Beim Vergleich der beiden Gruppen hinsichtlich ihrer Überlebenszeiten kann sich allein dadurch eine längere Überlebensdauer der gescreenten Bevölkerungsgruppe ergeben. Jedoch ist diese ein Resultat der früher entdeckten Erkrankung und nicht einer Verbesserung der Behandlungsmöglichkeit.

Stage migrations bias

Die Stadieneinteilung von Krebstumoren hängt unmittelbar mit den verwendeten Diagnoseverfahren zusammen. Durch modernere, intensivere und / oder systematischere Diagnosemethoden können mehr Erkrankungen in prognostisch günstigeren Phasen erkannt werden, als dies mit der älteren Technik möglich war. Wird dies nicht berücksichtigt, werden die stadienspezifischen Überlebenszeiten der Patientengruppen, bei denen bereits die neuere Methode angewandt wurde, günstigere Verläufe aufweisen, obwohl damit keine Auswirkung auf das populationsbezogene Gesamtergebnis zu erwarten ist.

3.5.1.5 Effektivitätskriterien

Häufig werden gescreente und nichtgescreente Populationen hinsichtlich der Anzahl entdeckter Erkrankungsfälle, der Stadienverteilung oder der Überlebenszeit miteinander verglichen. Diese Effektivitätskriterien werden jedoch oft in ihrer Aussagekraft nahezu unvermeidlich durch die soeben beschriebenen systematischen Fehler beeinträchtigt.

Die Krankheitsinzidenz kann zur Ermittlung der Effektivität von Screeningprogrammen, die auf die Entdeckung von Vorformen der Erkrankung abzielen, eingesetzt werden. Damit der Nutzen der Entdeckung und Elimination dieser Vorformen angemessen ermittelt werden kann, muss die Intensität der diagnostischen Nachverfolgung für gescreente und nichtgescreente Gruppen allerdings gleich sein.

Als das Effektivitätsmaß mit der höchsten Validität gilt die krankheitsspezifische Mortalität. Sie ist im Allgemeinen das vorrangige Ziel eines Screeningprogramms, wobei andere günstige Effekte wie der Möglichkeit einer weniger aggressiven Behandlung, die Verminderung von Leiden oder die Senkung der Kosten für adjuvante und palliative Therapie sowie Pflege nicht zu gering geschätzt werden sollten. Auch mit der Verwendung der krankheitsspezifischen Mortalität zur Quantifizierung der günstigen Auswirkungen des Screenings sind Validitäts- und / oder Interpretationsprobleme verbunden. Als grundlegende Voraussetzung für die Validität sei hier lediglich die Anwendung gleicher Grundsätze zur Ermittlung und Klassifikation der Todesursachen bei gescreenten und nichtgescreenten Populationen genannt²⁵.

3.5.2 Kritik zum Screening gegen Zervixkarzinom

Ziel des Screenings ist die Reduktion der Wahrscheinlichkeit für das Auftreten eines invasiven Karzinoms. Durch das Entdecken von eindeutigen Vorstadien kann die Wahrscheinlichkeit für ein späteres Auftreten der eigentlichen Zielerkrankung reduziert werden. Folglich ist auf die Entdeckung von Auffälligkeiten, die mit dem Zervixkarzinom in Zusammenhang gebracht werden können, jedoch kein Vorstadium im eigentlichen Sinn sind, besonderes Augenmerk zu lenken. Zu diesen Erkrankungen zählen verschiedene Läsionen, die Infektion mit dem HPV oder das Krebsvorstadium selbst. Es ist daher besonders problematisch, wenn Auffälligkeiten entdeckt und behandelt werden (z.B. Läsionen), von denen erwartet werden kann, dass sie ohne medizinisches Zutun von selbst ausheilen.

Kein Krebs-Screeningtest dürfte so häufig angewendet werden wie der Pap-Test. Umso überraschender erscheint es, dass nie eine geplante randomisierte klinisch kontrollierte Studie zum Zervixkarzinom-Screening durchgeführt wurde. Das Zervixkarzinom-Screening wird von den Gesundheitsbehörden als auch von der Öffentlichkeit als effektiv anerkannt, da das Absinken der Mortalität seit Einführung von Screeningprogrammen weltweit durchaus beeindruckend erscheint. Heutzutage ist aus ethischen Gründen die Durchführung einer für eine endgültige Beurteilung erforderliche randomisierte Untersuchung der epidemiologischen Effektivität des Pap-Tests nicht mehr durchführbar. Das daraus entstandene Informationsdefizit kann als Basis für die vielfältigen Ausgestaltungsmöglichkeiten des Zervixkarzinom-Screenings allein in Europa betrachtet werden. Die Tatsache, dass dieses Screening vor der Einführung nicht anhand von randomisierten klinischen Studien evaluiert wurde, hat die Implementation teils verzögert und gleichzeitig ewige Debatten bezüglich seiner „wahren“ Effizienz ausgelöst,

obwohl die Evidenzlage überzeugt, dass das Screening gegen Zervixkarzinom einen großen protektiven Effekt aufweist⁴⁹.

Trotz der ätiologischen Rolle des HPV-Infektionsscreenings zur Erkennung der Zielerkrankung kann das Screening mittels Pap-Test nach wie vor als der Goldstandard für das Screening gegen Zervixkarzinom betrachtet werden^{30, 50, 60}. Die folgenden Empfehlungen werden nach wie vor als gültig betrachtet:

Die International Union Against Cancer (UICC) hat 1990 festgestellt, dass ein organisiertes Screening gegen Zervixkarzinom mit einer hohen Teilnehmerate für die Zielgruppen 25 bis 60 Jahre, mit einem Intervall zwischen drei und fünf Jahren die maximale Wirksamkeit erzielt⁴⁰.

Das EU-Komitee von Krebsexperten aus dem Europe Against Cancer Programme hat 1993 konstatiert, dass ein optimales Screeningprogramm im Alter zwischen 25 und 65 Jahren mit einem drei- bis fünfjährigen Intervall stattfinden sollte¹⁰.

3.5.3 Das Zervixkarzinom-Screening in Deutschland

Das Screeningprogramm auf Zervixkarzinom wurde als eines der Krebsvorsorgeprogramme in West-Deutschland bereits 1971 eingeführt. 1991 wurde das Programm nach erfolgter Wiedervereinigung Deutschlands auf den Osten erweitert. Davor wurde seit den 50er-Jahren in opportunistischer Form mittels Pap-Test auf Zervixkarzinom untersucht, zunächst nur im Rahmen symptomatischer Gesundheitsversorgung in Universitätskliniken. Während der 60er-Jahre erweiterte sich dieses Screening rasch auf ganz West-Deutschland. Zwischen 1967 und 1972 wurden Zytologie-Ausbildungsstätten gegründet und einige lokale Screeningprogramme initiiert.

1971 wurden Pilotprogramme des Screenings auf Zervixkarzinom in den Bundesländern Bayern und Nordrhein-Westfalen gegründet. Das damit beginnende nationale Screeningprogramm stellte eine einmal-jährliche Untersuchung mittels Pap-Test für alle Frauen im Alter über 20 Jahren über die gesetzliche Krankenversicherung zur Verfügung (diesen Test konnten also 90 % der Frauen in Deutschland in Anspruch nehmen). Die gesetzlichen Krankenversicherungen übernahmen nicht nur die Kosten der Screeninguntersuchung, sondern auch für notwendige Nachuntersuchungen oder Behandlungen der durch das Screening gefundenen Auffälligkeiten.

Die Dokumentation dieses Screeningprogramms wurde während der vergangenen 30 Jahre durchgeführt. Es wurden nur Frauen, die eine private Krankenversicherung in Anspruch nahmen davon ausgenommen (ca. 10 %). In den Jahren von 1971 bis 1991 konnten nur westdeutsche Frauen erfasst werden, seit 1991 fand eine Erweiterung auf ostdeutsche Frauen statt.

3.5.3.1 Untersuchte Bevölkerungsgruppe und Methoden

Die **Einladung** zum Screening wurde bisher nicht durch ein aktives System betrieben, sondern erfolgte bis 1995 über die jährliche Erfassung von Frauen im Alter von 20 Jahren aufwärts, die einen Arzt mit einem Krankenkassenvertrag besuchten. Diese Informationen wurden auch später nicht für ein Einladungssystem genutzt. Die vorrangige Aufmerksamkeit galt den Ärzten, die die Screeninguntersuchungen durchführten. Sie sendeten nachträglich, am Ende eines Quartals, die erhaltenen Ergebnisse an eine Versicherungs-Abrechnungsstelle. Auf diese Weise wurden die praktizierenden Ärzte zu den Verbindungspersonen der Zielpopulation, und sollten diese zur jährlichen Untersuchung einladen. Die Abrechnung diente als Erinnerung an einen notwendigen Untersuchungstermin.

Diese Methode verschwand, als Mitte der 90-er Jahre für alle staatlich Versicherten die Abrechnung über Computerkarten eingeführt wurde. Die weiteren Einladungen erfolgten durch die medizinischen Angestellten oder die praktizierenden Gynäkologen bei der Routineuntersuchung.

Darüber hinaus wurde das Krebs-Screeningprogramm kontinuierlich in populären Magazinen, in der Zeitung sowie in Radio und Fernsehen bekannt gemacht. Auch wurden Informationen über

bundesweite Gesundheitsstellen, über die deutsche Krebsgesellschaft und den Krebsbund, in Form von Broschüren und Informationskampagnen verbreitet. Diese Anstrengungen variierten jedoch innerhalb Deutschlands stark und wurden nicht durch das Krebs-Screeningprogramm gesponsert oder koordiniert.

Als **Zielpopulation** galten alle deutschen Frauen im Alter über 20 Jahren, die staatlich krankenversichert waren (das waren über 90 % aller Frauen dieser Altersgruppe). Diese Gruppe schwankt durch Migration und Wechsel zwischen staatlicher und privater Krankenversicherung. Trotzdem wurde versucht, alle versicherten Frauen in allen Landesteilen in das Programm zu integrieren.

Wenn die Zielpopulation mit der altersgruppenspezifischen weiblichen Bevölkerung des gesamten Landes definiert wird, so sind das rund 33 Millionen Frauen im Alter über 19 Jahren. Die **Inanspruchnahme** erreichte 46 % der Zielpopulation innerhalb eines Jahres und über 80 % innerhalb von drei Jahren. Wenn die Zielpopulation mit der altersgruppenspezifischen weiblichen Bevölkerung, die staatlich versichert ist, definiert wird, sind das 30,5 Millionen Frauen im Alter über 19 Jahren. In dieser Gruppe nahmen 50 % der Frauen innerhalb eines Jahres das Screening wahr.

Etwa 90 % der Zervixkarzinom-Screeninguntersuchungen wurden durch niedergelassene Gynäkologen durchgeführt, 10 % durch Allgemeinmediziner. Das empfohlene Screeningintervall aufgrund eines negativen Pap-Tests liegt bei einem Jahr. Ungefähr 60 % der Frauen kommen innerhalb eines Jahres zur Untersuchung. Laut nationalen Richtlinien muss das Screeningprogramm eine gynäkologische Untersuchung und einen Pap-Test einschließen, jedoch keine Kolposkopie (Vaginalspiegelung). Es bleibt jedoch dem einzelnen Praktiker überlassen, ob er eine solche Untersuchung anschließen möchte oder nicht. Diese Minimaluntersuchung gilt für alle Frauen von 20 bis 29 Jahren, ab 30 Jahren enthält das Screening zusätzlich eine Brust- sowie eine Anleitung zur Selbstuntersuchung der Brust. Mit dem 45. Lebensjahr kommen zum Screening noch ein Test auf okkultes Blut im Stuhl und eine digital-rektale Untersuchung dazu. Für alle Leistungen ist ein zusätzlicher Tarifposten für die das Screening durchführenden Ärzte vorgesehen.

Es kann nicht genau angegeben werden, wie viele Frauen nur den Pap-Test gemacht haben und daher nicht in die gesamte Screeningdokumentation einfließen. Es kann jedoch ungefähr geschätzt werden, dass ca. 15 % aller Frauen das gesamte für ihr Alter vorgesehene Screeningprogramm jedes Jahr in Anspruch nehmen.

Die **Methoden** des Abstrichtests (Abnahmetechnik) schwanken ebenfalls. Es gibt keine genaue Regelung über die Abnahme des Abstrichs, es werden sowohl Wattetupfer als auch spezielle Spatel und Bürsten verwendet. Als Standard gilt jedoch, zwei Abstriche zu nehmen, einen von der Oberfläche und einen vom Gebärmutterhalskanal (von der Übergangszone).

Ein Fallmanagement für Personen mit suspekten Testergebnissen ist gegenwärtig nicht vorhanden. Die Ergebnisse der Abstriche erhält der durchführende Arzt. Die Laboratorien erheben zwar eine Falldokumentation der Abstriche, mit den betroffenen Personen kommunizieren jedoch nur die Ärzte. Im Fall eines auffälligen Testergebnisses werden die betroffenen Frauen vom behandelnden Arzt kontaktiert. Es wird in der Regel ein weiterer Termin für eine Untersuchung vereinbart. Über darüber hinausgehende erforderliche Untersuchungen, wie z.B. Biopsien, wird keine standardisierte Dokumentation geführt. Die Aufzeichnungen werden jedoch in den Laboratorien verwahrt, um später einer etwaigen Verwertung zur Verfügung zu stehen.

3.5.3.2 Qualitätssicherung

Die Richtlinien zum Zervixkarzinom-Screening wurden durch die Deutsche medizinische Gesellschaft und die Kammern der niedergelassenen Ärzte herausgegeben. Es gibt keine regionalen Regeln. Die nationalen Leitlinien enthalten ein Protokoll zur Qualitätssicherung, das

technische Aspekte, Befundung, Doppelbefundung, Prüfung unklarer Fälle, Training, Kommunikation der Resultate und eine Nachverfolgung der abnormalen Testergebnisse beinhaltet. Diese standardisierten Richtlinien basieren auf der Münchner Nomenklatur^{7, 52, 60}, die für die Berichterstattung verwendet wird. Für das Einladungssystem gibt es keine Regelung.

In Deutschland gibt es jedoch gravierende Mängel, die es nicht erlauben, die Richtlinien in der Praxis einzuhalten. Die vorhandenen Ressourcen können die durch die Qualitätsrichtlinien geforderten Maßnahmen nicht erreichen. Die systematischen Prozeduren wie Beobachtung und Dokumentation, wie sie in der Regelung gefordert werden, konnten nicht in der angestrebten Form etabliert werden. So hat z.B. das Dokumentationsprogramm, das Daten von allen beteiligten Zentren erhalten sollte, nie die vollständigen Informationen erhalten und konnte damit nicht den Forderungen entsprechen. Um dies auszugleichen hat das zentrale Forschungsinstitut, das die Resultate des Screenings mit standardisierten Methoden dokumentiert, ein Qualitätssicherungsprogramm entworfen, das die korrekt gesammelten Daten von Personen und Regionen für die Qualitätssicherung verfügbar macht. Auch eine Computerdokumentation wurde entwickelt, um die manuelle Dateneingabe zu verbessern.

3.5.3.3 Das durchführende medizinische Personal

In Deutschland werden die Pap-Abstriche durch niedergelassene Gynäkologen und Allgemeinmediziner vorgenommen. Oft werden die Laboruntersuchungen ebenfalls von diesen Ärzten, in den meisten Fällen in einem zytologischen Labor gemacht.

Gynäkologen erhalten eine postgraduelle Ausbildung in Zytologie. Anfang 1997 traten neue Qualitätssicherungsrichtlinien für alle Ärzte in zytologischen Laboratorien in Kraft, die ihre Fähigkeit, Pap-Testergebnisse zu befunden, verbessern sollten³¹. Dies hatte jedoch zur Folge, dass die Anzahl zytologischer Laboratorien zurückging. Die Deutsche zytologische Gesellschaft hat jedoch weitgehend die Trainingsrichtlinien in Verbindung mit den europäischen Regeln anerkannt.

Die meisten Zytologen des nationalen Screeningprogramms erhielten eine zweijährige Ausbildung in nichtgynäkologischer Zytologie und verwandten Gebieten der Pathologie. Es wurden jedoch keine zwingenden Vereinbarungen zur Weiterbildung der Spezialisten getroffen. Jedoch müssen sich alle das Screening durchführenden Ärzte verpflichten, mit den Entwicklungen auf diesem Gebiet Schritt zu halten. Dazu werden Fortbildungen in zytologischer Mikroskopie, Zytopathologie und -technologie angeboten.

Informationen zu diesem Abschnitt stammen von Schenck und Karsa⁷⁴, sowie aus der Gesundheitsberichterstattung des Bundes, Gesundheitsbericht für Deutschland⁷⁵.

3.5.4 Europäische Übersicht über das Zervixkarzinom-Screening

Eine Studie³⁵ verglich die Zervixkarzinom-Screeningprogramme innerhalb der 15 EU-Mitgliedstaaten hinsichtlich ihrer Ähnlichkeiten bzw. Unterschiede. Faktoren wie Screeningintervalle, Gründungszeitpunkte, Zielpopulationen, Organisation und Qualitätsmerkmale wurden beobachtet. In allen Programmen wurde der Pap-Test als die Screeningmethode der Wahl angewandt, wobei in den meisten Fällen diese Untersuchung mit einer allgemeinen gynäkologischen Untersuchung verbunden wird. In Italien wird der Pap-Test in Verbindung mit einer Kolposkopie verwendet.

Trotz aller Unterschiede in verschiedenen Aspekten des Zervixkarzinom-Screeningprogramms muss festgehalten werden, dass alle beteiligten Staaten eine Anzahl an Maßnahmen ergriffen haben, um eine frühzeitige Entdeckung dieser Karzinomform zu fördern.

In der folgenden Tabelle sollen Eigenschaften des Screenings der beteiligten Länder aufgelistet werden:

Tabelle 4: Art der Screeningprogramme und Zeitpunkt der Gründung in der EU.

Land	Art des Screeningprogramms	Gründungsjahr	Gründungsinstitution
Belgien	Regional (flämische Region).	1994.	1. flämische Gemeinschaft, EU (Organisation des Screenings). 2. Staatliches Krankenversicherungsinstitut (alle medizinischen Aktivitäten).
Dänemark	Regional organisiert in jedem der 15 Länder.	Drei Länder 1967 – 1968. Sieben Länder in den letzten zehn Jahren.	Staatliche Krankenversicherung.
Deutschland	Staatlich.	1971 (Westen, Ausbreitung in den Osten 1991).	Gesetzliche Krankenversicherung.
England	National organisiert, lokal koordiniert (acht Gesundheitsregionen).	1988.	Staatliches Gesundheitsservice (zentrale Verwaltung).
Finnland	Staatlich.	Pilot 1963, schrittweise Teilnahmesteigerung, auf fast bundesweit.	„Municipalities“, finnische Krebsorganisation.
Frankreich	Vier dauerhaft organisierte Programme in vier verschiedenen Bereichen (Départements).	1990.	Staatliche Krankenversicherungssysteme (keine speziellen Budgets für Screenings).
Griechenland	Zwei regionale Programme (eines in Ormylia Halkidiki, ein anderes in Messini und Ilia).	1991 (beide Programme).	EU-Jugendsekretariat, lokale Initiativen (Ormylia) EU, Anticancer fund raise, hellenisches Antikrebsinstitut, andere private Sponsoren (Messini und Ilia).
Niederlande	Staatlich.	1996.	Holländischer Versicherungsverband I. Regierung (Abteilung für Gesundheit).
Irland	Pilotphase (vor nationaler Ausdehnung).	Geplant ab dem Jahr 2000.	Staatliches Gesundheitsservice, EU.
Italien	National, regional organisiert, lokales Management, 73 aktive lokale Programme.	1980 (Florenz) 1992 (Turin) die meisten anderen 1995.	Staatlicher Gesundheitsfond.
Luxemburg	Staatlich.	1962.	Staatliche Krankenversicherung.
Österreich	Regional.	1970.	Staatlicher Onkologieplan (staatliches Gesundheitssystem).
Portugal	Regional (86 Länder der Zentralregion).	1990.	Swedish National Board of Health and Welfare (in 5 / 26 Ländern) oder € 5,95 bis € 14,88 zahlbar für die Frauen.
Schweden	Staatlich (regional organisiert).	Mitte der 60er-Jahre.	Zentralbudget von Castillia y León, EU.
Spanien	Regional (Castilla y León)	1986.	

EU = Europäische Union.

Quelle: Linoz, Riza³⁵, ARWIG eigene Übersetzung.

Es ist ersichtlich, dass in der **Art der Screeningprogramme** (staatlich / regional) und ihrer Dauer weit reichende Unterschiede existieren. Nationale Screeningprogramme wurden in sechs europäischen Staaten etabliert, in Finnland, England, Deutschland, Luxemburg, Niederlande und Schweden. In den 1960-er Jahren wurden die ersten nationalen Screeningprogramme in Finnland, Luxemburg und Schweden gestartet. Das letzte wurde zum Zeitpunkt der Publikation

im Jahr 2000 in Irland etabliert. In Österreich, Belgien, Frankreich, Griechenland, Italien, Portugal und Spanien wurden Zervixkarzinom-Screeningprogramme regional organisiert, in manchen dieser Staaten werden nur einzelne Teile der jeweiligen Bevölkerung erreicht. In Dänemark ist das Screeningprogramm in 15 lokal organisierten Zentren des Staates geregelt.

In den meisten Staaten hat bereits vor der Etablierung organisierter Screeningprogramme ein opportunistisches Screening stattgefunden. Diese **Screeningform** läuft zusätzlich zu den etablierten Screeningprogrammen, wobei es unklar ist, welches Verhältnis diese Methode zu den etablierten Formen einnimmt. Daher ist es auch unklar, inwieweit die Zielpopulation von den Screeningprogrammen erfasst wird und wie diese genau definiert ist. In manchen Ländern wird eine Registrierung verwendet (Belgien, Dänemark, Finnland, Niederlande, Schweden), in anderen Krankenversicherungs- oder nationale Gesundheitssystemlisten (England, Frankreich, Deutschland), Populationsregister oder lokale Gesundheitsadministrationen (Italien), Zensusdaten (Griechenland, Spanien) oder eine Kombination aus mehreren dieser administrativen Erfassungssysteme.

Die **Altersgruppen** der Programme reichen in den meisten Ländern von 20 bis 25 Jahren bis zu 59 bis 64 Jahren. Die europäischen Richtlinien empfehlen, dass Frauen im Alter zwischen 25 und 64 Jahren ein Zervixkarzinom-Screening in Anspruch nehmen sollten. Einige Länder akzeptieren diese Richtlinien nur hinsichtlich spezieller Risikofaktoren (sozial, kulturell). Finnland und die Niederlande sprechen eine kleinere Zielgruppe für das organisierte Screeningprogramm an, nämlich Frauen von 30 bis 60 Jahren. In Österreich werden Frauen ab 20 Jahren zum Screeningprogramm eingeladen und in Luxemburg sogar schon Frauen ab 15 Jahren, in diesen beiden Ländern bestehen keine oberen Altersgrenzen für das Screeningprogramm.

Diese Faktoren werden in der folgenden Tabelle aufgelistet:

Tabelle 5: Einladung zum Zervixkarzinom-Screening und Einschlusspopulation innerhalb der europäischen Union.

Land	Art der Einladung	Zielbevölkerung	Altersgruppen	Anteil organisiert	Anteil opportunistisch
Belgien	Persönliche Einladung.	Knapp 1,6 Millionen (flämische Region).	25 - 64	82,3 % organisiert und opportun.	Unbekannt.
Dänemark	Persönliche Einladung.	1,3 Millionen.	23 - 59	90 %	Unbekannt.
Deutschland	Aufforderung des praktischen Arztes und Personal im Zuge von Routinebesuchen.	25 Millionen (West). 5,5 Millionen (Ost).	≥ 20	46 - 50 %	Unbekannt.
England	Persönliche Einladung.	Beinahe 14 Millionen.	20 - 64	84 %	4 %
Finnland	Persönliche Einladung.	256.616	30 - 60	89,5 %	Unbekannt (93 % für alle Abstriche).
Frankreich	Persönliche Einladung.	477.301 (17 Millionen national).	25 - 65 (nicht in allen Projekten).	22 - 69 %	Unbekannt.
Griechenland	Persönliche Einladung (beide Programme).	Unter 17.000 (Ormylia).	25 - 64 (beide Programme).	87,8 % (Ormylia).	Unbekannt.

Fortsetzung Tabelle 5

Niederlande	Persönliche Einladung zum Besuch eines praktischen Arztes.	3,6 Millionen.	30 - 60	80 %	Inklusive opportunistisch.
Irland	Einladung und Selbstregistrierung.	67.000 (Pilot) 1 Million (national erreichbar).	25 - 60	Nicht verfügbar.	Unbekannt.
Italien	Persönliche Einladung.	230.000 (Florenz) 280.000 (Turin) 5,3 Millionen (aktiv organisierte Programme)	25 - 64	74 % (Turin) annähernd 70 % in den meisten aktiven Programmen.	Beide inklusive opportunistisch.
Luxemburg	Über praktische Ärzte und Gynäkologen.	180.982	< 15	38,92 %	
Österreich	Persönliche Einladung.	120.000	≥ 20	85 %	60 %
Portugal	Einladung über praktische Ärzte.	292.000	20 - 64	51 %	Nicht verfügbar.
Schweden	Persönliche Einladung.		20 - 59	50 - 70 % (in den meisten Bundesländern).	
Spanien	Persönliche Einladung.	627.788	25 - 65	41,47 %	Unbekannt.

Quelle: Linoz, Riza³⁵, ARWIG-eigene Übersetzung.

Weitere Unterschiedlichkeiten der Screeningprogramme beziehen sich auf den professionellen Hintergrund der **Abstrichnehmer**. In zwei Ländern werden die Abstriche nur von Allgemeinärzten genommen (Dänemark, Portugal), Allgemeinärzte und andere Ärzte, vor allem Gynäkologen werden in Belgien, Frankreich, Deutschland, Luxemburg und Spanien eingesetzt. Andere Gesundheitsberufe wie z.B. Assistenten, Hebammen oder geschulte Krankenschwestern werden in Finnland und Schweden eingesetzt. Kombinationen aus Allgemeinärzten und anderen Gesundheitsspezialisten nehmen in England, Griechenland, Irland, Italien und in den Niederlanden die Abstriche. In Österreich sind es ausschließlich Gynäkologen.

Die **Screeningintervalle** nach einem negativen Testresultat betragen in den meisten Ländern drei Jahre (Belgien, Dänemark, Frankreich, Italien, Spanien, Schweden). Andere Nationen bevorzugen einen längeren Abstand (drei bis fünf Jahre): England, Finnland, Irland, die Niederlande. In Österreich, Deutschland und Luxemburg ist ein viel kürzerer Zeitraum gebräuchlich (ein Jahr). In Griechenland und Portugal wird eine spezielle Art des Screeningintervalls verwendet: Die erste Wiederholung des Abstrichs erfolgt nach zwölf Monaten, wenn dann das Ergebnis negativ ist, folgen die weiteren Untersuchungen in zwei bis drei Jahren. Negative Testergebnisse werden per Informationsbrief an die Frauen versendet, unklare Testergebnisse werden durch den Abstrichnehmer mitgeteilt, da dieser die Verantwortung für die Information und die folgenden Untersuchungen hat.

Alle genannten Screeningprogramme folgen in Verbindung mit der **Qualitätssicherung** speziellen **Richtlinien** verschiedener Quellen. Die Länder mit nationalen Screeningprogrammen folgen nationalen Regeln: Dänemark, England, Finnland, Frankreich, Deutschland, Irland, Italien, die Niederlande, Portugal und Schweden. In Österreich sind nationale Leitlinien der österreichischen Gesellschaft für Pathologie, der österreichischen Gesellschaft für Zytologie und der staatlichen Krankenversicherung in Gebrauch. In manchen Fällen werden nationale Regelungen durch regionale ergänzt. In Luxemburg wurden seit 1990 spezielle Kriterien nach

dem National Health Fund zur Regulierung des Screeningprogramms eingehalten. Die **europäischen Richtlinien** werden allgemein anerkannt, enthalten jedoch nicht spezielle Inhalte und / oder Konditionen mancher genannter Staaten. In den meisten Ländern wurden die europäischen Richtlinien gemeinsam mit den bestehenden nationalen bzw. regionalen Regelungen eingebaut. Da in Griechenland keine nationalen oder regionalen Leitlinien existieren, gelten dort allein die europäischen Regeln.

Tabelle 6: Organisation des Zervixkarzinom-Screenings innerhalb der europäischen Union.

Land	Abstrichnehmer	Screening-intervall	Mitteilung des Resultats	
			Normal	Verdächtig
Belgien	Gynäkologen, Allgemeinärzte.	3 Jahre.	Bericht durch den Abstrichnehmer.	Bericht durch den Abstrichnehmer.
Dänemark	Allgemeinärzte.	3 Jahre.	Direkt an die Frauen beim Besuch des Arztes.	Bericht durch den Allgemeinarzt.
Deutschland	Niedergelassene Gynäkologen und Allgemeinmediziner.	1 Jahr.	Durch den Abstrichnehmer.	E-Mail oder Telefon durch den Abstrichnehmer.
England	Allgemeinärzte oder allgemein praktizierende Krankenschwestern	3 - 5 Jahre.	Bericht durch den Abstrichnehmer.	Bericht durch den Abstrichnehmer.
Finnland	Trainierte Krankenschwestern (Hebammen)	5 Jahre.	Brief direkt an die Frauen.	Wenn möglich per Telefon, oder per E-Mail.
Frankreich	Gynäkologen, Allgemeinärzte	3 Jahre (nach 2 negativen Abstrichen).	Nichts Bestimmtes.	Nichts Bestimmtes.
Griechenland	Gynäkologen (Ormylia), Gynäkologen, geschulte Landärzte und Hebammen (Messini und Illia).	1 Jahr nach dem ersten Abstrich, dann alle 3 Jahre (Ormylia) 2 Jahre (Messini und Illia).	Brief direkt an die Frauen (beide Programme).	Telefon und persönliches Treffen mit dem Screeningarzt (Ormylia), Telefon oder Hausbesuch (Messini und Illia).
Niederlande	Allgemeinärzte und deren Assistenten.	5 Jahre.	Über den Allgemeinarzt.	Über den Allgemeinarzt.
Irland	Allgemeinärzte, Familienplanungs- und Gemeindegemeinden, Krankenhäuser.	5 Jahre (Pilot).	Brief direkt an die Frauen.	Anweisung, den Abstrichnehmer zu kontaktieren.
Italien	Midwives (hauptsächlich), Gynäkologen.	3 Jahre.	Verschiedenes: meistens ein Brief, direkt an die Frauen.	Verschiedenes.
Luxemburg	Allgemeinärzte und Gynäkologen.	1 Jahr.	Nichts Bestimmtes.	Nichts Bestimmtes.

Fortsetzung Tabelle 6

Österreich	Gynäkologen.	1 Jahr.	E-Mail oder Telefon durch den Abstrichnehmer.	E-Mail oder Telefon durch den Abstrichnehmer.
Portugal	Allgemeinärzte.	1 Jahr nach dem ersten normalen Abstrich, dann alle 3 Jahre.	Brief durch den Allgemeinarzt.	Brief durch den Allgemeinarzt.
Schweden	Geschulte Hebammen.	3 Jahre.	Brief direkt an die Frauen (6 Bundesländer informieren die Frauen nicht bei negativem Testresultat).	Über die lokale gynäkologische Klinik.
Spanien	Familienärzte.	3 Jahre.	Brief durch den Arzt.	Durch den Arzt.

Quelle: Linoz, Riza 2000³⁵, ARWIG-eigene Übersetzung.

3.6 Ethische Aspekte

Die Medizinethik hat in den letzten Jahrzehnten an Bedeutung gewonnen, da die Medizin derartige technische Fortschritte gemacht hat, dass es notwendig wurde, die menschlichen Bedürfnisse der Patienten stärker zu berücksichtigen. Vor allem in den Bereichen des genetischen und des vorgeburtlichen Screenings treten ethische Aspekte in den Vordergrund, aber auch in jedem Screening, das zur Früherkennung von Erkrankungen mit lebensbedrohlicher Bedeutung eingesetzt wird³⁸.

Im Fall des Screenings auf Zervixkarzinom mittels Pap-Test wird mit einem Verfahren, das in breiter Form Anwendung findet, jedoch keine genaue Vorhersage für die Erkrankung liefern kann, nach Frühformen eines Karzinoms gesucht. Bei einer Einstufung von PAP-V müssen weitere Untersuchungen durchgeführt werden.

3.6.1 Grundsätze der medizinischen Ethik

Die Grundlagen jeder medizinischen Untersuchung und Behandlung sind das Selbstbestimmungsrecht sowie die Freiheit der Entscheidung der untersuchten bzw. behandelten Person. Von alters her hat daher jeder Arzt die Pflicht, die ihn konsultierenden Personen über alle Aspekte der Untersuchung bzw. Behandlung aufzuklären. Dies ist bereits in den ursprünglichen Persönlichkeitsrechten des deutschen Grundgesetzes (Strafrecht, 1871) verankert.

Die Aufklärungspflicht umfasst die Information über alle Gesichtspunkte der Untersuchung bzw. Behandlung, wie Art und Ablauf des Eingriffs, Sinn und Zweck, sowie verschiedene Risikoaspekte. Nach dieser umfassenden medizinischen Aufklärung (die in einer allgemein verständlichen Sprache erfolgen muss) ist die Patientin im Allgemeinen imstande, eine Entscheidung über Inanspruchnahme oder Ablehnung der Untersuchung bzw. Behandlung zu treffen (für unzurechnungsfähige oder minderjährige Personen gelten eigene Regelungen).

Im Fall einer Screeninguntersuchung werden zum Großteil gesunde Personen einem Testverfahren unterzogen, wobei der Zweck des Screeningprogramms ist, Personen mit der gesuchten Diagnose herauszufiltern; das Ziel der konsultierenden Personen ist vor allem, Sicherheit über ihren Gesundheitszustand zu erlangen. Diese unterschiedliche Auffassung zwischen Ärzten und Laien ist Grundlage für einen verantwortungsvollen Umgang der Ärzte mit

Diagnosen, vor allem im Hinblick auf psychologische und soziale Konsequenzen für die Betroffenen.

Ein Aufklärungsgespräch über eine Screeninguntersuchung sollte Informationen darüber enthalten, welchem Zweck die Untersuchung dienen soll, welche Risiken für den Untersuchten dabei eine Rolle spielen - hier sollten vor allem die Raten falsch-negativer und falsch-positiver Testergebnisse genannt werden - sowie die Wahrscheinlichkeit für den Untersuchten, mit einem solchen Ergebnis konfrontiert zu werden. Auch die Behandlungsmöglichkeiten im Fall einer Diagnose sollten erwähnt werden. Finanzielle Aspekte der Untersuchung sollten gegebenenfalls erwähnt werden.

3.6.2 Evidenz ethischer Aspekte bei Screeninguntersuchungen

Bezüglich der Ethik in Screeninguntersuchungen waren wenig Arbeiten zu finden. Es konnten jedoch zwei Übersichtsarbeiten zum Thema der falsch-negativen Testresultate und zum Thema der ethischen Herausforderungen von Screeningprogrammen und -tests gefunden werden. Die Inhalte werden im Folgenden dargestellt.

3.6.2.1 Falsch-negative Resultate

In einer Übersichtsarbeit, die 23 Studien zu Konsequenzen falsch-negativer Screeninguntersuchungsergebnisse zum Inhalt hatte, enthielten 13 Untersuchungen quantitative Informationen der medizinischen Konsequenzen, acht Publikationen Informationen über psychologische Konsequenzen und zwei Arbeiten Hinweise zu ökonomischen Folgen⁴⁵. Die Autoren des vorliegenden Berichts beurteilen die Evidenz dieser Studien als niedrig, es erwies sich dass falsch-negative Testresultate einen großen rechtlichen Einfluss haben.

Die Raten falsch-negativer Testergebnisse in retrospektiven Studien zum Pap-Test werden mit 15 % bis 22 % angegeben²². Diese Zahlen müssen jedoch unter dem Gesichtspunkt gesehen werden, dass die Sensitivität nicht genau geschätzt werden konnte, da in manchen Studien die Anzahl der richtig-negativen nicht bekannt war und einige Studien nicht die reale Screeningpraxis aufzeigen konnten²⁶.

In der erwähnten Übersichtsarbeit⁴⁵ stellte sich heraus, dass selbst gut etablierte und qualitativ hochwertige Screeningprogramme einen Anteil falsch-negativer Testergebnisse aufweisen. Dies kommt dadurch zustande, dass kein in einem Screeningprogramm verwendetes Testverfahren 100 % Sensitivität hat. Diese Mängel müssen in Beziehung zu dem Nutzen des Screenings gesetzt werden. In einer früheren Übersichtsarbeit wurden bereits die Schäden und Konsequenzen falsch-negativer Testresultate aufgezeigt⁴⁴.

Medizinische Konsequenzen falsch-negativer Resultate

Der Nachteil eines falsch-negativen Testergebnisses liegt vor allem in der falschen Sicherheit bezüglich der Gesundheit der Patienten, und den damit verbundenen Verhaltensweisen (Ernährung, Gesundheitsbewusstsein) der Betroffenen, sowie der späten Erkennung und Behandlung der Erkrankung. Im Fall von Krebserkrankungen tritt häufig der Fall ein, dass die Erkrankung später symptomatisch wird, meistens in einem Stadium, in dem eine invasivere und kostenintensivere Behandlung eingesetzt werden muss, die oftmals nicht so erfolgreich sein kann, da sie zu spät begonnen wurde.

Die Evidenz solcher Ereignisse ist schlecht belegt. Es konnte jedoch gezeigt werden, dass Verzögerungen bei der Inanspruchnahme weiterer Untersuchungen bei Brustkrebs durch die falsche Sicherheit aufgrund falsch-negativer Testergebnisse ausgelöst wurden^{1, 67}. Eine retrospektive Studie über die Entwicklung von tastbarem Brustkrebs von Frauen mit negativer Mammographie zeigte, dass die Verzögerung der Behandlung zu vermehrter Beteiligung axillärer Lymphknotenkarzinome bei Operationen beitrug².

Allgemein kann gesagt werden, dass in einigen Screeningprogrammen die Personen mit Intervallkarzinomen (die falsch-negative Testresultate beinhalten oder innerhalb eines Screenings-

intervalls auftreten) eine schlechtere Überlebensrate aufwiesen als Personen von nicht gescreenten Bevölkerungsgruppen¹³.

Psychologische Konsequenzen

Als größtes Problem der falsch-negativen Testresultate für die betroffene Person wird vor allem die falsche Sicherheit, als gesund eingestuft worden zu sein, genannt. Oft führt diese Sicherheit zu verspäteten weiteren Untersuchungen und zu Verhaltensweisen, die der Gesundheitsvorsorge nicht entsprechen⁶⁵. Dies kann auch als ein Resultat der widersprüchlichen medizinischen- und Laienperspektive über den Sinn von Screenings betrachtet werden. Das medizinische Ziel ist, Personen mit der Zielerkrankung herauszufinden, während die Laien Sicherheit über ihre Gesundheit haben wollen³⁸.

Eine andere psychologische Folge der falsch-negativen Testergebnisse könnte in dem Vertrauensverlust in die medizinische Vorsorge gesehen werden, die Verwandte und Freunde von Opfern falsch-negativer Testergebnisse erleiden. Es muss jedoch gesagt werden, dass dies aufgrund der tatsächlich sehr geringen Aufklärung über die Raten falsch-negativer Testergebnisse im Vorfeld von Screeninguntersuchungen eine gewisse Berechtigung hat. Diese Tatsachen können durch die Untersuchungen der Akzeptanz der Geburten von Kindern mit Down-Syndrom erhärtet werden. Die Eltern, die kein vorgeburtliches Screening in Anspruch nahmen, konnten die Störung ihres Kindes besser akzeptieren, als die Eltern, die ein solches Kind nach einem falsch-negativen Testergebnis bekamen^{24, 53}.

Die Literatur über die psychischen Folgen falsch-negativer Testresultate belegte durch hohe Übereinstimmung den starken Einfluss auf das Vertrauen der Bevölkerung in Screeningprogramme. Es konnte gezeigt werden, dass die Ergebnisse von Screeningprogrammen umfassender sind als alle anderen Typen der Gesundheitsversorgung. Jedoch könnte ihre Sicherheit noch erhöht werden.¹³

3.6.2.2 Falsch-positive Resultate

Bei Betrachtung der medizinethischen Literatur konnte noch weniger über die Folgen falsch-positiver Testergebnisse gefunden werden. Falsch-positive Testergebnisse kommen durch fehlende Testgüte der verwendeten Verfahren (zu geringe Spezifität) zustande, und sind ebenso wie falsch-negative Testergebnisse in allen Screeningprogrammen zu unterschiedlichen Anteilen vorhanden.

Von Seiten der Ethik muss jedoch betont werden, dass sich präventivmedizinisches Handeln für betroffene Personen nicht zusätzlich belastend, sondern konkret entlastend auswirken soll. Im Fall der falsch-positiven Testresultate werden gesunde Menschen mit einer oftmals lebensbedrohenden Diagnose konfrontiert, die im derzeit oft angewendeten Screeningrahmen nicht schonend, sondern direkt ausgesprochen wird, und sich die betroffenen Menschen mit ihrer psychischen Situation allein abfinden müssen. Die medizinische Seite der Screeningprogramme hat in der Regel zu wenige Ressourcen für eine ausreichende Arzt-Patienten-Kommunikation zur Verfügung.

Aufgrund der gefundenen Testergebnisse werden die Betroffenen weiteren (in diesen Fällen unnötigen) Untersuchungen zur Abklärung zugeführt. Dies stellt eine zusätzliche körperliche und psychische Belastung für die Betroffenen dar. In manchen Fällen kommt es zu Operationen, obwohl die behandelte Person gesund ist. Selbst wenn sich nach der ersten weiteren Untersuchung der Irrtum herausstellt, und die Diagnose rückgängig gemacht werden kann, durchleidet die behandelte Person bis zu diesem Zeitpunkt ein hohes Maß an Verunsicherung und Verängstigung. Außerdem sind damit für das Gesundheitssystem hohe Kosten verbunden, ohne irgendeinen positiven Zweck zu erreichen. In diesen Fällen wird durch die Präventivmaßnahme mehr geschadet als genützt.

In den Übersichtsarbeiten zum Thema ethischer Konsequenzen von Screeningprogrammen wird auf die Situation der Betroffenen falsch-positiver Testergebnisse nicht eingegangen^{38, 45}. Es

liegt jedoch auf der Hand, dass diese Personen vor allem unter psychischen Belastungen zu leiden haben, was zumindest psychosoziale Überlegungen im Rahmen von Screeninguntersuchungen zur Folge haben müsste.

3.6.2.3 Konsequenzen im ethischen Sinne

Eine Arbeit befasste sich mit der Rolle der Ethik im Lösen von den eben besprochenen Einflüssen die durch Sensitivität und Spezifität der Screeningtests ausgelöst werden⁵⁸. Die Autoren schlagen vor, dass die Höhe der Spezifität und Sensitivität der Testverfahren Konsequenzen für ihre Verwendung als Diagnosemethode haben sollte. Dabei muss zwischen hoher Sensitivität und hoher Spezifität unterschieden werden.

Eine **hohe Spezifität** wäre wünschenswert wenn:

- Die Diagnose mit großer Angst oder Stigmatisierung verbunden ist.
- Weitere Untersuchungen zeitaufwändig, schmerzhaft und / oder teuer wären.
- Es möglich wäre, die Fälle durch andere Methoden zu entdecken, bevor es „zu spät“ zum Behandeln ist.
- Eine Behandlung, speziell wenn sie schmerzvoll oder teuer ist, sollte ohne weitere Untersuchungen angeboten werden.

Im Gegensatz dazu wäre eine **hohe Sensitivität** wünschenswert wenn:

- Negative Konsequenzen einer fehlenden Diagnose für das Individuum zu erwarten sind, z.B. wenn spätere Behandlung signifikant schlechter wären als frühe.
- Negative Konsequenzen bei einer fehlenden Diagnose für die Gesellschaft, z.B. bei einer ernsthaften ansteckenden Krankheit.
- Die Diagnose sollte durch eine andere Untersuchungsmethode bestätigt werden, damit die Ängstigungsperiode kurz gehalten wird, bzw. die korrekte Diagnose gestellt werden kann, bevor die Behandlung beginnt.

Bei manchen Screeningtests ist die Wahl der vorrangigen Perspektive zwischen Sensitivität und Spezifität schwierig. Dies soll am Beispiel des HIV-Tests (Human-Immundefizienz-Virus-Test) verdeutlicht werden: Die Diagnose ist mit hoher Ängstigung und Stigmatisierung verbunden, es gibt jedoch keine anderen Untersuchungsmethoden, die Diagnose zu bestätigen. Die Behandlung ist teuer, mit vielen Nebenwirkungen und möglicherweise nicht effektiv. Es bestehen jedoch hohe negative Konsequenzen fehlender Diagnosen für die Gesellschaft.

Die Autoren bieten eine Kosten-Nutzen-Aufstellung im ethischen Sinn an, indem sie Kosten-Nutzen-Zuordnungen für die Betroffenen vornehmen:

Tabelle 7: Kosten und Nutzen in Verbindung mit tatsächlich positiven, falsch-positiven, falsch-negativen und tatsächlich negativen Testergebnissen.

Tatsächlich positiv		Falsch-positiv	
Kosten	Nutzen	Kosten	Nutzen
Steigen möglicherweise während Lebenserwartung nicht an, wenn die Behandlung ohne Effekt ist, jedoch haben Patienten möglicherweise länger zu leben, wenn die Diagnose früher gestellt wird (lead time bias). Ängstigung, Stigmatisierung reduziert die Lebensqualität z.B. durch Nebeneffekte oder Behandlungskomplikationen.	Frühere Behandlung (billiger, angenehmer, bessere Prognose), Nutzen durch „Krankenrolle“ (entschuldigte soziale Verantwortung begünstigt Suche und Einverständnis mit der Behandlung). Erklärung der Symptome: Wissen über die Herkunft der Symptome, so dass die Behandelten ihre Angst abbauen können.	Unnötige Ängstigung und Stigmatisierung bis zur korrekten Diagnose verursacht unnötige Untersuchungen und Behandlungen (zeitaufwändig, teuer, schmerzhaft, nebenwirkungs- und komplikationsreich). Schleichende Zweifel: Patienten tendieren dazu, Ärzten zu erlauben. Möglicherweise ist es schwierig, den Patienten zu versichern, dass die erste Diagnose falsch war.	Fördert die Fortsetzung eines ungesunden Lebensstils.

Fortsetzung Tabelle 7

Falsch-negativ		Tatsächlich negativ	
Kosten	Nutzen	Kosten	Nutzen
Falsche Sicherheit ermöglicht legitimierten ungesunden Lebensstil, spätere Behandlung (unangenehm, teurer, schlechte Prognosen) lässt die Kosten ansteigen, bei geringem Outcome.	Eingesparte Ängstigung bei Behandlung ohne Nutzen.	Ängstigung während Erwartung der Testergebnisse (worried well) begünstigt ungesunden Lebensstil, z.B. „wenn mein Lungenröntgen in Ordnung ist, dann kann mir Rauchen nichts anhaben“.	Sicherheit.

Quelle: Shickle, Chadwick⁵⁸.

3.6.3 Diskussion der ethischen Aspekte

Grundlage aller Screeningprogramme ist die möglichst frühzeitige Erkennung von im Frühstadium gut behandelbaren Erkrankungen, die einen hohen Einfluss auf die Lebensdauer und / oder -qualität haben. Dabei werden gesunde Menschen untersucht, wobei das medizinische Ziel ist, möglichst viele Personen mit der Diagnose zu identifizieren und einer geeigneten Behandlung zuzuführen. Die untersuchten Personen wollen hinsichtlich ihres Gesundheitszustands Sicherheit.

Grundsätzlich muss gesagt werden, dass kein Screening ohne Risiken ist. Sie bestehen hauptsächlich aus Fehlinterpretationen der Screeningtestresultate wie falsch-negativer und falsch-positiver Testergebnisse, die in einem Fall verzögerte Behandlung einer bereits bestehenden Erkrankung, im anderen Fall unnötige Untersuchungen und hohe Verunsicherung von gesunden Personen zur Folge haben³². Daraus sind Konsequenzen zur Verwendung von Testverfahren mit entweder hoher Sensitivität oder hoher Spezifität ableitbar. Es müssen im jeweiligen Screeningfall die Konsequenzen des Einsatzes des Testverfahrens berücksichtigt werden.

Das persönliche Recht der Personen, die an einem Screening teilnehmen, auf Aufklärung über den Sinn und Zweck, die Risiken und die Folgen eines solchen Programms und die darin verwendeten Testverfahren ist in den Persönlichkeitsrechten gesetzlich verankert, wird aber in der ärztlichen Praxis aus Zeitgründen selten oder unzureichend durchgeführt. Es bleibt jedoch anzumerken, dass die Einwilligung der Personen zu den Untersuchungen ethisch nur bei ausreichender Information über die Risiken rechtlich vertretbar ist, obwohl in der Praxis viele teilnehmende Personen kaum Aufklärung erhalten und im Fall falsch-positiver oder falsch-negativer Testresultate aufgrund Ressourcenmangels kaum entsprechende Begleitung erhalten.

3.7 Juristische Grundlagen

Die Richtlinien des Bundesausschusses der Ärzte und Krankenkassen, in der Fassung vom 26. April 1976, zuletzt geändert am 21. Juni 2002 beinhalten Richtlinien über die Früherkennung von Krebserkrankungen („Krebsfrüherkennungs-Richtlinien“, veröffentlicht im Bundesanzeiger Nr. 186 (S 23077) vom 5. Oktober 2002, in Kraft getreten am 1. Oktober 2002).

An dieser Stelle sollen die Inhalte dieser Richtlinien, das Zervixkarzinom betreffend, wiedergegeben werden:

A Allgemeines

1. Die nach diesen Richtlinien durchzuführenden ärztlichen Maßnahmen dienen

a) bei Frauen

Der Früherkennung von Krebserkrankungen des Genitales, vom Beginn des 20. Lebensjahres an sowie zusätzlich der Brust und der Haut vom Beginn des 30. Lebensjahres an sowie zusätzlich des Rektums und des übrigen Dickdarms vom Beginn des 50. Lebensjahres an,

[...]

B. Früherkennungsmaßnahmen bei Frauen

Die Maßnahmen zur Früherkennung von Krebserkrankungen des Genitales, der Brust, der Haut, des Rektums und des übrigen Dickdarms bei Frauen umfasst folgende Leistungen:

1) Klinische Untersuchungen

Vom Beginn des 20 Lebensjahres an:

- Gezielte Anamnese
- Spiegeleinstellung der Portio
- Entnahme von Untersuchungsmaterial von der Portio-Oberfläche und aus dem Zervikalkanal
- Fixierung des Untersuchungsmaterials für die zytologische Untersuchung
- Bimanuelle gynäkologische Untersuchung

[...]

In diesem Gesetzestext werden keine Angaben zur Häufigkeit der Untersuchungen gemacht (Anmerkung der Autoren).

3.8 Verfahrensbewertung des Papanicolaou-Tests

Da es keine randomisierten Studien vor Einführung des Pap-Test-Screenings gegen Zervixkarzinom gab, werden behelfsweise Studien zur Beurteilung hinsichtlich der folgenden Studieninhalte durchgeführt:

- Vergleich von Inzidenz und Mortalität.
- Inzidenz von Oberflächenkarzinom im Laufe der Zeit.
- Fall-Kontroll-Studien.
- Empirische Analysen.
- Simulationsstudien.

Beurteilungsgegenstände für das Screening sind:

1. Epidemiologie: Erkrankungswahrscheinlichkeit der Lebensspanne (Perioden-Prävalenz).
2. Tumorprogression: Progressionstempo der Läsionen zur Vorstufe weiter zum invasiven Karzinom (length time bias).
3. Testcharakteristika des Pap-Tests.
4. Teilnahmeverhalten.

Der Wirkungsgrad des Pap-Test-Screenings ist vor allem abhängig von der Teilnahmerate der Zielbevölkerung, vom Zeitintervall zwischen den Abstrichen, von der Laborqualität sowie von der Nachverfolgung der abnormalen Abstriche¹². Während die meisten dieser Faktoren von der Organisation und dem Monitoring des Screenings abhängen, sind die Zielgruppen und Screeningintervalle von der altersspezifischen Prävalenz, von präinvasiven und mikroinvasiven Erkrankungen sowie von der Länge der entdeckbaren präklinischen Phase^{9, 29} abhängig.

3.8.1 Testgütekriterien des konventionellen Pap-Tests sowie der neuen Technologien

„Der konventionelle Pap-Test besitzt - gestützt auf einige wenige Studien, die schwere Informationsverzerrungen vermieden - eine entgegen den Erwartungen überraschend geringe Sensitivität von 51 Prozent (95 % Konfidenzintervall von 37 % bis 66 %) und eine Spezifität von 98 Prozent (95 % Konfidenzintervall von 97 % bis 99 %)“ (Zitat McCrory et al.³⁷). Die von den zitierten Autoren erstellte Metaanalyse aus 84 Einzelstudien erklärte anhand spezieller Me-

thodik, die die Abhängigkeit von Sensitivität und Spezifität berücksichtigt, eine möglichst unverzerrte Punktschätzung für die Testcharakteristika.

Die Schwankungsbreite der Testkenngrößen ist beträchtlich. So betrug die Spannweite der Sensitivität in den 84 Einzelstudien 93 Prozentpunkte (6 % bis 99 %), ebenso die Spezifität 93 Prozentpunkte (6 % bis 99 %). Das bedeutet, dass der Pap-Test zwar in hohem Maße nur erkrankte Frauen (Zervixdysplasien und -karzinome) als positiv anzeigt, andererseits, aber die Hälfte übersieht. Diese Testgüte bezieht sich auf einen einzelnen Test.

Die Reproduzierbarkeit eines Tests ist ebenfalls ein wichtiges Kriterium seiner Güte. Dies zur Reliabilität (Zuverlässigkeit) zählende Charakteristikum bringt zum Ausdruck, wie reproduzierbar seine Ergebnisse sind. Die wachsende Besorgnis bezüglich technischer Unzulänglichkeiten wird durch die als moderat bezeichnete Reproduzierbarkeit untermauert. Die zytologischen Interpretationen wiesen einen Kappa (Übereinstimmungsparameter) von 0,46 auf. Diese interpretative Variabilität ist eine der wesentlichen Quellen für falsch-negative oder falsch-positive Befunde⁶¹.

Um die Sensitivität des Pap-Tests zu steigern, sind verschiedene technische Modifikationen vorgeschlagen worden. Die Dünnschichtpräparation (Synonym: flüssigkeitsbasierte Zytologie oder Monolayer-Zytologie) wurde entwickelt, um abnahme- und präparationsbedingte Fehler, die für etwa zwei Drittel der falsch-negativen Befunde beim Pap-Test verantwortlich gemacht werden, zu reduzieren. Durch die Verbesserung der Qualität der Präparate wird eine Steigerung der Sensitivität bzw. eine Reduktion des Anteils falsch-negativer Ergebnisse erwartet. Ob die dafür am Markt angebotenen Geräte (AutoCyte PREP (Tri Path Imaging) oder ThinPrep (Cytoc Corp) tatsächlich eine gesteigerte Testgenauigkeit gegenüber den herkömmlichen Pap-Tests aufweisen, ist nach wie vor umstritten.

Als Mindestforderung zur Beurteilung der Leistungsfähigkeit eines Tests gilt die Bestimmung der Testgüteparameter Sensitivität und Spezifität. Gerade darin besteht jedoch das Problem der diagnostischen Studien zur FBC. Zumeist berichten die Arbeiten über Findungsraten, also Kriterien der technischen Leistungsfähigkeit, die jedoch keine Aussagen über die Sensitivität und Spezifität zulassen. Genaue, detaillierte und kritische Betrachtungen der FBC sind einer vom New Zealand Health Technology Assessment Clearing House (NZHTA) durchgeführten Evaluation zu entnehmen^{4, 5}.

Insgesamt wird bei den neuen Technologien der Gewinn an Lebenserwartung mit fünf Stunden bis 1,6 Tagen angegeben, die Kosten pro Untersuchung werden mit 30 US-Dollar beschrieben⁶.

Eine Arbeit besagt jedoch, dass die Anwendung der FBC sowohl die richtig-positiven als auch die falsch-positiven im Vergleich zum Pap-Test erhöht. Die Möglichkeit, dass FBC die Sensitivität erhöht, jedoch die Spezifität senkt, sollte vor allem bei Frauen mit einem geringen Risiko für Zervixkarzinom zur Vorsicht bei der Annahme dieser Technologie mahnen. Bei Frauen mit normalen Resultaten bei früheren Pap-Tests (geschätzte Prävalenz: 0,05 %) wären mehr als 1.800 FBC nötig, um einen weiteren richtig-positiven Fall zu entdecken. Mehr als 50 zusätzliche falsch-positive Resultate würden jedem einzelnen richtig positiven Fall gegenüberstehen. In einer Bevölkerung mit einer Prävalenz von 3 % von Zervixerkrankungen würden 300 FBC-Tests nötig sein, um einen zusätzlichen abnormalen Fall zu entdecken, - dies wäre durch „nur“ 16 falsch-positive zusätzliche Fälle begleitet. Die Autoren weisen aufgrund der erhöhten Kosten dieser neuen Technologien darauf hin, dass vor allem Frauen mit geringem Einkommen weniger das Screening in Anspruch nehmen würden und im Fall eines breiteren Einsatzes dieser Technologien die paradoxe Situation entstünde, dass trotz des verbesserten Tests die Sterblichkeit an Zervixkarzinom steigen würde. Die FBC weist laut dieser Untersuchung eine Sensitivität von 90 % bei einer Spezifität von 85 % auf⁶².

3.8.2 Screeningintervall

Offen bleibt die Frage der Altersspanne, innerhalb derer auf Zervixkarzinom gescreent werden sollte. Es gibt Studien die nahe legen, dass die Wahrscheinlichkeit, dass es sich bei einem entdeckten Fall um eine Fehldiagnose handelt (falsch-positive Diagnose) im Alter ab 50 Jahren (postmenopausale Frauen) höher ist, als die Wahrscheinlichkeit einer korrekten Diagnose (richtig-positiv)⁵⁰.

Die empfohlenen Screeningmaßnahmen gegen Zervixkarzinom variieren stark zwischen verschiedenen Ländern hinsichtlich der gescreenten Altersspanne, Screeningintervall und daher der Gesamtzahl der innerhalb einer Lebensspanne vorgesehenen Pap-Abstriche.

Unabhängig von den verschiedenen nationalen Bevölkerungen und Annahmen in zahlreichen Modellrechnungen wird ein Screeningintervall von einem Jahr übereinstimmend als sehr ineffizient betrachtet. Die zwei Hauptgründe für die Ineffizienz sind die Kosten durch die große Anzahl der erforderlichen Screeningtests und die Kosten der Behandlung von entdeckten Dysplasien, die vor dem nächsten Screening verschwunden wären, würde ein längeres Screeningintervall angewendet werden. Screeningintervalle von mehr als zehn Jahren werden mit einem eingeschränkten Nutzen zur Zielerreichung, der Senkung der Inzidenz und Mortalität von Zervixkarzinom betrachtet. Sogar ein einziger Test innerhalb einer Lebensspanne (zwischen 36 und 45 Jahren) kann signifikante Reduktionen der Mortalität erzielen, wie anhand einer Reihe von Entwicklungsländern festgestellt wurde¹⁹.

Finnland hat anhand von 45.572 zytologischen Abstrichen das organisierte, mit Fünfjahresintervallen durchgeführte Screeningprogramm, einer Evaluation unterzogen. Aufgrund des anhand von Risikomaßzahlen analysierten Tempos der Tumorprogression zeigte sich, dass das relative Risiko einer präinvasiven Läsion nach einem negativen Abstrichtest für mehr als fünf Jahre verringert ist. Aus diesem Grund wurde das Screeningintervall von fünf Jahren als passend erachtet⁶⁹.

Eine umfassende Analyse des Screeningintervalls wurde 1986 von The International Agency for Research on Cancer (IARC, einer Teilorganisation der WHO) auf der Basis des damaligen Kenntnisstands vorgenommen. Die wesentlichsten Ergebnisse sind die Folgenden:

Die prozentuelle Reduktion der Inzidenz des Zervixkarzinoms bei Frauen im Alter von 35 bis 64 Jahren betrage bei einem einjährigen Screening 93,5 %, bei einem Screening alle zwei Jahre 92,5 %, alle drei Jahre 90,8 %, alle fünf Jahre 83,6 % und alle zehn Jahre 64,1 %. Es wird explizit darauf hingewiesen, dass der Unterschied zwischen jährlich und alle drei Jahre als verschwindend zu betrachten ist. Diese Arbeit berücksichtigt jedoch

1. Keine Non-Compliance.
2. Die Inzidenzreduktion basiert auf norwegischen Inzidenzdaten zu einem Zeitpunkt, zu dem das Screening noch nicht implementiert war.
3. Die absolute Inzidenzreduktion basiert auf der in der Bevölkerung zugrunde liegende Inzidenz, die wiederum von der Art der Untersuchungsorganisation abhängig ist.
4. Darüber hinaus hat die IARC keine Screeningkosten berücksichtigt, daher dem Nutzen keine Kosten gegenübergestellt.^{71, 72, 73}
5. Darüber hinaus differieren die Testcharakteristika (Sensitivität, Spezifität) der damaligen Publikationen und jenen der jüngeren wissenschaftlichen Literatur beträchtlich.
6. Therapieeffekte unberücksichtigt (Inzidenz ist nicht mit Mortalität gleichzusetzen).^{29, 71, 72}

3.9 Einfluss der Untersuchungsintervalle - Ergebnisse

Die zentrale Aufgabe dieses HTA-Berichts besteht in der Erarbeitung der Konsequenzen von unterschiedlichen Untersuchungsintervallen der Anwendung des Pap-Tests. Diese Analyse wird in der folgenden Form vorgenommen. Ein Basisszenario ermittelt die Ausgangssituation der Durchführung eines Screenings im Vergleich zu der es als gäbe es kein Screening. Diese

Sachlage wurde in der Realität nie untersucht, kann aber mittels publizierter epidemiologischer Kenngrößen retrospektiv simuliert werden. Diese als Basisszenario gekennzeichnete Modellrechnung (Screening gegen kein Screening) ermittelt den Einfluss der verschiedenen Untersuchungsintervalle auf vier verschiedene Problembereiche:

- Einfluss auf die kombinierten Testkennwerte.
- Einfluss für 100.000 Frauen.
- Einfluss für die Bundesrepublik Deutschland.
- Einfluss auf basisökonomische Zusammenhänge.

Neben diesem Basisszenario werden spezielle Berechnungen für die „neuen Technologien“ gemacht (FBC, Dünnschichtpräparation, Monolayer-Zytologie) sowie eine Sensitivitätsanalyse für den Einflussparameter Teilnehmerate.

3.9.1 Einfluss der Intervalle auf die Testkennwerte

3.9.1.1 Basisszenario

Das Basisszenario fußt auf den ermittelten Testkennwerten der Metaanalyse von McCrory et al.³⁷, sowie der erwarteten Prävalenz der ungescreenten deutschen Bevölkerung.

Tabelle 8: Kenngrößen der kombinierten Pap-Tests in Abhängigkeit des Screeningintervalls.

Sensitivität	Spezifität	Prävalenz	Teilnehmerate	Screeningspanne
51 %	98 %	2 %	-	40 Jahre
PPV	NPV	Falsch-positiv	Falsch-negativ	Anteil falsch Pos.
34,23 %	98,99 %	2 %	49 %	5,52 %

Kenngrößen	Intervall in Jahren						
	1	2	3	4	5	10	20
Anzahl der Untersuchungen	40	20	13	10	8	4	2
PPV (%)	100	99,98	99,63	98,49	96,5	81,29	56,74
NPV (%)	100	100	100	100	100	100	99,99
WS falsch-positiv (%)	55,43	33,24	23,61	18,29	14,92	7,76	3,96
WS falsch-negativ (%)	0	0	0,01	0,08	0,33	5,76	24,01

NPV = Negativer Vorhersagewert. PPV = Positiver Vorhersagewert. WS = Wahrscheinlichkeit.
Quelle: ARWIG-eigene Berechnungen Basisszenario für neue Technologien.

Bei gegenwärtigen Erkenntnissen über die Genauigkeit des Pap-Tests sowie einer Lebenszeitpanne für die Inanspruchnahme des Screenings von 40 Jahren (20 bis 60 Jahre) ergeben sich die in der Tabelle angegebenen Kenngrößen der Testcharakteristika, die beispielhaft interpretiert werden.

Bei einer Lebenszeitpanne von 40 Jahren für die Durchführung der Untersuchungen und einer jährlichen Screeninguntersuchung erlebt eine Frau im Laufe ihres Lebens 40 Screeninguntersuchungen. Diese Kombination ergibt, dass der positive PPV (Anteil der tatsächlich erkrankten bei positivem Testergebnis) dieser gesamten Testbatterie 100 % beträgt. Eine Person, die jedoch nicht 40, sondern nur 20 Screeninguntersuchungen erhalten würde, weil sie nur alle zwei Jahre untersucht wird, würde zu 99,98 % tatsächlich erkrankt sein, wenn die Tests dies feststellen. Im Fall eines dreijährigen Screenings würde die reale Erkennungswahrscheinlichkeit noch immer 99,63 % betragen, bei einer Untersuchung alle fünf Jahre 96,50 %.

Wesentlich für die Beurteilung eines Screenings ist die Gefahr des Übersehens von Erkrankungsfällen. Der NPV veranschaulicht die Wahrscheinlichkeit, dass die Person tatsäch-

lich nicht erkrankt ist, wenn der Test bzw. die Tests dies vermuten lassen. Bei diesem Kennwert kann festgestellt werden, dass das Untersuchungsintervall nahezu irrelevant ist und sich im Fall sämtlicher Intervalle der NPV nahezu 100 % beträgt.

Wesentlich für die realen Konsequenzen, vor allem in medizinischer, ökonomischer und ethischer Hinsicht ist die Wahrscheinlichkeit eines falschen Resultats. Wird jedes Jahr über die gesamte relevante Lebensspanne eine Screeninguntersuchung durchgeführt, erhält mehr als die Hälfte der teilnehmenden Frauen (55,43 %) zumindest einmal eine falsch-positive Diagnose mit allen Konsequenzen für weitere Maßnahmen, die von einer Untersuchungswiederholung bis zur Einleitung konkreter medizinischer Therapien reichen. Dieser Anteil reduziert sich stark auf 33,24 % bei zweijähriger Untersuchung oder auf 23,61 % bei einer Untersuchung alle drei Jahre, währenddessen die Kehrseite, die falsch-negativen Untersuchungsergebnisse („übersehene Fälle“) praktisch bei Null sind.

Für die Genauigkeit der FBC ergeben sich in Abhängigkeit der Untersuchungsintervalle die in der folgenden Tabelle angegebenen kombinierten Testkennwerte.

Tabelle 9: Kenngrößen der kombinierten FBC-Tests in Abhängigkeit des Screeningintervalls.

Sensitivität	Spezifität	Prävalenz	Teilnahme	Screeningspanne
90 %	85 %	2 %	-	40
PPV	NPV	Falsch-positiv	Falsch-negativ	Anteil falsch Pos.
10,91 %	99,76 %	15,00 %	10,00 %	57,89 %

Kenngrößen	Intervall in Jahren						
	1	2	3	4	5	10	20
Anz. Screen Leben	40	20	13	10	8	4	2
PPV (%)	99,02	90,08	78,57	68,5	60,31	37	20,63
NPV (%)	100	100	100	100	100	100	100
WS falsch-positiv (%)	99,85	96,12	88,55	80,31	72,75	47,8	27,75
WS falsch-negativ (%)	0	0	0	0	0	0,01	1

FBC = Flüssigkeitsbasierte Zytologie. PPV = Positiver Vorhersagewert. NPV = Negativer Vorhersagewert. WS = Wahrscheinlichkeit.

Beim Vergleich des konventionellen Pap-Tests mit der FBC ist eine Reihe von teils überraschenden Resultaten auffällig. Die stark erhöhte Sensitivität (verbesserte Erkennungsgenauigkeit) geht zu Lasten der Spezifität. Der Spezifität wird bei der Analyse eines Massenscreenings besondere Aufmerksamkeit geschenkt, da es um die Identifikation der Mehrheit der Untersuchten (die Gesunden) geht.

Der PPV sinkt bei der FBC von rund 34 % auf 11 %, der NPV steigt von rund 99 % auf 99,76 %. Der Anteil der falsch-positiven Resultate vergrößert sich um den Faktor 7,5, der Anteil der falsch-negativen (übersehenen Fälle) sinkt von 49 % auf 10 %.

Bei mehrfacher Anwendung dieser „genaueren Tests“ ergibt sich eine Verdeutlichung dieses Phänomens. Die Wahrscheinlichkeit eines falsch-positiven Resultats im Laufe eines Lebens steigt bei jährlicher Anwendung auf eine an Sicherheit grenzende Wahrscheinlichkeit (99,85 %). Diese Wahrscheinlichkeit sinkt auf rund 96 % bei zweijähriger Anwendung sowie 89 % bei dreijähriger Anwendung der FBC. Die Wahrscheinlichkeit eines falsch-negativen Resultats ist bei der FBC deutlich geringer, was sich jedoch erst ab einem Untersuchungsintervall von fünf Jahren oder länger quantitativ relevant zu Buche schlägt. Bei kürzeren Intervallen ist der Unterschied praktisch bedeutungslos.

Für kurze Untersuchungsintervalle (fünf Jahre oder kürzer) kann somit festgehalten werden, dass bei der Anwendung der FBC ein bedeutsames Mehr an falsch-positiven Resultaten praktisch keinem Unterschied bei den falsch-negativen gegenübersteht. Das Absinken des PPV im Vergleich zum konventionellen Pap-Test wird durch den starken Anstieg der falsch-positiven Resultate verursacht.

Vergleicht man diese beiden Technologien, kann festgehalten werden, dass den Vorteilen der verbesserten Sensitivität der FBC vor allem bei kurzen Untersuchungsintervallen Nachteile gegenüberstehen, die um ein Vielfaches schwerer wiegen als die klinisch bedeutsame Perspektive des Nichtübersehens von Erkrankten. Von einer Anwendung der FBC als Methode für ein Massenscreening kann nur abgeraten werden.

3.9.2 Einfluss der Intervalle auf 100.000 Frauen

Die folgenden Inhalte beleuchten den Einfluss des Untersuchungsintervalls aus der Betrachtungsweise der Betroffenen bzw. der Testresultate. Werden 100.000 Frauen anhand unterschiedlicher Untersuchungsintervalle diagnostiziert, so ergeben sich daraus Aussagen mit epidemiologischer Relevanz.

Werden 100.000 repräsentativ ausgewählte Frauen anhand eines einjährigen Untersuchungsintervalls diagnostiziert, werden mit dem konventionellen Pap-Test 100.000 Frauen korrekt erkannt. Bedeutsam ist bei dieser Betrachtung, dass es sich nicht um 100.000 unabhängige Resultate handelt, sondern um 100.000 Testergebnisse von Frauen, die im jeweiligen Untersuchungsintervall dem Diagnoseschema entsprechend das Screening in Anspruch genommen hatten.

Tabelle 10: Ergebnisse der Screeninguntersuchungen in Abhängigkeit des Untersuchungsintervalls für den Pap-Test auf 100.000 Frauen.

Sensitivität	Spezifität	Prävalenz	Teilnahme	Screeningspanne			
51 %	98 %	2 %	-	40 Jahre			
PPV	NPV	Falsch-positiv	Falsch-negativ	Anteil falsch Pos.			
34,23 %	98,99 %	2,00 %	49,00 %	5,52 %			
Intervall in Jahren							
Kenngrößen	1	2	3	4	5	10	20
Erkannte Fälle	100.000	99.977	99.625	98.485	96.498	81.286	56.741
Übersehene Fälle	0	0	7	80	332	5.765	24.010
Falsch-positiv	55.430	33.239	23.614	18.293	14.924	7.763	3.960
Fehldiagnosen	55.430	33.239	23.621	18.373	15.256	13.528	27.970

PPV = Positiver Vorhersagewert. NPV = Negativer Vorhersagewert.

Quelle: ARWIG-eigene Berechnungen.

Bei einem einjährigen Untersuchungsschema wird von 100.000 Untersuchungen im Rahmen eines einjährigen Untersuchungsintervalls kein einziger Fall übersehen. Bei einem zweijährigen Untersuchungsintervall entgeht bei 100.000 Untersuchungen ebenfalls keiner. Bei einem dreijährigen Untersuchungsintervall werden dagegen sieben Fälle übersehen. Stellt man diesen Ergebnissen die Resultate für die zu erwartenden falsch-positiven Testergebnisse gegenüber, so ist bei einem einjährigen Untersuchungsintervall auf 100.000 Frauen während der gesamten Screeningspanne eine Fehldiagnose für 55.430 Frauen zu erwarten (in diesem Fall sind sämtliche Fälle falsch-positive Diagnosen). Die übersehenen Fälle und die falsch-positiven Diagnosen haben einen streng monotonen, jedoch gegenläufigen Trend. Betrachtet man daher diese beiden Formen der Fehldiagnosen (falsch-negativ und falsch-positiv) gemeinsam, ergibt sich daraus ein Minimum von 13.528 Fällen bei einem zehnjährigen Untersuchungsintervall. Würde man ein falsch-positives Resultat eines Tests und ein falsch-negatives Ergebnis gleichsetzen (was medizinisch, ökonomisch und ethisch nicht vertretbar wäre), würde ein zehnjähriges

jähriges Untersuchungsintervall ein Optimum hinsichtlich des Kriteriums „Minimierung der Fehldiagnosen“ bedeuten.

3.9.2.1 Auf 100.000 Frauen für neue Technologien

Die folgende Tabelle stellt die Ergebnisse der FBC bezogen auf 100.000 Frauen in Abhängigkeit der Untersuchungsintervalle dar.

Tabelle 11: Fallkonstellation flüssigkeitsbasierte Zytologie in Abhängigkeit der Untersuchungsintervalle.

Sensitivität	Spezifität	Prävalenz	Teilnahme	Screeningspanne			
90 %	85 %	2 %	-	40			
PPV	NPV	Falsch-positiv	Falsch-negativ	Anteil falsch Pos.			
10,91 %	99,76 %	15 %	10 %	57,89 %			
Intervall in Jahren							
Kenngroßen	1	2	3	4	5	10	20
Erkannte Fälle	99.015	90.076	78.566	68.498	60.311	37.001	20.628
Übersehene Fälle	0	0	0	0	0	10	1.000
Falsch-positiv	99.850	96.124	88.547	80.313	72.751	47.799	27.750
Fehldiagnosen	99.850	96.124	88.547	80.313	72.751	47.809	28.750

PPV = Positiver Vorhersagewert. NPV = Negativer Vorhersagewert.

Quelle: ARWIG-eigene Berechnungen.

Im Fall der FBC ist die hohe Erkennungsrate – gekoppelt mit der gesunkenen Spezifität - für die große Anzahl an falsch-positiven Resultaten verantwortlich. Versucht man aus diesen Testkennwerten das Kriterium der Minimierung der Fehldiagnosen anzuwenden, resultiert daraus, dass das Untersuchungsintervall länger als 20 Jahre sein müsste, um einen „Trade off“ – ein Gleichgewicht zwischen übersehenen und falsch-positiven Fällen - zu erhalten.

Der oftmals klinisch betrachteten vorteilhaften Eigenschaft, keine Fälle zu übersehen steht im Fall der FBC eindrucksvoll der für das Gesundheitssystem und die Frauen erkennbare Nachteile gegenüber. Die Anzahl der im Zuge der Untersuchungsschemata zu erwartenden Fehldiagnosen (die praktisch nur falsch-positive Resultate sind) ist völlig inakzeptabel. Verantwortlich hierfür zeichnet die relativ geringe Spezifität. In Anbetracht dieser Resultate muss festgehalten werden, dass die Evidenzlage für die Feststellung der Testcharakteristika sehr schlecht ist. Weitere Untersuchungen zur Präzisierung (möglicherweise deutlichen Verbesserung) der Testkennwerte der einzelnen Technologien wären wesentlich.

3.9.3 Einfluss der Intervalle für die Bundesrepublik Deutschland

Die folgenden Darstellungen beziehen sich auf Berechnungen mit der Bevölkerungsverteilung der Bundesrepublik Deutschland von 2000. Die Daten stammen aus der Sterbetafel der weiblichen Bevölkerung, wie sie der WHO übermittelt wurde. Die Gesamtanzahl der laut der gesetzlichen Regelung vereinbarten Anspruchsberechtigten beträgt 33.443.770 Frauen. Selbstverständlich kann nicht davon ausgegangen werden, dass der Anspruchsberechtigte Bevölkerungsteil die angebotene Untersuchung auch tatsächlich in Anspruch nimmt, daher ist eine Teilnehmerate im Berechnungsmodell berücksichtigt worden. Die folgende Tabelle veranschaulicht ausgehend von der Gesamtanzahl der anspruchsberechtigten Frauen sowie der Berücksichtigung einer Teilnehmerate die Anzahl der erkannten und der übersehenen Fälle.

Tabelle 12: Erkannte und übersehene Fälle in Abhängigkeit des Untersuchungsintervalls für die Bundesrepublik Deutschland.

Sensitivität	Spezifität	Prävalenz	Teilnahme	Screeningspanne			
51 %	98 %	2 %	80 %	40			
PPV	NPV	Falsch-positiv	Falsch-negativ	Anteil falsch Pos.			
34,23 %	98,99 %	2,00 %	49,00 %	5,52 %			
Intervall in Jahren							
Kenngrößen	1	2	3	4	5	10	20
Anzahl Screen / Jahr	33.443.770	16.721.885	11.147.923	8.360.943	6.688.754	3.344.377	1.672.189
Erkannte Fälle / Jahr	16.722	16.718	16.659	16.469	16.136	13.593	9.488
Übersehene Fälle / Jahr	0	0	1	13	56	964	4.015
Real erkannte	13.378	13.374	13.327	13.175	12.909	10.874	7.590
Real übersehene	3.344	3.347	3.396	3.560	3.868	6.812	13.146
Erkannte minus übersehene	10.033	10.027	9.932	9.614	9.041	4.062	-5.556

PPV = Positiver Vorhersagewert. NPV = Negativer Vorhersagewert.
Quelle: ARWIG-eigene Berechnungen.

Betrachtet man die Anzahl der real erkannten bzw. der real übersehenen Fälle (die die Teilnahmerate berücksichtigt), ergibt sich daraus dass sich das ein-, zwei- oder dreijährige Untersuchungsintervall in der Zahl der epidemiologisch relevanten Erkrankten praktisch überhaupt nicht voneinander unterscheiden. Einen beträchtlichen Unterschied machen hier die Anzahl der Screeninguntersuchungen aus, die sich auch ökonomisch als bedeutsam niederschlagen. Einem beträchtlichen Einsparungspotenzial von einer Änderung eines ein- auf ein zweijähriges Untersuchungsintervall steht daher eine epidemiologisch (nicht klinisch!) irrelevante Größe an übersehenen Fällen gegenüber. Diese als Interpretationsanleitung zu verstehende Aussage kann nicht als eine Screeningempfehlung betrachtet werden, da es im Rahmen dieser Arbeit kein vereinbartes Kriterium gibt, das als Optimierungsmaßstab herangezogen werden könnte.

3.9.3.1 Für Bundesrepublik Deutschland neue Technologien

Im Fall der Anwendung der neuen Technologien ergeben sich die in der folgenden Tabelle angegebenen Kenngrößen.

Tabelle 13: Anwendung der neuen Technologien in Abhängigkeit der Untersuchungsintervalle für die Bundesrepublik Deutschland.

Sensitivität	Spezifität	Prävalenz	Teilnahme	Screeningspanne			
90 %	85 %	2 %	80 %	40			
PPV	NPV	Falsch-positiv	Falsch-negativ	Anteil falsch Pos.			
10,91 %	99,76 %	15 %	10 %	57,89 %			
Kenngrößen	Intervall in Jahren						
	1	2	3	4	5	10	20
Anzahl Screen / Jahr	33.443.770	16.721.885	11.147.923	8.360.943	6.688.754	3.344.377	1.672.189
Erkannte Fälle / Jahr	16.557	15.062	13.138	11.454	10.085	6.187	3.449
Übersehene Fälle / Jahr	0	0	0	0	0	2	167
Real erkannte	13.246	12.050	10.510	9.163	8.068	4.950	2.760
Real übersehene	3.311	4.507	6.047	7.394	8.489	11.609	13.965
Erkannte minus übersehene	9.934	7.543	4.463	1.770	-421	-6.659	-11.205

PPV = Positiver Vorhersagewert. NPV = Negativer Vorhersagewert.

Quelle: ARWIG-eigene Berechnungen.

Aus diesen Ergebnissen ist ersichtlich, dass die Teilnehmerate von 80 % für den wesentlichsten Teil der übersehenen Fälle sorgt. Im Vergleich zum konventionellen Pap-Test ist die Anzahl der richtig erkannten aufgrund des gesteigerten Anteils der falsch-positiven Resultate sogar geringer. Bei 100-prozentiger Teilnahme fällt die gesteigerte Sensitivität dadurch auf, dass die Zahl der übersehenen Fälle bei den neuen Technologien geringer wäre als beim konventionellen Pap-Test. Dieser Perspektive, die üblicherweise von klinischer Seite eingenommen wird, sind daher für eine umfassende Beurteilung die falsch-positiven Fälle gegenüberzustellen. Daraus ist erkennbar, dass die Teilnehmerate einen viel stärkeren Einfluss auf die Anzahl der erkannten Fälle hat als die Anwendung der flüssigkeitsbasierten Technologien.

3.9.3.2 Einfluss der Teilnehmerate

Die Teilnehmerate am Screening ist ein für den Ertrag eines Screenings entscheidender Parameter. Da die Ergebnisse verschiedener Untersuchungsintervalle dargestellt werden, soll auch eine Sensitivitätsanalyse der Teilnehmerate vorgenommen werden, um daraus Entscheidungsgrundlagen für eine Ertragsoptimierung des Pap-Test-Screenings ableiten zu können.

Die folgenden Tabellen zeigen die Anhängigkeit der real erkannten und der übersehenen Fälle in Abhängigkeit der Teilnehmerate aus dem Kreis der anspruchsberechtigten Frauen in der Bundesrepublik Deutschland.

Tabelle 14: Erkannte und übersehene Zervixkarzinomfälle in Abhängigkeit der Teilnahmerate.

Intervall in Jahren (1 bis 20 Jahre)							
Teilnahmerate 50 %							
Kenngrößen	1	2	3	4	5	10	20
Real erkannte	8.361	8.359	8.330	8.234	8.068	6.796	4.744
Real übersehene	8.361	8.363	8.394	8.501	8.709	10.890	15.993
Teilnahmerate 60 %							
Kenngrößen	1	2	3	4	5	10	20
Real erkannte	10.033	10.031	9.996	9.881	9.682	8.156	5.693
Real übersehene	6.689	6.691	6.728	6.854	7.096	9.530	15.044
Teilnahmerate 70 %							
Kenngrößen	1	2	3	4	5	10	20
Real erkannte	11.705	11.703	11.661	11.528	11.295	9.515	6.642
Real übersehene	5.017	5.019	5.062	5.207	5.482	8.171	14.095
Teilnahmerate 80 %							
Kenngrößen	1	2	3	4	5	10	20
Real erkannte	13.378	13.374	13.327	13.175	12.909	10.874	7.590
Real übersehene	3.344	3.347	3.396	3.560	3.868	6.812	13.146
Teilnahmerate 90 %							
Kenngrößen	1	2	3	4	5	10	20
Real erkannte	15.050	15.046	14.993	14.822	14.523	12.233	8.539
Real übersehene	1.672	1.676	1.730	1.914	2.255	5.453	12.198
Teilnahmerate 100 %							
Kenngrößen	1	2	3	4	5	10	20
Erkannte Fälle / Jahr	16.722	16.718	16.659	16.469	16.136	13.593	9.488
Übersehene Fälle / Jahr	0	0	1	13	56	964	4.015

Quelle: ARWIG-eigene Berechnungen.

Aus den Ergebnissen der Berechnungen zur Teilnahmerate ist erkennbar, dass der Einflussfaktor Teilnahmerate viel stärker auf den Screeningsertrag wirkt als die Frage des Untersuchungsintervalls bei kurzen Zeitspannen. Daraus ist abzuleiten, dass der möglichst vollständigen Teilnahme (auch bei einem Untersuchungsintervall von fünf Jahren) der Vorzug gegenüber einem Verlust auch nur an 10 % an Teilnehmerinnen bei jährlicher Teilnahme zu geben ist. Eine solche Strategie kann insofern angedacht werden, als der Druck durch die nicht jährliche Empfehlung zur Teilnahme verringert und dadurch eine größere Zahl an Frauen gewonnen werden könnte, die noch keine Untersuchung in Anspruch genommen hat.

3.9.4 Schlussfolgerungen

Aus den Berechnungen kann Folgendes abgeleitet werden:

- Die Unterschiede der Erträge des Screenings bei kurzen Untersuchungsintervallen ist epidemiologisch gering.
- Der finanzielle Aufwand ist stark von den Untersuchungsintervallen abhängig und läuft linear zur Anzahl der erforderlichen Untersuchungen.
- Die FBC kann für die Durchführung eines Massenscreenings nach gegenwärtigem Kenntnisstand nicht empfohlen werden.

- Die Teilnehmerate ist eine für den Ertrag des Screenings entscheidende Einflussgröße und sollte besondere Aufmerksamkeit bei der Ressourcenallokation erhalten.

3.10 Ökonomische Bewertung

In die Berechnungen werden zur Analyse der ökonomischen Konsequenzen Preise von 10 Euro für einen konventionellen Pap-Test sowie 30 Euro für eine flüssigkeitsbasierte Technologie berücksichtigt. Diese Preise ergeben sich aus einer Fülle unterschiedlicher Tarifkataloge, bei denen die Leistungs- und die Personalkomponente durch Annahmen getrennt wurden.

Die folgende Tabelle zeigt die ökonomischen Konsequenzen der Untersuchungsintervalle für den konventionellen Pap-Test.

3.10.1 Einfluss der Intervalle auf die Kostenwirksamkeit

3.10.1.1 Basisszenario

Tabelle 15: Ökonomische Konsequenzen des konventionellen Pap-Tests für die Bundesrepublik Deutschland in Abhängigkeit des Untersuchungsintervalls.

Sensitivität	Spezifität	Prävalenz	Teilnahme	Screeningspanne			
51 %	98 %	2 %	-	40			
PPV	NPV	Falsch-positiv	Falsch-negativ	Anteil falsch Pos.			
34,23 %	98,99 %	2 %	49 %	5,52 %			
Intervall in Jahren							
Kenngrößen	1	2	3	4	5	10	20
Kosten in Millionen-Euro	334	167	111	84	67	33	17
Reale Kosten in Millionen-Euro	134	67	45	33	27	13	7
Kosten pro richtig erkanntem Fall (Euro)	10.000	5.001	3.346	2.538	2.073	1.230	881

PPV = Positiver Vorhersagewert. NNP = Negativer Vorhersagewert.
Quelle: ARWIG-eigene Berechnungen.

Aus der ökonomischen Betrachtung ist die Bedeutung des Screeningintervalls ersichtlich. Die Kosten zwischen einem ein- und einem doppeljährigen Screening sind im Wesentlichen mit dem Faktor zwei anzunehmen, während diesen Aufwendungen kein bedeutsamer epidemiologischer Nutzen gegenübersteht. Die Kosten pro real erkannten Fall haben bei dieser Betrachtung noch kein Optimum, woraus aus der ökonomischen Perspektive erst bei Berücksichtigung von maßnahmenbezogenen direkten sowie indirekten Kosten eine Entscheidungsgrundlage zu erwarten ist.

Einem im Wesentlichen konstanten Nutzen während einem drei- oder fünfjährigen Untersuchungsintervall (je nach Betrachtungsweise) stehen Kosten gegenüber, die sich um das drei- bis vierfache unterscheiden. Die Entscheidung zugunsten des einen oder anderen Untersuchungsintervalls kann nur in Abstimmung mit anderen Maßnahmen, den zur Verfügung stehenden Ressourcen bzw. gesundheitspolitischen Mechanismen getroffen werden.

3.10.1.2 Neue Technologien

Die gleichen Berechnungen wie für den konventionellen Pap-Test werden auch für die FBC vorgenommen.

Tabelle 16: Ökonomische Konsequenzen des konventionellen Pap-Tests für die Bundesrepublik Deutschland in Abhängigkeit des Untersuchungsintervalls.

Sensitivität	Spezifität	Prävalenz	Teilnahme	Screeningspanne			
90 %	85 %	2 %	-	40			
PPV	NPV	Falsch-positiv	Falsch-negativ	Anteil falsch Pos.			
10,91 %	99,76 %	15,00 %	10,00 %	57,89 %			
Kenngrößen	Intervall in Jahren						
	1	2	3	4	5	10	20
Kosten in Millionen-Euro	1.003	102	334	251	201	100	50
Reale Kosten in Millionen-Euro	401	201	134	100	80	40	20
Kosten pro real erkannte Fälle (Euro)	30.298	16.653	12.728	10.949	9.948	8.108	7.272

PPV = Positiver Vorhersagewert- NPV = Negativer Vorhersagewert.
Quelle: ARWIG-eigene Berechnungen.

Die ökonomischen Konsequenzen der verschiedenen Untersuchungsintervalle sind beträchtlich und reichen real von rund 400 Millionen Euro pro Jahr bis zu 20 Millionen Euro (Faktor 20). Diesen bedeutsamen Unterschieden wären die epidemiologischen Kenngrößen gegenüberzustellen, was bei einfacher Betrachtung (Kosten / richtig erkannter Fall) noch kein Optimum ergibt.

Aus der kurzen ökonomischen Betrachtung ist ersichtlich, dass sich die Kosten pro real erkannten Krankheitsfall stetig durch die Vergrößerung des Untersuchungsintervalls senken lassen. Daraus lässt sich daher keine Empfehlung für ein optimales Untersuchungsintervall ableiten.

In diese Berechnungen fließen nur die direkten Kosten der Anwendung des Pap-Tests ein und nicht die jedoch damit direkt verbundenen Kosten für die weiteren Untersuchungen, Konisation oder anderer medizinischer Maßnahmen. Für eine genauere Analyse würde hier eine detailliertere Erhebung der Kosten und der Verzweigungsprozesse der medizinischen Maßnahmen erforderlich sein, was im Rahmen dieser Arbeit nicht beantwortet werden soll.

3.10.1.3 Nutzenproblem

Anhand zahlreicher Studien wurde belegt, dass ein Land den größten Nutzen hat, das überhaupt kein Screening durchführt und zu einem Screening einmal pro Jahr wechselt (ungefähr ab einem Alter von 35 Jahren)^{23, 41, 65}. Dieser einmalige Test zu diesem optimalen Alter soll einen 25 % höheren protektiven Effekt im Vergleich zu einem vollständigen, alle drei Jahre stattfindenden Screening besitzen. Je öfter ein Screening durchgeführt wird, desto mehr Fälle können entdeckt werden, die Anzahl der zusätzlich erkannten Fälle ist jedoch abnehmend. Der Nutzenzuwachs wird also mit der Anzahl der durchgeführten Screeningrunden kleiner, der Nutzen folgt dem Gesetz des streng monoton abnehmenden Grenznutzens der Mikroökonomie (Gesetz des abnehmenden Grenznutzens, Sättigungsgesetz, erstes Gossen'sches Gesetz). Der Nutzen jeder weiteren Einheit ist bis zur Vollerfassung durch jedes Screening positiv, jedoch jedes weitere Mal bringt einen immer kleineren Effekt.

Aus diesem Gesetz ist erklärbar, dass aus medizinischer oder aus klinischer Perspektive ein Mehr an Untersuchungen ein Mehr an früh erkannten Fällen bringt, die Frage ist nur, um welchen Preis. Mit dem Preis sind einerseits Kosten, andererseits falsch-positiv diagnostizierte Fälle gemeint. Das Verhältnis von richtig-positiven zu falsch-positiven Fällen hängt einerseits von der Untersuchung selbst (Sensitivität, Spezifität), andererseits von der Prävalenz in der Bevölkerung ab.

Je öfter ein Screening im Rahmen einer relevanten Lebensspanne stattfindet, desto geringer ist die Gefahr von falsch-negativen Befunden und desto größer die Gefahr von falsch-positiven Befunden. Mit epidemiologischen Worten ist ein Mehr an Sensitivität automatisch mit einem Weniger an Spezifität und umgekehrt - bei sonst gleichen Bedingungen - verbunden.

Für die Berechnungen im Rahmen dieser Arbeit lässt sich das Nutzenproblem in der folgenden Weise darstellen:

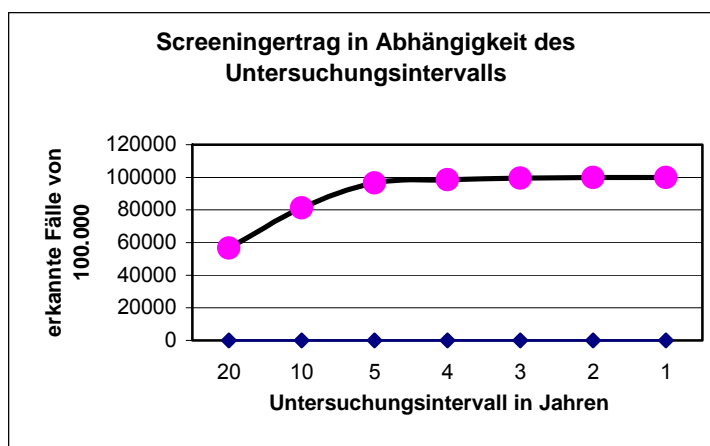
Tabelle 17: Erkannte Fälle sowie inkrementeller Nutzen in Abhängigkeit des Untersuchungsintervalls (Basisszenario, konventioneller Pap-Test).

	Intervall in Jahren						
Kenngößen	1	2	3	4	5	10	20
Erkannte Fälle	100.000	99.977	99.625	98.485	96.498	81.286	56.741
Differenzen	+ 23	+ 352	+ 1.140	+ 1.987	+ 15.212	+ 24.545	+ 56.741

Quelle: ARWIG-eigene Berechnungen.

Die grafische Darstellung verdeutlicht das Gesetz des abnehmenden Grenznutzens, der in Abhängigkeit des Untersuchungsintervalls die erkannten Fälle darstellt.

Abbildung 1: Ertrag des Screenings in Abhängigkeit des Untersuchungsintervalls.



Aus den Berechnungen ist der abnehmende Grenznutzen deutlich erkennbar. Je kürzer das Untersuchungsintervall wird, desto geringer ist der zusätzliche Nutzen, gemessen an der Anzahl der zusätzlich erkannten Fälle.

Aus Simulationsstudien konnte anhand der holländischen Bevölkerung ermittelt werden, dass sich 15 von 500 fiktiven Screeninganordnungen als effizient erwiesen. Darunter sind auch Anordnungen erfasst, die nur zwei Pap-Abstriche während der gesamten Lebensspanne einer Frau vorsahen, bis zu Arrangements 40 Abstrichen (im Alter von 20 bis 60 Jahren, einmal jährlich). Der Lebenserwartungszugewinn betrug bei nur zwei Abstrichen 11,6 Tage für die weibliche Bevölkerung, verglichen mit weiteren 20,8 Tagen (insgesamt 32,4 Tage) im Fall von 40 Abstrichen während der Lebenszeit. Die komplette Elimination des Zervixkarzinoms würde einen Lebenserwartungszugewinn von 46 Tagen für die weibliche Bevölkerung bringen (dies bedeutet, dass ungefähr jede 100-ste Frau 12,5 Jahre früher stirbt als im nicht erkrankten Fall).

Bei dieser Mikrosimulationsstudie wurde ein Zervixkarzinom-Screening mit einem einjährigen Intervall ausgeschlossen, da die inkrementellen Zugewinne und / oder die durchschnittliche Kosteneffektivität weit außerhalb jener Spannweite lag, die als akzeptierbar zu betrachten wäre. In detaillierten holländischen Untersuchungen zeigte sich, dass die Teilnehmerate mit mehr als 80 % unrealistisch erscheint. Diese 20 % der Frauen, teilen sich in 10 % überzeugte und persistierende Nichtteilnehmerinnen. Diese Frauen werden in die Hochrisikogruppe eingestuft

und stehen für 25 % der Mortalität, umgelegt wären nur 34 von theoretisch 46 Tagen Lebenserwartungszugewinn in der Bevölkerung tatsächlich realisierbar. Die aus 500 verschiedenen Anordnungen als effizient einzustufenden Ausgestaltungen resultieren in einem Lebenserwartungszugewinn zwischen 12 und 32 Tagen und variieren zwischen 2 und 40 Pap-Abstrichen pro Lebenszeit. Die negativen Seiten des Screenings, vor allem die Lebensqualität, sind weitestgehend proportional zur Anzahl der durchgeführten Untersuchungen pro Lebensspanne⁶⁵.

Tabelle 18: Ergebnisse verschiedener Screeningstrategien in Abhängigkeit des Untersuchungsintervalls.

Anzahl der Abstriche	Intervall in Jahren	Altersspanne	Kosten	Effekte	CER	ICER	
			US-\$ pro Frau pro Jahr	Jährliche Lebenszugewinn in Tagen			Verlängerte Lebenserwartung
2	12	40 - 52	0,99	0,14	11,6	6.700	6.700
3	9	35 - 53	1,42	0,19	15,5	7.300	9.100
4	8	32 - 56	1,79	0,23	18,3	7.700	10.200
5	9	32 - 68	1,94	0,23	18,7	7.900	11.800
6	7	32 - 67	2,27	0,25	20,3	8.500	15.200
7	6	32 - 68	2,55	0,27	21,4	8.900	17.300
8	7	27 - 76	2,86	0,29	22,9	9.400	16.800
9	6	27 - 75	3,17	0,30	23,9	9.900	20.800
10	5	27 - 72	3,54	0,31	25,2	10.600	26.300
12	5	27 - 82	3,79	0,32	25,4	11.000	32.300
15	4	22 - 78	4,93	0,35	28,1	13.000	35.000
20	3	22 - 79	6,35	0,37	29,6	15.500	55.700
25	2,5	20 - 80	7,43	0,38	30,8	17.500	79.500
30	2	22 - 80	8,49	0,39	31,4	19.400	119.000
40	2,5	20 - 79	10,88	0,41	32,4	23.900	173.700

CER = Cost-effectiveness ratio (Kosten-Wirksamkeits Verhältnis). ICER = Incremental cost-effectiveness (inkrementelles Kosten-Wirksamkeits Verhältnis). US-\$ = US-Dollar.
Quelle: van den Akker-van Marle et al.⁶⁶.

3.10.2 Kosten

3.10.2.1 Kosten / Nutzen-Analyse des Screenings

Gelegentlich wird ein Screening dadurch gerechtfertigt, dass eine einfache Bilanz gebildet wird: Der Test ist billig und hat Potenzial, erkrankte Personen frühzeitig zu erkennen und einer Therapie zuführen zu können. Bei der Berechnung der Gesamtkosten eines Verfahrens müssen jedoch auch die Kosten berücksichtigt werden, die deshalb entstehen, weil Personen falsch-positiv erkannt werden und weiteren Untersuchungen zugeführt werden müssen.

Bei der Kosten-Effektivitätsbilanz spielen darüber hinaus nicht nur die finanziellen Kosten eine Rolle, sondern auch die nichtfinanziellen Aufwendungen, die dem Patienten entstehen, wozu auch Angst, emotionaler Stress sowie Unannehmlichkeiten zu rechnen sind. Als Beurteilungskriterium kann die Rate falsch-positiver Ergebnisse herangezogen werden. Ist sie groß, ist der Anteil der Personen hoch, bei denen weitere Untersuchungen durchgeführt werden und / oder Ängste geschürt werden, obwohl sie in Wirklichkeit nicht an der betreffenden Krankheit leiden.

Also gehören zu den Kosten eines Tests nicht nur die Verfahrenskosten, sondern auch sämtliche Kosten der Nachuntersuchungen. Darüber hinaus sind gesundheitsökonomisch auch Lebensqualitätsbeeinträchtigungen Gegenstand ökonomischer Evaluationen.

Der Nutzen muss die Risiken übertreffen. Die Kostenwirksamkeit verschiedener Zervixkarzinom-Screeningprogramme ist sehr sensitiv hinsichtlich der Anzahl der während der Lebensspanne durchgeführten Pap-Abstriche. Als Hintergrund der Unsicherheit der Erkenntnisse der Literatur soll kurz die Erkennungsfähigkeit des Tests dargestellt werden. Die Kenngröße Spezifität verdeutlicht dabei die Wahrscheinlichkeit von falsch-positiven Testergebnissen. Eine genauere Darstellung wird bei der Beschreibung des Pap-Tests gegeben.

Tabelle 19: Sensitivität – Übersicht über die Variabilität der Kenngrößen.

Sensitivität	Spezifität	Quelle
51 (6 - 99)	49 (6 – 99)	Mc Crory et al. ³⁷
45 – 94		Shingleton, Patrick et al. ⁷⁸
80		Gay, Donaldson et al. (1995).
78,1 (mild to moderate)		
81,4 Oberflächenkarzinom		
82,3 (invasive carcinom)		Soost, Lange ⁵⁹
58 (11 – 99)		Fahey et al. ¹⁷
30 – 87	86 - 100	Nanda ⁴¹
67	77	Mitchell ²¹

Kostenwirksamkeit

Ein Bericht des Office of Technology Assessment 1981 stellte die Basis für mehrere Kostenwirksamkeitsanalysen dar, von denen Fahs et al.¹⁷ sowie Eddy¹⁵ die meistzitierten durchführten. Beide Modelle enthalten eine Fülle von Detailparametern (z.B. Markov-Übergangswahrscheinlichkeiten), deren Übertragbarkeit auf bundesdeutsche Verhältnisse gewagt erscheinen. Darüber hinaus sind diese beiden Modelle insofern unvergleichlich, da die Zielpopulationen einander ausschließen. Eddys Zielpopulation waren asymptomatische 20-jährige Frauen mit mittlerem Risiko, währenddessen Fahs et al.¹⁷ Frauen die älter als 65 Jahre waren, mit einbezog. In beiden Fällen wurde die private Finanzierung als die Bewertungsperspektive eingenommen, die ebenfalls für eine Beurteilung aus der Sicht der Anspruchsberechtigung aus der gesetzlichen Krankenversicherung unvergleichbar ist.

Eine Aussage bzw. eine Übertragbarkeit auf die bundesdeutsche Situation ist daher nicht möglich.

3.10.3 Diskussion

Die Untersuchung des Einflusses des Untersuchungsintervalls wird auf der Basis des gegenwärtigen wissenschaftlichen Wissensstands durchgeführt. Dabei konzentriert man sich auf die Beurteilung des Massenscreenings.

Folgende Punkte sind dabei für die Bewertung entscheidend:

- Die Variabilität der Sensitivität und Spezifität des konventionellen Pap-Tests ist beträchtlich und daher mit Unsicherheit behaftet.
- Angaben über die Sensitivität und Spezifität der in den Klassifikationssystemen als klinische Vorstufe einzustufenden Fälle sind äußerst dürftig und unscharf.
- Einflüsse von nicht testbezogenen Screeningmodalitäten (Einladesystem, Teilnahme) sind nur quantitativ möglich.
- Die Genauigkeit der FBC entspringt geringer wissenschaftlicher Evidenz.

- Die Rolle des HPV sowie der Einfluss auf die langfristige Entwicklung des Zervixkarzinoms ist gegenwärtig nicht abschätzbar.
- Die Frauen werden beim Massenscreening als eine Gruppe mit gleichem Risiko betrachtet. Eine ergebnisabhängige Form des Untersuchungsintervalls wäre anzustreben im Sinne der Möglichkeit eines effizienten Ressourceneinsatzes.
- Risikofaktoren für die Entwicklung des Zervixkarzinoms sollten in die tatsächliche Ausgestaltung des Screenings einfließen.
- Eine Optimierung des Screenings unter Berücksichtigung von Alter, Risikofaktoren, Ergebnisse der Tests und Untersuchungsintervall wäre dringend zu empfehlen.
- Der Aufbau eines entsprechenden Registers zur Dokumentation der Fälle zur Etablierung eines Qualitätssicherungssystems für das Screening ist für den Erfolg und die Ertragsmaximierung entscheidend.
- Dem Anteil der nie an einer Untersuchung teilnehmenden Frauen sollte besondere Aufmerksamkeit geschenkt werden. Eine einmalige Teilnahme ist viel besser als überhaupt keine.
- Die FBC erscheint zum gegenwärtigen Zeitpunkt nicht als massenstrategische Untersuchungsmethode geeignet.
- Die Kombination mit einem Test auf HPV birgt aufgrund der hohen Prävalenz des HPV die Gefahr einer stark steigenden falsch-positiven Ereigniseinschätzung in sich.
- Die Steigerung der Teilnehmerate ist ein viel bedeutender Einflussfaktor auf den Screeningertrag als die Variation der Untersuchungsintervalle im Rahmen zwischen einem und zehn Jahren.
- Aufgrund der epidemiologischen als auch der ökonomischen Situation ergibt sich kein eindeutig optimales Untersuchungsintervall – dies müsste aufgrund der Bewertung von falsch-negativen bzw. falsch-positiven Fällen speziell untersucht werden.

3.11 Zusammenfassende Diskussion aller Ergebnisse

Mit der vorliegenden Arbeit sollte festgestellt werden, welchen Einfluss das Untersuchungsintervall auf das System der Früherkennung auf Zervixkarzinom hat. Die neuesten Erkenntnisse aus internationaler Sicht sollten zu diesem Thema herangezogen und damit eine Überprüfung der Testgüte des Pap-Tests, sowie seiner bestmöglichen Einsatzmöglichkeiten erarbeitet werden. Empfehlungen, die ermöglichen, das deutsche Screening auf Zervixkarzinom den neuesten Erkenntnissen anzupassen, sollen gegeben werden.

Im Rahmen der EU unterscheiden sich bei Anwendung ein und desselben Tests die Modalitäten des Screenings auf Zervixkarzinom beträchtlich. So schwankt das empfohlene Untersuchungsintervall zwischen einem (Deutschland) und fünf Jahren. Auch die Altersspanne, innerhalb der untersucht werden soll, differiert zum Teil stark.

Anhand des zentralen Kriteriums, der Anzahl der erkennbaren Fälle kann festgestellt werden, dass sich das Untersuchungsintervall von einem bis drei Jahren im Ergebnis fast überhaupt nicht unterscheiden. Jedoch sind deutliche Unterschiede in der Zahl der zu erwartenden Fehldiagnosen (falsch-positive Resultate) als auch beträchtliche Differenzen im für das Screening erforderlichen finanziellen Aufwand erkennbar. Entsprechend der wissenschaftlichen Literatur als auch der Modellrechnungen kann die Aufrechterhaltung eines einjährigen Untersuchungsintervalls für das Massenscreening nicht empfohlen werden.

Die Teilnehmerate ist für das Resultat des Screenings ein wesentlicher Einflussfaktor. Ihre Steigerung um 10 % hat einen viel deutlicheren positiven Effekt auf die Früherkennung als die Diskussion über das Untersuchungsintervall im Rahmen zwischen einem und zehn Jahren.

Die amerikanischen Richtlinien (American Cancer Society (ACS))⁷⁶ für die Früherkennung des Zervixkarzinoms werden im Anhang aufgeführt. Diese an bestimmte Ereignisse oder das Alter gebundenen Handlungsanweisungen durchbrechen jedoch in dieser Form das deutsche Krankenversicherungssystem und sind in ihrem Zustandekommen hinsichtlich Evidenzorientierung und wissenschaftlichen Hintergrund nicht nachvollziehbar.

Ebenfalls im Anhang wiedergegeben sind die Richtlinien der US-Preventive Services Task Force (USPSTF)⁷⁷, die ein im Wesentlichen dreijähriges Untersuchungsintervall empfiehlt – geknüpft an andere Kriterien sowie andere Altersgrenzen als die ACS-Richtlinien.

3.12 Schlussfolgerung

Das Screening auf Früherkennung des Zervixkarzinoms sollte nicht generell auf ein Untersuchungsintervall festgelegt werden, sondern die spezielle Risikokonstellation der Frau miteinbeziehen. Generell ist ein einjähriges Untersuchungsintervall im Vergleich zu einem zwei- oder dreijährigen Screening sehr ineffizient und mit großem Aufwand verbunden. Dem stehen zudem die beträchtliche Gefahr von falsch-positiven Testergebnissen bei kurzen Untersuchungsintervallen gegenüber.

Insgesamt soll die Risikokonstellation auf der Ebene der Frau bzw. der Patientin in evidenzbasierter Form in das angebotene Screening aufgenommen werden. Dies führt zu einer Optimierung des Screenings vor dem Hintergrund der Anpassung der Testauswahl sowie der Untersuchungsintervalle an die Risikokonstellation sowie die vorangegangenen Testergebnisse. In einigen Ländern sind diese ergebnisabhängigen Untersuchungsintervalle realisiert.

Die Optimierung des Screenings vor dem Hintergrund der Personenbiographie ermöglicht beträchtliche Ressourcenschonung sowohl für das Gesundheitssystem als auch die betroffene Person. Die Optimierung muss daher die Erkennungswahrscheinlichkeit, als auch die Gefahr von falsch-positiven Resultaten, die zwangsläufig mit der Durchführung der Untersuchung verbunden ist – und zwar in einem beträchtlichem Ausmaß – miteinbeziehen. Dies ermöglicht auch, die persönlichen Präferenzen der betroffenen Frau (risikoaversives bis risikofreudiges Gesundheitsverhalten) zu berücksichtigen.

Anzustreben wäre auf Basis dieser Arbeit eine gemeinsame Aktion der Versicherungsverbände, die die Optimierung des Screenings zum Ziel hat. Diese Aufgabe wäre auf der Basis der gegenwärtigen Evidenzlage durchführbar und könnte die persönliche Situation der untersuchten Frau berücksichtigen. Eine solche Vorgehensweise wäre nicht nur innovativ im Sinne eines „Disease Managements“ oder eines modernen Gesundheitsmanagements, sondern würde zu Ressourceneinsparungen sowie zur gesamten Verringerung des seelischen Leids infolge von Fehldiagnosen führen. Darüber hinaus könnte durch die Berücksichtigung von individuellen Präferenzen die Teilnehmerate möglicherweise gesteigert werden, was als zentrales Kriterium für den Erfolg des Screenings betrachtet werden muss.

Die Empfehlungen, bzw. Richtlinien, in welchem Alter mit einem Screening auf Zervixkarzinom begonnen bzw. geendet werden soll, variieren beträchtlich. Da z.B. der Start an das Alter zum Zeitpunkt des ersten Geschlechtsverkehrs gebunden ist, sind hier keine generellen Grenzen festlegbar. In dieser Arbeit wurde daher auf die gesetzlichen Rahmenbedingungen – vor allem der Bundesrepublik Deutschland – eingegangen. Durch die Berücksichtigung der Anzahl der Abstriche während der gesamten Lebensspanne wird das Problem der tatsächlichen stark variierenden Beginnzeitpunkte des Screenings, umgangen. Analog gilt dies für das ebenso für Empfehlungen bezüglich des Endes der Screeningspanne.

4 Anhang

4.1 Richtlinien der American Cancer Society zur Früherkennung von Krebs (englisch)

- All women should begin cervical cancer screening about 3 years after they begin having vaginal intercourse, but no later than when they are 21 years old. Screening should be done every year with the regular Pap test or every 2 years using the newer liquid-based Pap test.
- Beginning at age 30, women who have had 3 normal Pap test results in a row may get screened every 2 to 3 years with either the conventional (regular) or liquid-based Pap test. Women who have certain risk factors such as diethylstilbestrol (DES) exposure before birth, HIV infection, or a weakened immune system due to organ transplant, chemotherapy, or chronic steroid use should continue to be screened annually.
- Another reasonable option for women over 30 is to get screened every 3 years (but not more frequently) with either the conventional or liquid-based Pap test, *plus* the HPV DNA test (see below for more information on this test).
- Women 70 years of age or older who have had 3 or more normal Pap tests in a row and no abnormal Pap test results in the last 10 years may choose to stop having cervical cancer screening. Women with a history of cervical cancer, DES exposure before birth, HIV infection or a weakened immune system should continue to have screening as long as they are in good health.
- Women who have had a total hysterectomy (removal of the uterus and cervix) may also choose to stop having cervical cancer screening, unless the surgery was done as a treatment for cervical cancer or precancer. Women who have had a hysterectomy without removal of the cervix should continue to follow the guidelines above.

Some women believe that they do not need examinations by a health care professional once they have stopped having children. This is not correct. They should continue to follow ACS guidelines.

Although the Pap test has been more successful than any other screening test in preventing a cancer, it is not perfect. One of its limitations is that Pap tests are examined by humans, so an accurate analysis of the hundreds of thousands of cells in each sample is not always possible. Engineers, scientists, and doctors are working together to improve this test. Because some abnormalities may be missed (even when samples are examined in the best laboratories), it is not a good idea to have this test less often than ACS guidelines recommend.

Increasing the Accuracy of Your Pap Tests

There are several things you can do to make your Pap test as accurate as possible:

- Try not to schedule an appointment for a time during your menstrual period.
- Do not douche for 48 hours before the test.
- Do not have sexual intercourse for 48 hours before the test.
- Do not use tampons, birth control foams, jellies, or other vaginal creams or vaginal medications for 48 hours before the test.

Quelle:

http://www.cancer.org/docroot/CRI/content/CRI_2_4_2X_Can_cervical_cancer_be_prevented_8.asp?sitearea= (Datum 02.02.2004)

4.2 Empfehlungen der US Preventive Services Task Force (englisch)

What's New From the USPSTF

Screening for Cervical Cancer

This series of fact sheets is based on the work of the U.S. Preventive Services Task Force (USPSTF). The USPSTF systematically reviews the evidence of effectiveness of a wide range of clinical preventive services—including screening, counseling, and chemoprevention (the use of medication to prevent diseases)—to develop recommendations for preventive care in the primary care setting.

This fact sheet presents highlights of USPSTF recommendations on this topic and should not be used to make treatment or policy decisions.

Select for PDF file (75.5 KB). PDF Help.

What Screening Is Recommended by the USPSTF?

- The U.S. Preventive Services Task Force (USPSTF) strongly recommends screening women for cervical cancer if they are sexually active and have a cervix.
- The USPSTF recommends against routinely screening women older than age 65 if they have had adequate recent screening with normal Pap smears and are not otherwise at increased risk for cervical cancer.
- The USPSTF recommends against routine Pap screening for women who have had a total hysterectomy for benign disease.
- The USPSTF concludes that the evidence is insufficient to recommend for or against new technologies (such as ThinPrep®) in place of conventional Pap tests.
- The USPSTF concludes that the evidence is insufficient to recommend for or against *human papillomavirus* (HPV) testing as a primary screening test for cervical cancer.

The Task Force concludes that screening should begin within 3 years of the start of sexual activity or age 21, whichever comes first, and should be done at least every 3 years. The risk for cervical cancer and the yield of screening decline through middle age.

For women older than 65 who have had normal Pap smears, the benefits of continued screening may not outweigh the potential harms, such as false-positive test results and invasive procedures. The Task Force also concludes that the yield of detecting vaginal neoplasms is too low to justify continuing screening after a total hysterectomy.

Most cases of cervical cancer occur in women who are not screened adequately. Clinicians, hospitals, and health plans should develop systems to identify and screen women, including older women, who have had no screening or who have been screened inadequately in the past.

Quelle: <http://www.ahrq.gov/clinic/3rduspstf/cervcan/cervcanwh.htm> (02.02.2004)

4.3 Abkürzungsverzeichnis

ACS	American Cancer Society
ASC	Atypical squamous cells; dt. atypische Zellen
EBM	Evidence-based Medicine, dt.: evidenzbasierte Medizin
EU	Europäische Union
FIGO	International Federation of Gynecology and Obstetrics
FBC	Flüssigkeitsbasierte Zytologie
HIV	Human-Immunodeficiency-Virus-Test
HPV	Humanes Papillomavirus
HPV-DNA	Humanes Papillomavirus – Desoxyribonuclein-acid
HTA	Health Technology Assessment
IARC	The International Agency for Research on Cancer
NPV	Negative Predicted Value, dt.: negativer Vorhersagewert
NZHTA	New Zealand Health Technology Assessment
PPV	Positive Predicted Value dt.: positiver Vorhersagewert
SGB	Sozialgesetzbuch
TNM	Tumor-Node-Metastasen Klassifikationssystem
UICC	International Union Against Cancer
USPSTF	US Preventive Services Task Force
WHO	Weltgesundheitsorganisation

4.4 Tabellenverzeichnis

Tabelle 1: Häufigkeit von atypischen Zellen (ASC) oder negative und positive Tests auf karzinogene humane Papillomavirus Typen.....	9
---	---

Tabelle 2: Die Papanicolaou-Test-Gruppen	10
Tabelle 3: Stadieneinteilungen des Zervixkarzinoms nach FIGO (Fédération Internationale de Gynécologie et d'Obstétrique, 1994) und TNM-Klassifikation (1997).	11
Tabelle 4: Art der Screeningprogramme und Zeitpunkt der Gründung in der EU.....	19
Tabelle 5: Einladung zum Zervixkarzinom-Screening und Einschlusspopulation innerhalb der europäischen Union.....	20
Tabelle 6: Organisation des Zervixkarzinom-Screenings innerhalb der europäischen Union.	22
Tabelle 7: Kosten und Nutzen in Verbindung mit tatsächlich positiven, falsch-positiven, falsch-negativen und tatsächlich negativen Testergebnissen.....	26
Tabelle 8: Kenngrößen der kombinierten Pap-Tests in Abhängigkeit des Screeningintervalls.	31
Tabelle 9: Kenngrößen der kombinierten FBC-Tests in Abhängigkeit des Screeningintervalls.	32
Tabelle 10: Ergebnisse der Screeninguntersuchungen in Abhängigkeit des Untersuchungsintervalls für den Pap-Test auf 100.000 Frauen.....	33
Tabelle 11: Fallkonstellation flüssigkeitsbasierte Zytologie in Abhängigkeit der Untersuchungsintervalle.	34
Tabelle 12: Erkannte und übersehene Fälle in Abhängigkeit des Untersuchungsintervalls für die Bundesrepublik Deutschland.....	35
Tabelle 13: Anwendung der neuen Technologien in Abhängigkeit der Untersuchungsintervalle für die Bundesrepublik Deutschland.....	36
Tabelle 14: Erkannte und übersehene Zervixkarzinomfälle in Abhängigkeit der Teilnehmerate.	37
Tabelle 15: Ökonomische Konsequenzen des konventionellen Pap-Tests für die Bundesrepublik Deutschland in Abhängigkeit des Untersuchungsintervalls.....	38
Tabelle 16: Ökonomische Konsequenzen des konventionellen Pap-Tests für die Bundesrepublik Deutschland in Abhängigkeit des Untersuchungsintervalls.....	39
Tabelle 17: Erkannte Fälle sowie inkrementeller Nutzen in Abhängigkeit des Untersuchungsintervalls (Basisszenario, konventioneller Pap-Test).	40
Tabelle 18: Ergebnisse verschiedener Screeningstrategien in Abhängigkeit des Untersuchungsintervalls.	41
Tabelle 19: Sensitivität – Übersicht über die Variabilität der Kenngrößen.....	42

4.5 Abbildungsverzeichnis

Abbildung 1: Ertrag des Screenings in Abhängigkeit des Untersuchungsintervalls.	40
--	----

5 Literaturverzeichnis

1. Basset L, Butler D: **Mammography and early breast-cancer detection.** In: Am Fam Physican (1991), 43, S. 546-557.
2. Basset I, Gold R: **False negative mammograms: Causes and effects.** In: Contemp Diagn Radiol (1988), 11, S. 1-6.
3. Böcker W (Hrsg): **Pathologie.** 2.Auflage, Urban & Fischer, München 2001.
4. Broadstock M: **Effectiveness and cost effectiveness of automated and semi automated cervical screening devices. A systematic review of the literature.** New Zealand Health Technology Assessment (NZHTA). The Clearing House for Health Outcomes and Health Technology Assessment. NZHTA Report. 3(1). 2000.
5. Broadstock M: **Effectiveness and cost effectiveness of automated and semi automated cervical screening devices.** The Clearing House for Health Outcomes and Health Technology Assessment. NZHTA Report 2000; 3(1). Kurzfassung in: Cytopathology (2001), 12, S. 141-143.
6. Brown AD., Garber AM: **Cost-effectiveness of 3 methods to enhance the sensitivity of Papanicolaou testing.** In: JAMA (1999), Jan 27;281(4), S. 347-353.
7. Bundesärztekammer: **Leitlinie der Bundesärztekammer zur Qualitätssicherung zytologischer Untersuchungen im Rahmen der Früherkennung des Zervixkarzinoms** In: Deutsches Ärzteblatt (1994), 91, B298-B300.
8. Castle PE, Wacholder S, Lorincz AT, Scott DR, Sherman ME, Glass AG, Rush BB, Schussler JE, Schiffman M: **A prospective study of high-grade cervical neoplasia risk among human papillomavirus-infected women.** In: J Natl Cancer Inst. (2002), Sep 18, 94(18), S. 1406-1614.
9. Cole P, Morrison AS: **Basic issues in cancer screening.** In: Miller AB (Hrsg.): Screening in cancer. UICC Tech Rep Ser; 40: 7-39; 1978.
10. Coleman D, Day N, Douglas G, Farmery E, Lynge E, Philip J, Segnan N: **European Guidelines for Quality Assurance in Cervical Cancer Screening. Europe against cancer programme.** In: Eur J Cancer (1993), 29A Suppl 4, S1-38.
11. Cruickshank ME, Angus V, Kelly M, McPhee S, Kitchener HC: **The case for stopping cervical screening at age 50.** In: Obstet Gynaecol (1997), May, 104(5), S. 586-589.
12. Day NE: **Screening for Cancer of the cervix.** In: J Epidemiol Community Health (1989), 43, S. 103-106.
13. Donaldson C: **Using economics to assess the place of screening.** In: J Med Screen (1994), 1, S. 124-128.
14. Earle C, Herbert P: **A reader's guide to the evaluation of screening-studies.** In: Postgrad Med J (1996), 72, S. 77-83.
15. Eddy DM: **Screening for cervical cancer.** In: Ann Intern Med (1990) Aug 1, 113(3), S. 214-226 [review].
16. Evander M, Edlund K, Gustafsson A, Jonsson M, Karlsson R, Rylander E, Wadell G: **Human papillomavirus infection is transient in young women: a population-based cohort study.** In: J Infect Dis (1995), Apr, 171(4), S. 1026-1030.
17. Fahey MT, Irwig L, Macasill P: **Meta analysis of Pap test accuracy.** In: Am J Epidemiol (1995), 141, S. 680 – 689.
18. Fahs MC, Mandelblatt J, Schechter C, Muller C: **Cost effectiveness of cervical cancer screening for the elderly.** In: Ann Intern Med (1992), Sep 15, 117(6), S. 520-527.
19. Fahs MC, Plichta SB, Mandelblatt JS: **Cost-effective policies for cervical cancer screening. An international review.** In: Pharmacoeconomics (1996), März, 9(3), S. 211-230 [review].

20. Fischer U, Raptis G, Geßner W, Roschlau U, Dressel K, Bilek K, Horn LC: **Epidemiologie und formale Pathogenese des Zervixkarzinoms.** In: Zentralbl Gynäkol (2001), 123, S. 198-205.
21. Mitchell MF, Cantor SB, Brookner C, Utzinger U, Schottenfeld D, Richards-Kortum R: **Screening for squamous intraepithelial lesions with fluorescence spectroscopy.** In: Obstet Gynecol (1999), 94, S. 889-896.
22. Frable W: **Does a zero error standard exist for the papanicolaou smear? A pathologist's perspective.** In: Arch. Pathol. Lab. Med. (1997), 121, S. 301-310.
23. Goldie SJ, Kuhn L, Denny L, Pollack A, Wright TC: **Policy Analysis of Cervical Cancer Screening Strategies in Low-Resource Settings Clinical Benefits and Cost-Effectiveness.** In: JAMA (2001), Jun 27, Vol 285, No 24.
24. Grayson A: **Fetal screening. The triple test decision.** In: Mod Midwife (1996), 6, S. 16-19.
25. Hennekens CH, Buring JE: **Epidemiology in Medicine.** Boston - Toronto: Little Brown 1987.
26. Herbert A: **Is proficiency testing in cervical cytology proficient?** In: J. Clin Path (1997), 50, S. 536-537.
27. Hildesheim A, Schiffman MH, Gravitt PE, Glass AG, Greer CE, Zhang T, Scott DR, Rush BB, Lawler P, Sherman ME: **Persistence of type-specific human papillomavirus infection among cytologically normal women.** In: J Infect Dis (1994) Feb, 169(2), S. 235-240.
28. Ho GY, Bierman R, Beardsley L, Chang CJ, Burk RD: **Natural history of cervicovaginal papillomavirus infection in young women.** In: N Engl J Med (1998), Feb 12, 338(7), S. 423-428.
29. IARC Working Group on Evaluation of Cervical Cancer Screening programmes: **Screening for squamous cervical cancer: duration of low risk after negative results of cervical cytology and its implication for screening policies.** In: BMJ (1986), 293, S. 659-664.
30. Jenkins D, Sherlaw-Johnson C, Gallivan S: **Can papilloma virus testing be used to improve cervical cancer screening?** In: Int J Cancer (1996), März 15, 65(6), S. 768-773.
31. Kassenärztliche Bundesvereinigung: **Neufassung der Qualitätsvoraussetzungen in der gynäkologischen Zytologie** In: Deutsches Ärzteblatt (1992b), 89, S. 1357-1358.
32. Kimberley A, Quaid PD: **Psychological and ethical Considerations in Screening for Disease.** In: Am J Cardiol (1993), Sept.30; 72, S. 64-67.
33. Koutsky L: **Epidemiology of genital human papillomavirus infection.** In: Am J Med (1997), May 5, 102(5A), S. 3-8.
34. Levi F, Lucchini F, Negri E, Franceschi S, la Vecchia C: **Cervical cancer mortality in young women in Europe patterns and trends.** In: Eur J Cancer (2000) Nov; 36(17), S. 2266-2271.
35. Linos A, Riza E: **Comparisons of cervical cancer screening programmes in the European Union.** In: Eur J Cancer (2000), 36, S. 2260-2265.
36. Mandelblatt JS, Lawrence WF, Womack SM, Jacobson D, Yi B, Hwang YT, Gold K, Barter J, Shah K: **Benefits and costs of using HPV testing to screen for cervical cancer.** In: JAMA (2002), May 8, 287(18), S. 2372-2381.
37. McCrory DC, Matchar DB, Bastian L, et al.: **Evaluation of Cervical Cytology. Evidence Report / Technology Assessment No. 5. (Prepared by Duke University under Contract No. 290-97-0014.)** AHCPR Publication No. 99-E010. Rockville, MD: Agency for Health Care Policy and Research; February 1999.
38. McQueen MJ: **Some ethical and designed challenges of screening programs and screening tests.** In: Clin Chim Acta (2002), Jan 315 (1-2), S. 41-48.
39. Milborn K, MacAskill S: **Cervical cancer screening: Continuing Concerns in the 1990's.** In: Health Ed. J. (1994), 54, S. 201-213.

40. Miller AB, Chamberlain J, Day NE, Hakama M, Prorok PC: **Report on a Workshop of the UICC Project on Evaluation of Screening for Cancer.** In: Int J Cancer (1990), Nov 15, 46(5), S. 761-769.
41. Myers ER, Mccrory DC, Subramanian S, Mccall N, Nanda K, Datta S, Matchar DB: **Getting the Target for a better Cervical Screening Test: Characteristics of a Cost-Effective Test for Cervical Neoplasia Screening.** In: Obstet Gynecol (2000), Vol. 96, Nov. No 5 P1.
42. Nanda K, McCrory DC, Myers ER, Bastian LA, Hasselblad V, Hickey JD, Matchar DB: **Accuracy of the Pap test in screening for and follow up of cervical cytologic abnormalities. A systematic review.** In: Ann Intern Med (2000), May 16, 132(10), S. 810-819.
43. Nygard JF, Skare GB, Thoresen SO: **The cervical cancer screening programme in Norway, 1992-2000: changes in Pap smear coverage and incidence of cervical cancer.** In: J Med Screen (2002), 9, S. 86-91.
44. Petticrew MP, Sowden AJ, Lister-Sharp D, Wright K: **False False-negative results in screening-Programs: Systematic review of impact and implications.** In: Health technol assess (2000), 4, S. 1-120.
45. Petticrew MP, Sowden AJ, Lister-Sharp D: **False-negative results in screening Programs.** In: Int J technol Assess Health Care (2001), 17:2, S. 164-170.
46. Pfleiderer A (Hrsg.): **Gynäkologie und Geburtshilfe.** Thieme Stuttgart, New York 2000.
47. Ponten J, Adami HO, Bergstrom R, Dillner J, Friberg LG, Gustafsson L, Miller AB, Parkin DM, Sparen P, Trichopoulos D: **Strategies for global control of cervical cancer.** In: Int J Cancer (1995), Jan 3, 60(1), S. 1-26.
48. Pschyrembel: **Klinisches Wörterbuch.** De Gruyter, Berlin 1998.
49. Raffle AE, Alden B, Mackenzie EF: **Detection rates for abnormal cervical smears: what are we screening for?** In: Lancet (1995), Jun 10, 345(8963), S. 1469-1473.
50. Richart RM, Masood S, Syrjanen KJ, Vassilakos P, Kaufman RH, Meisels A, Olszewski WT, Sakamoto A, Stoler MH, Vooijs GP, Wilbur DC: **Human papillomavirus.** International Academy of Cytology Task Force summary. Diagnostic Cytology Towards the 21st Century: An International Expert Conference and Tutorial. In: Acta Cytol (1998), Jan-Feb, 42(1), S. 50-58 [review].
51. Sawaya GF, Grady D, Kerlikowske K: **The positive predicted value of cervical smears in previously screened postmenopausal women: The Heart and Estrogen / progestin Replacement Study (HERS).** In: Ann Intern Med (2000), Vol. 133, S. 942-950.
52. Schenck U: **Befundwiedergabe in der Zytologie: Münchner Nomenklatur II und Bethesda System 1991.** In Schenck U (Hrsg.): Referateband der 13. Fortbildungstagung für klinische Zytologie München, 26. 11-02.12., S 224-233, 1995.
53. Schneider A, Durst M, Kaufmann AM: **HPV-Infektion und Zervixkarzinom.** In: Zentralbl Gynakol (2001), Apr;123(4), S. 179-185.
54. Seror V, Muller F, Moatti J, Legales C, Boue A: **Economic Assessment of maternal serum screening for Down's Syndrome using human chorionic gonadotropin.** In: Prenat Diagn (1993), 13, S. 281-292.
55. Sherlaw-Johnson C, Gallivan S, Jenkins D: **Withdrawing low risk women from cervical screening programmes: mathematical modelling study.** In: BMJ (1999), Feb 6, 318(7180), S. 356-360.
56. Sherlaw-Johnson C, Gallivan S, Jenkins D, Jones MH: **Cytological screening and management of abnormalities in prevention of cervical cancer: an overview with stochastic modelling.** In: J Clin Pathol (1994), May, 47(5), S. 430-435.
57. Sherman ME, Lorincz AT, Scott DR, Wacholder S, Castle PE, Glass AG, Mielzynska-Lohnas I, Rush BB, Schiffman M: **seline cytology, human papillomavirus testing, and risk for cervical neoplasia: a 10-year cohort analysis.** In: J Natl Cancer Inst. (2003), Jan 1, 95(1), S. 46-52.

58. Shickle D, Chadwick R: **The ethics of screening: Is 'screeningitis' an incurable disease?** In: J med ethics (1994), 20, S. 12-18.
59. Sigurdsson K, Arnadottir T, Snorraddottir M, Benediktsdottir K, Saemundsson H: **Human papillomavirus (HPV) in an Icelandic population: the role of HPV DNA testing based on hybrid capture and PCR assays among women with screen-detected abnormal Pap smears.** In: Int J Cancer (1997), Jul 29, 72(3), S. 446-452.
60. Soost HJ: **Nomenklatur und Befundwiedergabe in der gynäkologischen Zytologie.** In Schenck U, Soost HJ (Hrsg.): Referate der 11. Fortbildungstagung für klinische Zytologie in München, 01-07 Dezember, München. Klinische Zytologie u. Autoren S. 251-258; 1991.
61. Stoler MH, Schiffmann M: **Interobserver Reproducibility of Cervical Cytologic and Histologic Interpretations. Realistic Estimates From the ASCUS-LSIL Triage Study.** In: JAMA (2001), March 21, Vol 285, No 11.
62. Sulik SM, Kroeger K, Schultz JK, Brown JL, Becker LA, Grant WD: **Are fluid-based cytologies superior to the conventional Papanicolaou test? A systematic review.** In: Fam Pract (2001), Dec, 50(12), S. 1040-1046 [review].
63. Swedish Council on Technology Assessment in Health Care (SBU): **Mass Screening for Prostate Cancer.** In: Int J Cancer (1996), Suppl. 9: I-III, S. 1-72.
64. Tymstra T, Bieleman B: **The psychosocial impact of mass screening for cardiovascular risk factors.** In: Fam Pract (1987), 4, S. 287-290.
65. v.d.Akker-van Marle ME, van Ballegooijen M, van Oortmarssen GJ, Boer R, Habbema JDF: **Cost-Effectiveness of Cervical Cancer Screening: Comparison of Screening Policies.** In: J Nat Cancer Inst (2002), Feb 6, Vol 94, No 3.
66. Van Dijk J, Verbeek A, Hendriks J, Holland R: **The current detectability of breast cancer in a mammographic screening program. A review of the previous mammograms of interval and screen-detected cancers.** In: Cancer (Phila.) (1993), 72, S. 1933-1938.
67. van Doornum GJ, Prins M, Juffermans LH, Hooykaas C, van den Hoek JA, Coutinho RA, Quint WG: **Regional distribution and incidence of human papillomavirus infections among heterosexual men and women with multiple sexual partners: a prospective study.** In: Genitourin Med (1994), Aug, 70(4), S. 240-246.
68. Viikki M, Pukkala E, Hakama M: **Risk of cervical cancer after a negative Pap smear.** In: J Med Screen (1999), 6(2), S. 103-107.
69. Wagener C, Hossfeld DK: **Diagnostische Validität von Tumormarkern.** In: Urologe (B) (1997), 37, S. 203-208.
70. Waugh N, Robertson A: **Costs and benefits of cervical screening. II. Is it worthwhile reducing the screening interval from 5 to 3 years?** In: Cytopathology (1996), Aug, 7(4), S. 241-248.
71. Waugh N, Smith I, Robertson A, Reid GS, Halkerston R, Grant A: **Costs and benefits of cervical screening. III. Cost / benefit analysis of a call of previously unscreened women.** In: Cytopathology (1996), Aug, 7(4), S. 249-255.
72. Waugh N, Smith I, Robertson A, Reid GS, Halkerston R, Grant A: **Costs and benefits of cervical screening. I. The costs of the cervical screening programme.** In: Cytopathology (1996), Aug, 7(4), S. 231-240.
73. Wilson JMG, Jungner G: **Principles and Practice of Screening for Disease.** WHO Genf 1968.
74. Schenck U, von Karsa L: **Cervical cancer screening in Germany.** In: Eur J Cancer, Nov, 36(17), S. 2221-2226.
75. Gesundheitsberichterstattung des Bundes: **Gesundheitsbericht für Deutschland.** Nov. 1998.

76. American Cancer Society (ACS) **Guidelines for the early detection of Cervical Neoplasia and Cancer.**
http://www.cancer.org/docroot/CRI/content/CRI_2_4_2X_Can_cervical_cancer_be_prevented_8.asp?sitearea
77. **Recommendations and Rationale Screening for Cervix Cancer.** (Feb. 2004 – online).
<http://www.ahrq.gov/clinic/3rduspstf/cervcan/cervcanwh.htm>
78. Shingleton HM, Patrick RL, Johnston WW, Smith RA: **The current status of the Papanicolaou smear.** In: CA Cancer J Clin.(1995), 45(5), S. 305-320.