

Vaginale pH-Selbstmessung zur Verhinderung von Frühgeburten weiterhin gerechtfertigt

K. Friese, W. Kirschner, A. Scheffler

Im Deutschen Ärzteblatt (1) wurden kürzlich die Ergebnisse eines Modellprojekts zur pH-Selbstmessung nach § 63 SGB V vorgestellt. Die Publikation weist folgende Schlussfolgerung auf: „Die Effektivität der Screening-Intervention hinsichtlich der Frühgeburten <37+0 SSW konnte nicht belegt werden.“ Die DGGG bezieht zu dieser Studie und ihren Ergebnissen Stellung. Denn die vaginale pH-Messung bleibt zur Reduktion des Frühgeburtsrisikos nach wie vor eine erfolgversprechende Methode.

Zwischen 2004 und 2006 haben vier Krankenkassen ihre Versicherten über das Projekt informiert und interessierten Schwangeren auf Anforderung die Testhandschuhe kostenlos zur Verfügung gestellt. Modellvorhaben nach § 63 SGB V erfordern eine Evaluation der Interventionsmaßnahmen, da es u.a. Ziel der gesetzlichen Norm ist, Maßnahmen, die sich in der Evaluation als wirksam und wirtschaftlich erweisen, gegebenenfalls in die Regelversorgung zu überführen. Das Vorhaben hatte so einen nicht geringen versorgungspolitischen Impact. Mit der Intervention und Evaluation sollte belastbar geprüft werden, ob die pH-Messung tatsächlich zu einer relevanten und signifikanten Verringerung der Frühgeburten führt.

Die Publikation weist folgende Schlussfolgerung auf: „Die Effektivität der Screening-Intervention hinsichtlich der Frühgeburten <37+0 SSW konnte nicht belegt werden.“ Tatsächlich liegt in der Interventionsgruppe eine grenzwertig erhöhte Frühgeburtenrate vor (RR: 1,06; 95% KI 1,00/1,12).

Wie ist dieser Befund epidemiologisch-evaluativ und individualmedizinisch zu bewerten?

Das Studienergebnis kann u.E. grundsätzlich keinen Zweifel an der individualmedizinischen Eignung des pH-Testhandschuhs zur Erkennung von Störungen der Vaginalflora begründen, wenn die Applikation regelgerecht durchgeführt wird. Festgestellte Störungen des Scheidenmilieus sollten entsprechend therapiert werden. Auch die Eignung des Handschuhs als Instrument zum Screening auf möglicherweise bestehende oder sich entwickelnde vaginalinfektionen (z.B. bakterielle Vaginose) steht außer Frage, wenn die positiven Fälle im Screening fachgerecht differenzialdiagnostisch abgeklärt und diagnostizierte Krankheiten fachgerecht und wirksam therapiert werden. Hierin besteht zweifellos das prinzipielle präventive Potenzial des Instruments, um wichtige Risikofaktoren der Frühgeburt und Frühgeburten selbst zu vermindern. Diese Bewertung gilt u.E. trotz der Tatsache, dass Informationen zu wichtigen Kenngrößen zur Qualität des Screeninginstruments (z.B. falsch positive Fälle/falsch negative Fälle usw.) bisher nur begrenzt vorliegen.

Diese Bewertung gilt zunächst für die Handhabung des Instruments durch den Frauenarzt im Verlauf der

Schwangerschaft. Auch die pH-Selbstmessung durch Schwangere führt aber letztlich zu keiner anderen Bewertung, wenn die Applikation regelmäßig und regelgerecht durchgeführt wird und positive Befunde zu einer zeitnahen ärztlichen Abklärung führen. Unter Beachtung dieser Kriterien kann der Einsatz des pH-Testhandschuhs individuell zu einer Vermeidung einer infektionsbedingten Frühgeburt führen.

Zur Wirksamkeit des pH-Testhandschuhs in Kollektiven liegt in den letzten Jahren und Jahrzehnten eine Reihe von Publikationen vor, die die Plausibilität der Wirksamkeit stützen, einen statistisch belastbaren und sicheren Nachweis der Wirksamkeit aber nicht erbrachten, da die Studien teilweise erhebliche methodische Mängel aufweisen (2).

Vor diesem Hintergrund wurde im Jahr 2002/2003 mit Vorbereitungen und Verhandlungen zur Durchführung eines Modellprojekts „pH-Selbstmessung“ begonnen, wobei das Vorhaben im medizinischen Dienst der Krankenkassen in hohem Maße umstritten war (2).

Der Nachweis der Wirksamkeit der pH-Messung zur Verringerung von Frühgeburten ist vor dem Hintergrund eines prinzipiell gegebenen Wirksamkeitspotenzials (s.o.) durch gut geplante, gut durchgeführte und gut evaluierte Interventionsstudien an ausreichend großen Untersuchungskollektiven zu führen. Im Idealfall arbeitet jede Intervention mit randomisierten und wenn mög-

lich mit doppelt verblindeten Fallkontrollstudien.

Wenn von der Evaluation einer Intervention erwartet wird, dass sie in Grenzen des Konfidenzintervalls (üblicherweise 95%) auf der Grundlage eines Hypothesentests statistisch gesicherte Bewertungen der Wirksamkeit liefert, so muss hier zunächst darauf hingewiesen werden, dass auch die Evaluation analog der Vier-Felder-Tafel folgende Outcomes und die damit verbundenen zwei Fehlermöglichkeiten beinhaltet:

1. Der Intervention wird Wirksamkeit attestiert und sie ist in Wirklichkeit auch wirksam (richtig positive Entscheidung).
2. Der Intervention wird Wirksamkeit attestiert, sie ist in Wirklichkeit aber nicht wirksam (falsch positive Entscheidung).
3. Der Intervention wird keine Wirksamkeit attestiert, sie ist in Wirklichkeit auch nicht wirksam (richtig negative Entscheidung).
4. Der Intervention wird keine Wirksamkeit attestiert, sie ist in Wirklichkeit aber wirksam (falsch negative Entscheidung).

Logischerweise sind die Evaluationsoutcomes 2 und 4 kritisch, da sie gegebenenfalls zum Angebot unwirksamer oder zur Vorenthaltung wirksamer Maßnahmen führen.

Für die epidemiologisch-evaluative Bewertung der Ergebnisse der Evaluation des pH-Testhandschuhs stellt sich die Frage, wie wahrscheinlich vor dem Hintergrund der gegebenen Datenlage die Outcomes 3 oder 4 sind.

Die Evaluation von Interventionsmaßnahmen gehört zu den schwierigsten Tätigkeitsfeldern in der empirischen Forschung. Aufgrund der Komplexität der Aufgabenstellung erfolgt die Evaluation in der Regel in vier Phasen:

1. Ex-ante Evaluation (Interventionsplanung),
2. Strukturevaluation,
3. Produkt- und Prozessevaluation,
4. Ergebnisevaluation.

Zentrales Ziel der Interventionsplanung ist es, in einem Fall-Kontrollansatz zufällige, möglichst nicht selektierte Interventionsgruppen zu generieren, wobei Selektionen nach soziodemographischen, sozialen, gesundheitlichen und auch motivationalen Faktoren möglichst weitestgehend auszuschließen sind.

Sich selbst selektierende Stichprobe

De facto lag in der Studie in der Interventionsgruppe eine ausschließlich sich selbst selektierende „Stich-

probe“ vor, die das Untersuchungsangebot ihrer Krankenkasse kennen und selbst aktiv werden musste, um die Testhandschuhe durch eine Bestellung zu erhalten. In Bezug auf die Richtung der Selektionen ist einerseits zu vermuten, dass gesundheitsbewusste Schwangere mittlerer und höherer Schichten überproportional vertreten sind, es ist aber auch nicht unplausibel anzunehmen, dass z.B. Schwangere mit Komplikationen in früheren Schwangerschaften sich überproportional beteiligt haben.

Mit Blick auf Alternativen zur Gruppenbildung ist einzuräumen, dass es die Möglichkeit einer Zufallsauswahl auf Seiten der Krankenkassen nicht gab, da den Kassen die Schwangerschaft ihrer Versicherten im notwendig frühen Stadium nicht bekannt ist. Eine Zufallsauswahl wäre prinzipiell aber über die Beteiligung der niedergelassenen Frauenärzte erreichbar oder jedenfalls anstrebbare gewesen.

In der Interventionsplanung ist auch sicherzustellen, dass die Qualität der Intervention hinsichtlich der Strukturen, Produkte und Prozesse optimal verläuft und dass die notwendigen Daten für die Ergebnisevaluation vollständig und reliabel erhoben oder generiert werden.

Interventionen bedürfen einer sorgfältigen Steuerung und Kontrolle der Strukturen, in denen sie organisiert werden, der Produkte, die sie einsetzen und der Prozesse, in denen sie ablaufen. Im Rahmen der Studie war eine Steuerung und Kontrolle der Strukturen und Prozesse nicht möglich, die Interventionen verliefen nach Erhalt der Handschuhe in Eigenregie der teilnehmenden Schwangeren.

Ob das Produkt, der pH-Testhandschuh,

- überhaupt,
- in den empfohlenen Intervallen regelmäßig,
- richtig und regelgerecht angewendet wurde, konnte lediglich über einen Fragebogen erhoben werden, der an einer Teilstichprobe der Probanden eingesetzt wurde. Die Reliabilität und Validität der Angaben zu diesen Fragen ist ebenso unklar wie die Repräsentativität der Teilstichprobe für die gesamte Interventionsgruppe.

Auch Aussagen zu wichtigen Fragen wie

- dem Auftreten positiver pH-Befunde,
- der wiederholten Testung,
- der Arztbesuche zur Abklärung und
- der Infektionsdiagnostik und gegebenenfalls der Therapie basieren ausschließlich auf den Fragebogenangaben der Teilstichprobe.

Auch in Bezug auf die Organisation und Steuerung der Intervention und die Dokumentation hätte eine Alternative wahrscheinlich auch hier nur in der Einbeziehung von Frauenärzten bestanden.

Sind diese Limitierungen und Probleme in der Intervention letztlich dem gewählten Interventionsansatz im Zusammenspiel von Kassen und schwangeren Versicherten geschuldet, so fügt der gewählte methodische Ansatz zur Ergebnisevaluation weitere geradezu gravierende Prob-

leme hinzu. Ein Fragebogen kommt nur bei der Interventionsgruppe als Teilstichprobe zum Ansatz. Auf eine Befragung einer Stichprobe der Kontrollgruppe wird völlig verzichtet. Gegenstand dieser Befragung wären in der Fall- und Kontrollgruppe die Erfassung aller relevanten Risikofaktoren und protektiven Faktoren der Frühgeburt gewesen.

Offene Fragen bleiben ungeklärt

Nur auf diese Weise wäre es möglich gewesen, die (unterschiedlichen) Selektionen in den Gruppen zu erkennen und gegebenenfalls zu standardisieren. Die wenigen Daten aus den Perinatal- und Kassendaten reichen dafür keineswegs aus, z.T. sind diese Daten auch nicht valide (z.B. Rauchen). Das Argument der fehlenden Akzeptanz und der hohen Kosten dieses Ansatzes ist keinesfalls stichhaltig.

Ist so der Verzicht auf diese Datenerhebung der zentrale Mangel in der Ergebnisevaluation, so ist er auch der Grund, dass auf der Basis der in der Studie gewonnenen Daten nicht in Erfahrung zu bringen ist, warum die Intervention nicht wirksam ist.

Wenn man auf der Grundlage der in der Studie erhobenen Daten (Dokumentationsbogen) aber zumindest sieht, dass 1.887 von 7.469 Schwangeren (25%) mindestens einen abweichenden pH-Wert nannten, aber nur 803 (10,8%) eine ärztliche Kontrolle durchführten ließen und nur 642 (9%) wie auch immer medikamentös behandelt wurden, so ist offensichtlich, dass damit eine Reduktion der Frühgeburtenrate gar nicht erreichbar ist.

Die finanziell, organisatorisch und methodisch aufwändige Studie zur Frage der Wirksamkeit der pH-Messung hat keine der hierzu bestehenden nicht wenigen offenen Fragen einer Klärung nahegebracht. Alles spricht also für einen richtig negativen Evaluationsbefund.

Der Einsatz des pH-Instruments ist u.E. daher auf der Basis der Ergebnisse nur aufgrund ärztlicher Indikation und unter ärztlicher Anleitung sowie Verlaufskontrolle sinnvoll. Jeder weitere Einsatz des Instruments in Kollektiven sollte von einer belastbaren Evaluation begleitet sein.

Literatur

1. Bitzer EM, Schneider A, Wenzlaff P, Hoyme UB, Siegmund-Schultze E: Dtsch Arztebl Int 2011;108(6):81-6.DOI:10.3238/ärztebl.2011.0081
2. S. Bauer: Sozialmedizinische Stellungnahme zur Überprüfung des vaginalen pH-Wertes mit Hilfe des Einmalhandschuhs CAREPLAN, MDS, 17.04.2003, S.28/29, S.33

Für die Autoren



Dr. Wolf Kirschner
 Forschung, Beratung +
 Evaluation GmbH
 c/o Charité
 Frauenklinik CVK
 Augustenburger Platz 1
 13353 Berlin
 wolf.kirschner@fb-e.de