Evidence Brief: Accuracy of Self-report for Cervical and Breast Cancer Screening

Supplemental Materials

February 2019

Prepared for:
Department of Veterans Affairs
Veterans Health Administration
Health Services Research & Development Service
Washington, DC 20420

Prepared by:
Evidence Synthesis Program (ESP)
Coordinating Center
Portland VA Medical Center
Portland, OR
Mark Helfand, MD, MPH, MS, Director

Authors:
Investigators:
Johanna Anderson, MPH
Donald Bourne, MPH
Kim Peterson, MS
Katherine Mackey, MD, MPP

U.S. Department of Veterans Affairs
Veterans Health Administration
Health Services Research & Development Service

TABLE OF CONTENTS

APPENDIX A. CANCER SCREENING GUIDELINE RECOMMENDATIONS

Breast Cancer

Recommended By (Last updated)	Age to Start Mammograms	Screening Interval	Previous Recommendations
ACOG (2011)	40	Annual	2003: 40 (1-2 years) 50 (annually)
ACS (2015)	45 (individual choice 40-44)	Annual (45-54) 1-2 years (55+)	1976: 35 (only if history of breast cancer), 40 (only if mother or sisters had breast cancer), 50 (yearly) 1980: 35 (baseline mammogram), 40 (consult personal physician), 50 (yearly) 1983: 35 (baseline mammogram), 40 (every 1-2 years), 50 (yearly) 1992: 40 (every 1-2 years), 50 (yearly) 2003: 40 (yearly)
AMA (2012)	40	Annual	NA
CDC (2016)	50 (individual choice 40-49)	2 years	NA
NCCN (2015)	40	Annual	NA
USPSTF (2016)	50 (individual choice 40-49)	2 years	2002: 40 (every 1-2 years)
VHA (2015)	45 (individual choice 40-44)	Annual (45-54) 1-2 years (55+)	NA
WHO (2014)	50 (Individual choice 40-49)	2 years	NA

Abbreviations: ACOG = American College of Obstetrics and Gynecology, ACS = American Cancer Society, AMA = American Medical Association, CDC = Centers for Disease Control and Prevention, HPV = Human Papilloma Virus, NCCN = National Comprehensive Cancer Network, USPSTF = US Preventive Services Task Force, VHA = Veteran's Health Administration, WHO = World Health Organization

Cervical Cancer

Recommended By (Last updated)	Age to Start Pap test	Screening Interval	Previous Recommendations
ACOG (2016)	21-29 Pap test alone 30-65 Pap test and HPV test	3 years (21-29) 5 years Pap+HPV (preferred) or 3 years (just Pap) (30-65)	NA
ACS (2012)	21-29 Pap test alone 30-65 Pap test and HPV test	3 years (21-29) 5 years Pap+HPV (preferred) or 3 years (just Pap) (30-65)	Pre-1980: Not specified (as part of a regular check-up) 1980: When sexually active (yearly for 2 years, then every 3 years) 1987: 18+ and sexually active (yearly for 3 years, then discretion of doctor) 2003: 3 years after first vaginal intercourse or 21 (yearly) until age 30, then every 2-3 years.
AMA (2018)	21-29 Pap test alone 30-65 Pap test and HPV test	3 years (21-29) 5 years Pap+HPV (preferred) or 3 years (just Pap) (30-65)	NA
CDC (2018)	21-29 Pap test alone 30-65 Pap test and HPV test	3 years (21-29) 5 years Pap+HPV (preferred) or 3 years (just Pap) (30-65)	NA
NCCN (2012)	21-29 Pap test alone 30-65 Pap test and HPV test	3 years (21-29) 5 years Pap+HPV (preferred) or 3 years (just Pap) (30-65)	NA
USPSTF (2018)	21-29 Pap test alone 30-65 Pap test and HPV test	3 years (21-29) 5 years Pap+HPV (preferred) or 3 years (just Pap) (30-65)	2003: Recommends screening in “women who have been sexually active and have a cervix” (regardless of age)
VHA (2018)	21-29 Pap test alone 30-65 Pap test and HPV test	3 years (21-29) 5 years Pap+HPV (30-65))	NA

Abbreviations: ACOG = American College of Obstetrics and Gynecology, ACS = American Cancer Society, AMA = American Medical Association, CDC = Centers for Disease Control and Prevention, HPV = Human Papilloma Virus, NCCN = National Comprehensive Cancer Network, USPSTF = US Preventive Services Task Force, VHA = Veteran’s Health Administration

APPENDIX B. SEARCH STRATEGIES

1. Search for current systematic reviews **Date Searched: 11/5/2018**	
Sources:	**Search Strategy:**
AHRQ	Search: self report and screening; mammography; pap smear; breast cancer screening; cervical cancer screening
CADTH	Search: self report and screening; mammography; pap smear; breast cancer screening; cervical cancer screening
NICE (NHS Evidence)	Search: self report and screening; mammography; pap smear; breast cancer screening; cervical cancer screening
VA Products: VATAP, PBM, HSR&D publications, VA ART Database	A. http://www.hsrd.research.va.gov/research/default.cfm B. http://www.research.va.gov/research_topics/ C. http://art.puget-sound.med.va.gov/default.cfm Search: self report and screening; mammography; pap smear; breast cancer screening; cervical cancer screening
Cochrane Database of Systematic Reviews	Database: EBM Reviews - Cochrane Database of Systematic Reviews <2005 to October 31, 2018> Search Strategy: --- 1 (self report* or self-report*).mp. (1994) 2 (Mammograph* or Tomosynthes*).mp. (55) 3 breast cancer screen*.mp. (23) 4 (Pap smear or Pap test or Papanicolaou smear or Papanicolaou test).mp. (32) 5 cervical cancer screen*.mp. (26) 6 or/2-5 (96) 7 1 and 6 (34) 8 limit 7 to yr="2005 -Current" (29) **********

2. Systematic reviews currently under development (forthcoming reviews & protocols) **Date Searched: 11/5/2018**	
Sources:	**Search Strategy:**
PROSPERO (SR registry)	Search: self report and screening; mammography; pap smear; breast cancer screening; cervical cancer screening
DoPHER (SR Protocols)	Search: self report and screening; mammography; pap smear; breast cancer screening; cervical cancer screening

3. Current Guidelines **Date Searched: 11/5/2018**	
Sources:	**Search Strategy:**
VA/DoD Clinical Practice Guidelines	Search: N/A
Guideline Central	Search: self report and screening; mammography; pap smear; breast cancer screening; cervical cancer screening
American Cancer Society	Search: self report and screening; mammography; pap smear; breast cancer screening; cervical cancer screening

U.S. Preventive Services Task Force	Search: self report and screening; mammography; pap smear; breast cancer screening; cervical cancer screening
National Comprehensive Cancer Network	Search: self report and screening; mammography; pap smear; breast cancer screening; cervical cancer screening
American College of Obstetricians and Gynecologists	Search: self report and screening; mammography; pap smear; breast cancer screening; cervical cancer screening
International Agency for Research on Cancer	Search: self report and screening; mammography; pap smear; breast cancer screening; cervical cancer screening
American College of Radiology	Search: self report and screening; mammography; pap smear; breast cancer screening; cervical cancer screening
American College of Physicians	Search: self report and screening; mammography; pap smear; breast cancer screening; cervical cancer screening
American Academy of Family Physicians	Search: self report and screening; mammography; pap smear; breast cancer screening; cervical cancer screening

4. Current primary literature **Date Searched: 11/5/2018**	
Sources:	**Search Strategy:**
MEDLINE	Database: Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions(R) <1946 to November 02, 2018> Search Strategy: --- 1 exp Self Report/ (24544) 2 (self report* or self-report*).mp. (141351) 3 1 or 2 (141351) 4 exp MAMMOGRAPHY/ (28157) 5 exp Ultrasonography, Mammary/ (4808) 6 (Mammograph* or Tomosynthes*).mp. (37426) 7 breast cancer screen*.mp. (5546) 8 exp Papanicolaou Test/ (6197) 9 (Pap smear or Pap test or Papanicolaou smear or Papanicolaou test).mp. (9808) 10 cervical cancer screen*.mp. (5669) 11 or/4-10 (54106) 12 3 and 11 (805) 13 limit 12 to english language (797) 14 limit 13 to yr="2005 -Current" (548) **********
CINAHL	Database: EBSCOhost CINAHL Plus with Full Text Search Strategy: --- S1 (MH "Self Report+") (59908) S2 TX (self report* or self-report*) (148406) S3 S1 or S2 (149377) S4 (MH "Mammography") (10287) S5 TI ((Mammograph* or Tomosynthes*)) OR AB ((Mammograph* or Tomosynthes*)) (7916)

	S6 TI ((breast cancer screen*)) OR AB ((breast cancer screen*)) (2361) S7 (MH "Cervical Smears+") (5571) S8 TI ((Pap smear or Pap test or Papanicolaou smear or Papanicolaou test)) OR AB ((Pap smear or Pap test or Papanicolaou smear or Papanicolaou test)) (2819) S9 TI ((cervical cancer screen*)) OR AB ((cervical cancer screen*)) (2453) S10 S4 OR S5 OR S6 OR S7 OR S8 OR S9 (20357) S11 S3 AND S10 (1122) S12 Narrow by Language: - english (1122) S13 Limiters - Published Date: 20050101-20191231 (879) S14 Limiters – Source Types: Academic Journals (724) **********
PsycINFO	Database: PsycINFO <1806 to October Week 5 2018> Search Strategy: --- 1 exp Self Report/ (15600) 2 (self report* or self-report*).mp. (123082) 3 1 or 2 (123082) 4 exp MAMMOGRAPHY/ (1168) 5 exp Ultrasonography, Mammary/ (0) 6 (Mammograph* or Tomosynthes*).mp. (1882) 7 breast cancer screen*.mp. (982) 8 exp Papanicolaou Test/ (0) 9 (Pap smear or Pap test or Papanicolaou smear or Papanicolaou test).mp. (903) 10 cervical cancer screen*.mp. (844) 11 or/4-10 (3456) 12 3 and 11 (234) 13 limit 12 to english language (228) 14 limit 13 to yr="2005 -Current" (176) **********
CCRCT	Database: EBM Reviews - Cochrane Central Register of Controlled Trials <September 2018> Search Strategy: --- 1 exp Self Report/ (1766) 2 (self report* or self-report*).mp. (21717) 3 1 or 2 (21717) 4 exp MAMMOGRAPHY/ (749) 5 exp Ultrasonography, Mammary/ (70) 6 (Mammograph* or Tomosynthes*).mp. (1781) 7 breast cancer screen*.mp. (429) 8 exp Papanicolaou Test/ (219) 9 (Pap smear or Pap test or Papanicolaou smear or Papanicolaou test).mp. (663) 10 cervical cancer screen*.mp. (460) 11 or/4-10 (2803) 12 3 and 11 (106) 13 limit 12 to english language (88) 14 limit 13 to yr="2005 -Current" (66) **********

APPENDIX C. LIST OF EXCLUDED STUDIES

Exclude reasons: 1 = Ineligible population (*ie*, cancer patients or those with high risk of cancer), 2 = Ineligible intervention, 3 = Ineligible comparator, 4 = Ineligible outcome (*ie*, no diagnostic accuracy outcomes), 5 = Ineligible timing, 6 = Ineligible study design, 7 = Ineligible publication type, 8 = Outdated or ineligible systematic review

#	Citation	Exclude reason
1.	Barratt A, Cockburn J, Smith D, Redman S. Reliability and validity of women's recall of mammographic screening. *Australian and New Zealand Journal of Public Health.* 2000;24(1):79-81.	E4
2.	Bowman JA, Sanson-Fisher R, Redman S. The accuracy of self-reported pap smear utilisation. *Social Science & Medicine.* 1997;44(7):969-976.	E4
3.	Boyce GJ, Fruchter GR, Romanzi HL, Sillman HF, Maiman HM. The fallacy of the screening interval for cervical smears. *Obstetrics & Gynecology.* 1990;76(4):627-632.	E1
4.	Canfell K, Beral V, Green J, Cameron R, Baker K, Brown A. The agreement between self-reported cervical smear abnormalities and screening programme records. *Journal of Medical Screening.* 2006;13(2):72-75.	E1
5.	Craig BM, Quinn GP, Vadaparampil ST. Sensitivity of self-report mammography use in older women. *American Journal of Preventive Medicine.* 2009;37(5):441-444.	E4
6.	Cronin KA, Miglioretti DL, Krapcho M, et al. Bias associated with self-report of prior screening mammography. *Cancer Epidemiology, Biomarkers & Prevention.* 2009;18(6):1699-1705.	E4
7.	Dimatteo MR, Robinson JD, Heritage J, Tabbarah M, Fox SA. Correspondence among patients' self-reports, chart records, and audio/videotapes of medical visits. *Health Communication.* 2003;15(4):393-413.	E4
8.	Etzi S, Lane DS, Grimson R. The use of mammography vans by low-income women: The accuracy of self-reports. *American journal of public health.* 1994;84(1):107.	E4
9.	Fehringer G, Howlett R, Cotterchio M, Klar N, Majpruz-Moat V, Mai V. Comparison of papanicolaou (pap) test rates across ontario and factors associated with cervical screening. *Canadian Journal of Public Health Revue Canadienne de Sante Publique.* 2005;96(2):140-144.	E4
10.	Fiscella K, Holt K, Meldrum S, Franks P. Disparities in preventive procedures: Comparisons of self-report and medicare claims data. *BMC Health Services Research.* 2006;6:122.	E4
11.	Hampton T, Hampton T. Women have fewer mammograms than they report. *JAMA: Journal of the American Medical Association.* 2006;296(2):160-160.	E7
12.	Hoyo C, Ostbye T, Skinner CS, Yarnall KS, Chowdhary J. Reproducibility of self-reported pap test utilization in middle-aged african-american women. *Ethnicity & Disease.* 2005;15(1):84-89.	E4
13.	Larouche G, Bouchard K, Chiquette J, Desbiens C, Simard J, Dorval M. Self-reported mammography use following brca1/2 genetic testing may be overestimated. *Familial Cancer.* 2012;11(1):27-32.	E1
14.	Lofters AK, Moineddin R, Hwang SW, Glazier RH. Does social disadvantage affect the validity of self-report for cervical cancer screening? *International Journal of Women's Health.* 2013;5:29-33.	E7
15.	McKenna MT, Speers M, Mallin K, Warnecke R. Agreement between patient self-reports and medical records for pap smear histories. *American Journal of Preventive Medicine.* 1992;8(5):287-291.	E1

#	Citation	Exclude reason
16.	McPhee SJ, Nguyen TT, Shema SJ, et al. Validation of recall of breast and cervical cancer screening by women in an ethnically diverse population. *Preventive Medicine.* 2002;35(5):463-473.	E4
17.	Michielutte R, Dignan M, Bradley W, et al. Errors in reporting cervical screening among public health clinic patients. *J Clin Epidemiol.* 1991;44(4/5):403-408.	E4
18.	Mojica CM, Bastani R. Receipt of diagnostic tests for breast cancer: Validity of self-reports among low-income, mostly latina, indigent women. *Evaluation & the Health Professions.* 2010;33(4):437-451.	E1
19.	Montaño DE, Phillips WR. Cancer screening by primary care physicians: A comparison of rates obtained from physician self-report, patient survey, and chart audit. *American journal of public health.* 1995;85(6):795.	E4
20.	Nandy K, Menon U, Szalacha LA, Park H, Lee J, Lee EE. Self-report versus medical record for mammography screening among minority women. *Western Journal of Nursing Research.* 2016;38(12):1627-1638.	E4
21.	Pijpe A, Mulder RL, Manders P, Hebon, van Leeuwen FE, Rookus MA. Validation study suggested no differential misclassification of self-reported mammography history in brca1/2 mutation carriers. *Journal of Clinical Epidemiology.* 2011;64(12):1434-1443.	E1
22.	Powe BD, Cooper DL. Self-reported cancer screening rates versus medical record documentation: Incongruence, specificity, and sensitivity for african american women. *Oncology Nursing Forum.* 2008;35(2):199-204.	E4
23.	Rauscher GH, Johnson TP, Cho YI, Walk JA. Accuracy of self-reported cancer-screening histories: A meta-analysis. *Cancer Epidemiology, Biomarkers & Prevention.* 2008;17(4):748-757.	E7
24.	Rivera S, Vernon SW, Tiro JA, et al. Test-retest reliability of self-reported mammography in women veterans. *Preventive Medicine.* 2006;42(4):320-326.	E4
25.	Tiro JA, Sanders JM, Shay LA, et al. Validation of self-reported post-treatment mammography surveillance among breast cancer survivors by electronic medical record extraction method. *Breast Cancer Research & Treatment.* 2015;151(2):427-434.	E1
26.	Walker MJ, Chiarelli AM, Mirea L, et al. Accuracy of self-reported screening mammography use: Examining recall among female relatives from the ontario site of the breast cancer family registry. *Isrn Oncology Print.* 2013;2013:810573.	E4
27.	Warnecke RB, Sudman S, Johnson TP, O'Rourke D, Davis A, Jobe J. Cognitive aspects of recalling and reporting health-related events: Papanicolaou smears, clinical breast examinations, and mammograms. *Am J Epidemiol.* 1997;148(11).	E4
28.	Zapka JG, Bigelow C, Hurley T, et al. Mammography use among sociodemographically diverse women: The accuracy of selfl-report. *The American Journal of Public Health.* 1996;86(7):1016.	E4

APPENDIX D. EVIDENCE TABLES

Data Abstraction of Diagnostic Accuracy Studies

Author, Year (N)	Setting: Clinic or pop. based Part of referral program?	Patient characteristics: Age (yrs) Minority or low-income?	Survey method Timeframe of recall	Study design Pap, mam., both
Allgood, 2014[1] 1221	Population No	40+ Yes: low income, Black	Face-to-face interview 2 years	Cohort Mammogram
Armstrong, 2004[2] 399	Clinic No	50-75 Yes: low income	Phone interview Ever	Cohort Mammogram
Baron-Epel, 2008[3] 1550	Clinic No	52-74 Yes: Jewish, Arab	Phone interview 2 years	Cohort Mammogram
Bowman, 1991[4] 157	Population No	18-70 No	Face-to-face interview 3 years	Cohort Pap
Brown, 1992[5] 189	Clinic No	17-79 No	Phone interview 1 month	Cohort Mammogram
Caplan, 2003[6] 480	Clinic No	40-74 No	Phone interview Ever	Cohort Both
Caplan, 2003a[7] 949	Clinic Yes	50-80 No	Phone interview Ever	Cohort Mammogram
Champion, 1998[8] 229	Clinic Yes	45-64 Yes: low income, Black	Face-to-face interview 3 years	Cohort Mammogram
Clark, 2009[9] 411	Clinic Yes	40-75 Yes: black	Face-to-face survey 2 years	Cohort Mammogram
Crane, 1996[10] 576	Clinic Yes	50+ Yes: Hispanic and Black	Phone interview Last screen	Cohort Mammogram
Degnan, 1992[11] 456	Population No	50-74 No	Phone interview Last screen	Cohort Mammogram
Eaker, 2001[12] 944	Population No	25-60 No	Phone interview Ever	Case-control Pap
Fowles, 1997[13] 400	Clinic No	19-75 No	Phone interview 3 years	Cohort Pap
Fruchter, 1992[14] 263	Clinic No	NR Yes: Black and Latino	Face-to-face interview 3 years	Cohort Pap
Fulton-Kehoe, 1992[15] 78	Clinic Yes	50-75 No	Mail survey 1 year	Cohort Mammogram

Author, Year (N)	Setting: Clinic or pop. based Part of referral program?	Patient characteristics: Age (yrs) Minority or low-income?	Survey method Timeframe of recall	Study design Pap, mam., both
Giorgi Rossi, 2006[16] 641	Population Yes	25-64 No	Phone interview Ever	Cohort Pap
Gordon, 1993[17] 431	Clinic No	40-74 No	Mail survey/phone interview 2 years	Cohort Both
Hiatt, 1995[18] 691	Clinic No	35-75 Yes: Hispanic	Phone interview 5 years	Cohort Both
Holt, 2006[19] 5461	Population No	65+ No	Face-to-face interview 1 year	Cohort Mammogram
Johnson, 1995[20] 251	Population No	35-65 Yes: Native American	Face-to-face interview Ever	Cohort Both
Johnson, 2005[21] 588	Population No	50+ No	Phone/computer-assisted interview 3 years	Cohort Both
King, 1990[22] 199	Clinic Yes	50-74 No	Phone interview 1 year	Case-control Mammogram
Klungsoyr, 2009[23] 16574	Population No	18-45 No	Phone or paper or web 3 years	Cohort Pap
Lawrence, 1999[24] 93	Clinic No	50-70 Yes: Mexican-American	Phone interview 1 year	Cohort Mammogram
Lofters, 2015[25] 39027 (P) 15877 (M)	Population No	24-69 (P) 52-69 (M) No	Computer-assisted telephone and personal interviews Ever	Cohort Both
Mahnken, 2007[26] 199	Population No	50-74 Yes: Mexican-American	Face-to-face interview 2 years	Cohort Mammogram
Martin, 2000[27] 599	Clinic No	21+ No	Phone interview 2 years mam./3 years pap	Cohort Both
McGovern, 1998[28] 477	Clinic No	40-92 Yes: low income	Face-to-face interview Ever	Cohort Both
Newell, 2000[29] 423	Population No	18-81 No	Face-to-face interview Last screen	Cohort Pap
Norman, 2003[30] 2495	Both No	40-64 No	Phone interview Last screen	Case-control Mammogram
Paskett, 1996[31] 555	Population Yes	40+ Yes: low income	Face-to-face interview Ever	Cohort Both
Pizzaro, 2002[32] 161	Clinic No	18-89 Yes: low income	Phone interview 6 and 12 months	Cohort Pap

Author, Year (N)	Setting: Clinic or pop. based Part of referral program?	Patient characteristics: Age (yrs) Minority or low-income?	Survey method Timeframe of recall	Study design Pap, mam., both
Sawyer, 1989[33] 98	Population No	16-75 Yes: rural, Black	Face-to-face interview 3 years	Cohort Pap
Son, 2013[34] 155	Clinic No	18+ Yes: intellectual disability	Face-to-face interview Ever	Cohort Both
Suarez, 1995[35] 450	Population Yes	40+ Yes: Mexican-American	Face-to-face interview Ever	Cohort Both
Thompson, 1999[36] 360	Clinic No	50-69 Yes: low income	Mail survey Ever	Cohort Mammogram
Tsurda, 2018[37] 411294	Population Yes	50-69 No	Mail survey Ever	Cohort Mammogram
Tumiel-Berhalter, 2004[38] 314	Clinic No	40+ Yes: Black, Puerto Rican	Face-to-face interview 3 years	Cohort Both
Walter, 1988[39] 750	Clinic No	20-69 No	Face-to-face interview 5 years	Case-control Pap

Meta-analysis Statistics

Author, Year N	TP	FP	FN	TN	Sensitivity	Specificity	PPV	NPV	LR+	LR-	Rep/Rec Ratio
Pap Smear											
Bowman, 1991[4] 111	64	19	5	23	92.753	54.761	77.108	82.142	2.050	0.132	1.202
Caplan, 2003[6] 440	339	54	3	44	99.122	44.897	86.259	93.617	1.798	0.019	1.149
Eaker, 2001[12] 944	896	21	6	21	99.334	50.000	97.709	77.777	1.986	0.013	1.016
Fowles, 1997[13] 259	209	34	4	12	98.122	26.086	86.008	75.000	1.327	0.071	1.140
Fruchter, 1992[14] 138	109	19	6	4	94.782	17.391	85.156	40.000	1.147	0.300	1.113
Georgi Rossi, 2006[16] 641	325	46	58	212	84.856	82.170	87.601	78.518	4.759	0.184	0.968

Gordon, 1993[17] 352	239	69	7	37	97.154	34.905	77.597	84.090	1.492	0.081	1.252
Hiatt, 1995[18] 691	331	127	54	179	85.974	58.496	72.270	76.824	2.071	0.239	1.189
Johnson, 1995[20] 215	118	37	4	56	96.721	60.215	76.129	93.333	2.431	0.054	1.270
Johnson, 2005[21] 588	267	90	33	198	89.000	68.750	74.789	85.714	2.848	0.160	1.190
Klungsoyr, 2009[23] 15474	12613	1126	380	1355	97.075	54.615	91.804	78.097	2.138	0.053	1.057
Lofters, 2015[25] 39027	26892	5627	983	5525	96.473	49.542	82.696	84.895	1.911	0.071	1.166
Martin, 2000[27] 175	150	12	2	11	98.684	47.826	92.592	84.615	1.891	0.027	1.065
McGovern, 1998[28] 281	91	48	20	122	81.981	71.764	65.467	85.915	2.903	0.251	1.252
Newell, 2000[29] 146	84	33	2	27	97.674	45.000	71.794	93.103	1.775	0.051	1.360
Paskett, 1996[31] 438	268	137	8	25	97.101	15.432	66.172	75.757	1.148	0.187	1.467
Pizarro, 2002[32] 174	63	41	14	56	81.818	57.731	60.576	80.000	1.935	0.314	1.350
Sawyer, 1989[33] 98	63	17	3	15	95.454	46.875	78.750	83.333	1.796	0.096	1.212
Son, 2013[34] 149	90	32	5	22	94.736	40.740	73.770	81.481	1.598	0.129	1.284
Suarez, 1995[35] 215	110	54	6	45	94.827	45.454	67.073	88.235	1.738	0.113	1.413
Tumiel-Berhalter, 2004[38] 251	225	16	7	3	96.982	15.789	93.361	30.000	1.151	0.191	1.038
Walter, 1988[39] 576	317	80	31	148	90.600	63.000	79.848	82.681	2.450	0.150	1.140
Mammogram											
Allgood, 2014[1] 1204	510	369	29	296	94.619	44.511	58.020	91.076	1.705	0.120	1.630
Armstrong, 2004[2] 399	199	84	16	100	92.558	54.347	70.318	86.206	2.027	0.136	1.316

Brown, 1992[5] 189	65	18	6	100	91.549	84.745	78.313	94.339	6.001	0.099	1.169
Caplan, 2003[6] 433	391	26	4	12	98.987	31.578	93.764	75.000	1.446	0.032	1.055
Caplan, 2003a[7] 949	645	121	43	140	93.750	53.639	84.203	76.502	2.022	0.116	1.113
Champion, 1998[8] 268	168	67	6	27	96.551	28.723	71.489	81.818	1.354	0.120	1.350
Clark, 2009[9] 411	233	81	6	91	97.489	52.906	74.203	93.814	2.070	0.047	1.313
Crane, 1996[10] 403	209	46	58	90	78.277	66.176	81.960	60.810	2.314	0.328	0.955
Degnan, 1992[11] 456	245	41	1	169	99.593	80.476	85.664	99.411	5.101	0.005	1.162
Fulton-Kehoe, 1992[15] 78	59	8	1	10	98.333	55.555	88.059	90.909	2.212	0.030	1.116
Gordon,1993[17] 386	243	58	5	80	97.983	57.971	80.730	94.117	2.331	0.034	1.213
Hiatt, 1995[18] 687	276	153	45	213	85.981	58.196	64.335	82.558	2.056	0.240	1.336
Holt, 2006[19] 5461	2010	640	212	2599	90.459	80.240	75.849	92.458	4.578	0.118	1.192
Johnson, 1995[20] 124	92	22	3	7	96.842	24.137	80.701	70.000	1.276	0.130	1.200
Johnson, 2005[21] 587	425	56	15	91	96.590	61.904	88.357	85.849	2.535	0.055	1.093
King, 1990[22] 199	94	6	0	99	100.000	94.285	94.000	100.000	17.500	0	1.063
Lawrence, 1999[24] 232	122	34	18	58	87.142	63.043	78.205	76.315	2.357	0.203	1.114
Lofters, 2015[25] 15877	9650	2103	337	3787	96.625	64.295	82.106	91.828	2.706	0.052	1.176
Mahnken, 2007[26] 199	105	37	8	49	92.920	56.976	73.943	85.964	2.159	0.124	1.256
McGovern, 1998[28] 456	168	64	21	203	88.888	76.029	72.413	90.625	3.708	0.146	1.227

Norman, 2003[30] 747	385	78	13	271	96.733	77.650	83.153	95.422	4.328	0.042	1.163
Paskett, 1996[31] 441	212	88	12	129	94.642	59.447	70.666	91.489	2.333	0.090	1.339
Son, 2013[34] 82	68	7	4	3	94.444	30.000	90.666	42.857	1.349	0.185	1.041
Suarez, 1995[35] 215	67	20	4	124	94.366	86.111	77.011	96.875	6.794	0.065	1.225
Thompson, 1999[36] 360	207	56	9	88	95.833	61.111	78.707	90.721	2.464	0.068	1.217
Tsurda, 2018[37] 411,294	354008	8879	412	47995	99.883	84.388	97.553	99.148	6.398	0.001	1.023
Tumiel-Berhalter, 2004[38] 180	102	62	1	15	99.029	19.480	62.195	93.750	1.229	0.049	1.592

Quality Assessment of Included Systematic Reviews

Author Year	Study Eligibility Criteria	Identification and Selection of Studies	Data Collection and Study Appraisal	Synthesis and Findings	Overall Risk of Bias
Howard, 2009[40]	Low Predefined and reasonable criteria with appropriate restrictions	Low Searched multiple databases with reasonable search terms, searched bibliographies of included studies, 2 authors reviewed studies	Low Dual data abstraction and quality assessment using predefined criteria	Low Dual data abstraction and quality assessment using predefined criteria	Low

Quality Assessment of Diagnostic Accuracy Studies

Author Year	Could the selection of patients have introduced bias?	Could the conduct or interpretation of the index test have introduced bias?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Could the patient flow have introduced bias?	Overall Quality (good, unclear, poor)
Allgood, 2014[1]	Unclear Convenience sample at several venues	No Survey completed prior to medical record review. Validated all reports.	No Abstractors blinded to self-report. Dual abstraction.	Unclear Excluded 18% of eligible sample with documentation at a different facility.	Unclear
Armstrong, 2004[2]	No Consecutive sampling at a clinic	Unclear Unclear if self-report collection blinded to medical record	Unclear Unclear if data collectors blinded to self-report. Unclear if multiple observers.	No All interviewed women had medical record review.	Unclear
Baron-Epel, 2008[3]	No Random telephone sampling	Unclear Unclear if data collectors blinded to claims data	Unclear Unclear if data collectors blinded to self-report data. Unclear if multiple observers.	No All participants medical records searched. Included all participants responding to survey. 1% had mammograms privately outside of claims records.	Unclear
Bowman, 1991[4]	No Random household survey	No	Unclear	Unclear	Unclear

Author Year	Could the selection of patients have introduced bias?	Could the conduct or interpretation of the index test have introduced bias?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Could the patient flow have introduced bias?	Overall Quality (good, unclear, poor)
		Interviews conducted prior to accessing medical records	Unclear if medical record data collection blinded to self-report. Unclear if multiple observers.	29.3% of interviewed did not have medical records checked.	
Brown, 1992[5]	Unclear Medical record audit of new HMO members	Unclear Unclear if self-report collection blinded to medical record	No Medical audit conducted prior to interviews	No 88% of audit sample had interviews completed.	Unclear
Caplan, 2003[6]	No Random telephone interview (BRFSS)	No Interviews conducted prior to medical record review	Unclear Unclear if medical record data collection blinded to self-report. Unclear if multiple observers.	No 93% of survey sample had medical records reviewed.	Unclear
Caplan, 2003a[7]	Unclear Women enrolled in a breast cancer screening program (85% of invited women enroll)	No Interviews conducted prior to linking with medical records	Unclear Unclear if medical record data collection blinded to self-report. Unclear if multiple observers.	No 8% of survey sample excluded with no medical record (not at group health for past 5 years).	Unclear
Champion 1998[8]	Unclear Women enrolled in a breast cancer screening intervention. Unclear why excluded patients reporting mammogram > 3 years ago.	No Interviews conducted prior to searching medical records	Unclear Unclear if medical record data collection blinded to self-report. Unclear if multiple observers.	No 15% “not able to verify” which included closed sites, sites unable to be contacted, and sites unable to verify record.	Unclear
Clark, 2009[9]	No Women at risk for fragmented care	Unclear Unclear if data collectors blinded to medical record	Unclear Unclear if data collectors blinded to self-report data and unclear if multiple observers	No Excluded 6% of eligible sample for no records or not having regular providers at primary care sites.	Unclear
Crane, 1996[10]	No Random sample of screening referrals	Unclear Unclear if self-report collection blinded to medical record	Unclear Unclear if medical record data collection blinded to self-report. Unclear if multiple observers.	No Incomplete information for 4% of survey respondents.	Unclear

Author Year	Could the selection of patients have introduced bias?	Could the conduct or interpretation of the index test have introduced bias?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Could the patient flow have introduced bias?	Overall Quality (good, unclear, poor)
Degnan, 1992[11]	No Random telephone survey	No Survey done prior to collection of medical record data	Unclear Unclear if medical record data collection blinded to self-report. Unclear if multiple observers.	No Included 94% of initial survey sample.	Unclear
Eaker, 2001[12]	Unclear Random samples, case control design	Unclear Unclear if self-report collection blinded to medical record	Unclear Unclear if medical record data collection blinded to self-report. Unclear if multiple observers.	Unclear 15% of non-attendees excluded without medical record or who had a pap in between notification and interview.	Poor
Fowles, 1997[13]	No Random selection from claims data file	Unclear Unclear if self-report collection blinded to medical record	Unclear Unclear if medical record data collection blinded to self-report. Dual medical record abstraction.	No 88% of sample had complete data.	Unclear
Fruchter, 1992[14]	Unclear Women attending various clinics, unclear what type of clinic or how sample was taken	No Interviews conducted prior to searching medical records	Unclear Unclear if medical record data collection blinded to self-report. Two screeners, but unclear if first data was checked.	Yes 53% of surveyed did not have medical records searched. Unclear if medical records of those reporting no pap smear were searched.	Poor
Fulton-Kehoe, 1992[15]	No Random subset of larger breast cancer screening program	Unclear Postal survey. No info on blinding of data collectors.	Unclear Unclear if medical record data collection blinded to self-report. No info on audit methods.	No 92% of records searched.	Unclear
Giorgi Rossi, 2006[16]	Unclear Unclear how the sample was created from list of female residents	Unclear Unclear if data collectors blinded to medical record	No Medical record reviewed prior to survey	Unclear Surveyed 53.2-84.2% of those with medical record review.	Unclear
Gordon, 1993[17]	No	Unclear	Unclear	No	Unclear

Author Year	Could the selection of patients have introduced bias?	Could the conduct or interpretation of the index test have introduced bias?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Could the patient flow have introduced bias?	Overall Quality (good, unclear, poor)
	Random sample of health plan members	Unclear if self-report data collection blinded to medical record	Unclear if medical record data collection blinded to self-report. 2 reviewers. Unclear if dual review.	82-90% of respondent's medical records reviewed.	
Hiatt, 1995[18]	No Random telephone survey	No Survey conducted prior to medical record review	Unclear Unclear if medical record data collection blinded to self-report. Single reviewer.	No 95-97% of respondent's medical records reviewed.	Unclear
Holt, 2006[19]	No Medicare Benefits Survey sample during specific time period	No Medicare Current Beneficiary Survey done prior to accessing claims data	Unclear Unclear if data collectors blinded to self-report data. Unclear if multiple observers.	No Excluded 11% for absence of corresponding claims.	Unclear
Johnson, 2005[21]	No Random telephone survey	No Surveys conducted prior to medical record abstraction	No Medical record data abstracted by the individual medical facilities. Unclear if multiple observers.	Unclear 58% of interviews had medical records extracted.	Unclear
Johnson, 1995[20]	No Random sample from tribe	No Interview conducted prior to medical record review	Unclear Unclear if medical record data collection blinded to self-report. Second reviewer check for every 10th record.	No 86% of interviewed had medical records checked.	Unclear
King, 1990[22]	Unclear Case-control out of random survey	No Surveys conducted prior to medical record review	Unclear Unclear if data collectors blinded to self-report. Unclear number of observers.	Yes Unclear if medical records checked for those reporting no mammogram	Poor
Klungsoyr, 2009[23]	No Random survey sample	No	No	No All national cytology records included. <1% of survey	Good

Author Year	Could the selection of patients have introduced bias?	Could the conduct or interpretation of the index test have introduced bias?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Could the patient flow have introduced bias?	Overall Quality (good, unclear, poor)
		Surveys conducted separately from cytology register	Cytology register independently collects data. Unclear if multiple observers.	respondents did not respond to pap question.	
Lawrence, 1999[24]	No Random sample of women in health clinics/hospital	No Interviewers blinded to screening status	No Medical record data collection blinded to self-report	No 2-6% eligible participants unable to verify screening mammogram done outside study system.	Good
Lofters, 2015[25]	No National survey sample	No Survey conducted separately from claims database	No Claims database with 95% of physician claims. Unclear if multiple observers.	No Appears all with survey had medical record linkage, report missing data as separate variable.	Good
Mahnken, 2007[26]	No Block random sampling of Mexican-American population	No Survey conducted prior to medical record review	Unclear Unclear if note to medical offices indicated response of patient. Unclear if multiple observers.	Unclear 44% of interview respondents did not have medical records checked.	Unclear
Martin, 2000[27]	No Random telephone survey (BRFSS)	No Survey conducted prior to medical record abstraction	No Medical record data collection blinded to self-report	No 99.5% of interviewed had chart audit completed	Good
McGovern, 1998[28]	Unclear Interviewed women in waiting rooms. Unclear how sample was determined.	No Interviews conducted prior to searching medical records.	Unclear Unclear if medical record data collection blinded to self-report. Unclear number of observers.	No 4-5% unable to check medical records.	Unclear
Newell, 2000[29]	No Random household survey	No Interviews conducted prior to collecting medical records	Unclear Unclear if medical record data collection blinded to self-report. Unclear number of observers.	Unclear Unclear how many were not eligible for pap screening adequacy and reasons for ineligibility	Unclear

Author Year	Could the selection of patients have introduced bias?	Could the conduct or interpretation of the index test have introduced bias?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Could the patient flow have introduced bias?	Overall Quality (good, unclear, poor)
Norman, 2003[30]	No Controls random digit dialing	No Interviews conducted prior to collecting medical records	Unclear Unclear if medical record data collection blinded to self-report. Unclear number of observers.	Yes 9% of controls did not have medical record information. Did not validate negative reports.	Poor
Paskett, 1996[31]	No Random sample of target population	No Interviews conducted prior to collecting medical records.	Unclear Unclear if medical record data collection blinded to self-report. Unclear number of observers.	Unclear 79-80% of survey respondents had data for verification	Unclear
Pizzaro, 2002[32]	Unclear Random sample of women at a clinic. Selected patients only who had medical record available. Unclear how many excluded.	Unclear Unclear if self-report blinded to medical record	Unclear Unclear if medical record data collection blinded to self-report. Two reviewers, unclear if dual review.	Unclear Discord between number of included patients between text and tables.	Poor
Sawyer, 1989[33]	Yes Door-to-door canvasing, social worker contacts, and friends and neighbors	No Interview conducted prior to contacting providers	Unclear Unclear if medical record data collectors blinded to self-report. Single observer.	No 11% interviewed did not have medical record checked.	Poor
Son, 2013[34]	Unclear Sampling from various partner sites	No Survey conducted prior to medical record review	No Medical practices asked to fill out dates for any receipt of pap or mammogram during study period. Unclear if multiple observers.	Unclear Unclear if any excluded for not having medical records checked.	Unclear
Suarez, 1995[35]	No Random household survey	No Interviews conducted prior to collection of medical records	Unclear Unclear if medical record data collection blinded to self-report. Unclear number of observers.	Unclear 32% of sample did not have medical records checked.	Unclear
Thompson, 1999[36]	Unclear	Unclear	Unclear	No	Unclear

Author Year	Could the selection of patients have introduced bias?	Could the conduct or interpretation of the index test have introduced bias?	Could the reference standard, its conduct, or its interpretation have introduced bias?	Could the patient flow have introduced bias?	Overall Quality (good, unclear, poor)
	Unclear how sample was determined	Unclear if self-report data collection blinded to medical record	Unclear if medical record data collection blinded to self-report. Unclear number of observers.	3% missing data	
Tsurda, 2018[37]	Unclear Only 67% of the potentially eligible population filled out the self-report questionnaire (women coming to screening program).	No All data collected and linked by Cancer Registry unit.	No All data collected and linked by Cancer Registry unit. Unclear if multiple observers.	No All participant records matched with the Cancer Registry databases.	Unclear
Tumiel-Berhalter, 2004[38]	No Consecutive women presenting for care at clinics	Unclear Unclear if self-report data collection blinded to medical record	Unclear Unclear if medical record data collection blinded to self-report. Unclear number of observers.	No 97% interviewed had medical record data	Unclear
Walter, 1988[39]	Unclear Case-control with cases consecutive and controls randomly sampled	No Patient interviews conducted prior to physician interviews	No Physicians blinded to patient response. Unclear if multiple observers.	Unclear More physicians contacted in cases compared to controls.	Unclear

APPENDIX E. META-ANALYSIS

Meta-regression

Cervical Cancer Screening

Sensitivity and Specificity

Parameter	category	nstudies	Sensitivity	p1	Specificity	p2
Year		22	0.96 [0.94 - 0.97]	0.97	0.49 [0.41 - 0.57]	0.95
Total_N		22	0.96 [0.93 - 0.97]	1.00	0.48 [0.41 - 0.56]	1.00
Design	Yes	20	0.95 [0.93 - 0.97]	0.03	0.48 [0.40 - 0.56]	0.57
	No	2	0.97 [0.94 - 1.00]	.	0.57 [0.31 - 0.82]	.
Setting	Yes	11	0.96 [0.94 - 0.98]	0.00	0.52 [0.41 - 0.63]	0.44
	No	11	0.95 [0.92 - 0.98]	.	0.45 [0.33 - 0.56]	.
Program	Yes	3	0.94 [0.88 - 1.00]	0.01	0.48 [0.27 - 0.68]	0.99
	No	19	0.96 [0.94 - 0.98]	.	0.49 [0.40 - 0.57]	.
Admin	Yes	11	0.95 [0.92 - 0.98]	0.00	0.44 [0.33 - 0.55]	0.57
	No	11	0.96 [0.94 - 0.98]	.	0.53 [0.42 - 0.63]	.
Recall		22	0.96 [0.94 - 0.97]	0.94	0.49 [0.41 - 0.57]	0.89
Min_Low	Yes	10	0.94 [0.90 - 0.97]	0.00	0.43 [0.32 - 0.55]	0.59
	No	12	0.97 [0.95 - 0.98]	.	0.52 [0.42 - 0.62]	.

Joint Model

Parameter	category	LRTChi2	Pvalue	I2	I2lo	I2hi
Year		1.52	0.47	0	0	100
Total_N		0.67	0.71	0	0	100
Design	Yes	2.88	0.24	31	0	100
	No	.	.	.	.	.
Setting	Yes	3.15	0.21	37	0	100
	No	.	.	.	.	.
Program	Yes	0.71	0.70	0	0	100
	No	.	.	.	.	.
Admin	Yes	3.52	0.17	43	0	100
	No	.	.	.	.	.
Recall		1.59	0.45	0	0	100
Min_Low	Yes	12.81	0.00	84	67	100
	No	.	.	.	.	.

Breast Cancer Screening

Sensitivity and Specificity

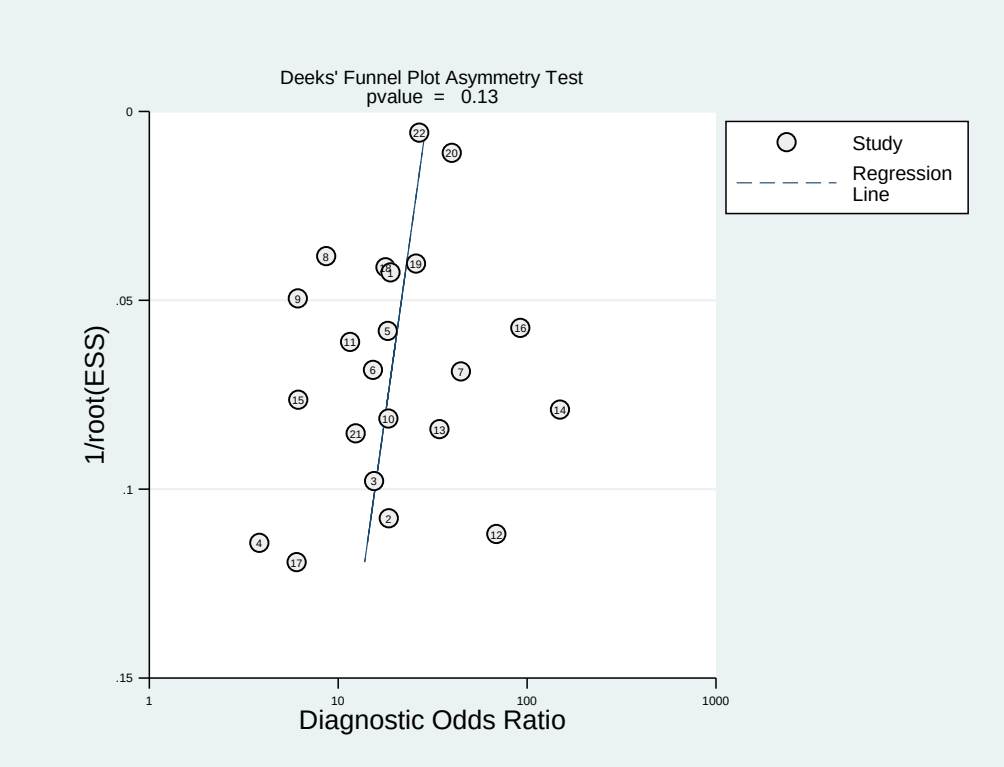
Parameter	category	nstudies	Sensitivity	p1	Specificity	p2
Year		29	0.96 [0.95 - 0.98]	0.95	0.61 [0.53 - 0.68]	0.91
Total_N		29	0.96 [0.95 - 0.97]	1.00	0.61 [0.54 - 0.69]	1.00
Design	Yes	28	0.96 [0.95 - 0.98]	0.86	0.61 [0.53 - 0.69]	0.33
	No	1	0.97 [0.90 - 1.00]	.	0.78 [0.48 - 1.00]	.
Setting	Yes	10	0.97 [0.95 - 0.99]	0.00	0.67 [0.55 - 0.79]	0.78
	No	19	0.96 [0.94 - 0.98]	.	0.58 [0.49 - 0.68]	.
Program	Yes	9	0.97 [0.95 - 0.99]	0.00	0.68 [0.56 - 0.81]	0.86
	No	20	0.96 [0.94 - 0.98]	.	0.58 [0.49 - 0.68]	.
Admin	Yes	11	0.95 [0.92 - 0.98]	0.00	0.53 [0.40 - 0.66]	0.06
	No	18	0.97 [0.95 - 0.98]	.	0.66 [0.57 - 0.75]	.
Recall		29	0.96 [0.95 - 0.98]	0.95	0.61 [0.52 - 0.68]	0.89
Min_Low	Yes	16	0.94 [0.91 - 0.97]	0.00	0.53 [0.43 - 0.64]	0.01
	No	13	0.98 [0.97 - 0.99]	.	0.70 [0.61 - 0.80]	.

Joint Model

Parameter	category	LRTChi2	Pvalue	I2	I2lo	I2hi
Year		2.17	0.34	8	0	100
Total_N		13.72	0.00	85	70	100
Design	Yes	0.83	0.66	0	0	100
	No	.	.	.	.	.
Setting	Yes	1.81	0.40	0	0	100
	No	.	.	.	.	.
Program	Yes	2.21	0.33	9	0	100
	No	.	.	.	.	.
Admin	Yes	3.37	0.19	41	0	100
	No	.	.	.	.	.
Recall		0.80	0.67	0	0	100
Min_Low	Yes	14.14	0.00	86	71	100
	No	.	.	.	.	.

Publication Bias

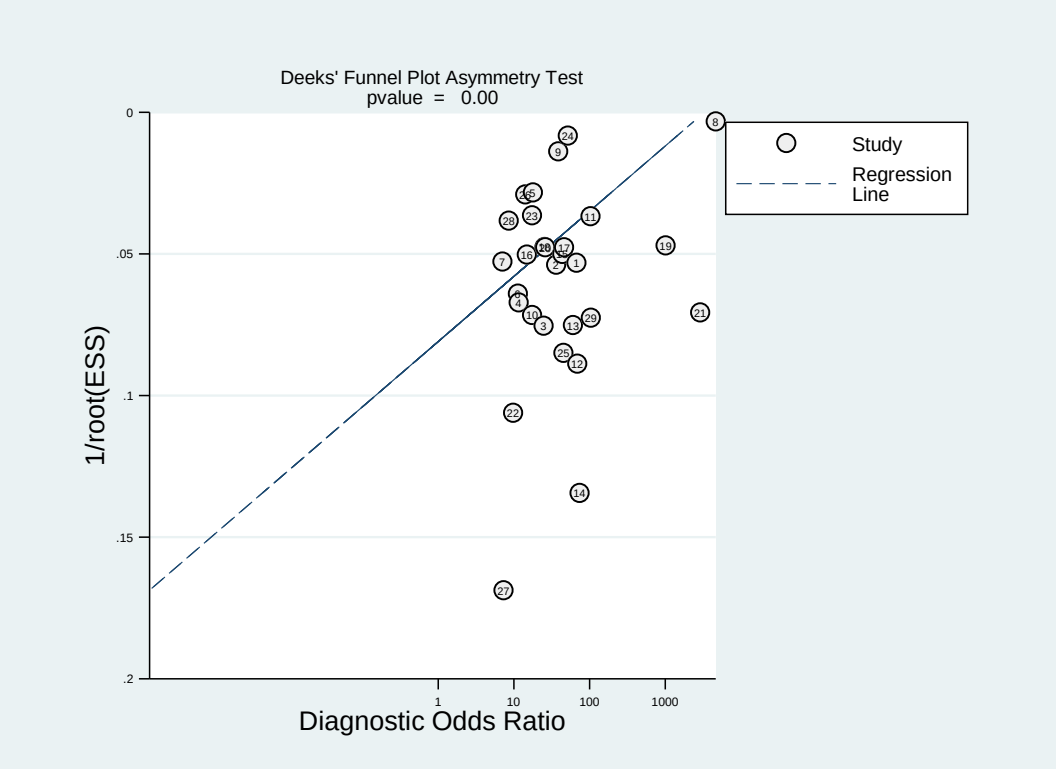
Cervical Cancer Screening



STATISTICAL TESTS FOR SMALL STUDY EFFECTS/PUBLICATION BIAS

yb	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Bias	-6.461954	4.102605	-1.58	0.131	-15.01984	2.095931
Intercept	3.396988	.0901981	37.66	0.000	3.208838	3.585138

Breast Cancer Screening



STATISTICAL TESTS FOR SMALL STUDY EFFECTS/PUBLICATION BIAS

yb	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
Bias	-100.0387	23.45697	-4.26	0.000	-148.1684	-51.90893
Intercept	8.097446	.3527632	22.95	0.000	7.373636	8.821256

APPENDIX F. DEFINITIONS

2x2 Table		Evidence of screening in medical record	
		Yes	No
Self-report screening	Yes	True Positive (TP)	False Positive (FP)
	No	False Negative (FN)	True Negative (TN)

- **Sensitivity:** probability that a person will report screening when there is evidence of screening in the medical record (true positive rate).
  = TP / (TP+FN)
  - Adjusted sensitivity = modified via bivariate random-effects meta-analysis
- **Specificity:** probability that a person will not report screening when there is no evidence of screening in the medical record (true negative rate).
  = FN / (FN+FP)
  - Adjusted sensitivity = modified via bivariate random-effects meta-analysis
- **Positive predictive value (PPV):** probability that there is evidence of screening in the medical record when a patient self-reports that screening occurred. (Excluded case-control studies)
  = TP / (TP+FP)
  - Adjusted PPV = $\frac{sensitivity \times prevalence}{sensitivity \times prevalence + (1-specificty) \times (1-prevalence)}$
- **Negative predictive value (NPV):** probability that there is no evidence of screening in the medical record when a patient self-reports that screening has not occurred. (Excluded case-control studies)
  = TN / (TN+FN)
  - Adjusted NPV = $\frac{specificity \times (1-prevalence)}{(1-sensitivity) \times prevalence + specificity \times (1-prevalence)}$
- **Concordance (Accuracy):** overall probability that a patient will correctly self-report receiving or not receiving screening.
  = (TP+TN) / (TP+FP+TN+FN)
  - Pooled concordance = average prevalence between studies
- **Report-to-Records Ratio:** ratio between the number of people self-reporting screening and the number of people with evidence of screening in their medical records.
  = (TP+FP) : (TP+FN)
  - Pooled Report-to-Records Ratio = average Report-to-Records Ratio between studies

- **Prevalence of screening:** the proportion of people who have evidence of screening as verified by medical records.
  = (TP+FN) / (TP+FP+TN+FN)
    - Pooled Prevalence of screening = overall numbers for (TP+FN) / (TP+FP+TN+FN)

APPENDIX G. PEER REVIEW COMMENT TABLE

Comment #	Reviewer #	Comment	Author Response
Are the objectives, scope, and methods for this review clearly described?			
1	1	Yes	*None*
2	2	Yes	*None*
3	3	Yes	*None*
4	4	Yes	*None*
5	5	Yes	*None*
Is there any indication of bias in our synthesis of the evidence?			
6	1	No	*None*
7	2	No	*None*
8	3	No	*None*
9	4	No	*None*
10	5	No	*None*
Are there any published or unpublished studies that we may have overlooked?			
11	1	No	*None*
12	2	No	*None*
13	3	No	*None*
14	4	No	*None*
15	5	No	*None*
Additional suggestions or comments can be provided below. If applicable, please indicate the page and line numbers from the draft report.			
16	1	The report is well written. Statistical methods are clear and limitations, future research, and applicability are described effectively.	*None*
17	2	This Evidence Brief is conducted based on standard systematic review methods. The background is brief and clear, and provides the context for the review. The scope is clearly defined. My comments mainly focused on the statistical methods and interpretation of results.	*None*
18	2	1) Methods: It is appropriate to use a bivariate random-effects model to estimate pooled effects and this method is often used to	*We have added details on our methods to obtain pooled PPV, NPV, concordance, and Rep/Rec ratio to*

		combine sensitivity and specificity. Please provide the methods used to obtain pooled PPV, NPV, concordance and Rep/Rec ratio.	*the methods section of the report and added definition and formulas to the supplemental materials.* *"Positive and negative predictive values were calculated using the MIDAS command (generates summary predictive values) and adjusted for overall prevalence. Concordance and Rep/Rec ratio were averaged across studies. Dentitions and formulas for these calculations is provided in the supplemental materials."*
19		Estimates of PPV and NPV need information of the prevalence of disease (here screening) in the population. When the study populations of the included studies are representative samples of the general population, it would be fine to estimate PPV and NPV using the study sample; otherwise, PPV and NPV could not be estimated based on the study sample only.	*We agree that the prevalence should be given as context for interpreting the PPV and NPV. The prevalence of screening in the study population, as verified by medical records, was 74.13% for cervical cancer and 84.11% for breast cancer. For prevalence of screening via self-report, the numbers were 84.07% for cervical cancer and 86.83% for breast cancer. The above numbers are comparable to self-reported prevalence's in the VA and non-VA settings: 93% and 82% for cervical cancer screening; and 85% and 65% for breast cancer screening. As these numbers are comparable, we felt it was adequate to estimate the PPV and NPV using the study sample.* *We have added this information to the results and discussion sections of the report.*
20	2	2) Results: The main results are based on PPV and NPV; the interpretation could be more clear by more explicitly stating the denominator of each measure. For example, PPV is the proportion of women who had a pap test in the medical record among those who self-reported had a pap test. Similarly, NPV is the proportion of women who did not had a pap test in the medical record among those who did NOT self-reported had a pap test.	*We have added a clearer, plain-language explanation for PPV and NPV to the report:* *"Overall, 84% (95% CI 82 to 86) of self-reports of screening were verified in the medical record (PPV), while 83% (95% CI 82 to 84) of self-reports of no screening were verified in the medical record (NPV)."- page 15* *"Overall, 80% (95% CI 79 to 81) of self-reports of screening were verified in the medical record (PPV), while 86% (95% CI 85 to 87) of self-reports of no*

			screening were verified in the medical record (NPV).”- page 18
21	2	2) Results: (cont.) At the same time, provide some interpretation for sensitivity and specificity, since the proportion of false positive is still very high based on specificity, and over 50% of women reported they had a pap test while there is no medical record to support it.	*We have added a clearer, plain-language explanation for sensitivity and specificity to the report in on page 13 and 15.* *“Among unscreened women, according to their medical record, 48% (95% CI 41 to 56) accurately reported no screening (ie, specificity). This indicates that over 50% of women without screenings in their medical record inaccurately reported having a screening. Among screened women, according to their medical record, 96% (95% CI 94 to 97) accurately reported this screening (ie, sensitivity).” -page 15*
22	2	3) Are there measures to quantify heterogeneity among studies for NPV and PPV? There is no forest plot, test of heterogeneity/I2 for PPV and NPV so they might be calculated based on summary measures of sensitivity and specificity?	*As we clarified in comment #18, since we did not do a meta-analysis for the measures of NPV and PPV, we did not test for heterogeneity, as functions to test for heterogeneity are lacking from the MIDAS (Meta-analytical Integration of Diagnostic Accuracy Studies) command in STATA.*
23	2	3) (cont.) However, it is misleading to interpret accuracy of self-reported results based on PPV and NPV in KQ 1 while investigating how the study characteristics impact accuracy of self-reported results based on sensitivity and specificity in KQ 3. That is, while KQ 1 and KQ3 used the same term when interpreting results (accuracy of self-reported results), they are based on different measures (PPV/NPV vs. Sensitivity/specificity). While the values of sensitivity were similar across studies, the values of PPV and NPV are not. Most included studies are cohort studies in this report – if the study samples are representative of the underlying study populations, meta-regression may be conducted directly on PPV and NPV.	*We agree and have reorganized the results to consistently emphasize the results of our meta-analysis on sensitivity and specificity, as these are standard characteristics of a test and do not vary based on the prevalence.*
24	2	4) Supplemental materials, table “Meta-analysis Statistics”: It is not appropriate to calculate PPV and NPV for case-control studies.	*Agreed. We have removed data from the case-control studies from our PPV and NPV calculations and updated our original numbers based on your recommendation to adjust for the prevalence of screening in the respective study populations. The*

			report and supplemental materials are updated with the new numbers.
25	2	5) What does testing for publication bias mean here? The symmetry of funnel plot and its associated tests are developed for comparative studies where small and negative studies may be less likely to be published. Here there is no intervention and it is hard to assume a favored direction beforehand. Not sure that the funnel plot is expected to be symmetric based on the diagnostic ratio ratio (a measure not discussed in the text)? Or Diagnostic odd ratio is a good measure for this purpose, that is, the authors will decide whether or not to publish a paper based on diagnostic ratio ratio.	*Publication bias here may suggest that small studies that did not find a significant difference between self-report and what was reported in the medical record may have been less likely to be published as those that found a significant difference. We agree that the assessment of the potential for publication bias is more complicated for diagnostic accuracy studies than for studies of interventions because it is difficult to assume there would be a favored direction of publication. Evidence for the existence of publication bias in this area of literature is scarce, but we felt it was reasonable to test for publication bias and report our results because, regardless of study design, one would expect some level of symmetry in the results (as discussed in this meta-analysis).[41] We have reworded our interpretation of the test to reflect the level of uncertainty surrounding publication bias in this field: "However, the assessment of the potential for publication bias is more complicated for diagnostic accuracy studies than for studies of interventions, because it is difficult to assume there would be a favored finding for publication. Empirical evidence for the existence of publication bias in this area of literature is scarce."* *We agree and used the diagnostic odds ratio to detect publication bias, because bivariate methods for the detection of publication bias are not currently available for diagnostic accuracy studies. We used Deeks' test, which is recommended by The Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy, because this test has been developed especially for test accuracy reviews and uses effective sample size, instead of total sample size, to create more accurate estimates. Additionally, a review on testing methods for detecting publication bias in meta-analyses of diagnostic accuracy studies also found Deeks' test to*

			be the most robust.[42] We have added details on these methods to our report.
26	2	5) (cont.) On Page 14, line 20, what does “with insignificant effects?” mean here? What effects are insignificant?	*By “insignificant effects”, we meant to emphasize that publication bias is often the result of studies with small effects being under-reported, compared to studies showing larger effects. In this case, that would mean studies that did not find a significant difference between self-report, and what was reported in the medical record may not have been published at the same rate as studies that found a significant difference. We have added wording to better explain this.*
27	2	6) Conclusions: P20, line 44: it is more accurate to say that “Women tend to over-report having had a mammogram or pap test, but MORE accurately report when they have not had a test.”	*Corrected*
28	2	Without the method to obtain pooled PPV and NPV, it is difficult to evaluate the appropriateness of the pooled estimates of these two measures.	*Methods for calculating NPV and PPV have been added to the methods section.*
29	3	Table 1, page 9. The headers of the columns and the data below are a little hard to follow. For example, the column labeled "Clinic or Pop. Based...Screening Program" looks like all one phrase when it actually refers to two different characteristics. Suggest maybe graphically finding a way to distinguish between the characteristics when they are grouped together in one column. It is hard to follow.	*We have created a new column to separate the information for “Clinic or Population-based Screening Program?” and “Was the study a part of a screening or referral program?” to alleviate confusion and simplify the reading of the table.*
30	3	p.7 I'd suggest defining "Clinic or population based" and "Screening Program" a bit more. They are briefly defined but I am still not sure what each means.	*We have added further descriptions at the bottom of Table 1.*
31	3	Outstanding work with a professional team that was easy to work with. This will help in our discussions about prevention policy and performance measurement.	*None*
32	4	Pg 17, lines 8-9. this is picky and it really does not matter one way or the other to me, but should "that was larger than every other study combined" be changed to "that was larger than all other studies combined"?	*Changed*

33	4	Pg 18, starting at line 14. Either there is a run-on sentence or it is difficult to read and understand. I think there should be a period after the word "test" and before the "(" on line 16. Overall, among 39 studies examining the accuracy of self-report for cervical and/or breast cancer screening compared to medical records, women tended to over-report having had a mammogram or pap test, but accurately reported when they had not had a test (61% to 66% (PPV) of women accurately reported having a pap test or mammogram, and 85% to 89% of women accurately reported NOT having a pap test or mammogram respectively).	*Corrected*
34	5	Excellent and impressive work. Appreciate this thorough and thoughtful review. My time was redirected so Dr. [ ] will provide more detail as a reviewer. The only correction is last credential for [ ]	*Corrected*
35	5	No further comments as [ ] will take lead on review as I was redirected to another time-sensitive deadline.	*None*

REFERENCES

1. Allgood KL, Rauscher GH, Whitman S, Vasquez-Jones G, Shah AM. Validating self-reported mammography use in vulnerable communities: findings and recommendations. *Cancer Epidemiology, Biomarkers & Prevention.* 2014;23(8):1649-1658.
2. Armstrong K, Long JA, Shea JA. Measuring adherence to mammography screening recommendations among low-income women. *Preventive Medicine.* 2004;38(6):754-760.
3. Baron-Epel O, Friedman N, Lernau O. Validity of self-reported mammography in a multicultural population in Israel. *Preventive Medicine.* 2008;46(6):489-491.
4. Bowman JA, Redman S, Dickinson JA, Gibberd R, Sanson-Fisher RW. The accuracy of Pap Smear utilization self-report: a methodological consideration in cervical screening research. *Health Services Research.* 1991;26(1):97.
5. Brown BJ, Adams EM. Patients as reliable reporters of medical care process: recall of ambulatory encounter events. *Medical Care.* 1992;30(5):400-411.
6. Caplan LS, McQueen DV, Qualters JR, Leff M, Garrett C, Calonge N. Validity of women's self-reports of cancer screening test utilization in a managed care population. *Cancer Epidemiology, Biomarkers & Prevention.* 2003;12(11 Pt 1):1182.
7. Caplan LS, Mandelson MT, Anderson LA. Validity of self-reported mammography: examining recall and covariates among older women in a health maintenance organization. *American Journal of Epidemiology.* 2003 157(3):267-272.
8. Champion VL, Menon U, McQuillen DH, Scott C. Validity of self-reported mammography in low-income African-American women. *American Journal of Preventive Medicine.* 1998;14(2):111-117.
9. Clark CR, Baril N, Kunicki M, et al. Mammography use among Black women: the role of electronic medical records. *Journal of Women's Health.* 2009;18(8):1153-1162.
10. Crane LA, Kaplan CP, Bastani R, Scrimshaw SCM. Determinants of adherence among health department patients referred for a mammogram. *Women's Health.* 1996;24(2):43-64.
11. Degnan D, Harris R, Ranney J, Quade D, Earp JA, Gonzalez J. Measuring the use of mammography: two methods compared. *American Journal of Public Health.* 1992;82(10):1386.
12. Eaker S, Adami H-O, Sparén P. Reasons women do not attend screening for cervical cancer: a population-based study in Sweden. *Preventive Medicine.* 2001;32(6):482-491.
13. Fowles JB, Fowler E, Craft C, McCoy CE. Comparing claims data and self-reported data with the medical record for pap smear rates. *Evaluation & The Health Professions.* 1997;20(3):324-342.
14. Fruchter RG, Rones K, Roth T, Webber CA, Camilien L, Boyce JG. Pap smear histories in a medical clinic: accuracy of patients' self-reports. *New York state journal of medicine.* 1992;92(10):421.
15. Fulton-Kehoe D, Burg MA, Lane DS. Are self-reported dates of mammograms accurate? *Public Health Reviews.* 1992;20(3-4):233.
16. Giorgi Rossi P, Esposito G, Brezzi S, Brachini A, Raggi P, Federici A. Estimation of Pap-test coverage in an area with an organised screening program: challenges for survey methods. *BMC Health Services Research.* 2006;6:36.
17. Gordon NP, Hiatt RA, Lampert DI. Concordance of self-reported data and medical record audit for six cancer screening procedures. *Journal of the National Cancer Institute.* 1993;85(7):566.

18. Hiatt RA, Pérez-Stable EJ, Quesenberry C, Sabogal F, Otero-Sabogal R, McPhee SJ. Agreement between self-reported early cancer detection practices and medical audits among Hispanic and non-Hispanic white health plan members in northern California. *Preventive Medicine.* 1995;24(3):278.
19. Holt K, Franks P, Meldrum S, Fiscella K. Mammography self-report and mammography claims: racial, ethnic, and socioeconomic discrepancies among elderly women. *Medical Care.* 2006;44(6):513-518.
20. Johnson CS, Archer J, Campos-Outcalt D. Accuracy of Pap smear and mammogram self-reports in a southwestern Native American tribe. *American Journal of Preventive Medicine.* 1995;11(6):360-363.
21. Johnson TP, O'Rourke DP, Burris JE, Warnecke RB. An investigation of the effects of social desirability on the validity of self-reports of cancer screening behaviors. *Medical Care.* 2005;43(6):565-573.
22. King ES, Rimer BK, Trock B, Balshem A, Engstrom P. How valid are mammography self-reports? *American Journal of Public Health.* 1990;80(11):1386.
23. Klungsoyr O, Nygard M, Skare G, Eriksen T, Nygard JF. Validity of self-reported Pap smear history in Norwegian women. *Journal of Medical Screening.* 2009;16(2):91-97.
24. Lawrence VA, De Moor C, Glenn ME. Systematic differences in validity of self-reported mammography behavior: a problem for intergroup comparisons? *Preventive Medicine.* 1999;29(6):577-580.
25. Lofters A, Vahabi M, Glazier RH. The validity of self-reported cancer screening history and the role of social disadvantage in Ontario, Canada. *BMC Public Health.* 2015;15:28.
26. Mahnken JD, Freeman DH, DiNuzzo AR, Freeman JL. Mammography use among older Mexican-American women: correcting for over-reports of breast cancer screening. *Women's Health.* 2007;45(3):53-64.
27. Martin LM, Leff M, Calonge N, Garrett C, Nelson DE. Validation of self-reported chronic conditions and health services in a managed care population *American Journal of Preventive Medicine.* 2000;18(3):215-218.
28. McGovern PG, Lurie N, Margolis KL, Slater JS. Accuracy of self-report of mammography and Pap smear in a low-income urban population. *American Journal of Preventive Medicine.* 1998;14(3):201.
29. Newell S, Girgis A, Sanson-Fisher R, Ireland M. Accuracy of patients' recall of Pap and cholesterol screening. *American Journal of Public Health.* 2000;90(9):1431.
30. Norman SA, Localio AR, Zhou L, et al. Validation of self-reported screening mammography histories among women with and without breast cancer. *American Journal of Epidemiology.* 2003;158(3):264-271.
31. Paskett ED, Tatum CM, Mack DW, Hoen H, Case LD, Velez R. Validation of self-reported breast and cervical cancer screening tests among low-income minority women. *Cancer Epidemiology, Biomarkers & Prevention.* 1996;5(9):721.
32. Pizarro J, Schneider T, Salovey P. A source of error in self-reports of pap test utilization. *Journal of Community Health.* 2002;27(5):351-356.
33. Sawyer JA, Earp JA, Fletcher RH, Daye FF, Wynn TM. Accuracy of women's self-report of their last Pap smear. *American Journal of Public Health.* 1989;79(8):1036.
34. Son E, Parish SL, Swaine JG, Luken K. Accuracy of self-reported cervical and breast cancer screening by women with intellectual disability. *American Journal on Intellectual and Developmental Disabilities.* 2013;118(4):327-336.

35. Suarez L, Goldman DA, Weiss NS. Validity of pap smear and mammogram self-reports in a low-income Hispanic population. *American Journal of Preventive Medicine.* 1995;11(2):94-98.
36. Thompson B, Taylor V, Goldberg H, Mullen M. Mammography status using patient self-reports and computerized radiology database. *American Journal of Preventive Medicine.* 1999;17(3):203-206.
37. Tsuruda KM, Sagstad S, Sebuodegard S, Hofvind S. Validity and reliability of self-reported health indicators among women attending organized mammographic screening. *Scandinavian Journal of Public Health.* 2018:1403494817749393.
38. Tumiel-Berhalter LM, Finney MF, Jaén CR. Self-report and primary care medical record documentation of mammography and Pap smear utilization among low-income women. *Journal of the National Medical Association.* 2004;96(12):1632.
39. Walter SD, Clarke EA, Hatcher J, Stitt LW. A comparison of physician and patient reports of pap smear histories. *Journal of Clinical Epidemiology.* 1988;41(4):401-410.
40. Howard M, Agarwal G, Lytwyn A. Accuracy of self-reports of Pap and mammography screening compared to medical record: a meta-analysis. *Cancer Causes & Control.* 2009;20(1):1-13.
41. Dodou D, de Winter JC. Agreement between self-reported and registered colorectal cancer screening: a meta-analysis. *European Journal of Cancer Care.* 2015;24(3):286-298.
42. van Enst WA, Ochodo E, Scholten RJ, Hooft L, Leeflang MM. Investigation of publication bias in meta-analyses of diagnostic test accuracy: a meta-epidemiological study. *BMC Medical Research Methodology.* 2014;14:70.