

---

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

---

February 2019

### Prepared for:

Department of Veterans Affairs  
Veterans Health Administration  
Quality Enhancement Research Initiative  
Health Services Research & Development Service  
Washington, DC 20420

### Prepared by:

Evidence Synthesis Program (ESP)  
Coordinating Center  
Portland VA Health Care System  
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

## PREFACE

The VA Evidence Synthesis Program (ESP) was established in 2007 to provide timely and accurate syntheses of targeted healthcare topics of importance to clinicians, managers, and policymakers as they work to improve the health and healthcare of Veterans. These reports help:

- Develop clinical policies informed by evidence;
- Implement effective services to improve patient outcomes and to support VA clinical practice guidelines and performance measures; and
- Set the direction for future research to address gaps in clinical knowledge.

The program is comprised of four ESP Centers across the US and a Coordinating Center located in Portland, Oregon. Center Directors are VA clinicians and recognized leaders in the field of evidence synthesis with close ties to the AHRQ Evidence-based Practice Center Program and Cochrane Collaboration. The Coordinating Center was created to manage program operations, ensure methodological consistency and quality of products, and interface with stakeholders. To ensure responsiveness to the needs of decision-makers, the program is governed by a Steering Committee comprised of health system leadership and researchers. The program solicits nominations for review topics several times a year via the [program website](#).

Comments on this evidence report are welcome and can be sent to Nicole Floyd, Deputy Director, ESP Coordinating Center at [Nicole.Floyd@va.gov](mailto:Nicole.Floyd@va.gov).

**Recommended citation:** Anderson J, Bourne D, Peterson, K, Mackey K. Evidence Brief: Accuracy of Self-report for Cervical and Breast Cancer Screening. VA ESP Project #09-199; 2019. Posted final reports are located on the ESP [search page](#).

This report is based on research conducted by the Evidence Synthesis Program (ESP) Coordinating Center located at the **Portland VA Health Care System, Portland, OR**, funded by the Department of Veterans Affairs, Veterans Health Administration, Health Services Research and Development. The findings and conclusions in this document are those of the author(s) who are responsible for its contents; the findings and conclusions do not necessarily represent the views of the Department of Veterans Affairs or the United States government. Therefore, no statement in this article should be construed as an official position of the Department of Veterans Affairs. No investigators have any affiliations or financial involvement (eg, employment, consultancies, honoraria, stock ownership or options, expert testimony, grants or patents received or pending, or royalties) that conflict with material presented in the report.

## TABLE OF CONTENTS

Executive Summary .....	1
Key Findings .....	1
Evidence Brief .....	3
Introduction.....	3
Purpose.....	3
Background .....	3
Scope.....	5
Key Questions.....	5
Eligibility Criteria .....	5
Methods.....	6
Results.....	8
Literature Flow.....	8
Key Question 1: What is the accuracy of self-reported measures of cervical and breast cancer screening compared to medical records? .....	12
Cervical Cancer Screening.....	12
Breast Cancer Screening .....	15
Key Question 2: What are the potential adverse effects and unintended consequences of using or not using self-reported measures of cervical and breast cancer screening? .....	17
Key Question 3: Do the accuracy and/or adverse effects of using self-reported measures of cervical and breast cancer screening vary by patient or measure characteristics or setting? ...	17
Cervical Cancer Screening.....	17
Breast Cancer Screening .....	17
Summary and Discussion.....	19
Key Findings and Clinical Considerations .....	19
Limitations .....	20
Future Research .....	21
Conclusions.....	21
Acknowledgments.....	22
Operational Partners.....	22
Peer Reviewers.....	22
References.....	23

## FIGURES AND TABLES

ES Figure. Summary of Findings.....	2
Figure 1: Literature Flowchart.....	8
Table 1. Characteristics of Included Studies.....	9
Figure 2. Risk of Bias of Included Studies.....	12
Table 2. Pap Self-report Meta-analysis Results.....	13
Figure 3. Sensitivity and Specificity for Pap Self-report.....	14
Table 3. Mammogram Self-report Meta-analysis Results.....	15
Figure 4. Sensitivity and Specificity for Mammogram Self-report.....	16

## EXECUTIVE SUMMARY

### Key Findings

- Unscreened women tend to over-report having had a mammogram or pap test, but screened more accurately report their screening.
  - 48% to 61% of unscreened patients according to their medical record accurately reported no screening (39% to 52% over-reported screening).
  - 96% of screened patients according to their medical record accurately reported their screening.
- We have moderate confidence in these findings, as there are a large number of mostly fair-quality studies directly assessing the accuracy of self-report compared to medical records.
- Future research should focus on assessing the impact of accepting self-report on clinical and system-level outcomes.

Guideline-based breast and cervical cancer screening are considered essential health benefits and are fundamental components of high-quality primary care services in the US. The aim of cancer screening is to identify cancers in an early stage, often before symptoms present, when treatment is more likely to be effective. Accurate measurement of cancer screening rates is vital in order to understand if women are adequately screened and to determine if disparities exist in receipt of screening. Most measurement of cancer screening rates relies on patient self-report, but there are trade-offs to consider when utilizing self-report versus medical record documentation. Although self-report data may reduce administrative burden compared to medical record documentation, its accuracy is often questioned. Concerns about self-report data accuracy and potential for adverse outcomes, including the potential for missed screenings, missed diagnoses, or duplicative screening, have led to consideration of increasing the requirements for reporting documentation of cancer screenings. Prior to considering changing documentation requirements, it is important to understand benefits and harms of accepting self-report, including the potential for over- or under-screening, missed diagnoses, and impacts on provider, administrative, or patient burden.

We included 39 studies assessing the accuracy of self-reported cervical and breast cancer screening compared to medical record. None were in Veterans. Overall, unscreened women tended to over-report screening, but screened women more accurately reported their screening (ES Table). Forty-eight to 61% (specificity) of unscreened women, according to their medical records, accurately reported not having a screening (39% to 52% over-reported screening). However, screened women were more accurate in reporting they had a pap test or mammogram (96% (sensitivity) of screened women, according to their medical records, accurately reported their screening). Additionally, 80% to 84% (positive predictive value (PPV)) of women's self-reports of having a screening were verified in the medical record, and 83% to 86% (negative

### Background

The ESP Coordinating Center (ESP CC) is responding to a request from the VHA Performance Workgroup for an evidence brief on the accuracy of self-report for cervical and breast cancer screening. Findings from this evidence brief will be used to inform collection and use of self-report measures in cervical and breast cancer screening in the VHA.

### Methods

To identify studies, we searched MEDLINE®, Cochrane Database of Systematic Reviews, Cochrane Central Register of Controlled Trials, and other sources up to November 2018. We used prespecified criteria for study selection, data abstraction, and rating internal validity and strength of the evidence. See our PROSPERO protocol for our full methods.

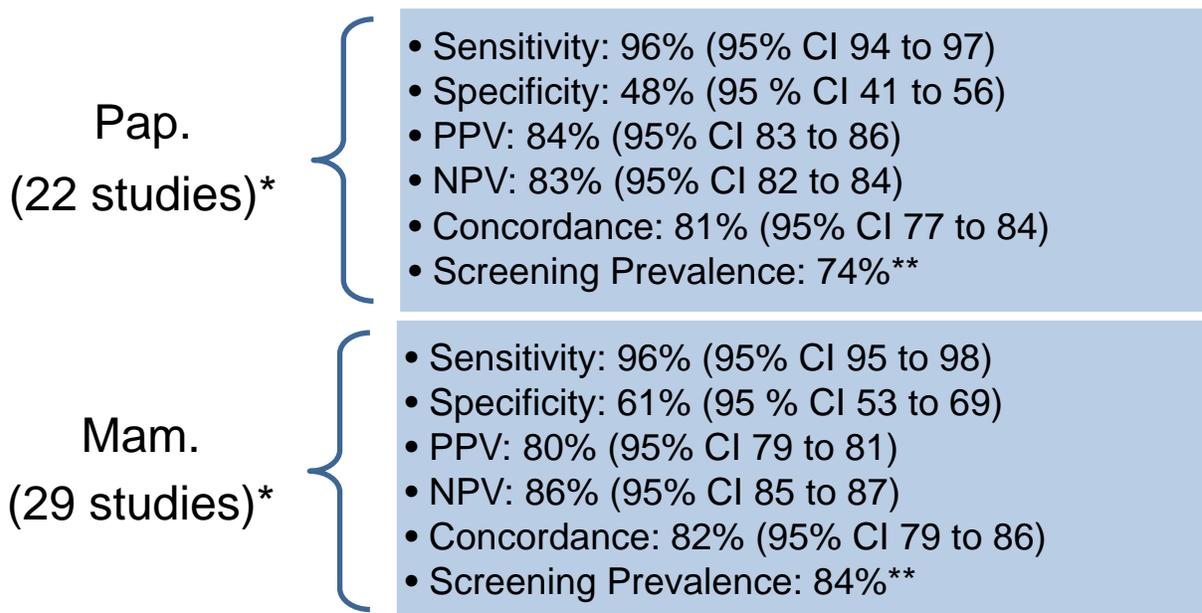


predictive value (NPV)) of women's self-reports of not having screening were verified in the medical record. These results indicate that self-report is an accurate measure for patients who have had a mammogram or pap test. However, the lower specificity values suggest that there is over-reporting of mammograms and pap tests, with patients reporting they have had a test, but no medical record documentation of the test being found.

We have moderate confidence in these findings, as they come from a large number of mostly fair-quality studies which directly assessed the accuracy of self-report compared to medical records. Furthermore, these findings are consistent with a previous meta-analysis published in 2009. There was significant heterogeneity among studies, but this is likely due to differences between the populations studied (*eg*, a study in low-income Black women versus a population-based study in a Scandinavian country). Subgroup analyses revealed a significant association of low-income or minority population with self-report accuracy for both cervical and breast cancer screening, with studies in minority or low-income populations having lower sensitivity than those in general populations. Although heterogeneity remained in these subgroup analyses, this may be due to our inability to adjust for differences in age and between low-income or minority populations because of inadequate reporting in the studies.

We did not find any studies reporting clinical outcomes or adverse events or unintended consequences of accepting self-report. With over-reporting of screening by self-report there is the potential to overestimate the success of screening interventions, mask disparities in screening prevalence, under-screen individuals, and miss cancer diagnoses. Future research should investigate these outcomes as well as system-level outcomes (*eg*, patient and provider burden) to weigh the potential harms of accepting self-report against the time and resource burden of tracking medical records.

### ES Figure. Summary of Findings



Abbreviations: PPV = positive predictive value; NPV = negative predictive value;

\*Not mutually exclusive; \*\*As verified by medical record

# EVIDENCE BRIEF

## INTRODUCTION

### PURPOSE

The ESP Coordinating Center (ESP CC) is responding to a request from the VHA Performance Workgroup for an evidence brief on the accuracy of patient self-report for cervical and breast cancer screening. Findings from this evidence brief will help the VHA Performance Workgroup decide whether to continue the current practice of accepting patient self-reported data on cervical and/or breast cancer screening or require medical record documentation of prior screening, which is currently the standard for colorectal cancer screening.

### BACKGROUND

Guideline-based breast and cervical cancer screening are considered essential health benefits and are fundamental components of high-quality primary care services in the US.<sup>1,2</sup> The aim of cancer screening is to identify cancers in an early stage, often before symptoms present, when treatment is more likely to be effective. This timely treatment can result in reducing cancer incidence and disease burden and improving chances of survival.<sup>3-5</sup> Current recommendations for cervical cancer screening in average-risk women are the Papanicolaou (pap) test every 3 years from ages 21 to 29, and from ages 30 to 65 either a pap test every 3 years, a pap test plus a human papilloma virus (HPV) test every 5 years, or HPV testing alone every 5 years (see Supplemental Materials for further detail). Breast cancer screening recommendations in average-risk women (*ie*, those without increased risk for breast cancer, such as a positive family history or predisposing genetic mutation) vary by guideline and range from annual mammograms starting at age 40 to mammograms every 2 years starting at age 50 (with individual choice to start screening at age 40). The Veterans Health Administration (VHA) currently follows the American Cancer Society (ACS) and US Preventive Services Task Force (USPSTF) guidelines for cervical cancer screening (*ie*, pap test every 3 years starting at age 21, with the option to screen every 5 years, with pap plus HPV or HPV alone from ages 30 to 65)<sup>6</sup> and the ACS guidelines for mammography (*ie*, annual mammograms starting at age 45 and every other year after age 55).<sup>7</sup>

Despite initiatives and interventions to improve adherence to these guideline recommendations, breast and cervical cancer screening rates still fall short of the Healthy People 2020 goals to have 81.1% of women receiving breast cancer screening and 93% of women receiving cervical cancer screening in concordance with recent guidelines.<sup>8</sup> Since the 1980s, mammography screening rates in the US have increased from 30% to 64% (percentage of women 40 years of age and older having received a mammogram within the past 2 years) from 1987 to 2015, respectively.<sup>9</sup> US cervical cancer screening rates have remained relatively stable, with a slight decline, since the 1980s, going from 74% to 70% (percentage of women aged 18 years and older having received a pap test within the past 3 years) from 1987 to 2015, respectively.<sup>10</sup> Minority or low-income women may face more barriers to screening and be particularly vulnerable to screening nonadherence (*eg*, for women below 100% of the poverty level, 43% to 55% had mammograms and 55% to 69% had pap tests in 2013).<sup>11</sup> Racial and economic disparities in receipt of care also exist,<sup>12,13</sup> and it has been hypothesized that minority or low-income women may over-report screening more than others (potentially due to lower health literacy, cultural factors, or the

quality of medical record documentation in these populations),<sup>14,15</sup> which may mask this disparity in population-based surveys.

Accurate measurement of cancer screening rates is vital to understand if women are adequately screened and determine if disparities exist in receipt of screening. These data can be used to further develop initiatives to increase the uptake of screening (*ie*, public education campaigns, patient reminders, and provider and health system-level practices)<sup>16,17</sup> and to assess hospital and provider performance through quality performance measures.<sup>18-20</sup> There are several methods to collect cancer screening data ranging from population-based mail or telephone surveys to track trends over time<sup>21</sup> to specific medical record documentation requirements for quality reporting.<sup>22</sup> Surveys generally rely on patient self-report, while medical record documentation can vary from physician-recorded data (entered from patient self-report) to pathology, lab, or radiology reports confirming that a test was performed.

There are several trade-offs to consider when determining whether to utilize self-report or medical record documentation for measurement of cancer screening rates. Given that most individuals in the US receive health care from multiple providers, clinics, and hospitals, obtaining specific medical record documentation of prior screening may be time-consuming and costly. Collecting specific medical record documentation typically requires patient consent and may entail calling or faxing various clinics to request records, scanning or filing these records, and review by a physician or other clinical team member. For these reasons, most cancer screening data rely on patient self-report.<sup>21,23</sup> However, although self-report data may reduce administrative burden, its accuracy is often questioned compared to medical record documentation.<sup>24-26</sup> Self-reported screening data is subject to potential conscious and unconscious biases including telescoping (remembering an event occurring more recently than it actually occurred),<sup>27</sup> acquiescence bias (the tendency to respond “yes” when in doubt),<sup>28</sup> and social desirability response bias (over-reporting of events that are socially desirable).<sup>29-32</sup> The extent of these potential biases can depend on the data collection method (*ie*, mailed survey vs telephone interview vs face-to-face interview) and how questions are phrased (*ie*, “How long has it been since you had your last mammogram?” vs “Have you had a mammogram in the past 2 years?”) and are important considerations when utilizing self-report for measurement of cancer screening.<sup>33,34</sup>

Concerns about the accuracy of self-report data for cervical and breast cancer screening and the potential for adverse outcomes, including missed screenings, missed diagnoses, or duplicative screening, have led to policy variations in requirements for reporting documentation of these screenings. The Healthcare Effectiveness Data and Information Set (HEDIS) is a performance tool set developed and sustained by the National Committee for Quality Assurance (NCQA) that is utilized by health plans, hospitals, and health systems for quality metrics and performance improvement.<sup>35</sup> HEDIS accepts patient-reported information, collected by primary care physicians and reported in the medical record, for cancer screening quality measures.<sup>22</sup> The VHA guidelines for collecting and reporting cancer screening data generally agree with the HEDIS requirements. One exception to this is that the VHA has historically required more rigorous documentation (*ie*, does not allow for patient self-report) for certain colorectal cancer screening tests.<sup>36</sup> The VHA is currently considering requiring more rigorous documentation for cervical and breast cancer screening as well.

Prior to considering changing documentation requirements, it is important to understand the accuracy of self-report compared to medical records, as well as the benefits and harms of

accepting self-report, including the potential for over- or under-screening, missed diagnoses, and impacts on provider, administrative, or patient burden. A 2009 systematic review<sup>25</sup> on the accuracy of self-report for breast and cervical cancer screening found that women tend to over-report cancer screenings (38% to 53% of unscreened women, according to medical records, inaccurately reported having a screening), but these findings may be outdated, as several subsequent studies have emerged since publication of this review. The impacts of accepting self-report on other outcomes (*ie*, missed diagnoses, administrative burden) were not assessed. The purpose of our evidence review is to provide an updated synthesis of the evidence on the accuracy of self-reported measures of cervical and breast cancer screening compared to medical records and the benefits and potential harms of relying on patient self-report instead of medical record documentation.

## SCOPE

This evidence brief will address the following key questions and inclusion criteria:

### Key Questions

Key Question 1: What is the accuracy of self-reported measures of cervical and breast cancer screening compared to medical records?

Key Question 2: What are the potential adverse effects and unintended consequences of using or not using self-reported measures of cervical and breast cancer screening?

Key Question 3: Do the accuracy and/or adverse effects of using self-reported measures of cervical and breast cancer screening vary by patient or measure characteristics or setting?

### Eligibility Criteria

The ESP included studies that met the following criteria:

- **P**opulation: women reporting cervical or breast cancer screening, not identified as high risk (*ie*, genetic mutation carriers, patients with abnormal previous screens) and without a history or diagnosis of breast or cervical cancer
- **I**ntervention: Self-reported cervical (Papanicolaou smear and/or HPV testing) or breast cancer (mammogram) screening
- **C**omparator: Medical records
- **O**utcomes:
  - Accuracy (sensitivity, specificity, positive and negative predictive value, *etc*)
  - Adverse effects or unintended consequences: Any (missed diagnoses, over- or under-screening, patient or provider burden, *etc*)
- **T**iming: Any
- **S**etting: Any
- **S**tudy design: Any, but may prioritize to accommodate timeline using a best-evidence approach

## METHODS

To identify articles relevant to the key questions, our research librarian searched Medline, CINAHL, PsycINFO, CCRCT, and the Cochrane Database of Systematic Reviews, using terms for self-report and screening, mammography, pap smear, breast cancer screening, and cervical cancer screening (see Supplemental Materials for complete search strategies). Additional citations were identified from hand-searching reference lists and consultation with content experts. We limited the search to published and indexed articles involving human subjects available in the English language. Study selection was based on the eligibility criteria described above. Titles and abstracts were independently reviewed by 2 investigators. Full-text articles were reviewed by one investigator and checked by another. All disagreements were resolved by consensus.

We used predefined criteria to rate the internal validity of all included studies. We used the Cochrane ROBIS tool to rate the internal validity of systematic reviews.<sup>37</sup> We used the QUADAS-2 tool to rate the internal validity of diagnostic accuracy studies.<sup>38</sup> We abstracted data from all studies for prespecified study and patient characteristics of interest and results for each included outcome. We calculated sensitivity, specificity, positive predictive value, negative predictive value, concordance, and report to record (Rep/Rec) ratio for all included studies. Studies which did not provide enough data for calculation of these accuracy characteristics were excluded. All data abstraction and internal validity ratings were first completed by one reviewer and then checked by another. All disagreements were resolved by consensus.

We informally graded the strength of the evidence based on the AHRQ Methods Guide for Comparative Effectiveness Reviews<sup>39</sup> by considering risk of bias (includes study design and aggregate quality), consistency, directness, and precision of the evidence. Ratings typically range from high to insufficient, reflecting our confidence that the evidence reflects the true effect. For this review, we applied the following general algorithm: evidence comprised of multiple large studies with moderate risk of bias and inconsistent finding (significant heterogeneity) received a rating of “low strength”, and this same type of evidence but with consistent findings (limited heterogeneity) received a rating of “moderate strength”.

Where studies were appropriately homogenous, we synthesized outcome data quantitatively using STATA v.15 (StataCorp. College Station, TX). Overall numbers for sensitivity and specificity were calculated and adjusted using a bivariate random-effects meta-analysis. Positive predictive values (PPV) and negative predictive values (NPV) were calculated using the midas command (generates summary predictive values) in STATA v.15<sup>40</sup> and adjusted for overall prevalence. Concordance and Rep/Rec ratio were averaged across studies. Definitions and formulas for these calculations are provided in the Supplemental Materials. We also conducted a univariable bivariate and multivariable univariate meta-regression to examine the impact of moderator variables. Variables were prespecified and included age, publication year, sample size, study design (case-control vs cohort), setting (population- vs clinic-based), referral or screening program (whether the patients were a part of a screening or referral program vs not), survey administration method (in-person vs telephone or mail survey), timeframe of recall ( $\leq 1$  year to ever), minority status, and low-income status. Heterogeneity was assessed using the Q statistic and the  $I^2$  statistic. Publication bias was examined by Deeks' test.<sup>41</sup> Where meta-analysis was not suitable due to limited data or heterogeneity, we synthesized the evidence qualitatively.

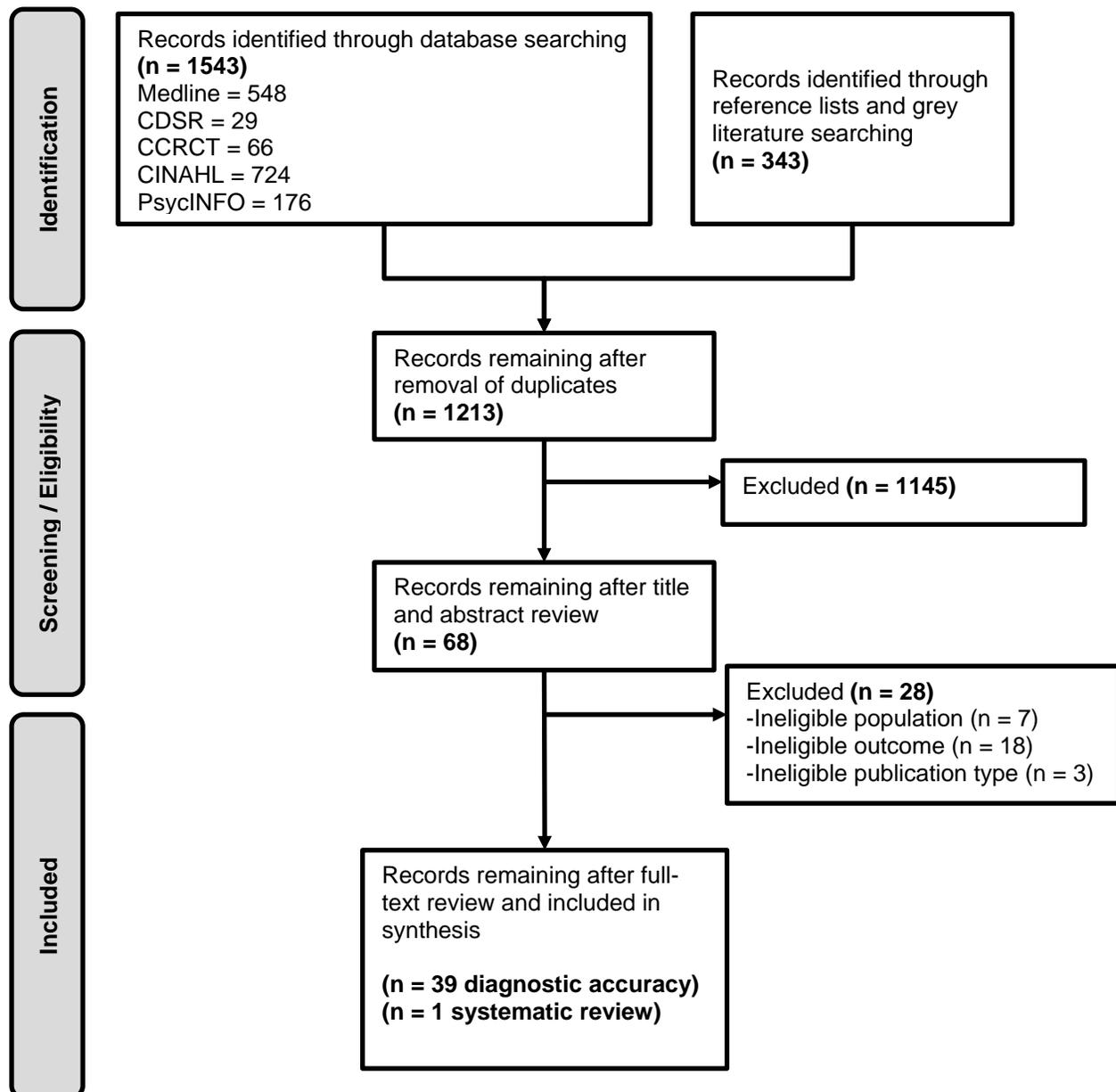
A draft version of this report was reviewed by peer reviewers as well as clinical leadership (see Supplemental Materials for disposition of peer review comments). The complete description of our full methods can be found on the PROSPERO international prospective register of systematic reviews (<http://www.crd.york.ac.uk/PROSPERO/>; registration number CRD42019116781).

## RESULTS

### LITERATURE FLOW

The literature flow diagram (Figure 1) summarizes the results of the search and study selection processes (see Supplemental Materials for full list of excluded studies).

**Figure 1: Literature Flowchart**



Our search identified 1,213 unique, potentially relevant articles. We included 39 diagnostic accuracy studies: 22 in cervical cancer screening,<sup>15,34,42-61</sup> 29 in breast cancer screening<sup>15,33,34,51-60,62-77</sup> (not mutually exclusive), and 1 systematic review.<sup>25</sup> Among the diagnostic accuracy studies, 28 were included in the previous systematic review, and 11 were new. All studies

included for cervical cancer screening assessed the pap test, and none specifically examined HPV testing. Most studies (62%) included fewer than 500 participants, but a few large population-based studies ranged from over 10,000 to over 400,000 participants (Table 1) (see Supplemental Materials for full data tables). Most studies were old, with only 6 studies published in the past 10 years. Overall, 19 studies (49%) included minority or low-income populations, 23 studies (59%) were clinic-based (*ie*, used data from patients who attended a specific hospital or clinic), which is likely most applicable to the VA setting, and 23 studies (59%) conducted surveys face-to-face.

**Table 1. Characteristics of Included Studies**

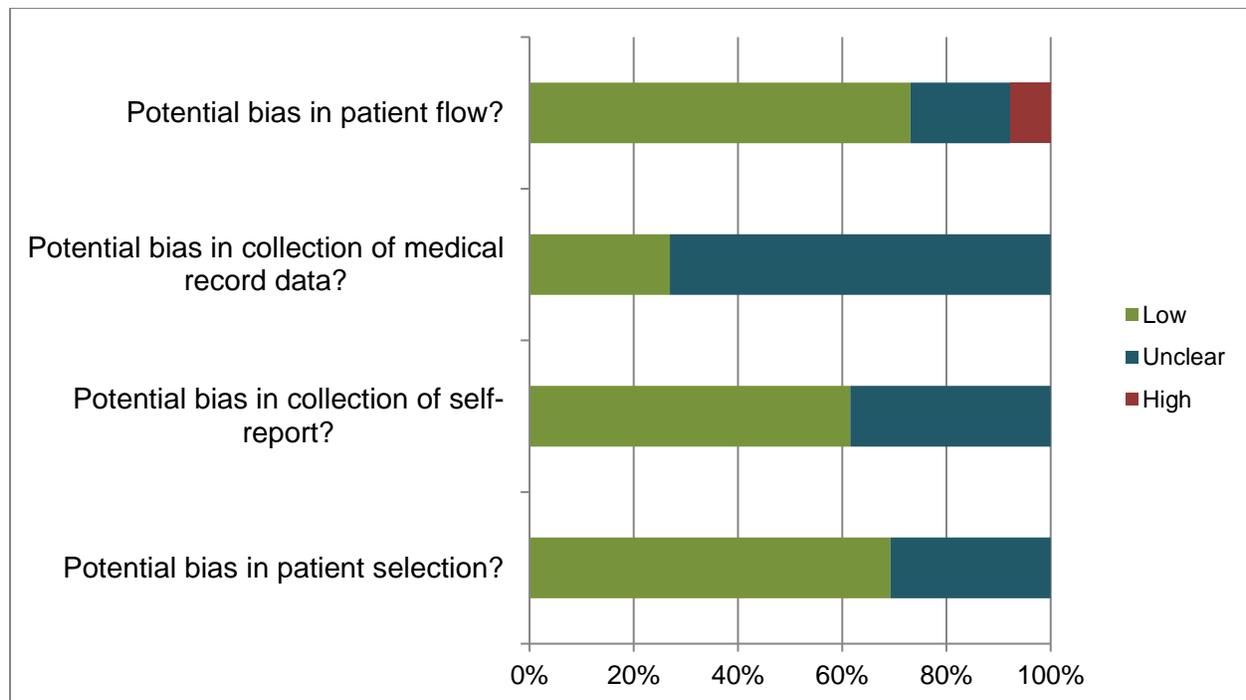
Author, Year N	Age range Minority/ Low Income?	Clinic- or Population-based Screening? <sup>†</sup>	Study part of a screening or referral program? <sup>‡</sup>	Survey Method Timeframe of Recall
<i>Pap Smear</i>				
Bowman, 1991 <sup>*42</sup> 157	18-70 None	Population-based	No	Face-to-face 3 yrs
Eaker, 2001 <sup>*43</sup> 944	25-60 None	Population-based	No	Phone Ever
Fowles, 1997 <sup>*61</sup> 400	19-75 None	Clinic-based	No	Phone 3 yrs
Fruchter, 1992 <sup>*44</sup> 263	Age NR Black/Latino	Clinic-based	No	Face-to-face 3 yrs
Georgi-Rossi, 2006 <sup>45</sup> 641	25-64 None	Population-based	Yes	Phone Ever
Klungsoyr, 2009 <sup>46</sup> 16,574	18-45 None	Population-based	No	Phone/mail/ computer 3 yrs
Newell, 2000 <sup>*47</sup> 423	18-81 None	Population-based	No	Face-to-face Last screen
Pizzaro, 2002 <sup>*48</sup> 161	18-89 Low income	Clinic-based	No	Phone 1 yr
Sawyer, 1989 <sup>*49</sup> 98	16-75 Rural, Black	Population-based	No	Face-to-face 3 yrs
Walter, 1988 <sup>*50</sup> 750	20-69 None	Clinic	No	Face-to-face 5 yrs
<i>Mammogram</i>				
Allgood, 2014 <sup>62</sup> 1,221	40+ yrs Low income/Blac k	Population-based	No	Face-to-face 2 yrs
Armstrong, 2004 <sup>*33</sup> 399	50-75 yrs Low income	Clinic-based	No	Phone Ever
Baron-Epel, 2008 <sup>63</sup> 1,550	52-74 Jewish/Arab	Clinic-based	No	Phone 2 yrs
Brown, 1992 <sup>*64</sup>	17-79	Clinic-based	No	Phone

189	None			1 month
Caplan (a), 2003* <sup>65</sup> 949	50-80 No	Clinic-based	Yes	Phone Ever
Champion, 1998* <sup>66</sup> 229	45-64 Low income/Black	Clinic-based	Yes	Face-to-face 3 yrs
Clark, 2009* <sup>67</sup> 411	40-75 Black	Clinic-based	Yes	Face-to-face 2 yrs
Crane, 1996* <sup>68</sup> 576	50+ Hispanic/Black	Clinic-based	Yes	Phone Last screen
Degnan, 1992* <sup>69</sup> 456	50-74 None	Population-based	No	Phone Last screen
Fulton- Kehoe, 1992* <sup>70</sup> 78	50-75 None	Clinic-based	No	Mail 1 yr
Holt, 2006* <sup>71</sup> 5,461	65+ None	Population-based	No	Face-to-face 1 yr
King, 1990* <sup>72</sup> 199	50-74 None	Clinic-based	Yes	Phone 1 yr
Lawrence, 1999* <sup>73</sup> 93	50-70 Mexican- American	Clinic-based	No	Phone 1 yr
Mahnken, 2007* <sup>74</sup> 199	50-74 Mexican- American	Population-based	No	Face-to-face 2 yrs
Norman, 2003* <sup>75</sup> 2,495	40-64 None	Clinic/Population- based	No	Phone Last screen
Thompson, 1999* <sup>76</sup> 360	50-69 Low income	Clinic-based	No	Mail Ever
Tsurda, 2018* <sup>77</sup> 411,294	50-69 No	Population-based	Yes	Mail Ever
<i>Pap Smear + Mammogram</i>				
Caplan, 2003* <sup>51</sup> 480	40-74 None	Clinic-based	No	Phone Ever
Gordon, 1993* <sup>52</sup> 431	40-74 None	Clinic-based	No	Mail/phone 2 yrs
Hiatt, 1995* <sup>53</sup> 691	35-75 Hispanic	Clinic-based	No	Phone 5 yrs
Johnson, 1995* <sup>54</sup> 251	35-65 Native American	Population-based	No	Face-to-face Ever
Johnson, 2005* <sup>34</sup> 588	50+ None	Population-based	No	Phone/ computer- assisted 3 yrs

Lofters, 2015 <sup>55</sup> 39,027 (Pap) 15,877 (Mam)	24-69 (Pap) 52-69 (Mam) None	Population-based	No	Phone/ computer- assisted Ever
Martin, 2000 <sup>*56</sup> 599	21+ None	Clinic-based	No	Phone 3 yrs (Pap) 2 yrs (Mam)
McGovern, 1998 <sup>*57</sup> 477	40-92 Low income	Clinic-based	No	Face-to-face Ever
Paskett, 1996 <sup>*58</sup> 555	40+ Low income	Population-based	Yes	Face-to-face Ever
Son, 2013 <sup>59</sup> 155	18+ Intellectual disability	Clinic-based	No	Face-to-face Ever
Suarez, 1995 <sup>*15</sup> 450	40+ Mexican- American	Population-based	Yes	Face-to-face Ever
Tumiel-Berhalter, 2004 <sup>*60</sup> 314	40+ Black/Puert o Rican	Clinic-based	No	Face-to-face 3 yrs

\*Included in Howard 2009; †Clinic-based screening used data from patients who attended a specific hospital or clinic. Population-based screening used data from multiple clinics or a nationally representative database; ‡Participants who are a part of a screening or referral program, which recruits individuals to attend an event or come into the clinic/hospital to take part in a specially organized program.  
Abbreviations: Mam. = mammography, Yrs = years

We have moderate confidence in these findings, as they come from a large number of mostly fair-quality studies which directly assessed the accuracy of self-report compared to medical records. A total of 61,383 patients were included in studies of cervical cancer screening, and 444,055 patients were included in studies of breast cancer screening. Most studies had unclear risk of bias (Figure 2) due to unclear blinding and incomplete medical record review. Lack of or unclear blinding of data collectors may have led to potential bias in collection of self-report or medical record data. However, this potential for bias is likely low, as the data were reported by patients and providers and not directly by those collecting the data. Incomplete verification of the medical records for all study participants (*ie*, only a single medical record source searched or out-of-the-area medical records not verified) also contributed to the potential for bias in the medical record review, since patients where a medical record was not found were often assumed to be unscreened, which may have led to a higher rate of false positives. There was significant heterogeneity among studies, but this is likely due to differences among the populations studied (*eg*, a study in low-income Black women vs a population-based study in a Scandinavian country). Although heterogeneity remained in subgroup analyses, this may be due to our inability to adjust for age or differences between low-income or minority populations because of inadequate reporting of detailed patient characteristics in the studies.

**Figure 2. Risk of Bias of Included Studies**

## KEY QUESTION 1: What is the accuracy of self-reported measures of cervical and breast cancer screening compared to medical records?

### Cervical Cancer Screening

Women tended to over-report receipt of cervical cancer screening among the 22 studies reporting the accuracy of pap test self-report compared to medical record (Table 2, Figure 3). 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). These findings are consistent with the previous meta-analysis (specificity: 47% and sensitivity: 95%).<sup>25</sup> Taken together, the average overall agreement (concordance) between self-report (both positive and negative) and medical record was 81% (95% CI 78 to 84).

Overall, 84% (95% CI 82 to 86) of self-reports of screening were verified in the medical record (PPV), and 83% (95% CI 82 to 84) of self-reports of no screening were verified in the medical record (NPV). These findings (PPV and NPV) are dependent upon the screening rate in the population, with a higher screening prevalence leading to higher PPV and lower NPV. The screening prevalence in the included studies (as verified by medical record) was fairly high at 74% in the included studies.

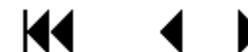
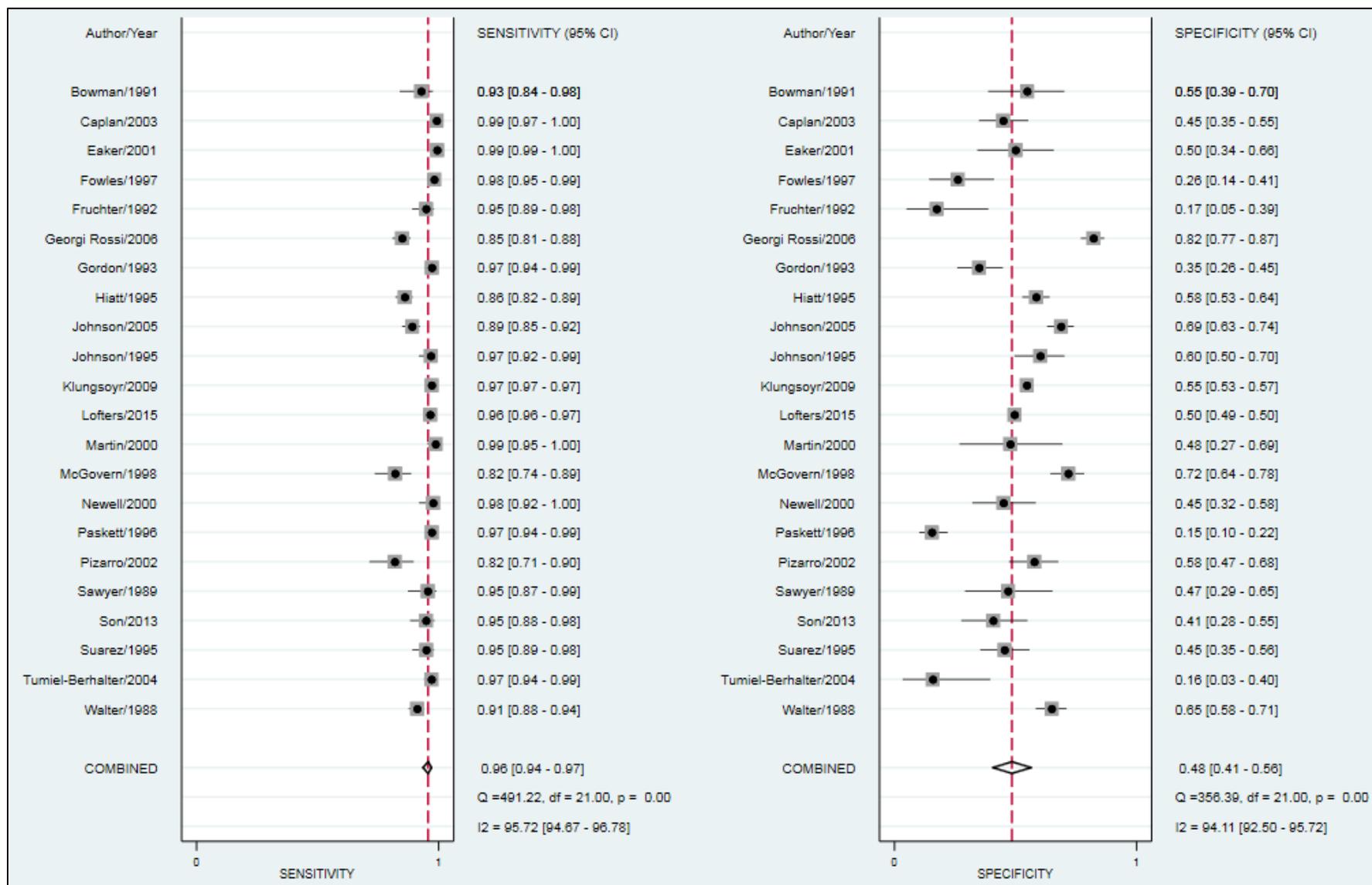
There was considerable heterogeneity and inconsistency for both sensitivity ( $\chi^2 P < .001$ ,  $I^2 = 95.7$ ) and specificity ( $\chi^2 P < .001$ ,  $I^2 = 94.1$ ) outcomes. There was no evidence of publication bias ( $P = .131$ ) (see Supplemental Materials).

**Table 2. Pap Self-report Meta-analysis Results**

<b>Statistic</b>	<b>Pooled Estimate (95% CI)</b>
Sensitivity	0.96 (0.94 to 0.97)
Specificity	0.48 (0.41 to 0.56)
PPV*	0.84 (0.82 to 0.86)
NPV*	0.83 (0.82 to 0.84)
Concordance	0.81 (0.78 to 0.84)
Rep/Rec Ratio	1.20 (0.97 to 1.47)

\*Numbers for PPV and NPV do not include data from case-control studies and are adjusted for prevalence  
Abbreviations: CI = confidence interval; PPV = positive predictive value; NPV = negative predictive value;  
Rep/Rec = report-to-record

**Figure 3. Sensitivity and Specificity for Pap Self-report**



## Breast Cancer Screening

Similar to the cervical cancer screening findings, women tended to over-report mammograms, among the 29 studies reporting the accuracy of mammogram self-report compared to medical record (Table 3, Figure 4). Among unscreened women, according to their medical record, 61% (95% CI 53 to 69) accurately reported no screening (*ie*, specificity). This indicates that almost 40% of women without screenings in their medical record inaccurately reported having a screening. The level of inaccurate reporting among unscreened women was lower for mammogram than for pap test recall. Among screened women, according to their medical record, 96% (95% CI 95 to 98) accurately reported this screening (*ie*, sensitivity). These findings are consistent with the previous meta-analysis (specificity: 62% and sensitivity: 95%).<sup>25</sup> Taken together, the average overall agreement (concordance) between self-report (both positive and negative) and medical record was 82% (95% CI 79 to 86).

Overall, 80% (95% CI 79 to 81) of self-reports of screening were verified in the medical record (PPV), and 86% (95% CI 85 to 87) of self-reports of no screening were verified in the medical record (NPV). These findings (PPV and NPV) are dependent upon the screening rate in the population, with a higher screening prevalence leading to higher PPV and lower NPV. The screening prevalence in the included studies (as verified by medical record) was fairly high at 84% in the included studies.

As was seen with the cervical cancer screening results, there was considerable heterogeneity and inconsistency for both sensitivity ( $\chi^2 P < .001$ ,  $I^2 = 99.5$ ) and specificity ( $\chi^2 P < .001$ ,  $I^2 = 99.7$ ) outcomes. There was evidence of publication bias ( $P < .001$ ), which 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 (see Supplemental Materials). 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.<sup>41</sup>

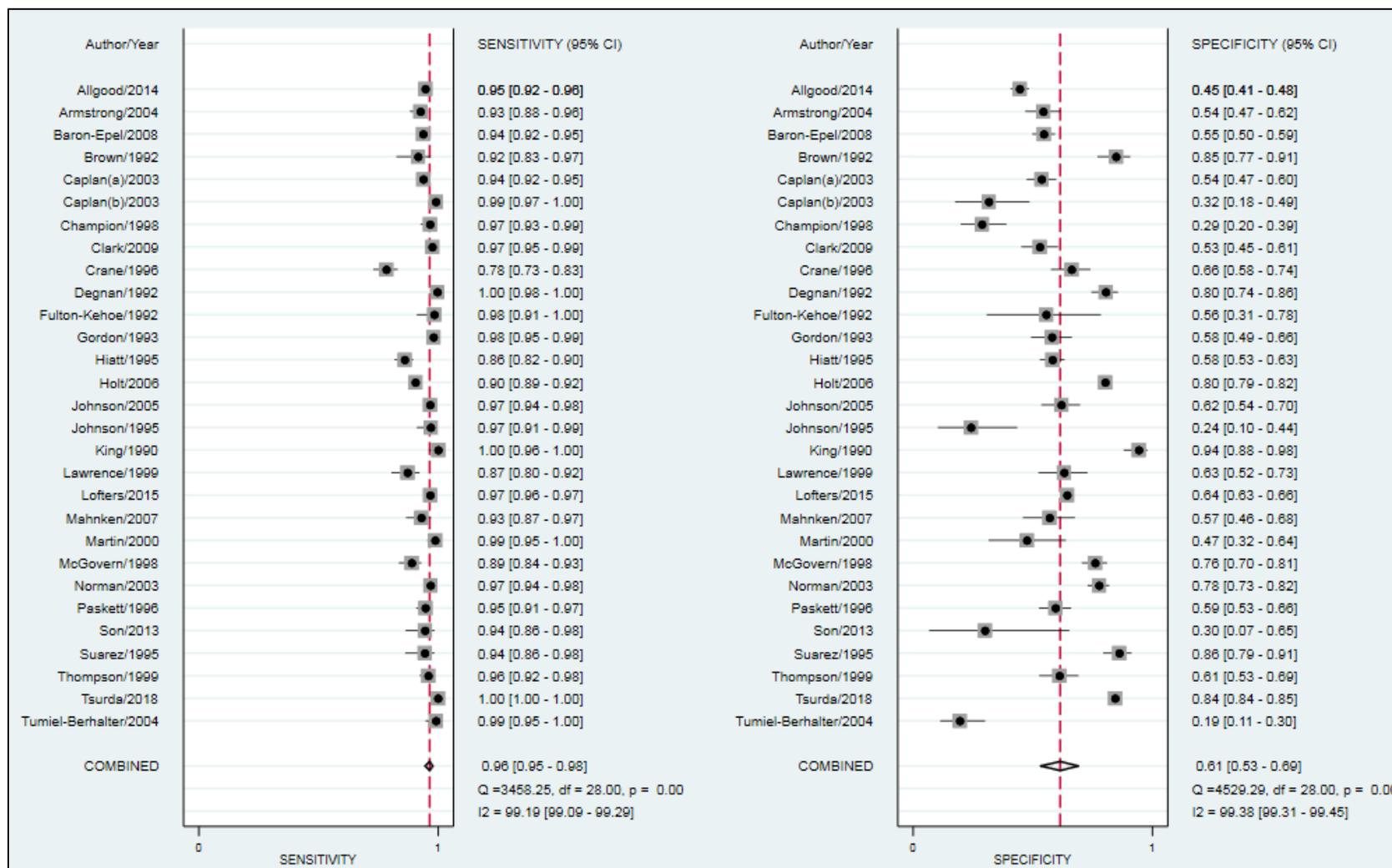
**Table 3. Mammogram Self-report Meta-analysis Results**

Statistic	Pooled Estimate (95% CI)
Sensitivity	0.96 (0.95 to 0.98)
Specificity	0.61 (0.53 to 0.69)
PPV*	0.80 (0.79 to 0.81)
NPV*	0.86 (0.85 to 0.87)
Concordance	0.82 (0.79 to 0.86)
Rep/Rec Ratio	1.20 (0.96 to 1.63)

\*Numbers for PPV and NPV do not include data from case-control studies

Abbreviations: CI = confidence interval; PPV = positive predictive value, NPV = negative predictive value; Rep/Rec = report-to-record

**Figure 4. Sensitivity and Specificity for Mammogram Self-report**



## **KEY QUESTION 2: What are the potential adverse effects and unintended consequences of using or not using self-reported measures of cervical and breast cancer screening?**

No studies reported adverse events or unintended consequences as outcomes. However, study authors frequently hypothesized potential adverse events or unintended consequences of using self-report, including overestimating of the success of screening interventions, masking of disparities in screening prevalence, less-frequent screening, and risk of missed cancer diagnoses. Mention of adverse events or unintended consequences of using medical record was less frequent, but included time and resource burden of tracking medical records and potential for inaccuracies in medical records.

## **KEY QUESTION 3: Do the accuracy and/or adverse effects of using self-reported measures of cervical and breast cancer screening vary by patient or measure characteristics or setting?**

We evaluated the effects of prespecified variables (publication year, sample size, study design (case-control vs cohort), setting (population- vs clinic-based), referral or screening program (whether the patients were a part of a screening or referral program vs not), survey administration method (in-person vs telephone or mail survey), timeframe of recall ( $\leq 1$  year to ever), and population (minority or low-income population vs not)) using univariable bivariate (which looks at the effect of each variable on sensitivity and specificity) and multivariable univariate meta-regression (which looks at the effect of multiple variables and accounts for any existing correlation between the effects of the significant factors) models (see Supplemental Materials for full details). Due to a lack of comprehensive reporting of specific moderators in primary studies, we did not include a variable for age in our meta-regressions, and we combined minority status and low-income status into a single variable.

### **Cervical Cancer Screening**

In the multivariable univariate model, only population (minority or low-income population vs not) was significantly associated with the accuracy of self-report for cervical cancer screening. This indicates that minority or low-income women may have lower self-report accuracy than the general population. When sensitivity and specificity were evaluated using a univariable bivariate model, several factors (study design, setting, referral or screening program, survey administration method, and population) were significantly associated with sensitivity, while no factors showed a statistically significant effect on specificity. The implications of the impact of these factors on sensitivity is unclear, as the results between studies for sensitivity were similar (tightly clustered), so minor differences resulted in statistical significance. In contrast, results for specificity were highly varied, so larger difference were necessary to reach statistical significance. Additionally, the sensitivity among studies was high (sensitivity range 94% to 97%), and the absolute differences were equal to or less than 3%, which may not be clinically significant.

### **Breast Cancer Screening**

In the multivariable univariate model, only population (minority or low-income population vs not) and sample size were significantly associated with the accuracy of self-report for breast cancer screening. Similar to cervical cancer screening, this indicates that minority or low-income

women may have lower self-report accuracy than the general population. The reason for the appearance of sample size is likely due a single study ( $N > 400,000$ ) that was larger than all other studies combined. When sensitivity and specificity were evaluated using a univariable bivariate model, several factors (setting, screening or referral program, survey administration method, and population) were significantly associated with sensitivity, while only population had a significant effect on specificity. Again, the reason more variables were found to be moderators for sensitivity and not specificity is likely due to how similar (tightly clustered) the overall results were for sensitivity.

## SUMMARY AND DISCUSSION

### Key Findings and Clinical Considerations

Concerns about the accuracy of self-report of cervical and breast cancer screening have led to the VHA's consideration of increasing the requirements for reporting documentation of cervical and breast cancer screenings. However, prior to considering changing documentation requirements, it is important to understand benefits and harms of accepting self-report, including the potential for over- or under-screening, missed diagnoses, and impacts on provider, administrative, or patient burden. Overall, among 39 studies examining the accuracy of self-report for cervical and/or breast cancer screening compared to medical records, unscreened women tended to over-report having had a mammogram or pap test, but screened women more accurately reported their screening. Among unscreened women, according to their medical record, only 48% to 61% (specificity) accurately reported not having a screening, but 96% (sensitivity) of screened women, according to their medical record, accurately reported their screening. In comparison, patients accurately reported having a colorectal cancer screening 30% to > 90% of the time (depending on the test performed).<sup>78</sup> The specificity was lower for pap test recall than for mammograms. This may be due to the fact that pap tests are often done as a part of routine gynecological exams, and patients may falsely assume that a pap test was done when it was not,<sup>48</sup> while mammograms are often separately scheduled appointments specifically for the purpose of completing a mammogram.

Overall, despite adding 11 new studies, these findings are consistent with the most recent meta-analysis,<sup>25</sup> which found over-reporting of cervical and breast cancer screenings with self-report compared to medical records. However, in contrast to Howard et al,<sup>25</sup> which did not find any statistically significant factors associated with diagnostic accuracy, we found that several factors had a significant effect on the accuracy of self-report for both cervical and breast cancer screenings. In adjusted models, studies in minority or low-income populations were more likely to have less accurate self-report compared to medical records, but absolute differences were equal to or less than 3%, which may not be clinically significant. This agrees with studies finding greater over-reporting of pap and mammogram in low-income or minority populations.<sup>49,62,66,74</sup> This greater over-reporting in minority or low-income populations may be due to lower health literacy, cultural factors, or the quality of medical record documentation in these populations. The National Assessment of Adult Literacy has consistently found lower health literacy scores among minority and low-income populations.<sup>14</sup> This may impact a patient's ability to accurately understand when specific procedures have been performed or their ability to interpret survey questions. Additionally, low-income and minority populations may rely more on public health care, which may include multiple providers from varying clinics, and may impact the ability for complete medical record review compared to a patient receiving care within the private sector.<sup>15</sup>

The potential for over-reporting of cervical and breast cancer screening may lead to delayed or missed screenings and should be considered when accepting self-report for these screening tests. The high rate of false positives (low specificity) in these studies may be due to potential conscious or unconscious biases including telescoping (remembering an event occurring more recently than it actually occurred), acquiescence bias (the tendency to respond "yes" when in doubt), and social desirability response bias (over-reporting of events that are socially desirable). Telescoping was evident in several of the included studies in which women recalled screening tests as occurring more recently than indicated by medical record.<sup>51,57,58</sup> The accuracy of self-

report may be improved by taking these potential biases into consideration when designing questionnaires and surveys for self-reported data collection. Patient recall may be improved by ensuring respondents understand the questions (*ie*, including detailed explanations of what the tests entail), phrasing survey questions to minimize social desirability bias (*ie*, asking about barriers or future plans prior to asking about past behaviors),<sup>34</sup> and having clear, exhaustive, mutually exclusive response categories. Additionally, patient recall may become more accurate as timeframes are extended (*ie*, a woman is more likely to have actually had a screening if timeframes are extended).<sup>79</sup> Expanding the timeframe in which patients are asked to recall a screening may improve self-report accuracy by allowing room for error in the self-reported timeline of screening. However, the recommended timing of screenings also needs to be considered (*ie*, if the timing of recall is expanded too far, it may be outside the recommended screening timeframe, and patients may still be at risk for under-screening).

We did not find any studies reporting clinical outcomes or adverse events or unintended consequences of accepting self-report. With over-reporting of screening by self-report, there is the potential to overestimate the success of screening interventions, mask disparities in screening prevalence, under-screen individuals, and miss cancer diagnoses. It is important to consider these potential impacts of over-reporting alongside the feasibility, cost, and administrative burden of searching medical records.

## Limitations

There are important limitations for the evidence in this review. There was significant heterogeneity among the studies for all outcomes assessed. Although we found a significant effect of studies in low-income or minority populations on diagnostic accuracy outcomes, heterogeneity remained in these subgroup analyses. This is most likely due to our inability to include all potential moderators in our meta-regression and lack of comprehensive reporting of specific moderators in primary studies. For instance, we were unable to include a variable for age due to the limited reporting in primary studies, and we combined low-income and minority status into a single variable termed “Population” because of how they were reported in primary studies. As the variable “Population” was significant in our meta-regression, to better target initiatives to improve reporting accuracy, it would be useful to know the contribution attributable specifically to low-income status and different racial/ethnic minority subgroups separately. It is possible that, if all studies reported on age and specific low-income status and different racial/ethnic minority subgroups minority status, these variables may have accounted for a majority of the heterogeneity observed in our results.

Common methodological limitations of the studies included lack of or unclear reporting of blinding of the medical record and/or self-report data collection and incomplete medical record review. Lack of blinding of patient interviewers or medical record data collectors may influence survey responses or the thoroughness of medical chart review. Incomplete medical record review is likely in population-based surveys or in clinic settings where patients are likely to seek care outside of a single clinic. It can be challenging to identify all possible locations where a test may have been performed, and this may lead to studies reporting higher false positives (a patient reports a test was done, but it cannot be confirmed in the medical record) and reduce the overall specificity of self-report found in the study. Studies varied in the extent of ensuring complete medical record review by combining medical record data with other administrative data, utilizing cytology records, or searching multiple hospitals or clinics for records of received screenings.

Limitations of our review methods include our literature search and our use of sequential instead of independent dual assessment. For our literature search, we limited the timeframe and used existing systematic reviews to identify earlier studies, which may have led to missed studies. Additionally, although widely used, sequential dual review has not been empirically compared to independent dual review and may increase the risk of error and bias.

## Future Research

Overall, we identified a large number of studies assessing the accuracy of self-report compared to medical record for cervical and breast cancer screening, and our findings are consistent with the most recent meta-analysis.<sup>25</sup> Further studies will likely not change the overall findings, but studies with improved reporting of variables, such as age and low-income or minority status, may improve the ability to assess reasons for heterogeneity in these findings. However, perhaps most importantly, no studies reported clinical outcomes or potential adverse events or unintended consequences. Future research should focus on assessing the impact of accepting self-report on clinical (missed diagnoses, under- or over-screening, mortality, *etc*) and system-level (provider, patient, or administrative burden) outcomes. Ideally, studies would directly compare the use of self-report versus provider documentation on these outcomes. However, directly assessing clinical outcomes may not be feasible due to the necessary lengthy follow-up to identify cancer outcomes and patients' fragmented health care use. Even so, it would be possible to track system-level outcomes to weigh the potential harms of accepting self-report against the time and resource burden of tracking medical records.

We did not identify any studies in Veterans, although the studies covered a wide range of populations and may be applicable to Veterans. The self-reported prevalence of screening in the VHA is 85% for breast cancer<sup>80</sup> and 93% for cervical cancer.<sup>81</sup> These rates are similar to the overall self-reported screening rates in the study population for breast (87%) and cervical cancer (84%). Overall, both of these self-reported screening rates are higher than those reported nationally (64% breast cancer<sup>9</sup> and 70% cervical cancer).<sup>10</sup> The population of women Veterans is growing rapidly, having more than doubled since 2000,<sup>82</sup> and more women Veterans are accessing care through the VHA.<sup>83</sup> Studies within the VHA would be valuable to understand the system-level factors associated with accepting self-report within a large, integrated health system. Although integrated medical records within a large health care system may ease provider documentation, it may be challenging to locate and verify screening records outside of the health care system without specific systems or methods in place.

## CONCLUSIONS

Unscreened women tend to over-report having had a mammogram or pap test, but screened women more accurately report their screening. There was significant heterogeneity among the studies, and questions remain regarding the impact of variables such as age and specific low-income and/or minority status on this heterogeneity. Future research should focus on assessing the impact of over-reporting on cancer-related and system-level outcomes, including overestimation of the success of screening interventions, masking disparities in screening prevalence, under-screening of individuals, missed cancer diagnoses, and patient or provider burden.

## ACKNOWLEDGMENTS

This topic was developed in response to a nomination by the VHA Performance Workgroup, for the purpose of informing collection and use of self-report measures in cervical and breast cancer screening in the VHA. The scope was further developed with input from the topic nominators (*ie*, Operational Partners) and the ESP Coordinating Center review team.

The authors gratefully acknowledge Emilie Chen, BA for editorial support and the following individuals for their contributions to this project:

### Operational Partners

Operational partners are system-level stakeholders who have requested the report to inform decision-making. They assure VA relevance; help develop and approve final project scope and timeframe for completion; provide feedback on draft report; and provide consultation on strategies for dissemination of the report to field and relevant groups.

#### **Jane Kim, MD, MPH**

Chief Consultant for Preventive Medicine  
National Center for Health Promotion and Disease Prevention

#### **Joe Francis, MD, MPH**

Chair, VHA Performance Workgroup  
Director, Clinical Analytics and Reporting

#### **Kathleen Pittman, RN, MPH**

National Program Manager for Health Promotion and Disease Prevention Programs  
National Center for Health Promotion and Disease Prevention

#### **Lori Hoffman-Hogg, MS, RN, CNS, AOCN**

National Program Manager for Prevention Policy  
National Center for Health Promotion and Disease Prevention

#### **Sally Haskell, MD**

Deputy Chief Consultant  
Women's Health Services

### Peer Reviewers

The Coordinating Center sought input from external peer reviewers to review the draft report and provide feedback on the objectives, scope, methods used, perception of bias, and omitted evidence. Peer reviewers must disclose any relevant financial or non-financial conflicts of interest. Because of their unique clinical or content expertise, individuals with potential conflicts may be retained. The Coordinating Center and the ESP Center work to balance, manage, or mitigate any potential nonfinancial conflicts of interest identified.

## REFERENCES

1. American Cancer Society. The health care law: How it helps people with cancer and their families. 2018; <https://www.cancer.org/treatment/finding-and-paying-for-treatment/understanding-health-insurance/health-insurance-laws/the-health-care-law.html>. Accessed December 30, 2018.
2. US Centers for Medicare & Medicaid Services. Preventive care benefits for women. <https://www.healthcare.gov/preventive-care-women/>. Accessed December 30, 2018.
3. Centers for Disease Control and Prevention. Cancer prevention and control. 2018; <https://www.cdc.gov/cancer/dcpc/prevention/screening.htm>. Accessed December 30, 2018.
4. National Cancer Institute. Screening tests: Screening tests that have been shown to reduce cancer deaths. 2015; <https://www.cancer.gov/about-cancer/screening/screening-tests>. Accessed December 30, 2018.
5. Nelson H, Fu R, Cantor A, Pappas M, Daeges M, Humphrey L. Effectiveness of breast cancer screening: Systematic review and meta-analysis to update the 2009 US preventive services task force recommendation. *Annals of Internal Medicine*. 2016;164(4).
6. US Department of Veterans Affairs. National center for health promotion and disease prevention: Screening for cervical cancer. 2016; [https://www.prevention.va.gov/Preventing\\_Diseases/Screening\\_for\\_Cervical\\_Cancer.asp](https://www.prevention.va.gov/Preventing_Diseases/Screening_for_Cervical_Cancer.asp). Accessed December 30, 2018.
7. US Department of Veterans Affairs. VA announces adoption of American Cancer Society mammogram screening guidelines. 2017; <https://www.va.gov/opa/pressrel/pressrelease.cfm?id=2899>. Accessed December 30, 2018.
8. US Department of Health and Human Services, Office of Disease Prevention and Health Promotion. HealthyPeople 2020 topics & objectives: Cancer. 2018; <https://www.healthypeople.gov/2020/topics-objectives/topic/cancer/objectives>. Accessed December 30, 2018.
9. Centers for Disease Control and Prevention. National Center for Health Statistics: Mammography use. 2017; <https://www.cdc.gov/nchs/fastats/mammography.htm>. Accessed December 30, 2018.
10. Centers for Disease Control and Prevention. National Center for Health Statistics: Pap tests. 2017; <https://www.cdc.gov/nchs/fastats/pap-tests.htm>. Accessed December 30, 2018.
11. National Center for Health Statistics. *Health, United States, 2015, with special feature on racial and ethnic health disparities*. Hyattsville, MD: National Center for Health Statistics (US);2016.
12. Zaslavsky AM, Hochhiemer J, Schneider E, et al. Impact of sociodemographic case mix on the HEDIS measures of health plan quality. *Medical Care*. 2000;38(10).
13. Virnig B, Lurie N, Huang Z, Musgrave D, McBean A, Dowd B. Racial variation in quality of care among medicare+choice enrollees. *Health Affairs*. 2002;21(6).
14. Rikard R, Thompson M, McKinney J, Beauchamp A. Examining health literacy disparities in the United States: A third look at the national assessment of adult literacy (NAAL). *BMC Public Health*. 2016;16(1).
15. 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.

16. Agide F, Sadeghi R, Garmaroudi G, Tigabu B. A systematic review of health promotion interventions to increase breast cancer screening uptake: From the last 12 years. *European Journal of Public Health*. 2018;28(6).
17. Rees I, Jones D, Chen H, Macleod U. Interventions to improve the uptake of cervical cancer screening among lower socioeconomic groups: A systematic review. *Preventive Medicine*. 2018;111:323-335.
18. National Committee for Quality Assurance. Breast cancer screening (BCS): The HEDIS measure. 2018; <https://www.ncqa.org/hedis/measures/breast-cancer-screening/>. Accessed December 30, 2018.
19. National Committee for Quality Assurance. Cervical cancer screening (CCS). 2018; <https://www.ncqa.org/hedis/measures/cervical-cancer-screening/>. Accessed December 30, 2018.
20. National Quality Forum. Qps tool. 2018; <http://www.qualityforum.org/QPS/QPSTool.aspx>. Accessed December 30, 2018.
21. Centers for Disease Control and Prevention. Behavioral risk factor surveillance system. 2018; <https://www.cdc.gov/brfss/index.html>. Accessed December, 2018.
22. National Committee for Quality Assurance. 2019 HEDIS summary table of measures, product lines and changes. 2018; [https://www.ncqa.org/wp-content/uploads/2018/08/20190000\\_HEDIS\\_Measures\\_SummaryofChanges.pdf](https://www.ncqa.org/wp-content/uploads/2018/08/20190000_HEDIS_Measures_SummaryofChanges.pdf). Accessed December 30, 2018.
23. Centers for Disease Control and Prevention. National Center for Health Statistics. 2019; <https://www.cdc.gov/nchs/>. Accessed January 17, 2019.
24. 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.
25. 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.
26. 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. *American Journal of Epidemiology*. 1997;148(11).
27. Barsky AJ. Forgetting, fabricating, and telescoping: The instability of the medical history. *Archives of Internal Medicine*. 2002;162(9):981-984.
28. Hurd M. *Anchoring and acquiescence bias in measuring assets in household surveys*. National Institute on Aging; 1999.
29. Ham SA, Ainsworth BE. Disparities in data on Healthy People 2010 physical activity objectives collected by accelerometry and self-report. *American Journal of Public Health*. 2010;100 Suppl 1:S263-268.
30. Vernon SW, Abotchie PN, McQueen A, White A, Eberth JM, Coan SP. Is the accuracy of self-reported colorectal cancer screening associated with social desirability? *Cancer Epidemiology, Biomarkers & Prevention*. 2012;21(1):61-65.
31. West R, Zatonski W, Przewozniak K, Jarvis MJ. Can we trust national smoking prevalence figures? Discrepancies between biochemically assessed and self-reported smoking rates in three countries. *Cancer Epidemiology, Biomarkers & Prevention*. 2007;16(4):820-822.
32. Furnham A. Response bias, social desirability and dissimulation. *Personality and Individual Differences*. 1986;7(3):385-400.

33. Armstrong K, Long JA, Shea JA. Measuring adherence to mammography screening recommendations among low-income women. *Preventive Medicine*. 2004;38(6):754-760.
34. 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.
35. National Committee for Quality Assurance. HEDIS and performance measurement. 2018; <https://www.ncqa.org/hedis/>. Accessed December 30, 2018.
36. US Department of Veterans Affairs. VHA National Center for Health Promotion and Disease Prevention, colorectal cancer screening. 2018; [http://vaww.prevention.va.gov/CPS/Colorectal\\_Cancer\\_Screening.asp](http://vaww.prevention.va.gov/CPS/Colorectal_Cancer_Screening.asp). Accessed December 30, 2018.
37. Whiting P, Savović J, Higgins JP, et al. Robis: A new tool to assess risk of bias in systematic reviews was developed. *Journal of Clinical Epidemiology*. 2016;69:225-234.
38. Whiting PF, Rutjes AW, Westwood ME, et al. Quadas-2: A revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine*. 2011;155(8):529-536.
39. Berkman ND, Lohr KN, Ansari M, et al. *Grading the strength of a body of evidence when assessing health care interventions for the effective health care program of the agency for healthcare research and quality: An update methods guide for effectiveness and comparative effectiveness reviews*. Rockville, MD; 2013.
40. Dwamena B. Midas: Stata module for meta-analytical integration of diagnostic test accuracy studies. *Statistical Software Components S456880*. 2009; Boston College Department of Economics.
41. 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.
42. 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.
43. 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.
44. Fruchter RG, Roness 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.
45. 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.
46. 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.
47. 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.
48. 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.
49. 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.
50. 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.

51. 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.
52. 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.
53. 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.
54. 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.
55. 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.
56. 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.
57. 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.
58. 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.
59. 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 & Developmental Disabilities*. 2013;118(4):327-336.
60. 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.
61. 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.
62. 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.
63. 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.
64. Brown BJ, Adams EM. Patients as reliable reporters of medical care process: Recall of ambulatory encounter events. *Medical Care*. 1992;30(5):400-411.
65. Caplan (a) 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.
66. 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.
67. 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.

68. Crane LA, Kaplan CP, Bastani R, Scrimshaw SCM. Determinants of adherence among health department patients referred for a mammogram. *Women & Health*. 1996;24(2):43-64.
69. 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.
70. Fulton-Kehoe D, Burg MA, Lane DS. Are self-reported dates of mammograms accurate? *Public Health Reviews*. 1992;20(3-4):233.
71. 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.
72. King ES, Rimer BK, Trock B, Balslem A, Engstrom P. How valid are mammography self-reports? *American Journal of Public Health*. 1990;80(11):1386.
73. 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.
74. Mahnken JD, Freeman DH, DiNuzzo AR, Freeman JL. Mammography use among older Mexican-American women: Correcting for over-reports of breast cancer screening. *Women & Health*. 2007;45(3):53-64.
75. 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.
76. 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.
77. 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.
78. Pignone M, Rich M, Teutsch SM, Berg AO, Lohr KN. Screening for colorectal cancer in adults at average risk: A summary of the evidence for the U.S. Preventive Services Task Force. *Annals of Internal Medicine*. 2002;137(2):132-141.
79. Bowman JA, Sanson-Fisher R, Redman S. The accuracy of self-reported pap smear utilisation. *Social Science & Medicine*. 1997;44(7):969-976.
80. US Department of Veterans Affairs. Quality of care: Prevention - breast cancer screening. <https://www.va.gov/QUALITYOFCARE/initiatives/compare/prevention-breast-cancer-screening.asp>. Accessed February 4, 2019.
81. US Department of Veterans Affairs. Quality of care: Prevention - cervical cancer screening. <https://www.va.gov/QUALITYOFCARE/initiatives/compare/prevention-cervical-cancer-screening.asp>. Accessed February 4, 2019.
82. US Department of Veterans Affairs. Women veterans health care: Fact sheet. 2011; [https://www.womenshealth.va.gov/WOMENSHEALTH/docs/Women\\_Veterans\\_Health\\_Care\\_factsheet.pdf](https://www.womenshealth.va.gov/WOMENSHEALTH/docs/Women_Veterans_Health_Care_factsheet.pdf). Accessed January 17, 2019.
83. US Department of Veterans Affairs. Increase in VA health care use by women veterans of OEF/OIF/OND. 2015; <https://www.publichealth.va.gov/epidemiology/reports/oefoifond/health-care-utilization/women-veterans.asp>. Accessed January 17, 2019.