| **Author, year** | **Study design** | **Age, *years* (mean or** **%); setting; population** | **Comparisons** | **Measures** | **Outcomes** | **Quality rating; limitations** |
| --- | --- | --- | --- | --- | --- | --- |
| Bredal et al, 2013207 | Before-after study | 57.7; women recalled in a screening program in Oslo, Norway | FP (n=560) and TP (n=80) at recall vs. 4 weeks later  | HADS (score ≥11)  | **Recall vs. 4 weeks later:** anxiety (% cases): FP 15% vs. 5.5% (NS), TP 19% vs. 16.7% (NS); depression (% cases): FP 1.4% vs. 1.3% (NS), TP 1.3% vs. 6.9% (NS). Factors predicting anxiety or depression in multivariate models: low general life expectations, previous history of anxiety and/or depression, anxiety at baseline, satisfaction with information (predicts depression only). | NA\* |
| Brodersen and Siersma, 2013203 | Nested case-control | 28% 50-54, 32% 55-59, 23% 60-64, 17% ≥65; women in screening programs in Copenhagen and Funen, Denmark; cases=recalled; controls=normal results in the same clinic and day as cases | FP (n=272) vs. Normal screen (n=864) vs. TP (n=174) | COS-BC | **After screening mammography:** Normal screen vs. FP and TP had significantly better scores on subscales for sense of dejection, anxiety, negative impact on behavior, sleep, or sexuality, breast examination, and on single items of feeling less attractive and keeping mind off things (p<0.001 for all outcomes); no differences between FP and TP on any subscales.**3 year followup:** TP vs. FP and Normal screen had significantly worse scores on subscales of sense of dejection, anxiety, negative impact on behavior, sleep, or sexuality, social network, and on single items of feeling less attractive and keeping mind off things (p<0.001 for all outcomes) and additional differences vs. Normal screen on subscales of inner calm, social networking, and existential values (p<0.001 for all outcomes); FP vs. Normal screen had significantly worse scores on subscales for sense of dejection, anxiety, negative impact on behavior, sleep, or sexuality, breast examination, inner calm, social network, existential values, and on single items of feeling less attractive and keeping mind off things (p<0.05).  | Good |
| Espasa et al, 2012206 | Case-control | 55% 50-59, 45% 60-69; women in screening program in Spain; cases=FP; controls=TN matched on age, education, marital and working status, and previous mammograms | FP (n=100) vs. Normal screen (n=50) | HADS, structured interview | **After 22 days of followup:** FP vs. Normal screen worried about having breast cancer (49% vs. 10%, p<0.0001) and had worries that affected mood or daily activities (31% vs. 2%, p<0.0001); but no differences in anxiety (11% vs. 14%, p=0.83) or depression (2% vs. 2%). | Fair; enrolled selected group of women; 2:1 ratio of cases to controls; did not control for confounders |
| Fitzpatrick et al, 2011200 | Retrospective cohort | Mean age: NR, range: 50-62; women screened through the National Breast Screening Programme in Ireland | FP (n=9,746) vs. Normal screen (n=148,589) | Re-attendance | **Rate of re-attendance:** 90.7% vs. 89.0%, p<0.001; age group 50-54 years: 91.0% vs. 89.6%, p<0.001; age group 55-59 years: 90.4% vs. 88.7%, p<0.001; age group 60-62 years: 90.4% vs. 87.4%, p<0.001**Adjusted OR of predictors of re-attendance (95% CI):** 0.8 (0.7 to 0.9) for age group 55-59 years and 0.8 (0.6 to 0.9) for age group 60-62 years vs. age group 50-54; 1.8 (1.5 to 2.2) for subsequent screen vs. initial screen; 0.9 (0.8 to 1.1) for core biopsy and 0.4 (0.3 to 0.6) for open benign biopsy vs. no tissue sampling; 0.997 (0.994 to 0.999) for every additional day from recall to assessment to non-malignant diagnosis | Fair, unclear if random or consecutive sample; baseline data not provided; did not control for confounders |
| Gibson et al, 2009198 | Prospective cohort | 6% <50, 32% 50-59, 34% 60-69, 22% 70-79, 6% ≥80; women registered in the New Hampshire Mammography Network and the New Hampshire Women for Health study | FP (n=2,107) vs. Normal screen (n=11,384) reference group | WHQ | **OR for depression (95% CI)**: overall FP 0.96 (0.72 to 1.28); white FP 0.84 (0.62 to 1.15); non-white FP 3.23 (1.32 to 7.91). | Fair; unclear how women were selected; baseline data not provided for groups of interest; outcomes self-reported  |
| Hafslund et al, 2012205 | Nested case-control | 57 (SD 5.8) for FP vs. 58 (SD 5.5) for TN; women from Hordaland, Sogn, and Fjordane Counties, Norway; cases=FP; controls=TN | FP (n=128) vs. Normal screen (n=195) | SF-36, HADS | **6 months followup:** FP vs. Normal screen clinical anxiety (mean HADS-A) 4.1 vs. 4.0, p=0.81; clinical depression (mean HADS-D) 3.2 vs. 2.4, p=0.045; mental function (mean SF-36) 80.6 vs. 85.0; p=0.03; vitality (mean SF-36) 70.3 vs. 77.0; p=0.02. | Fair; enrolled selected group of women; higher response rate in control group |
| Keyzer-Dekker, 2012199 | Prospective cohort | 50 (SD 0.8) for 1st screen recalls vs. 61 (SD 5.9) for repeat screen recalls, p<0.001; women with abnormal results referred to hospitals in The Netherlands | 1st screen recalls (n=186) vs. repeat screen recalls (n=296) | STAI, NEO-FFI, CES-D, WHOQOL | **After recall before diagnosis:** anxiety (mean STAI) 13.3 vs. 12.8, p=0.209; depression (mean CES-D) 8.9 vs. 9.0, p=0.836).**6 month followup:** anxiety (mean STAI estimated from graph) 10.6 vs. 10.3, p<0.001 for change over time for both groups; depression p<0.001 for change over time for both groups (data not shown), with no differences between groups. | Fair; older women in repeat screen group; outcomes were self-reported; did not report attrition |
| Klompen-houwer et al, 2014202 | Retrospective cohort | Mean age: NR, range: 50-75; women being screened in one of the specialized screening units in The Netherlands | Normal screen (n=373,474) vs. 1st screen recalls (n=6,672) vs. repeat screen recalls for different lesion (n=161) vs. repeat screen recalls for same lesion (n=89) | Re-attendance rates | **Rate of re-attendance:** 93.2% (95% CI, 93.1% to 93.3%) vs. 65.4% (95% CI, 64.0% to 66.8%) vs. 56.7% (95% CI, 47.1% to 66.4%) vs. 44.3% (95% CI, 31.4% to 57.1%); and 52.1% (95% CI, 44.4% to 59.8%) for all recalled groups combined | Fair, baseline data not provided for groups; did not control for confounders |
| Maxwell et al, 2013201 | Retrospective cohort | Mean age: NR, range: 49-66; women screened at 1 of 5 breast screening programs in the United Kingdom | FP (n-9,367) vs. Normal screen (n=243,650)andPrevalent screen (n=54,716) vs. incident screen (n=198,301) | Re-attendance rates | **Rate of re-attendance:** 87.7% of prevalent FP screen vs. 86.0% of prevalent normal screen, difference of 1.61% (95% CI, 0.54% to 2.62%); 92.0% of incident FP vs. 92.4% of incident normal screen, difference of -0.04% (95% CI, -1.18% to 0.31%); 86.2% of all prevalent screens vs. 92.4% of all incident screens**OR (95% CI) of re-attendance after additional procedures (reference is normal screen):** needle sampling only after prevalent screen 1.06 (0.90 to 1.24); needle sampling only after incident screen 0.88 (0.84 to 0.92); open biopsy after prevalent screen 0.64 (0.31 to 1.33); open biopsy after incident screen 0.40 (0.25 to 0.66); no tissue sampling after prevalent screen 1.20 (1.10 to 1.30); no tissue sampling after incident screen 1.00 (0.91 to 1.09)**OR (95% CI) of re-attendance by age:** 0.89 (0.86 to 0.93) for older age at prevalent screen with a reduction in the odds of re-attendance of 11% for each year’s increase in a women’s age; 0.99 (0.98 to 0.99) for older age at incident screen with a reduction in the odds of re-attendance of 1% for each year’s increase in a women’s age | Fair, baseline data not provided for groups; did not control for confounders |
| Tosteson et al, 2014202 | Nested case-control | 41% <50, 45% 50-64, 14% ≥65 years; women participating in DMIST in the United States; cases=FP; controls=TN matched by institution and age | FP (n=494) vs. Normal screen (n=534) | STAI, EuroQOL EQ-5D | **After mammography:** FP vs. Normalscreen anxiety (mean STAI) 35 vs. 33, p=NR; QOL (mean EQ-5D) 0.90 vs. 0.90, p=NR.**1 year followup:** FP anxiety (STAI mean difference) -1.53 (SD 13.14), p=0.01; QOL (EQ-5D mean difference) 0.001 (SD 0.13), p=0.13); Normal screen anxiety and QOL did not change over time. | Good |

\*Quality rating criteria not available for this study design.

**Abbreviations:** CES-D=Center for Epidemiological Studies-Depression Scale; CI=confidence interval; COS-BC=Consequences of Screening in Breast Cancer; DMIST=Digital Mammographic Imaging Screening Trial; FP=false-positive; HADS=Hospital Depression and Anxiety Scale; HADS-A=HADS-Anxiety Subscale; HADS-D=HADS-Depression Subscale; n=number; NA=not available; NEO-FFI=Neuroticism-Extraversion-Openness-Five Factor Inventory; NR=not reported; NS=not significant; OR=odds ratio; QOL=quality of life; SD=standard deviation; SF-36=Short-form 36 Health Survey; STAI=Spielberger State-Trait Anxiety Inventory; TP=true positive; vs.=versus; WHOQOL=World Health Organization Quality of Life Assessment Instrument; WHQ=Women's Health Questionnaire.