Recreational physical activity is associated with reduced breast cancer risk in adult women at high risk for breast cancer: a cohort study of women selected for familial and genetic risk

Running Title: Physical activity and breast cancer risk

Rebecca D Kehm, Jeanine M Genkinger, Robert J MacInnis, Esther M John, Kelly-Anne Phillips, Gillian S Dite, Roger L Milne, Nur Zeinomar, Yuyan Liao, Julia A Knight, Melissa C Southey, Wendy K Chung, Graham G Giles, Sue-Anne McLachlan, Kristen D Whitaker, Michael Friedlander, Prue C Weideman, Gord Glendon, Stephanie Nesci, kConFab Investigators, Irene L Andris, Saundra S Buys, Mary B Daly, John L Hopper, Mary Beth Terry

1. Department of Epidemiology, Mailman School of Public Health, Columbia University, USA; rk2967@cumc.columbia.edu

2. Department of Epidemiology, Mailman School of Public Health, Columbia University, USA; Herbert Irving Comprehensive Cancer Center, Columbia University Medical Center, USA; jg3081@cumc.columbia.edu

3. Centre for Epidemiology and Biostatistics, The University of Melbourne, Australia; Cancer Epidemiology Division, Cancer Council Victoria, Australia; Robert.MacInnis@cancervic.org.au

4. Department of Medicine and Stanford Cancer Institute, Stanford University School of Medicine, USA; emjohn@stanford.edu

5. Centre for Epidemiology and Biostatistics, The University of Melbourne, Australia; Division of Cancer Medicine, Peter MacCallum Cancer Centre, Australia; Sir Peter MacCallum
Department of Oncology, The University of Melbourne, Australia;
Kelly.Phillips@petermac.org

6. Centre for Epidemiology and Biostatistics, The University of Melbourne, Australia;
g.dite@unimelb.edu.au

7. Centre for Epidemiology and Biostatistics, The University of Melbourne, Australia; Cancer Epidemiology Division, Cancer Council Victoria, Australia; Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Australia;
Roger.Milne@cancervic.org.au

8. Department of Epidemiology, Mailman School of Public Health, Columbia University, USA;
nz2255@cumc.columbia.edu

9. Department of Epidemiology, Mailman School of Public Health, Columbia University, USA;
yl2265@cumc.columbia.edu

10. Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Canada; Dalla Lana School of Public Health, University of Toronto, Canada; knight@lunenfeld.ca

11. Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Australia; Cancer Epidemiology Division, Cancer Council Victoria, Australia; Department of Clinical Pathology, The University of Melbourne, Australia; msouthey@unimelb.edu.au

12. Department of Pediatrics and Medicine, Vagelos College of Physicians and Surgeons, Columbia University, USA; Herbert Irving Comprehensive Cancer Center, Columbia University Medical Center, USA; wkc15@cumc.columbia.edu

13. Centre for Epidemiology and Biostatistics, The University of Melbourne, Australia; Cancer Epidemiology Division, Cancer Council Victoria, Australia; Precision Medicine, School of
Clinical Sciences at Monash Health, Monash University, Australia;

Graham.Giles@cancervic.org.au

14. Department of Medicine, St Vincent’s Hospital, The University of Melbourne, Australia;
Department of Medical Oncology, St Vincent’s Hospital, Australia; Sue-
Anne.MCLACHLAN@svha.org.au

15. Department of Clinical Genetics, Fox Chase Cancer Center, USA; Kristen.Whitaker@fccc.edu

16. Prince of Wales Clinical School, University of New South Wales, Australia; Department of Medical Oncology, Prince of Wales Hospital, Australia; m.friedlander@unsw.edu.au

17. Centre for Epidemiology and Biostatistics, The University of Melbourne, Australia;
prue.weideman@unimelb.edu.au

18. Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Canada;
Gord.Glendon@uhnresearch.ca

19. Department of Medical Oncology, Peter MacCallum Cancer Centre, Australia;
Stephanie.Nesci@petermac.org

20. Sir Peter MacCallum Department of Oncology, The University of Melbourne, Australia; The Research Department, The Peter MacCallum Cancer Centre, Australia;
heather.thorne@petermac.org

21. Lunenfeld-Tanenbaum Research Institute, Sinai Health System, Canada; Departments of Molecular Genetics and Laboratory Medicine and Pathobiology, University of Toronto, Canada; aNDRULiS@lunenfeld.ca

22. Department of Medicine and Huntsman Cancer Institute, University of Utah Health Sciences Center, USA; saundra.buys@hsc.utah.edu
23. Department of Clinical Genetics, Fox Chase Cancer Center, USA; mary.daly@fccc.edu

24. Centre for Epidemiology and Biostatistics, The University of Melbourne, Australia;

j.hopper@unimelb.edu.au

25. Department of Epidemiology, Mailman School of Public Health, Columbia University, USA;
Herbert Irving Comprehensive Cancer Center, Columbia University Medical Center, USA;

mt146@columbia.edu

Address correspondence to:

Mary Beth Terry, PhD
Professor of Epidemiology,
Mailman School of Public Health, Columbia University
722 W. 168th Street, Room 1611
New York, NY 10032
Tel: 212-305-4915
Email: mt146@columbia.edu

Financial Support: The six sites of the Breast Cancer Family Registry were supported by grant U01 CA164920 from the USA National Cancer Institute. This work was also supported by grants to kConFab and the kConFab Follow-Up Study from Cancer Australia [grant numbers 809195, 1100868], the Australian National Breast Cancer Foundation [grant number IF 17 kConFab], the National Health and Medical Research Council [grant numbers 454508, 288704, 145684], the National Institute of Health U.S.A. [grant number 1RO1CA159868], the Queensland Cancer
Fund, the Cancer Councils of New South Wales, Victoria, Tasmania and South Australia, and the Cancer Foundation of Western Australia [grant numbers not applicable]. RDK is supported by the National Institutes of Health, National Cancer Institute, Cancer Epidemiology Training Grant [grant number T32-CA009529]. KAP is an Australian National Breast Cancer Foundation Practitioner Fellow.

Conflicts of interest: The authors declare no conflict of interest.

Word Count: 4,367

Total Number of Figures and Tables: 2 tables and 4 figures (6 total)
Abstract

While physical activity is associated with lower breast cancer risk for average-risk women, it is not known if this association applies to women at high familial/genetic risk. We examined the association of recreational physical activity (self-reported by questionnaire) with breast cancer risk using the Prospective Family Study Cohort (ProF-SC), which is enriched with women who have a breast cancer family history (N=15,550). We examined associations of adult and adolescent recreational physical activity (quintiles of age-adjusted total metabolic equivalents (METs) per week) with breast cancer risk using multivariable Cox proportional hazards regression, adjusted for demographics, lifestyle factors, and body mass index. We tested for multiplicative interactions of physical activity with predicted absolute breast cancer familial risk based on pedigree data and with BRCA1 and BRCA2 mutation status. Baseline recreational physical activity level in the highest 4 quintiles compared with the lowest quintile was associated with a 20% lower breast cancer risk (HR=0.80, 95% CI=0.68, 0.93). The association was not modified by familial risk or BRCA mutation status (p-interactions>0.05). No overall association was found for adolescent recreational physical activity. Recreational physical activity in adulthood may lower breast cancer risk for women across the spectrum of familial risk.

Significance

Findings suggest that physical activity might reduce breast cancer risk by about 20% for women across the risk continuum, including women at higher-than-average risk due to their family history or genetic susceptibility.
Introduction

Women at higher than average risk of breast cancer because of their family history or underlying genetic susceptibility often enquire about non-surgical strategies to reduce their breast cancer risk, such as physical activity and other lifestyle modifications (1). Findings from meta-analyses suggest that physical activity is associated with a breast cancer risk reduction of about 20% when comparing the most and least physically active women in a given study sample (2-5). Informed by these findings, the World Cancer Research Fund has concluded that physical activity probably protects against breast cancer (6). However, most of the epidemiologic evidence on physical activity and breast cancer risk comes from studies that were conducted using samples that were unselected for familial or genetic risk of breast cancer, and thus underpowered to detect associations for women in the highest tail of the absolute risk distribution (7). It has yet to be established if physical activity is associated with breast cancer risk for women across the spectrum of absolute breast cancer risk, in particular for women at high familial or genetic risk.

Four studies have examined the association of physical activity with breast cancer risk for women with a *BRCA1* or *BRCA2* mutation, but with mixed results (8-11). These studies were all retrospective, limited by numbers of events (N=89-443), and combined *BRCA1* and *BRCA2* mutation carriers in analyses; no study has examined the association of physical activity with breast cancer risk separately for *BRCA1* and *BRCA2* mutation carriers. Other studies have considered whether the association of physical activity with breast cancer risk is modified by having a family history of breast cancer, although usually as a secondary aim and using samples unselected for familial or genetic risk (2, 12). Most of these studies used a categorical construct...
of family history (generally defined as yes/no based on breast cancer in any first-degree relatives), which discounts the fact that there is a strong gradient in underlying familial risk based on factors including number of affected relatives and their age(s) at diagnosis (7, 13). The Sister Study, a cohort of over 50,000 women with a sister affected with breast cancer, recently used a more comprehensive definition of family history characterized by a Bayesian score that incorporated characteristics of the family structure, and found that the association of recreational physical activity with breast cancer risk (≥7 vs. 1hr/wk: hazard ratio (HR) = 0.77, 95% confidence interval (CI) = 0.66, 0.90) was not modified by family history (14). However, this study did not consider associations by BRCA1 or BRCA2 mutation status, and 66% of the cohort was postmenopausal at baseline (15). Therefore, further research on the association between physical activity and breast cancer risk is warranted, using a younger cohort that includes more high risk women.

The Prospective Family Study Cohort (ProF-SC) is a large international family cohort that is enriched with women who have a family history of breast cancer (N=18,853), including over 1,200 women with a known BRCA1 or BRCA2 mutation. More than half of the cohort was premenopausal at baseline. We used data from ProF-SC to examine the association of recreational physical activity, at baseline and during adolescence, with breast cancer risk. We examined whether associations were modified by underlying familial risk (based on multigenerational pedigree data) or by genetic risk (BRCA1 or BRCA2 mutation), thus generating new insight into the association of physical activity with breast cancer risk for women across the absolute risk continuum.

Materials and Methods
Study Sample

ProF-SC comprises baseline and follow-up data from the Breast Cancer Family Registry (BCFR), a collaboration of six breast cancer family studies from the United States, Canada, and Australia (16), and the Kathleen Cuningham Foundation Consortium for Research into Familial Breast Cancer (kConFab) Follow-Up Study, an Australian and New Zealand breast cancer family study (17, 18); additional details are available elsewhere (7). Participants and their family members were followed prospectively for up to 20 years, accumulating cancer and other health outcomes. The BCFR and kConFab were approved by the institutional review board at each participating study center; all participants provided written informed consent.

For this study, we included women who were enrolled before 30 June 2011 and unaffected with breast cancer (N=18,854). To be eligible, women had to be aged 18–79 years at baseline (excludes N=528) because breast cancer familial risk scores were only calculated up until age 80. Women also had to not have had a bilateral risk-reducing mastectomy at baseline (excludes N=108), and to have had at least two months of follow-up (excludes N=445). We excluded 247 women without sufficient pedigree data to allow calculation of breast cancer familial risk scores, 14 women with missing date of breast cancer diagnosis, and 1,962 women with incomplete data on recreational physical activity at baseline. The final sample for analysis was 15,550 women from 6,503 families, including 659 BRCA1 and 526 BRCA2 mutation carriers; 59% of the sample was premenopausal at baseline.

Baseline Data
The BCFR and kConFab used the same baseline questionnaire to collect information on demographics, education, height and weight, menstrual and reproductive history, and lifestyle factors, including recreational physical activity. Participants also completed a questionnaire on personal and family history of breast and other cancers, including cancers in first- and second-degree relatives. We used multi-generational pedigree data to estimate each participant’s absolute 1-year risk (from baseline) and lifetime familial risk to age 80 years (from birth) of invasive breast cancer using the Breast Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA) risk prediction model, which was developed in 2004 (19) and updated in 2008 (20) (we used the 2008 model in this study).

At enrollment, women reported by questionnaire their average hours per week of moderate (e.g., brisk walking) and strenuous (e.g., running) recreational physical activity during the previous three years (hereafter referred to as baseline recreational physical activity), as well as during adolescence (12–17 years of age). Response options included none, ½ hour, 1 hour, 1½ hours, 2 hours, 3 hours, 4–6 hours, 7–10 hours, or ≥11 hours. We converted hours per week of moderate and strenuous recreational physical activity to total metabolic equivalents (METs) per week (1 hour moderate=4 METs; 1 hour strenuous=7 METs) (21); the midpoint was used if a range of hours per week was reported (e.g., 4–6 hours moderate converted to 5 hours=20 METs) and ≥11 was coded as 11 hours per week.

**Outcome**

Information on diagnoses of incident breast cancer was obtained through self- or relative-report and confirmed pathologically for 81% of women (pathology reports were not
obtained for 19% of the sample). We calculated person-years from two months after completion of the baseline questionnaire (to ensure that prevalent cases were excluded) to age at first breast cancer diagnosis (100 cases were pathologically confirmed as ductal carcinoma *in situ*). Women were censored at the earliest of the following events: bilateral risk-reducing mastectomy, age 80 years, loss to follow-up, or death.

**Statistical Analysis**

Baseline age of women in our sample ranged from 18-79 years (mean=46.3 years; standard deviation (SD)=14.9), and baseline age was negatively correlated with both baseline (ρ= -0.20, p<0.001) and adolescent (ρ= -0.12, p<0.001) recreational physical activity. We assessed age-adjusted recreational physical activity in our analysis. We regressed log-transformed average METs per week on baseline age to obtain age-adjusted residuals, which we then analyzed as continuous and categorical variables. We assessed age-adjusted recreational physical activity categorized into quintiles (quintile 1 (Q1) = least physically active), as well as dichotomized as active (defined as highest 4 quintiles, Q2-Q5) versus inactive (defined as Q1). The minimum number of METs required to be classified as active for a given age at baseline is provided in the supplemental materials (eTable1). We also considered the joint association of baseline and adolescent recreational physical activity with breast cancer risk by comparing women who were classified as active at one or both time points (baseline and adolescence) to women who were classified as inactive at both time points.

We used multivariable Cox proportional hazards regression to estimate associations of recreational physical activity (baseline and adolescence) with breast cancer risk. The
proportionality assumption was assessed by evaluating Schoenfeld residuals. All models used age as the time scale and a robust variance estimator to account for multiple family members within the cohort (55% of the sample had at least one participating family member). Models were stratified by birth cohort (in 10-year categories) and adjusted for race/ethnicity (non-Hispanic white vs. other), and study center (Model 1). In addition to covariates included in Model 1, we examined models further adjusted for absolute predicted breast cancer risk (continuous) based on the BOADICEA model (20) (Model 2). Findings were similar when we used lifetime (from birth) or 1-year (from baseline) breast cancer risk predictions; results from lifetime risk models are reported. We also considered potential confounding by highest education level (≤high school degree vs. higher), parity and breastfeeding (nulliparous, 1-2 live births without breastfeeding, 1-2 live births with breastfeeding, ≥3 live births without breastfeeding, ≥3 live births with breastfeeding), and baseline health behaviors (never, former, current) including cigarette smoking, alcohol consumption, and use of hormonal birth control and menopausal hormone therapy (Model 3). Finally, we fitted models further adjusted for body mass index (BMI) categorized as <25 kg/m², 25-29.99 kg/m², and ≥30 kg/m² (Model 4). We tested for linear trends across quintiles of recreational physical activity based on the Wald test statistic for quintiles modeled as a continuous term using the median value for each recreational physical activity quintile. Statistical significance was determined as p<0.05 for a 2-sided hypothesis test.

We specified cross-product terms to test for multiplicative interaction by baseline characteristics including menopausal status (pre/post), lifetime breast cancer familial risk estimated by BOADICEA (modeled continuously and by tertiles of risk), race/ethnicity,
education, BMI category, and use of menopausal hormone therapy; statistical significance was based on the Wald test. We also tested for multiplicative interaction by mutation carrier status defined as *BRCA1* mutation carrier, *BRCA2* mutation carrier, or non-carrier (non-carriers included women who were tested and not known to carry pathogenic mutations, as well as women who did not undergo genetic testing). Evidence of effect modification was based on the Wald test statistic.

**Sensitivity Analyses**

We examined individual associations of moderate and strenuous recreational physical activity (categorized as none, 0.5-2, 2-3, and ≥4 hours per week) with breast cancer risk. We examined associations of recreational physical activity with breast cancer risk stratified by estrogen receptor (ER) status. We examined associations excluding pathologically confirmed ductal carcinoma *in situ* cases (N=100) and non-pathologically confirmed breast cancer cases (N=167), and we conducted a sensitivity analysis excluding the first two years of follow-up to assess potential reverse causation. We also examined associations including the first two months of follow-up after baseline (54 cases diagnosed during this time window).

Because breast cancer diagnosis was earlier on average for *BRCA1* and *BRCA2* mutation carriers, particularly for *BRCA1* carriers, it is possible that older mutation carriers unaffected with breast cancer at study enrollment are less susceptible (22, 23). Therefore, we conducted a sensitivity analysis examining associations of adolescent recreational physical activity with breast cancer risk using an expanded sample that included women who were diagnosed with breast cancer within 5 years prior to enrollment (N=7,905 including 368 *BRCA1* carriers and 336
BRCA2 carriers), in addition to the original prospective cohort. This allowed us to examine the association of adolescent recreational physical activity with breast cancer risk by gene mutation carrier status using a younger, and higher risk for age, cohort of women because prevalent breast cancer cases were diagnosed at a younger age on average compared with incident cases. This was true for non-carriers (average age at diagnosis for prevalent vs. incident cases: 48.9 vs. 58.9 years), BRCA1 carriers (41.1 vs. 48.3 years), and BRCA2 carriers (44.4 vs. 49.1 years). For this combined (retrospective and prospective) analysis, we used Cox models with person-years calculated from age 12 years, corresponding to the exposure period of interest. We did not assess baseline recreational physical activity using the combined cohort, given that this exposure occurred after diagnosis for prevalent cases. All analyses were conducted using Stata 15.1 (College Station, TX) (24).

Results

We observed 896 incident cases of breast cancer over 160,893 person-years of follow-up (median follow-up=10.3 years). Of these incident cases, 110 (12%) had a BRCA1 mutation and 69 (8%) had a BRCA2 mutation, while 324 (36%) were diagnosed before the age of 50 years. Eighty-six percent of women reported engaging in recreational physical activity during the three years prior to baseline (mean=24.0 METs per week; SD=24.3); 89% of women reported engaging in recreational physical activity during adolescence (mean=43.4 METs per week; SD=32.8). We present sample characteristics by quintiles of baseline recreational physical activity in Table 1. No clear differences in baseline characteristics were observed across quintiles of baseline recreational physical activity with the exception of BMI and current
smoking, which both decreased with increasing baseline METs per week (mean BMI: Q1=27.3±6.6 kg/m² vs. Q5=24.9±4.7 kg/m²; current smokers: Q1=19.4% vs. Q5=10.4%). There was a negative correlation between baseline METs per week and BMI (Pearson ρ = -0.17, p<0.001), and a positive correlation between baseline and adolescent METs per week (Pearson ρ=0.33, p<0.001).

As shown in Table 2, adjustment for full lifetime breast cancer familial risk (model 2) and other potential confounders including BMI (models 3-4) did not substantively alter HRs or 95% CIs. A one standard deviation increase in age-adjusted baseline METs per week was associated with a 10% reduced breast cancer risk using the fully adjusted model (Model 4: HR=0.90, 95% CI=0.85, 0.96). Each of the four highest quintiles of recreational physical activity was associated with lower breast cancer risk when compared with Q1 (p-trend=0.01), and comparing physically active (Q2-Q5) with inactive (Q1) at baseline was associated with a 20% lower breast cancer risk (Model 4: HR=0.80, 95% CI=0.68, 0.93). No association with breast cancer risk was found for adolescent recreational physical activity, modeled either continuously (Model 4: HR=1.00, 95% CI=0.93, 1.07) or categorically (Model 4: active vs. inactive HR=0.92, 95% CI=0.77, 1.09). Similar patterns of association were observed when strenuous and moderate recreational physical activity were assessed separately (see supplemental materials eTable2).

In Figure 1, we present the joint association of baseline and adolescent recreational physical activity with breast cancer risk by comparing women classified as active at baseline and adolescence (mean age=45.9 years), active at baseline but not adolescence (mean age=45.6 years), and active in adolescence but not baseline (mean age=48.1 years) with women classified as inactive at both time points (mean age=47.3 years). Being physically active at baseline was
consistently associated with lower breast cancer risk, regardless of being inactive (HR=0.74, 95% CI=0.54, 0.99) or active (HR=0.73, 95% CI=0.57, 0.94) in adolescence. For those who were inactive at baseline, there was no association between being active (vs. inactive) in adolescence with breast cancer risk (active in adolescence and inactive at baseline vs. inactive in adolescence and baseline: HR=0.89, 95% CI=0.66, 1.19).

We did not find evidence of multiplicative interaction between recreational physical activity (baseline or adolescent) and baseline age predicting breast cancer risk (p-values for cross-product terms > 0.05). As shown in Figure 2, the lower breast cancer risk associated with being active compared with inactive at baseline was consistently observed across strata of several baseline characteristics including menopausal status, race and ethnicity, education, BMI, menopausal hormone use, and tertiles lifetime breast cancer risk (all interaction p≥0.24), although with varying degrees of precision (Figure 3). We did not find evidence of a multiplicative interaction between lifetime breast cancer risk and baseline recreational physical activity when modeled continuously (p=0.39), categorically by quintiles (p=0.71), or dichotomized as active/inactive (p=0.19).

We found evidence for a multiplicative interaction between full lifetime breast cancer familial risk and adolescent recreational physical activity dichotomized as active/inactive (p=0.03). As illustrated in Figure 3, being active compared with inactive in adolescence was associated with a reduced breast cancer risk for women below the median level of full lifetime breast cancer familial risk (median = 18.5%), but not for women above the median level of risk. For example, being active in adolescence was associated with a 22% reduced breast cancer risk.
for women at the 10\textsuperscript{th} percentile of full lifetime familial risk (10\textsuperscript{th} percentile=13.3% risk), but was only associated with a 3\% reduced breast cancer risk for women at the 90\textsuperscript{th} percentile of full lifetime familial risk (90\textsuperscript{th} percentile=37.6% risk). We did not find evidence for a multiplicative interaction with full lifetime breast cancer familial risk when we modeled adolescent recreational physical activity continuously (p=0.10) or categorically by quintiles (p=0.13).

As shown in Figure 4, the association of baseline recreational physical activity with breast cancer risk was nearly 2-fold greater for \textit{BRCA2} carriers (HR=0.41, 95\% CI=0.20, 0.83) compared with non-carriers (HR=0.84, 95\% CI=0.70, 1.00), although we did not find evidence for multiplicative interaction when we formally tested for effect modification by carrier status (overall interaction p-value=0.15; \textit{BRCA2} vs non-carrier p=0.06; \textit{BRCA1} vs non-carrier p=0.42; \textit{BRCA2} vs \textit{BRCA1} p=0.42). Adolescent recreational physical was associated with a 14\% reduced breast cancer risk for non-carriers using the prospective cohort (HR = 0.86, 95\% CI = 0.71, 1.04) and a 9\% reduced breast cancer risk using the combined cohort that included retrospective cases diagnosed within 5 years of baseline (HR = 0.91, 95\% CI = 0.85, 0.96); only the estimate using the combined cohort was statistically significant (p<0.01). For \textit{BRCA1} carriers, a positive but not statistically significant association was found in the prospective cohort (HR = 1.47, 95\% CI = 0.77, 2.82), which was attenuated to the null in the combined cohort (HR = 1.01, 95\% CI = 0.76, 1.36). For \textit{BRCA2} carriers, no association was found between adolescent recreational physical activity and breast cancer risk using the prospective cohort (HR = 1.01, 95\% CI = 0.49, 2.08), while a negative but not statistically significant association was found using the combined cohort (HR = 0.88, 95\% CI = 0.68, 1.14). We did not find evidence for multiplicative interaction
between adolescent recreational physical activity and carrier status when we formally tested for effect modification using the prospective (p-value=0.25) or combined (p-value=0.54) cohort.

Similar associations were estimated for baseline and adolescent recreational physical activity when we stratified by ER-positive and ER-negative breast cancer, restricted to pathologically confirmed invasive breast cancer cases, excluded the first two years of follow-up, or included the first two months of follow-up.

**Discussion**

Using a prospective cohort enriched with women who have a family history of breast cancer and across a wide range of absolute predicted familial breast cancer risk (lifetime familial risk: mean=24.1%; range 8.1–96.4%), we found evidence suggesting that recreational physical activity during adulthood is associated with lower breast cancer risk. Specifically, we found that attaining at least 10.75 METs per week, which is the minimum amount of METs required to be classified as active for a given age at baseline (see eTable1), was associated with a 20% lower breast cancer risk. This equates to 2.7 hours of moderate or 1.5 hours of strenuous physical activity per week. Although we did not find a clear dose-response relationship between increasing quintiles of baseline recreational physical activity and breast cancer risk, our estimate comparing the highest with lowest quintile of baseline recreational physical activity was comparable to prior estimates from studies of women unselected for family history (2, 25).

We did not find evidence that the association of baseline recreational physical activity with breast cancer risk is modified by underlying breast cancer familial risk based on multigenerational pedigree data or by BRCA1 and BRCA2 mutation carrier status. Therefore, our
findings support that – in terms of absolute number of breast cancer cases prevented - physical activity interventions could have a greater absolute effect if targeted to women at higher familial/genetic risk (26). The association of baseline recreational physical activity with lower breast cancer risk was consistently observed across strata of different baseline characteristics, although with varying degrees of precision, providing support that findings were not driven by residual confounding. We also found that the association of baseline recreational physical activity with breast cancer risk was consistently observed regardless of recreational physical activity in adolescence, suggesting that behavior later in life reduces breast cancer risk independently of early life habits.

Overall, adolescent recreational physical activity was not associated with breast cancer risk. This is contrary to findings from some, but not all, previous studies that assessed the association of early life recreational physical activity with breast cancer risk, including two studies of BRCA1 and BRCA2 carriers (2, 8, 11). We recognize that our null findings could be driven by misclassification due to retrospective reporting of the adolescent exposure but inconsistent findings across studies could also reflect differences in the exposure window. For example, while the Sister Study found an association of recreational physical activity between ages 5 and 19 years with breast cancer risk (≥7 vs. <1 hour/week: HR= 0.75, 95% CI: 0.57, 0.99), the association was no longer observed (≥7 vs. <1 hour/week: HR=0.88, 95% CI: 0.72, 1.07) when they just considered physical activity between ages 13 and 19 years (26). There may also be differences in how recreational physical activity was defined and measured across studies. For example, some studies captured more detailed information on participation in team sports (e.g., type and duration) (9, 27), a major source of recreational physical activity in adolescence.
Differential findings could also be explained by differences in study design (e.g., prospective cohort vs. case-control), as well as characteristics of study sample.

We found that the association of adolescent recreational physical activity (dichotomized as active vs. inactive) with breast cancer risk was modified by underlying breast cancer familial risk, with the association increasing in magnitude with decreasing full lifetime familial risk. Further, being physically active in adolescence was associated with a 14% reduced breast cancer risk when we restricted to non-carriers using the prospective cohort. No such reduced association of adolescent recreational physical activity was found for BRCA1 and BRCA2 carriers using the prospective cohort. While this could suggest that early life physical activity only provides a protective benefit for women with low breast cancer familial risk, these findings might reflect bias stemming from the earlier age at breast cancer diagnosis for affected mutation carriers (29). When we expanded our study sample to include prevalent cases, which were diagnosed at a younger age on average than incident cases, adolescent recreational physical activity was associated with a 12% decrease in risk for BRCA2 carriers (compared to an estimated 6% increase in risk using the prospective cohort) and a 1% increase in breast cancer risk for BRCA1 carriers (compared to an estimated 47% increase in risk using the prospective cohort). Prospective studies of younger cohorts enriched with high risk women are needed to further explore these associations.

The underlying mechanisms through which recreational physical activity influences breast cancer risk are not fully understood and could vary by individual-level factors such as age. For example, regulation of body fat through higher levels of physical activity could be
associated with reduced breast cancer risk for postmenopausal women (2, 4), but because adiposity is associated with a lower risk of premenopausal breast cancer (2, 4), physical activity might operate through different mechanisms to reduce breast cancer risk for younger women. Mechanisms that might occur independently of change in adiposity include physical activity effects on estrogen metabolism, insulin sensitivity, chronic low-level inflammation, oxidative stress, and immune function (2, 4, 5). Physical activity-induced transcripational changes are also possible (11, 30). Given the findings of this study, future studies should also consider whether the biological mechanisms of physical activity differ by absolute breast cancer risk in conjunction with other factors such as menopausal status. Exercise intervention trials conducted on gene mutation carriers would be highly informative for this area of research.

Our study has several strengths. Most notably, we used data from a large prospective cohort of women enriched for familial and genetic risk of breast cancer, and we were able to compute continuous measures of familial risk from our detailed pedigree data using BOADICEA (19, 20). This allowed us to test associations of recreational physical activity across a wide range of underlying breast cancer familial risk, including for women in the high-risk tail of absolute lifetime risk. We tested associations of recreational physical activity with breast cancer risk separately for BRCA1 and BRCA2 mutation carriers making this, to our knowledge, the first study to do so. Further, we assessed both moderate and strenuous types of recreational physical activity, which we converted to a combined measure of total METs per week, and we considered recent exposure to recreational physical activity at baseline, as well as recreational physical activity in adolescence.
Our study was limited by the use of self-reported recreational physical activity, which is prone to misclassification, particularly for more distant adolescent recreational physical activity. However, given our prospective study design, exposure misclassification is likely to be non-differential, and prior research suggests that recall bias is not likely to fully explain associations of recreational physical activity with breast cancer risk (31). Further, while self-reported physical activity is known to be overestimated (32), prior studies have demonstrated the reliability (33-35) and validity (33, 35) of using of self-reported measures of recreational physical activity for rank ordering physical activity levels (i.e., stratifying more physically active individuals from less physically active individuals). Self-reported recreational physical activity levels in our study were comparable to those reported by the general population of US women (36). We also note that baseline recreational physical activity was correlated with baseline BMI, providing support for the validity of our measure. Another limitation is that we were unable to account for other types of physical activity (e.g., occupational, household) or other potential confounders, such as sedentary behavior. We also did not consider recreational physical activity prospectively after baseline in this analysis, and thus cannot draw conclusions about behavior change.

In conclusion, our findings provide further support for an association between recreational physical activity in adulthood and breast cancer risk and suggest that even a modest level of recreational physical activity in adulthood is associated with reduced breast cancer risk. Importantly, we found that this association exists for women across the absolute familial risk continuum, including for women at high familial and genetic risk. Therefore, physical activity interventions could be an effective primary prevention strategy for all women,
and be especially beneficial for women at higher than average familial risk of breast cancer who stand to benefit most from such efforts.
Acknowledgements

We thank the entire team of Breast Cancer Family Registry (BCFR) past and current investigators as well as the kConFab investigators. We also thank Heather Thorne, Eveline Niedermayr, Lucy Stanhope, Sandra Picken, all the BCFR and kConFab research nurses and staff, the heads and staff of the Family Cancer Clinics, and the many families who contribute to the BCFR and kConFab for their contributions to this resource.
References


Table 1. Baseline characteristics by recreational physical activity in the Prospective Family Study Cohort (N=15,550)

<table>
<thead>
<tr>
<th>Baseline characteristic</th>
<th>Quintiles of Age-Adjusted Baseline Recreational Physical Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Q1</td>
</tr>
<tr>
<td>Age, years, mean (sd)</td>
<td>47.8 (15.3)</td>
</tr>
<tr>
<td>Premenopausal, n (%)</td>
<td>1,704 (54.7)</td>
</tr>
<tr>
<td>Race and ethnicity, n (%)</td>
<td></td>
</tr>
<tr>
<td>non-Hispanic white</td>
<td>2,350 (75.4)</td>
</tr>
<tr>
<td>other</td>
<td>745 (23.9)</td>
</tr>
<tr>
<td>missing</td>
<td>23 (0.7)</td>
</tr>
<tr>
<td>Education, n (%)</td>
<td></td>
</tr>
<tr>
<td>high school/GED or less</td>
<td>1,392 (44.6)</td>
</tr>
<tr>
<td>College or more</td>
<td>1,708 (54.8)</td>
</tr>
<tr>
<td>Missing</td>
<td>18 (0.6)</td>
</tr>
<tr>
<td>Parity/breastfeeding, n (%)</td>
<td></td>
</tr>
<tr>
<td>nulliparous</td>
<td>568 (18.2)</td>
</tr>
<tr>
<td>parous 1–2/No BF</td>
<td>454 (14.6)</td>
</tr>
<tr>
<td>parous 1–2/BF</td>
<td>813 (26.1)</td>
</tr>
<tr>
<td>parous 3+/No BF</td>
<td>330 (10.6)</td>
</tr>
<tr>
<td>parous 3+/BF</td>
<td>953 (30.6)</td>
</tr>
<tr>
<td>Cigarette use, n (%)</td>
<td></td>
</tr>
<tr>
<td>never</td>
<td>1,694 (54.3)</td>
</tr>
<tr>
<td>former</td>
<td>790 (25.3)</td>
</tr>
<tr>
<td>current</td>
<td>606 (19.4)</td>
</tr>
<tr>
<td>missing</td>
<td>28 (0.9)</td>
</tr>
<tr>
<td>Alcohol use, n (%)</td>
<td></td>
</tr>
<tr>
<td>never</td>
<td>1,760 (56.5)</td>
</tr>
<tr>
<td>former</td>
<td>525 (16.8)</td>
</tr>
<tr>
<td>current</td>
<td>818 (26.2)</td>
</tr>
<tr>
<td>missing</td>
<td>15 (0.5)</td>
</tr>
<tr>
<td>Menopausal hormone therapy, n (%)</td>
<td></td>
</tr>
<tr>
<td>never</td>
<td>2,325 (74.6)</td>
</tr>
<tr>
<td>former</td>
<td>389 (12.5)</td>
</tr>
<tr>
<td>current</td>
<td>366 (11.7)</td>
</tr>
<tr>
<td>missing</td>
<td>38 (1.2)</td>
</tr>
<tr>
<td>Hormonal birth control, n (%)</td>
<td></td>
</tr>
<tr>
<td>never</td>
<td>777 (24.9)</td>
</tr>
<tr>
<td>former</td>
<td>1,955 (62.7)</td>
</tr>
<tr>
<td>current</td>
<td>357 (11.5)</td>
</tr>
<tr>
<td>missing</td>
<td>29 (0.9)</td>
</tr>
<tr>
<td>Body mass index, n (%)</td>
<td></td>
</tr>
<tr>
<td>&lt;25 kg/m²</td>
<td>1,342 (43.0)</td>
</tr>
<tr>
<td>25–&lt;30 kg/m²</td>
<td>832 (26.7)</td>
</tr>
<tr>
<td>≥30 kg/m²</td>
<td>858 (27.5)</td>
</tr>
<tr>
<td>missing</td>
<td>86 (2.8)</td>
</tr>
<tr>
<td>Lifetime familial BC risk, mean (sd)</td>
<td></td>
</tr>
<tr>
<td>none</td>
<td>2,449 (78.5)</td>
</tr>
<tr>
<td>any</td>
<td>528 (16.9)</td>
</tr>
<tr>
<td>Adolescent RPA, n (%)</td>
<td></td>
</tr>
<tr>
<td>none</td>
<td>141 (4.5)</td>
</tr>
</tbody>
</table>

Notes: BC=breast cancer; sd=standard deviation; GED=general education degree; BF=breast feeding; BOADICEA=Breast Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm; RPA=recreational physical activity. Refer to eTable 1 in the supplemental materials for the minimum number of metabolic equivalents (METs) required to be classified as active for a given age at baseline.
<table>
<thead>
<tr>
<th>Continuous Age-Adjusted METs per Week</th>
<th>Quintiles of Age-Adjusted Recreational Physical Activity</th>
<th>Active (Q2-Q5) vs. Inactive (Q1)&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>HR (95% CI)&lt;sup&gt;a&lt;/sup&gt;</td>
<td>HR (95% CI)</td>
<td>HR (95% CI)</td>
</tr>
<tr>
<td>Q1</td>
<td>Q2</td>
<td>Q3</td>
</tr>
<tr>
<td>Model 1&lt;sup&gt;c&lt;/sup&gt;</td>
<td>0.92 (0.86, 0.98)</td>
<td>1.00 (ref.)</td>
</tr>
<tr>
<td>Model 2&lt;sup&gt;d&lt;/sup&gt;</td>
<td>0.92 (0.87, 0.98)</td>
<td>1.00 (ref.)</td>
</tr>
<tr>
<td>Model 3&lt;sup&gt;e&lt;/sup&gt;</td>
<td>0.90 (0.85, 0.96)</td>
<td>1.00 (ref.)</td>
</tr>
<tr>
<td>Model 4&lt;sup&gt;f&lt;/sup&gt;</td>
<td>0.90 (0.85, 0.96)</td>
<td>1.00 (ref.)</td>
</tr>
</tbody>
</table>

**Baseline Recreational Physical Activity**

- Model 1<sup>c</sup> adjusted for race and ethnicity; study center; and baseline age; model stratified by birth cohort.
- Model 2<sup>d</sup> adjusted for race and ethnicity; study center; baseline age; and lifetime familial breast cancer risk; model stratified by birth cohort.
- Model 3<sup>e</sup> adjusted for race and ethnicity; study center; baseline age; lifetime familial breast cancer risk; education; parity and breastfeeding; and use of alcohol, cigarettes, hormonal birth control and menopausal hormone therapy; model stratified by birth cohort.
- Model 4<sup>f</sup> adjusted for race and ethnicity; study center; baseline age; lifetime familial breast cancer risk; education; parity and breastfeeding; use of alcohol, cigarettes, hormonal birth control, and menopausal hormone therapy; and body mass index; model stratified by birth cohort.

**Adolescent Recreational Physical Activity**

- Model 1<sup>c</sup> adjusted for race and ethnicity; study center; baseline age; and lifetime familial breast cancer risk; model stratified by birth cohort.
- Model 2<sup>d</sup> adjusted for race and ethnicity; study center; baseline age; and lifetime familial breast cancer risk; education; parity and breastfeeding; and use of alcohol, cigarettes, hormonal birth control and menopausal hormone therapy; model stratified by birth cohort.
- Model 3<sup>e</sup> adjusted for race and ethnicity; study center; baseline age; lifetime familial breast cancer risk; education; parity and breastfeeding; use of alcohol, cigarettes, hormonal birth control, and menopausal hormone therapy; and body mass index; model stratified by birth cohort.
- Model 4<sup>f</sup> adjusted for race and ethnicity; study center; baseline age; lifetime familial breast cancer risk; education; parity and breastfeeding; use of alcohol, cigarettes, hormonal birth control, and menopausal hormone therapy; and body mass index; model stratified by birth cohort.

<sup>a</sup>Hazard ratio (HR) reflects association for a one standard deviation (SD) change in continuous log-transformed and age-adjusted metabolic equivalents (METs) per week (baseline SD=1.98; adolescent SD=1.62)

<sup>b</sup>Refer to eTable 1 in the supplemental materials for the minimum number of metabolic equivalents (METs) required to be classified as active for a given age at baseline.

<sup>c</sup>Adjusted for race and ethnicity; study center; and baseline age; model stratified by birth cohort.

<sup>d</sup>Adjusted for race and ethnicity; study center; baseline age; and lifetime familial breast cancer risk; model stratified by birth cohort.

<sup>e</sup>Adjusted for race and ethnicity; study center; baseline age; lifetime familial breast cancer risk; education; parity and breastfeeding; and use of alcohol, cigarettes, hormonal birth control and menopausal hormone therapy; model stratified by birth cohort.

<sup>f</sup>Adjusted for race and ethnicity; study center; baseline age; lifetime familial breast cancer risk; education; parity and breastfeeding; use of alcohol, cigarettes, hormonal birth control, and menopausal hormone therapy; and body mass index; model stratified by birth cohort.

<sup>g</sup>Quintile 1 (Q1) includes the least physically active women in the sample during adolescence with METs ranging from 0-14 per week, Q2=15-27 METs per week, Q3=28-44 METs per week, and Q5=47-72 METs per week. N=14,619 (cases=852) because 931 participants missing data on recreational physical activity during adolescence.
Figure 1 Legend: N=14,619 (cases=852) because 931 participants were missing data on recreational physical activity during adolescence. Hazard ratios (HRs) and 95% CIs (confidence intervals) are adjusted for race/ethnicity, study center, lifetime familial BC risk, education, parity and breastfeeding, and use of alcohol, cigarettes, hormonal birth control and menopausal hormone therapy, and body mass index; stratified by birth cohort. Reference group = inactive (defined as lowest quintile (Q1) of age-adjusted recreational physical activity) in adolescence and baseline. Active is defined as highest four quintiles (Q2–Q5) of age-adjusted recreational physical activity for given exposure period. Refer to eTable 1 in the supplemental materials for the minimum number of metabolic equivalents (METs) required to be classified as active for a given age at baseline.
Figure 2 Legend: GED=general education degree. Reference group = inactive at baseline, defined as lowest quintile (Q1) of age-adjusted recreational physical activity. Active is defined as highest four quintiles (Q2–Q5) of age-adjusted recreational physical activity. Refer to eTable 1 in the supplemental materials for the minimum number of metabolic equivalents (METs) required to be classified as active for a given age at baseline. Hazard ratios (HRs) and 95% confidence intervals (CI) are adjusted for race/ethnicity, study center, lifetime familial breast cancer risk, education, parity and breastfeeding, and use of alcohol, cigarettes, hormonal birth control and menopausal hormone therapy, and body mass index; stratified by birth cohort. The interaction p-value is the Wald test statistic used for the interaction term between physical activity and the baseline characteristic predicting breast cancer risk.
Figure 3 Legend: Point estimates reflect hazard ratios comparing active (defined as highest four quintiles (Q2–Q5) of age-adjusted recreational physical activity) to inactive (defined as lowest quintile (Q1) of age-adjusted recreational physical activity) in adolescence by percentiles of full lifetime risk (interaction p-value = 0.03). The percent full lifetime familial breast cancer risk corresponding to each percentile is provided below the X axis. Refer to eTable 1 in the supplemental materials for the minimum number of metabolic equivalents (METs) required to be classified as active for a given age at baseline. Hazard ratios and 95% confidence intervals are adjusted for race/ethnicity, study center, education, parity and breastfeeding, use of alcohol, cigarettes, hormonal birth control and menopausal hormone therapy, and body mass index; stratified by birth cohort.
Figure 4 Legend: Point estimates reflect hazard ratios comparing active (defined as highest four quintiles (Q2–Q5) of age-adjusted recreational physical activity) to inactive (defined as lowest quintile (Q1) of age-adjusted recreational physical activity) for the given exposure period. Refer to eTable 1 in the supplemental materials for the minimum number of metabolic equivalents (METs) required to be classified as active for a given age at baseline. Hazard ratios and 95% confidence intervals are adjusted for race/ethnicity, study center, education, parity and breastfeeding, use of alcohol, cigarettes, hormonal birth control and menopausal hormone therapy, and body mass index; stratified by birth cohort. The interaction p-value is the Wald test statistic used for the interaction term between physical activity and mutation carrier status. Non-carriers include women who were tested and not known to carry pathogenic mutations, as well as women who did not undergo genetic testing. The combined cohort includes prevalent breast cancer cases diagnosed within 5 years prior to study enrollment.
Figure 1. Joint association of baseline and adolescent recreational physical activity with breast cancer risk in the Prospective Family Study Cohort (N=14,619)

Reference = Inactive in adolescence and inactive at baseline; N=1,086

- Active in adolescence and inactive at baseline: Hazard Ratio = 0.89 (0.66, 1.19), N=1,891
- Inactive in adolescence and active at baseline: Hazard Ratio = 0.74 (0.54, 0.99), N=1,838
- Active in adolescence and active at baseline: Hazard Ratio = 0.73 (0.57, 0.94), N=9,804
Figure 2. Association of baseline recreational physical activity with breast cancer risk by baseline characteristics of the Prospective Family Study Cohort (N=15,550)

<table>
<thead>
<tr>
<th>Category</th>
<th>Active vs. Inactive HR (95% CI)</th>
<th>Interaction P-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Menopausal Status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>premenopausal</td>
<td>0.81 (0.65, 1.01)</td>
<td>0.58</td>
</tr>
<tr>
<td>postmenopausal</td>
<td>0.80 (0.63, 1.01)</td>
<td></td>
</tr>
<tr>
<td><strong>Race and Ethnicity</strong></td>
<td></td>
<td>0.36</td>
</tr>
<tr>
<td>NH White</td>
<td>0.83 (0.69, 0.99)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>0.70 (0.48, 1.02)</td>
<td></td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td>0.57</td>
</tr>
<tr>
<td>≤ HS Degree or GED</td>
<td>0.77 (0.59, 1.00)</td>
<td></td>
</tr>
<tr>
<td>&gt; HS Degree or GED</td>
<td>0.81 (0.67, 1.00)</td>
<td></td>
</tr>
<tr>
<td><strong>Body Mass Index</strong></td>
<td></td>
<td>0.65</td>
</tr>
<tr>
<td>&lt; 25 kg/m²</td>
<td>0.87 (0.68, 1.11)</td>
<td></td>
</tr>
<tr>
<td>25-29.99 kg/m²</td>
<td>0.76 (0.56, 1.04)</td>
<td></td>
</tr>
<tr>
<td>≥ 30 kg/m²</td>
<td>0.76 (0.56, 1.04)</td>
<td></td>
</tr>
<tr>
<td><strong>Menopausal Hormone Therapy Use</strong></td>
<td></td>
<td>0.24</td>
</tr>
<tr>
<td>Never</td>
<td>0.83 (0.68, 1.01)</td>
<td></td>
</tr>
<tr>
<td>Former</td>
<td>0.71 (0.46, 1.09)</td>
<td></td>
</tr>
<tr>
<td>Current</td>
<td>0.68 (0.46, 0.99)</td>
<td></td>
</tr>
<tr>
<td><strong>Lifetime Breast Cancer Risk</strong></td>
<td></td>
<td>0.91</td>
</tr>
<tr>
<td>Low (&lt;16.44%)</td>
<td>0.83 (0.61, 1.13)</td>
<td></td>
</tr>
<tr>
<td>Med (16.44-21.74%)</td>
<td>0.77 (0.56, 1.06)</td>
<td></td>
</tr>
<tr>
<td>High (&gt;21.74%)</td>
<td>0.82 (0.65, 1.03)</td>
<td></td>
</tr>
</tbody>
</table>
Figure 3. Adolescent recreational physical activity and breast cancer risk by percentiles of full lifetime familial breast cancer risk estimated by the Breast and Ovarian Analysis of Disease Incidence and Carrier Estimation Algorithm (BOADICEA) in the Prospective Family Study Cohort (N=14,619)
Figure 4. Associations of baseline and adolescent recreational physical activity with breast cancer risk by BRCA1 and BRCA2 mutation carrier status in the Prospective Family Study Cohort.
Recreational physical activity is associated with reduced breast cancer risk in adult women at high risk for breast cancer: a cohort study of women selected for familial and genetic risk

Rebecca D Kehm, Jeanine M Genkinger, Robert J MacInnis, et al.

Cancer Res Published OnlineFirst October 2, 2019.

Updated version
Access the most recent version of this article at:
doi:10.1158/0008-5472.CAN-19-1847

Supplementary Material
Access the most recent supplemental material at:
http://cancerres.aacrjournals.org/content/suppl/2019/10/02/0008-5472.CAN-19-1847.DC2
http://cancerres.aacrjournals.org/content/suppl/2019/10/01/0008-5472.CAN-19-1847.DC1

Author Manuscript
Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts
Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions
To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions
To request permission to re-use all or part of this article, use this link
http://cancerres.aacrjournals.org/content/early/2019/10/03/0008-5472.CAN-19-1847.
Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.