Supplementary table 1. Methodological quality score – Cohort and Case-control studies

Criteria	Score	Comments
Selection bias		
1a. Percentage loss to follow-up? *Only applicable to **cohort** studies*		
> 20% or unknown or unclassifiable	0	The percentage is unclassifiable if the total eligible cohort is not clear.
5-20%	4	
< 5%	8	
1b. Percentage response of the cases **and** controls? *Only applicable to **case-control** studies*		
< 75% or unknown or unclassifiable	0	Response = (1-(refusal of subject/physician, contact problems, death before interview)/eligible subjects)*100. Judging the response of cases and controls, the lowest percentage counts. Unclassifiable if the complete eligible group is not clear.
75-90%	4	
> 90%	8	
2. Was the absolute difference in percentage response < 20% between cases and controls?		
Not applicable (cohort study)	10	
No or unknown	0	
Yes	*6*	
3. Percentage incident cases?		
< 75%	0	Some studies include prevalent and/or fatal cases without information about the date of diagnosis, which may introduce bias, as physical activity may also be associated with survival.
75-99% or unknown	4	
100%	7	
4. Did the cases and controls originate from the same source population?[i]		
Not applicable (cohort study)	10	
No or unknown	0	
Yes	10	
5. Were the same exclusion/inclusion criteria applied to cases and controls?		
Not applicable (cohort study)	7	
No or unknown	0	
Yes	*7*	
Maximal selection bias score	**42**	

Misclassification bias		
Determination of physical activity		
6. Was the measure of leisure time activities that was analysed complete enough?		
No	0	Activity may have been assessed extensively whereas only few components where included in the analysis. Judgement of the item was based on reviewers' consensus.
Yes	4	
7. Was total activity assessed?		
No	0	Total activity means leisure time activity and job/household activity. These should be combined in one effect measure.
Yes	4	
8. Did the measure of physical activity include intensity, frequency and duration?		
One component or unknown	0	
Intensity + frequency/duration per week	5	
9. Type of administration of physical activity questionnaire		
By proxy	0	By proxy means that physical activity is not individually assessed e.g. a family member was asked; classification was based on college registration of athletics.
Self-administered	3	
Interview-administered	4	
10. Was the operationalisation of the physical activity score understandable?		
No or partly	0	
Yes	2	
11. Did the physical activity measure include past physical activity?		
No, only recent	0	A physical activity measure covering more periods of life is supposed to be more accurate.
Yes, more life periods	4	
12. Did the authors consider changes over time in physical activity pattern in the analyses?		
No	0	Yes: e.g. when two measurements of physical activity several years apart were used to classify participants in consistently active or inactive.
Yes	4	
13. Was the physical activity questionnaire validated or was reliability tested?		
No or unknown	0	
Yes	2	
14. Was physical activity level assessed before breast cancer diagnosis?		
No, and different methods for physical activity assessment for cases and controls	0	When physical activity is measured after the diagnosis of breast cancer, there is a possibility of recall bias.
No, and the same methods for physical activity assessment for cases and controls	4	
Yes	7	
Outcome		
15. Was the case diagnosis valid?		
No or unknown	0	Breast cancer self-report may be valid.
Yes	4	

16. Could DCIS in any way have influenced the results?		
Yes or unknown	0	Any influence seems unlikely if < 5% of cases are DCIS and/or if separate analyses/exclusion did not result in different estimates.
No	2	
Maximal misclassification bias score	42	
Confounding bias		
17. Were confounders adjusted for in a correct way (statistically)?		
No or unknown	0	
Yes	4	
18. Could residual confounding be a problem?		
Yes	0	Potential confounders: age, body mass index (BMI), parity, age at menopause, hormone replacement therapy (HRT). Residual confounding could also be a problem when 1) continuous variables were crudely categorised or 2) BMI or HRT were not measured within 5 years of diagnosis.
Partly	4	
No	9	
19. Were the effects of leisure time activities adjusted for occupational/household activities?		
No or unknown	0	
Yes	8	
Maximal confounding bias score	**21**	

Maximal total score: 105

[i] Patients and controls are not from the same source populations when:

- hospital based controls are used and it is unlikely that the controls would be referred to the same hospital in case of breast/colon cancer .
- cases are recruited from a specialised cancer hospital (e.g. subgroup of patients with more advanced disease) and controls are population based and it is quite uncertain whether they would be referred to the same hospital if they had become a case.
- controls are not a random sample of the source population.
- cases and controls are selected from different areas/ countries.

Remark: item 12 and item 16 were excluded in the present meta-analysis.

Reference: Monninkhof EM, Elias SG, Vlems FA, van dT, I, Schuit AJ, Voskuil DW, van Leeuwen FE. Physical activity and breast cancer: a systematic review. Epidemiology 2007 January;18(1):137-57.