Table 4. Relative Risk of Death From Cardiovascular Disease, Cancer, and All Other Causes According to Leisure Time Spent Sitting and Physical Activity Among Women and Men, Cancer Prevention Study II Nutrition Cohort, 1993–2006

	Cardiovascular Disease			Cancer			Other Causes		
	No. of Deaths	Relative Risk[a]	95% CI	No. of Deaths	Relative Risk[a]	95% CI	No. of Deaths	Relative Risk[a]	95% CI
					Women				
Sitting in 1992, hours/day									
0–<3	833	1.00	Referent	1,284	1.00	Referent	921	1.00	Referent
3–5	1,196	1.20	1.10, 1.32	1,413	1.07	0.99, 1.16	1,172	1.13	1.04, 1.24
≥6	331	1.33	1.17, 1.52	411	1.30	1.16, 1.46	362	1.41	1.25, 1.60
P_{trend}		<0.0001			<0.0001			<0.0001	
Total physical activity in 1992, MET-hours/week									
<17.5	370	1.00	Referent	413	1.00	Referent	374	1.00	Referent
17.5–<24.5	224	0.98	0.83, 1.15	260	1.02	0.88, 1.20	215	0.93	0.78, 1.10
24.5–<31.5	400	0.74	0.64, 0.86	536	0.88	0.78, 1.00	446	0.82	0.71, 0.94
31.5–<42	484	0.76	0.67, 0.87	608	0.82	0.73, 0.94	496	0.76	0.67, 0.87
42–<52.5	369	0.72	0.62, 0.84	499	0.83	0.73, 0.95	387	0.74	0.64, 0.85
52.5–<63	224	0.71	0.60, 0.84	328	0.87	0.75, 1.01	220	0.68	0.57, 0.80
≥63	289	0.66	0.56, 0.77	464	0.86	0.75, 0.99	317	0.69	0.59, 0.80
P_{trend}		<0.0001			0.03			<0.0001	
					Men				
Sitting in 1992, hours/day									
0–<3	1,413	1.00	Referent	1,457	1.00	Referent	1,160	1.00	Referent
3–5	1,911	1.06	0.99, 1.14	1,853	1.05	0.98, 1.12	1,649	1.13	1.04, 1.22
≥6	685	1.18	1.08, 1.30	571	1.04	0.94, 1.15	608	1.33	1.20, 1.47
P_{trend}		0.0007			0.29			<0.0001	
Total physical activity in 1992, MET-hours/week									
<17.5	435	1.00	Referent	379	1.00	Referent	372	1.00	Referent
17.5–<24.5	353	0.87	0.75, 1.00	314	0.91	0.78, 1.06	324	0.92	0.79, 1.07
24.5–<31.5	496	0.81	0.71, 0.92	488	0.96	0.84, 1.10	413	0.77	0.67, 0.89
31.5–<42	818	0.86	0.77, 0.97	781	0.98	0.87, 1.11	685	0.81	0.71, 0.92
42–<52.5	638	0.76	0.68, 0.86	636	0.91	0.80, 1.03	583	0.77	0.67, 0.87
52.5–<63	516	0.78	0.68, 0.88	511	0.92	0.81, 1.05	426	0.70	0.61, 0.80
≥63	753	0.77	0.68, 0.87	772	0.95	0.84, 1.07	614	0.68	0.60, 0.78
P_{trend}		0.0001			0.52			<0.0001	

Abbreviations: CI, confidence interval; MET, metabolic equivalent.

[a] Adjusted for age at interview, race, marital status, education, smoking status, body mass index in 1992, alcohol use, total caloric intake, comorbidities score, and total physical activity (for sitting) and hours sitting (for total physical activity).

First, time spent sitting might be more easily measured than physical activity and/or may reflect a different aspect of inactivity than other indices usually used in epidemiologic studies. However, this potential misclassification of exposure is unlikely to fully explain our findings, because time spent sitting was significantly associated with mortality even among men and women with the highest levels of physical activity.

Second, time spent sitting might be associated with other unhealthy behaviors that are either not captured or incompletely captured through questionnaires. Total energy expenditure is reduced among individuals who are sedentary.

However, consistent with previous studies, the present study found no correlation between physical activity and time spent sitting ($r = -0.03$). Time spent sitting is also associated with greater food consumption and subsequent weight gain, especially when watching television (16, 34, 35). Time spent sitting was previously shown to be associated with increased weight gain in this cohort (18). While residual confounding by obesity could contribute to the association between sitting time and mortality, this association was attenuated but not eliminated by controlling for or stratifying on body mass index.

Third, prolonged time spent sitting, independent of physical activity, has important metabolic consequences that may influence specific biomarkers (such as triglycerides, high density lipoprotein cholesterol, fasting plasma glucose, resting blood pressure, and leptin) of obesity and cardiovascular and other chronic diseases (8–11). Animal studies have also shown that sedentary time substantially suppresses enzymes centrally involved in lipid metabolism within skeletal muscle, and low levels of daily life activity are sufficient to improve enzyme activity (36–38). Furthermore, substantial evidence in both adults and children from observational studies and randomized clinical trials shows that reducing time spent sitting lowers the risk of obesity and type II diabetes (19, 39–42).

Over the past century, a number of technologic changes have contributed to a decrease in total daily energy expenditure. For example, during the 2006–2007 broadcast year, the average US household reported 8 hours of television watching per day, which is an increase of 1 hour per day of television watching from only a decade ago (43). Although leisure-time physical activity levels have remained relatively constant over the past few decades (44, 45), it is well recognized that technologic advances in the workplace have also greatly reduced occupational physical activity. This reduction in overall physical activity, in conjunction with increased time spent sitting and higher caloric intake, has contributed in large part to the rise in obesity and likely influenced temporal trends in cardiovascular disease, type 2 diabetes, and some cancers.

The strengths of our study include the large sample size, prospective design, and ability to control for many potential confounding factors. The lack of occupational physical activity data is a potential limitation; however, we believe this to have minimal impact on daily physical activity levels because the majority of study participants were retired/homemakers (57% of men and 80% of women) and, among those that were not retired, few worked in jobs that involved any activity (21% of men and 7% of women). Because we measured only leisure time spent sitting, the lack of occupational sitting time may have underestimated sitting time among working individuals, since much of their sitting time may have occurred at work. However, adjusting for employment status (employed, retired, or homemaker) did not change risk estimates for time spent sitting or physical activity. Furthermore, we conducted a sensitivity analysis among only men and women who were retired or homemakers, and results were virtually identical to those in the overall cohort. Another limitation is the use of self-reported measures of time spent sitting, physical activity, and all other covariates including height and weight. Although the physical activity and sitting time questions we used are subject to misreporting, they are very similar to those used and validated in the Nurses' Health Study II, a prospective study with similar participant characteristics, which found a correlation of 0.79 between activity reported on recalls and questionnaire (46). These measures have also been associated with various cancers in this cohort (47–50). Finally, we were not able to differentiate between types of sitting (i.e., while watching television, reading, driving), and the energy expenditure and other behaviors may vary with different types of sitting.

In conclusion, we found that both leisure time spent sitting and physical activity are independently associated with total mortality. Associations were stronger for cardiovascular disease mortality than for cancer mortality. Public health messages and guidelines should be refined to include reducing time spent sitting in addition to promoting physical activity. Because a sizeable fraction of the population spends much of their time sitting, it is beneficial to encourage sedentary individuals to stand up and walk around as well as to reach optimal levels of physical activity.

ACKNOWLEDGMENTS

Author affiliations: Epidemiology Research Program, American Cancer Society, Atlanta, Georgia (Alpa V. Patel, Ausila Deka, Peter T. Campbell, Susan M. Gapstur, Michael J. Thun); City of Hope, Duarte, California (Leslie Bernstein); Kaiser Permanente, Denver, Colorado (Heather Spencer Feigelson); and Washington University Siteman Cancer Center, St. Louis, Missouri (Graham A. Colditz).

The authors would like to acknowledge the late Drs. Eugenia E. Calle and Carmen Rodriguez who were instrumental in conducting and guiding this research. The preparation of this manuscript would not have been possible without them.

Conflict of interest: none declared.

REFERENCES

1. Leitzmann MF, Park Y, Blair A, et al. Physical activity recommendations and decreased risk of mortality. *Arch Intern Med.* 2007;167(22):2453–2460.
2. Nocon M, Hiemann T, Müller-Riemenschneider F, et al. Association of physical activity with all-cause and cardiovascular mortality: a systematic review and meta-analysis. *Eur J Cardiovasc Prev Rehabil.* 2008;15(3):239–246.
3. Paffenbarger RS Jr, Hyde RT, Wing AL, et al. Physical activity, all-cause mortality, and longevity of college alumni. *N Engl J Med.* 1986;314(10):605–613.
4. World Cancer Research Fund/American Institute for Cancer Research. *Food, Nutrition, Physical Activity, and the Prevention of Cancer: A Global Perspective.* Washington, DC: American Institute for Cancer Research; 2007.
5. Blair SN, Morris JN. Healthy hearts—and the universal benefits of being physically active: physical activity and health. *Ann Epidemiol.* 2009;19(4):253–256.

6. Kushi LH, Fee RM, Folsom AR, et al. Physical activity and mortality in postmenopausal women. *JAMA*. 1997;277(16): 1287–1292.
7. Rockhill B, Willett WC, Manson JE, et al. Physical activity and mortality: a prospective study among women. *Am J Public Health*. 2001;91(4):578–583.
8. Fung TT, Hu FB, Yu J, et al. Leisure-time physical activity, television watching, and plasma biomarkers of obesity and cardiovascular disease risk. *Am J Epidemiol*. 2000;152(12): 1171–1178.
9. Ford ES, Kohl HW III, Mokdad AH, et al. Sedentary behavior, physical activity, and the metabolic syndrome among U.S. adults. *Obes Res*. 2005;13(3):608–614.
10. Healy GN, Wijndaele K, Dunstan DW, et al. Objectively measured sedentary time, physical activity, and metabolic risk: the Australian Diabetes, Obesity and Lifestyle Study (AusDiab). *Diabetes Care*. 2008;31(4):369–371.
11. Jakes RW, Day NE, Khaw KT, et al. Television viewing and low participation in vigorous recreation are independently associated with obesity and markers of cardiovascular disease risk: EPIC-Norfolk population-based study. *Eur J Clin Nutr*. 2003;57(9):1089–1096.
12. Haskell WL, Lee IM, Pate RR, et al. Physical activity and public health: updated recommendation for adults from the American College of Sports Medicine and the American Heart Association. *Med Sci Sports Exerc*. 2007;39(8):1423–1434.
13. Nelson ME, Rejeski WJ, Blair SN, et al. Physical activity and public health in older adults: recommendation from the American College of Sports Medicine and the American Heart Association. *Circulation*. 2007;116(9):1094–1105.
14. 2008 Physical activity guidelines for Americans. Washington, DC: US Department of Health and Human Services; 2008. (http://www.health.gov/paguidelines/pdf/paguide.pdf).
15. Department of Health, Physical Activity, Health Improvement, and Prevention. *At Least Five Days a Week: Evidence on the Impact of Physical Activity and Its Relationship to Health.* Norwich, United Kingdom: The Stationery Office; 2004.
16. Hu FB, Li TY, Colditz GA, et al. Television watching and other sedentary behaviors in relation to risk of obesity and type 2 diabetes mellitus in women. *JAMA*. 2003;289(14): 1785–1791.
17. Manson JE, Greenland P, Lacroix AZ, et al. Walking compared with vigorous exercise for the prevention of cardiovascular events in women. *N Engl J Med*. 2002;347(10):716–725.
18. Blanck HM, McCullough ML, Patel AV, et al. Sedentary behavior, recreational physical activity, and 7-year weight gain among postmenopausal U.S. women. *Obesity (Silver Spring)*. 2007;15(6):1578–1588.
19. Hu FB, Leitzmann MF, Stampfer MJ, et al. Physical activity and television watching in relation to risk for type 2 diabetes mellitus in men. *Arch Intern Med*. 2001;161(12):1542–1548.
20. Lank NH, Vickery CE, Cotugna N, et al. Food commercials during television soap operas: what is the nutrition message? *J Community Health*. 1992;17(6):377–384.
21. Katzmarzyk PT, Church TS, Craig CL, et al. Sitting time and mortality from all causes, cardiovascular disease, and cancer. *Med Sci Sports Exerc*. 2009;41(5):998–1005.
22. Dunstan DW, Barr EL, Healy GN, et al. Television viewing time and mortality: the Australian Diabetes, Obesity and Lifestyle Study (AusDiab). *Circulation*. 2010;121(3): 384–391.
23. Inoue M, Iso H, Yamamoto S, et al. Daily total physical activity level and premature death in men and women: results from a large-scale population-based cohort study in Japan (JPHC Study). *Ann Epidemiol*. 2008;18(7):522–530.
24. Calle EE, Rodriguez C, Jacobs EJ, et al. The American Cancer Society Cancer Prevention Study II Nutrition Cohort: rationale, study design, and baseline characteristics. *Cancer*. 2002;94(2):500–511.
25. Garfinkel L. Selection, follow-up, and analysis in the American Cancer Society prospective studies. *Natl Cancer Inst Monogr*. 1985;67:49–52.
26. Calle EE, Terrell DD. Utility of the National Death Index for ascertainment of mortality among Cancer Prevention Study II participants. *Am J Epidemiol*. 1993;137(2):235–241.
27. World Health Organization. *International Classification of Diseases*. Ninth Revision. Vol 1. Geneva, Switzerland: World Health Organization; 1977.
28. World Health Organization. *International Statistical Classification of Diseases and Related Health Problems*. Tenth Revision. Vol 1. Geneva, Switzerland: World Health Organization; 1992.
29. Ainsworth BE, Haskell WL, Whitt MC, et al. Compendium of physical activities: an update of activity codes and MET intensities. *Med Sci Sports Exerc*. 2000;32(9 suppl):S498–S504.
30. Cox DR. Regression models and life tables (with discussion). *J R Stat Soc (B)*. 1972;34(2):187–220.
31. Block G, Coyl L, Smucker R, et al. *Health Habits and History Questionnaire: diet history and other risk factors* [personal computer system documentation]. Bethesda, MD: National Cancer Institute Division of Cancer Prevention and Control, National Institutes of Health; 1989.
32. Flagg EW, Coates RJ, Calle EE, et al. Validation of the American Cancer Society Cancer Prevention Study II Nutrition Survey Cohort food frequency questionnaire. *Epidemiology*. 2000;11(4):462–468.
33. Hu FB, Willett WC, Li T, et al. Adiposity as compared with physical activity in predicting mortality among women. *N Engl J Med*. 2004;351(26):2694–2703.
34. Ching PL, Willett WC, Rimm EB, et al. Activity level and risk of overweight in male health professionals. *Am J Public Health*. 1996;86(1):25–30.
35. Coakley EH, Rimm EB, Colditz G, et al. Predictors of weight change in men: results from the Health Professionals Follow-up Study. *Int J Obes Relat Metab Disord*. 1998;22(2): 89–96.
36. Bey L, Hamilton MT. Suppression of skeletal muscle lipoprotein lipase activity during physical inactivity: a molecular reason to maintain daily low-intensity activity. *J Physiol*. 2003;551(pt 2):673–682.
37. Hamilton MT, Hamilton DG, Zderic TW. Exercise physiology versus inactivity physiology: an essential concept for understanding lipoprotein lipase regulation. *Exerc Sport Sci Rev*. 2004;32(4):161–166.
38. Hamilton MT, Hamilton DG, Zderic TW. Role of low energy expenditure and sitting in obesity, metabolic syndrome, type 2 diabetes, and cardiovascular disease. *Diabetes*. 2007;56(11): 2655–2667.
39. Hu FB. Sedentary lifestyle and risk of obesity and type 2 diabetes. *Lipids*. 2003;38(2):103–108.
40. Knowler WC, Barrett-Connor E, Fowler SE, et al. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med*. 2002;346(6):393–403.
41. Lynch J, Helmrich SP, Lakka TA, et al. Moderately intense physical activities and high levels of cardiorespiratory fitness reduce the risk of non-insulin-dependent diabetes mellitus in middle-aged men. *Arch Intern Med*. 1996;156(12): 1307–1314.
42. Tuomilehto J, Lindström J, Eriksson JG, et al. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects

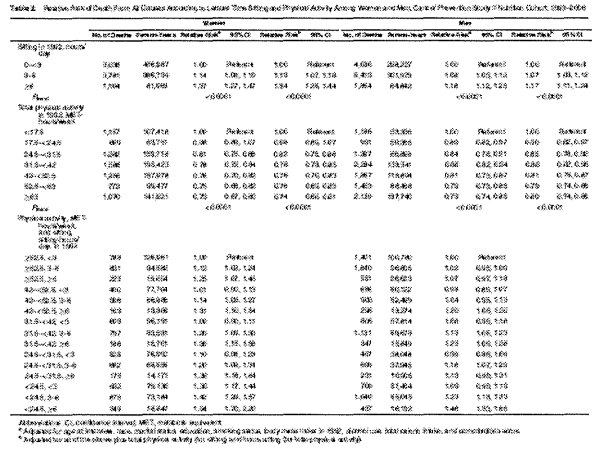
with impaired glucose tolerance. *N Engl J Med.* 2001; 344(18):1343–1350.
43. *Nielsen Reports Television Tuning Remains at Record Levels.* New York, NY: The Nielsen Company; 2007.
44. Trends in leisure-time physical activity by age , sex, and race/ethnicity—United States 1994–2004. *MMWR Morb Mortal Wkly Rep.* 2005;54(39):991–994.
45. Steffen LM, Arnett DK, Blackburn H, et al. Population trends in leisure-time physical activity: Minnesota Heart Survey, 1980–2000. *Med Sci Sports Exerc.* 2006;38(10):1716–1723.
46. Wolf AM, Hunter DJ, Colditz GA, et al. Reproducibility and validity of a self-administered physical activity questionnaire. *Int J Epidemiol.* 1994;23(5):991–999.
47. Chao A, Connell CJ, Jacobs EJ, et al. Amount, type, and timing of recreational physical activity in relation to colon and rectal cancer in older adults: the Cancer Prevention Study II Nutrition Cohort. *Cancer Epidemiol Biomarkers Prev.* 2004; 13(12):2187–2195.
48. Patel AV, Calle EE, Bernstein L, et al. Recreational physical activity and risk of postmenopausal breast cancer in a large cohort of US women. *Cancer Causes Control.* 2003;14(6): 519–529.
49. Patel AV, Feigelson HS, Talbot JT, et al. The role of body weight in the relationship between physical activity and endometrial cancer: results from a large cohort of US women. *Int J Cancer.* 2008;123(8):1877–1882.
50. Patel AV, Rodriguez C, Pavluck AL, et al. Recreational physical activity and sedentary behavior in relation to ovarian cancer risk in a large cohort of US women. *Am J Epidemiol.* 2006;163(8):709–716.

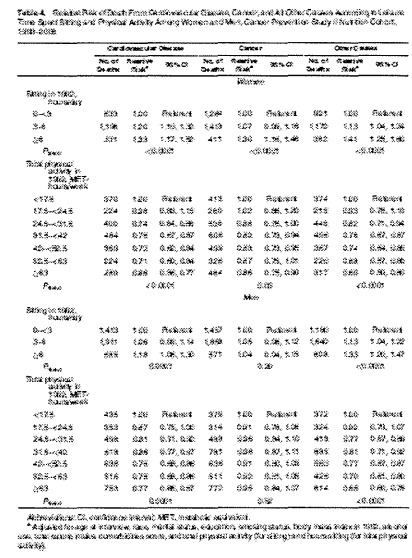
論文名	Leisure time spent sitting in relation to total mortality in a prospective cohort of US adults
著 者	Patel AV, Bernstein L, Deka A, Feigelson HS, Campbell PT, Gapstur SM, Colditz GA, Thun MJ
雑誌名	Am J Epidemiol
巻・号・頁	172(4) 419-429
発行年	2010
PubMedリンク	http://www.ncbi.nlm.nih.gov/pubmed/20650954

対象の内訳	ヒト	動物	地 域	欧米	研究の種類	縦断研究
対象	一般健常者	空白		( )		コホート研究
性別	男女混合	( )		( )		( )
年齢	男63.6歳(±6.0)	[redacted]	5)	( )		前向き研究
対象数	10000以上			( )		( )

調査の方法							
	質問紙	( )					
アウトカム	予 防	心疾患予防	なし	ガン予防	なし	死亡	( )
	維持・改善	なし	なし	なし	なし	( )	( )

図 表





図表掲載箇所	P423, Table2, P426, Table4
概 要 (800字まで)	本研究は、Americal Cancer SocietyによるNutrition Studyに参加している123,216名(男性53,440名、女性69,776名)を対象に、14年間の追跡調査を行い、余暇時間の不活動または身体活動量と総死亡における関連を検討したものである。余暇時間活動を評価するために、「過去一年間、仕事以外で、テレビや読書などの座位時間に一日何時間費やしたか。」「過去一年間、次のような活動を週あたり何時間行ったか(ウォーキング、ジョギング、水泳、テニス、自転車、エアロビクス、健康体操など)。」「過去一年間、次のような家事活動を週あたり何時間行ったか(ガーデニング、掃除、修繕、買い物など)。」といった質問を用いた。総死亡に関して、座位時間(不活動)3時間未満/日のグループと比較すると、座位時間6時間以上/日のグループは、女性でリスクが1.37(95%信頼区間:1.27-1.47)、男性で1.17(1.12-1.25)に上昇した。さらに、座位時間と身体活動量の複合効果として、座位時間3時間未満/日かつ身体活動量52.5MET時以上/週のグループと比較すると、座位時間6時間以上/日かつ身体活動量24.5MET時未満/週のグループは、女性で1.94(1.70-2.20)、男性で1.48(1.33-1.65)に上昇した。また、座位時間の長さは、男女共に心血管疾患による死亡リスクを有意に引き上げることがわかった。さらに、身体活動量の増加と心血管疾患による死亡リスクは逆相関を示すことがわかった。
結 論 (200字まで)	余暇時間における不活動(座位時間)は、身体活動量に関わらず総死亡のリスクを有意に引き上げることが明らかとなった。特に、心血管疾患による死亡との関連が最も有意であった。
エキスパートによるコメント (200字まで)	身体活動基準の策定に用いられた研究の1つである。身体不活動と死亡や疾患発症との関係については、近年非常に注目されている課題である。この身体不活動と死亡のリスクとの関係を12万以上の集団で追跡調査を行った非常に価値ある研究である。日本でも、このような不活動と死亡や様々な疾患発症との関係を明らかにすることが必要である。

担当者 久保絵里子・村上晴香・宮地元彦

International Journal of Obesity (2004) 28, 105–112
 0307-0565/04 $25.00
www.nature.com/ijo

PAPER

Longitudinal study of the long-term relation between physical activity and obesity in adults

L Petersen[1], P Schnohr[2] and TIA Sørensen[1]*

[1]*Danish Epidemiology Science Centre at Institute of Preventive Medicine, Copenhagen University Hospital, Copenhagen K DK-1399, Denmark; and* [2]*Copenhagen City Heart Study, Bispebjerg University Hospital, Copenhagen NV DK-2400, Denmark*

BACKGROUND: Earlier observational studies of the relation between physical activity and obesity are inconsistent and ambiguous, showing a clear cross-sectional inverse relation, and a prospective association only when physical activity at the time of follow-up is included.
OBJECTIVE: To examine the long-term effect of leisure time physical activity (LTPA) on subsequent development of obesity and the effect of body weight on later physical inactivity in a population-based longitudinal setting taking into account the effects of historical changes on future changes as well as pertinent confounders.
DESIGN: The study included 3653 women and 2626 men aged 20–78 y selected at random within sex–age strata from the general population of Copenhagen. At two surveys, 5 y apart, LTPA, body mass index (BMI) (weight/height^2, kg/m^2), several possible confounders and modifying factors were assessed. Obesity (defined as BMI $\geq$30 kg/m^2) and LTPA was assessed at the 3rd survey 10 y later. Odds ratios (with 95% confidence limits) for developing obesity between the last two surveys were estimated by logistic regression analysis, taking into account baseline and preceding changes in BMI and LTPA. A similar analysis of odds ratios for physical inactivity as outcome at the 3rd survey was conducted.
RESULTS: Compared to physical inactivity, the odds ratios of development of obesity among women with medium and high level of activity were 0.81 (0.53, 1.25) and 1.16 (0.73, 1.84), respectively, and among men, the odds ratios were 1.28 (0.71, 2.33) and 1.65 (0.91, 2.99), respectively. Compared to median BMI, the odds ratio of later physical inactivity among women with high BMI was 1.91 (1.39, 2.61), and among men the odds ratio was 1.50 (1.01, 2.22). The associations were not confounded or modified by age, pre-existing diseases, smoking, alcohol intake, educational level, occupational physical activity or by familial predisposition to obesity.
CONCLUSION: This study did not support that physical inactivity as reported in the freely living adult population in the long term is associated with the development of obesity, but the study indicates that obesity may lead to physical inactivity.
International Journal of Obesity (2004) **28**, 105–112. doi:10.1038/sj.ijo.0802548
Published online 25 November 2003

Keywords: physical activity; obesity; BMI; prevention; longitudinal

Introduction

The prevalence of obesity is rapidly increasing, and obesity has considerable adverse health effects. Since treatment often fails and since it is unfeasible to offer it in the magnitude requested, preventive measures are urgently needed.[1] Numerous cross-sectional studies have shown an inverse association between leisure time physical activity (LTPA) and obesity,[2,3] suggesting that physical inactivity may precede the development of obesity. Increased physical activity, and particularly avoidance of a sedentary lifestyle, is considered to be of paramount importance for prevention of obesity,[1] as well as a general health measure, because of the clear beneficial long-term effects on morbidity and mortality.[4–6] Recent thorough reviews have addressed the quantitative relationship between physical activity and weight gain or development of obesity and have come to the conclusion that there is evidence to support that physical activity levels that increase the total energy expenditure to above 1.7–1.8 times the basal metabolic rate are needed.[7,8] However, prospective observational population studies of adults, from the last 20 y with physical activity measured at baseline are few and have given inconsistent results with regard to the

*Correspondence: Dr TIA Sørensen, Institute of Preventive Medicine, Kommunehospitalet, Copenhagen K DK-1399, Denmark.
E-mail: tias@ipm.hosp.dk
Received 18 November 2002; revised 15 April 2003 and 30 June 2003; accepted 14 October 2003

effect of physical activity on body weight change and development of obesity.[2,7,8] It is also conceivable that the inverse cross-sectional relation may be due to a reduction of physical activity as a consequence of obesity, assuming that the discomfort of physical activity is worse the greater the overweight.

In investigations of these relationships, it is crucial to assess physical activity and obesity with the appropriate temporal sequence. Despite having available longitudinal, prospective or retrospective data on physical activity from baseline and body weight at follow-up, several authors report associations between concurrent changes.[2,7] The recent reviews of the problem do not make a clear distinction between studies respecting the temporal sequence, and studies reporting concurrent changes or inclusion of physical activity, assessed at the same time as obesity.[2,7–9] Any analysis that includes concurrent measurements of obesity and physical activity as outcome at the end of the observation period is essentially limited in a similar way as cross-sectional analyses. Associations between concurrent changes do not give information on the possible causal direction, as one or the other change could have arisen first and caused the other. A similar argument is valid for concurrent stability; if stable high activity is associated with concurrent lower risk of development of obesity, this can emerge from physical activity preventing obesity or it can emerge from stable body mass index (BMI) enabling the maintenance of high level of activity. However, the results of studies with an informative temporal sequence are inconsistent.[2,3,10,11] Moreover, future weight changes are dependent on current BMI and earlier changes in BMI,[12] and the physical activity habits at, and after, baseline may depend on preceding changes in physical activity. Fluctuations over time, in both physical activity and body weight, giving rise to the so-called regression-to-the mean phenomenon, may lead to misinterpretations of the relationships. In view of the well-documented concurrent inverse relation between BMI and physical activity, this means that both baseline BMI and preceding changes in BMI and physical activity should be taken into account in the analysis.

In the present study, the population was examined three times, providing the opportunity to consider the first two as combined baseline, with information of both level and changes of risk factors, of possible modifiers and confounders as baseline information. In this longitudinal setting, we analyzed the relationship between physical activity and BMI as a measure of obesity with each of the two being defined as an outcome and the other as a determinant. The focus was on LTPA, which, in contrast to occupational physical activity, may be easier to modify.

Materials and methods

The study population

For the Copenhagen City Heart Study,[13] a sex- and age-stratified (20–93 y of age) random sample was drawn from the Copenhagen Population Register among people living within defined areas in Copenhagen. For the 1st survey, which occurred between 1976 and 1978, 19 329 men and women were invited; 14 151 participated. After 5 y (1981–1983), 11 085 returned for a 2nd survey; 929 had died, 26 emigrated and 2111 did not respond. After 10 y (1992–1993), 6542 returned for the 3rd survey; 2329 had died, 37 emigrated and 2177 did not respond.

The study sample

Among the subjects examined in all three surveys, complete information on weight, height and LTPA was available in a sample of 6279 subjects. We excluded 725 who were obese (BMI $\geq 30\,kg/m^2$) at 2nd survey from analyses of obesity at the 3rd survey as outcome, leaving 5554 subjects. For the analyses of physical inactivity, the corresponding samples were used without the exclusion of those already obese at baseline; 433 did not respond to the question on LTPA at 3rd survey, leaving 5846 subjects for the analyses.

Anthropometric data and covariates

Height was measured without shoes, to the nearest half centimeter. Body weight was measured to the nearest decimal in kilograms on a fixed balance scale with the subject wearing light indoor clothing, but without shoes. BMI was calculated as the weight (kg) per height squared (m^2).

From the surveys, we have information from self-administered questionnaires about factors known or assumed to be related to obesity,[14,15] including LTPA and occupational physical activity, smoking habits, length of education and, from the 3rd survey, parents' height and weight.

Physical activity in leisure time was graded in four levels based on a questionnaire constructed by Saltin and Grimby[16] with minor modifications: (1) *Physical inactivity*: almost entirely sedentary (reading, TV, cinema) or light physical activity less than 2 h per week; (2) *Light physical activity*: 2–4 h per week, for example, walking, cycling, light gardening; (3) *Moderate physical activity*: more than 4 h per week or more vigorous activity 2–4 h per week, for example, brisk walking, fast cycling, heavy gardening, sports where you get sweaty or exhausted; and (4) *Highly vigorous physical activity*: more than 4 h per week or regular heavy exercise or competitive sports several times per week. The questionnaire has been validated with respect to maximal oxygen uptake, which is increasing significantly from low to high level.[17] In this study, the group at level 4 was too small to be kept separate, and it was therefore included in the group of those at level 3.

Smoking status was categorized as never smoker, ex-smoker and three levels of current smokers. Alcohol consumption in drinks per week was grouped into: less than 1, 1–6, 7–13 and 14 or above for women and less than 1, 1–6, 7–13, 14–27 and 28 or above for men. Educational level was grouped into: less than 8 y, 8–11 y and 12 or more years. Predisposition to obesity was assessed on the basis of

reported height and weight of parents at the time when the participants were in school: nonpredisposed had no parents with BMI $\geq 30\,kg/m^2$, predisposed had at least one parent with BMI $\geq 30\,kg/m^2$, and a third group, too numerous to be excluded, did not report on parents' height and weight. Occupational physical activity is used at three levels; low, medium and high activity, corresponding to sitting, standing/walking and walking/lifting or more physically exacting activity.

Chronic disease

Data on chronic diseases were obtained from the surveys and from linkage to an external register and used as described earlier.[18] We identified heart disease, stroke, chronic pulmonary disease, intermittent claudication and hypertension, occurring before the 3rd survey from the questionnaires, physical examinations and hospital discharges. From these combined sources of information, we identified subjects with pre-existing disease, defined as disease that had occurred before the 3rd survey.

Statistical methods

Logistic regressions were used in all analyses. Estimated odds ratios are given with 95% confidence limits (CIs). Data analyses were performed using Statistical Analysis System (SAS, version 8). In all analyses both sexes were included in the same model, but always as an interaction between sex and LTPA, which provided separate estimates for the two sexes. In the analyses of the cross-sectional relation between LTPA and obesity in each survey, we adjusted for age in three levels, allowing the age effect to depend on gender. In a logistic regression model, the odds of developing obesity between 2nd and 3rd survey (having BMI $\geq 30\,kg/m^2$ at 3rd survey among those with BMI below 30 at 2nd survey), was modelled. After having excluded any major differences in effects across age strata, we adjusted for age at the 2nd survey, at three levels (–50 y, 51–60 y, 61 y–), and allowed the age effect to depend on gender. We adjusted for BMI at the 2nd survey (continuous variable) and allowed its effect to vary depending on the quintiles of changes in BMI from 1st to 2nd survey. By including the change in BMI rather than the BMI at each of the two surveys, we avoided colinearity between the two BMI measures and allow for the possibility to include an interaction term. The choice of the most parsimonious model with regard to the combinations of the various variables constructed was based on the Akaike Information Criterion (AIC),[19] AIC is a measure used to compare models, which are not hierachical submodels of each other.

Models adjusted for sex, age and BMI are referred to in the tables as *basic models*. As possible modifying and/or confounding variables we tested for age, earlier BMI, smoking status, alcohol consumption, educational level, occupational physical activity and familial predisposition to obesity. For further exploration higher threshold values of obesity than BMI of $30\,kg/m^2$, was used as outcome in the final model, namely BMI $\geq 32\,kg/m^2$. For analyses of BMI and later physical inactivity, we used methods equivalent to those described, except that we did not exclude those who were obese or who were physically inactive at 2nd survey. The explanatory variables of main interest, that is BMI at 2nd survey, were used in quintiles. Cut points were BMI of 21.4, 23.1, 25.0 and $27.7\,kg/m^2$ for women and 23.0, 24.8, 26.5 and $28.7\,kg/m^2$ for men. Further confounder adjustments are indicated in the respective tables.

Results

Table 1 shows the distribution of age, BMI and proportion of obese in the sample used in the longitudinal analyses.

The three surveys each showed a concurrent inverse association between LTPA and obesity (Table 2). Odds ratios of obesity for active subjects were half of those for less active and the difference was highly significant.

As preparation for the model-based prospective analysis, we calculated the percentage becoming obese between 2nd and 3rd survey by level of LTPA at 2nd survey within strata of gender and overweight. Among men with BMI between 25 and $30\,kg/m^2$ in the 2nd survey the percentage developing obesity before 3rd survey was 11.7, 15.2 and 16.7% for those with low, medium and high LTPA, respectively, at 2nd survey. The corresponding values for women were 26.4, 23.1 and 27.7%. For men and women with a BMI below $25\,kg/m^2$, neither showed a consistent trend, with percentages ranging between 0.0 and 1.9. These crude data did not lend support to an inverse association between LTPA and later obesity. Odds ratios of obesity at 3rd survey crosstabulated by LTPA from 1st and 2nd survey are presented in Table 3. There is no major difference for fixed level of activity in the 2nd survey,

Table 1 Distribution of age and BMI, percentage distribution of LTPA and prevalence of obesity (BMI $\geq 30\,kg/m^2$)

	Women, N = *3653*	*Men*, N = *2626*
Age (y) at 2nd, median (range)	55 (26–79)	54 (25–83)
BMI (kg/m^2) at 2nd, median (range)	24.0 (15.2–45.1)	25.6 (16.3–47.2)
BMI (kg/m^2) at 2nd, median (range)[a]	23.5 (15.2–30.0)	25.1 (16.3–30.0)
LTPA at 2nd		
Inactive	13.2	12.6
Medium	56.3	44.4
High	30.4	43.0
LTPA at 3rd		
Missing	6.8	7.0
Inactive	12.1	11.4
Medium	56.5	45.8
High	24.6	35.9
Stable activity level (%)	54.7	51.4
Decreased activity level (%)	24.3	26.4
Increased activity level (%)	20.9	22.2
Obese at 3rd (%)	16.8	16.6
Obese at 3rd (%)[a]	9.2	8.1

[a]Only the subjects not obese at 2nd survey.

Table 2 Odds ratios with 95% CIs for obesity from the cross-sectional analysis of LTPA at each survey

			LTPA			
	Survey		*Low*	*Medium*	*High*	*P-value for trend*
Women						
	1st	7595	1	0.70 (0.59, 0.83)	0.51 (0.40, 0.64)	0.0001
	2nd	6935	1	0.75 (0.63, 0.91)	0.58 (0.47, 0.72)	0.0001
	3rd	4986	1	0.61 (0.49, 0.76)	0.36 (0.27, 0.47)	0.0001
	3rd[a]	3019	1	0.70 (0.49, 1.00)	0.41 (0.27, 0.63)	0.005
Men						
	1st	6395	1	0.71 (0.58, 0.85)	0.65 (0.52, 0.80)	0.0001
	2nd	5606	1	0.87 (0.70, 1.08)	0.76 (0.61, 0.95)	0.01
	3rd	3907	1	0.71 (0.54, 0.92)	0.52 (0.39, 0.68)	0.0001
	3rd[a]	2123	1	0.57 (0.37, 0.88)	0.39 (0.25, 0.63)	0.005

Adjusted for age, occupational physical activity, length of education, smoking and alcohol habits. [a]Only the subjects present at all three surveys and not obese at 2nd survey.

Table 3 Odds ratios of becoming obese between 2nd and 3rd survey and 95% CIs

	LTPA 2nd survey		
1-2-3 survey	*Low*	*Medium*	*High*
Women			
LTPA 1st survey			
Low	1	0.90 (0.39, 2.09)	1.65 (0.59, 4.61)
Medium	1.15 (0.50, 2.60)	0.91 (0.46, 1.80)	1.38 (0.66, 2.88)
High	0.38 (0.07, 2.07)	0.88 (0.40, 1.96)	1.05 (0.48, 2.26)
All	1	0.93 (0.59, 1.45)	1.35 (0.83, 2.18)
Men			
LTPA 1st survey			
Low	1	1.45 (0.51, 4.12)	2.47 (0.70, 8.76)
Medium	0.88 (0.26, 3.04)	1.28 (0.52, 3.16)	1.79 (0.70, 4.58)
High	1.12 (0.26, 4.76)	1.22 (0.45, 3.29)	1.64 (0.66, 4.11)
All	1	1.35 (0.73, 2.50)	1.98 (1.03, 3.60)

Adjusted for age and BMI at 1st and 2nd survey, occupational physical activity, length of education, smoking, alcohol habits and familial predisposition to obesity, cross tabulated by LTPA at 1st and 2nd survey.

whereas odds ratios for fixed level at 1st survey are increasing for increased activity at 2nd survey, especially in men. The interaction between the two measures of LTPA was, however, statistically insignificant ($P=0.70$), and so was the relation to the LTPA at 1st survey ($P=0.52$), wherefore the results mainly reflect the effects of LTPA at 2nd survey.

There was a significant direct association between level of LTPA at 2nd survey and development of obesity before 3rd survey ($P=0.03$), and there was no significant modifying effect of gender ($P=0.47$), but since the results appear different they are presented separately for men and women (Table 4). Among women, odds ratios were close to one with regard to the medium as well as high activity with no trend, whereas among men there was an insignificant tendency to increased odds ratio with a higher level of LTPA.

Neither age nor earlier BMI modified the effects of LTPA (data not shown). The effects of smoking status, alcohol consumption, educational level, occupational physical activity and familial predisposition to obesity were analyzed allowing for interaction with LTPA to find possible

Table 4 Odds ratios of becoming obese between 2nd and 3rd survey and corresponding 95% CIs

		LTPA			
	N	*Low*	*Medium*	*High*	*P-value for trend*
Women					
Basic model	3256	1	0.81 (0.53, 1.25)	1.16 (0.73, 1.84)	0.23
Adjusted[a]	3256	1	0.84 (0.55, 1.30)	1.23 (0.77, 1.98)	0.15
Fully adjusted[b]	3246	1	0.93 (0.59, 1.45)	1.35 (0.83, 2.18)	0.09
Free of disease[c]	821	1	1.15 (0.36, 3.64)	1.99 (0.60, 6.62)	0.14
Men					
Basic model	2298	1	1.28 (0.71, 2.33)	1.65 (0.91, 2.99)	0.06
Adjusted[a]	2298	1	1.33 (0.73, 2.42)	1.78 (0.97, 3.26)	0.04
Fully adjusted[b]	2284	1	1.35 (0.73, 2.50)	1.93 (1.03, 3.60)	0.02
Free of disease[c]	674	1	5.55 (0.61, 50.7)	6.32 (0.70, 57.3)	0.15

Adjusted for sex, age and earlier BMI, according to LTPA at 2nd survey. [a]Adjusted also for LTPA at 1st survey. [b]Adjusted as above and also for occupational physical activity, length of education, smoking, alcohol habits and familial predisposition to obesity at 2nd survey. [c]Free of pre-existing diseases at 3rd survey and fully adjusted.

Table 5 Odds ratios of physical inactivity at 3rd survey and corresponding 95% CIs

		BMI quintile at 2nd survey					*P-value for trend*[d]
	N	*1*	*2*	*3*	*4*	*5*	
Women							
Basic model	3403	0.74 (0.51, 1.08)	1.19 (0.84, 1.69)	1	1.09 (0.77, 1.55)	1.91 (1.39, 2.61)	<0.0001
Adjusted[a]	3403	0.74 (0.51, 1.08)	1.21 (0.85, 1.71)	1	1.09 (0.77, 1.55)	1.85 (1.35, 2.55)	<0.0001
Fully adjusted[b]	3392	0.72 (0.49, 1.06)	1.25 (0.88, 1.77)	1	1.12 (0.79, 1.59)	1.87 (1.35, 2.59)	<0.0001
Free of disease[c]	842	0.89 (0.39, 2.07)	1.21 (0.54, 2.72)	1	0.61 (0.24, 1.54)	1.35 (0.60, 3.02)	0.57
Men							
Basic model	2443	0.93 (0.61, 1.43)	1.02 (0.67, 1.56)	1	1.35 (0.91, 2.02)	1.50 (1.01, 2.22)	0.01
Adjusted[a]	2443	0.93 (0.61, 1.43)	1.04 (0.68, 1.59)	1	1.36 (0.91, 2.04)	1.47 (0.99, 2.19)	0.01
Fully adjusted[b]	2429	0.90 (0.58, 1.39)	1.08 (0.70, 1.67)	1	1.38 (0.92, 2.07)	1.48 (0.99, 2.22)	0.01
Free of disease[c]	708	0.93 (0.35, 2.46)	1.08 (0.39, 2.96)	1	2.36 (0.91, 6.12)	2.67 (1.02, 7.02)	0.01

Adjusted for age and earlier LTPA, according to BMI at 2nd survey. [a]Adjusted as above and also for BMI change between 1st and 2nd survey. [b]Adjusted as above and also for occupational physical activity, length of education, smoking, alcohol habits and familial predisposition to obesity at 2nd survey. [c]Free of pre-existing diseases at 3rd survey and fully adjusted. [d]Tests for trend were carried out by including the median BMI for each quintile as a continuous variable.

modifying effects on the relation between LTPA and later obesity, but none of them were statistically significant (all $P>0.10$). To explore possible residual confounding from age, the three age groups were each split into two, which did not change the results. None of the other risk factors confounded the association of major interest, since estimated odds of becoming obese did not change noticeably when adjusting for either of them separately (data not shown) or when adjusting for all of them (Table 4).

Excluding those with pre-existing diseases at 3rd survey reduces the material considerable, and leaves only 1495 subjects for the analysis, and hence wider confidence limits, but the odds ratios still suggest a direct association between LTPA and becoming obese among healthy subjects, particularly among men.

Using BMI $\geq 32.0\,kg/m^2$ as the definition of obesity, 5.8% women and 4.7% men developed obesity between 2nd and 3rd survey. Odds ratios were unchanged among women, whereas among men the higher threshold weakened the positive association (data not shown).

The crude relation between physical inactivity at the 3rd survey and earlier BMI showed an increasing percentage through the quintiles of BMI at the 2nd survey in both sexes, for women from 8.0 to 21.4% and for men from 10.0 to 15.6%. Model-based estimates of odds ratios of physical inactivity at 3rd survey by level of BMI quintile at 2nd survey are presented in Table 5. We found consistent tendencies that low BMI is associated with lower odds of inactivity, whereas higher BMI was significantly associated with higher odds of inactivity irrespective of which type of adjustment was applied in the analysis.

Discussion

From the 15-y longitudinal analyses with the first two surveys as combined baseline, we found no evidence that physical inactivity promotes the development of obesity. On the contrary, a reverse tendency was present, namely that among the more active subjects there were more obese later on. This finding was not explained by pre-existing diseases. Actually, estimates were strengthened in those free of pre-existing disease, and it was not modified by age, earlier BMI, alcohol consumption, level of education, occupational physical activity nor familiar predisposition to obesity. We found an insignificant gender difference, suggesting no association between LTPA and obesity in women and a weak direct association among men.

As expected, the cross-sectional inverse association between LTPA and obesity was strong and statistically significant: the more the activity the lower the odds ratios of concurrent obesity. The results were consistent in all three cross-sectional analyses as well as in the selected material of those present at all three surveys and nonobese at 2nd survey. Our findings indicate that BMI is a strong determinant of later LTPA; the greater the BMI the greater the risk of being physically inactive 10 y later, also when adjustments were made for previous LTPA and possible confounders of the relation.

As recently reviewed by Fogelholm and Kukkonen-Hajula,[2] our findings are compatible with the other published studies on the relation between physical activity and later obesity that respect the temporal sequence of possible cause and effect. There are three studies showing no association.[3,20,21] Two studies show direct relations,[10,11] one among men, the other in both genders, but only with regard to sports activity. In three studies, the expected inverse relation[11,22,23] was found, the two of them in women only, and the third study without adjustment for baseline BMI, which implies that the observed effect could also be a result of the effect of BMI on physical activity. There was no systematic relation between the duration of follow-up and the direction and strength of the association of interest. Thus, in earlier studies, when disregarding the number of years of follow-up, an inverse relation in women cannot be excluded, whereas in men the

relation, if any, surprisingly may be direct. In a new study, using labelled water and indirect calorimetry for the assessment of energy expended by physical activity and level of physical activity in Pima Indians, neither measures were correlated with later changes in body weight.[24] None of these studies took into account the preceding changes in BMI or physical activity, and none have explicitly addressed the reverse analysis of BMI as a possible determinant of later LTPA. A few studies did use LTPA at follow-up as 'predictor' for preceding weight changes, and found that low activity was associated with higher weight gain,[3,21] but this type of analysis does not allow an elucidation of the temporal sequence of the relations. These fundamental problems in study design were unfortunately not dealt with in the more recent otherwise thorough reviews of the evidence for a protective effect of various levels of physical activity on the risk of development of obesity.[2,7,8]

Our study has the advantages of being longitudinal in design with prospective data collection in a fairly large population sample, and measures of height and weight are objective; thus, there cannot be recall bias in the LTPA measure and no misclassification in BMI because of possible differential misreporting. The main limitations of the study are the possible selection bias due to the attrition of the cohort over time, the possible dilution of the effects due to measurement errors and the possible mutual irrelevance of variables measured at such long time intervals as in this study.

Any population-based study running over so many years as the present one, will be subject to attrition, which may introduce a selection bias in the results. When the subjects are called for a new examination, obviously attendance will be related to characteristics of the subjects—including vital status, illnesses, general health status, lifestyle and distance of current residence. On the other hand, proper evaluation of the risk of obtaining a biased estimate of the associations between characteristics at different points in time requires careful consideration of which type of sample attrition could produce a bias. If the expected relationship between physical inactivity and later development of obesity does exist, and our finding of no association is spurious, then this would require that the participation at the 3rd survey, for a given level of physical activity 10 y earlier, is different for subjects who developed obesity compared to those who did not. This could be the case, not least because of the long-term positive health effect of physical activity. On the other hand, the consistent and clear inverse relation between physical activity at all three surveys makes this bias less likely. It is also worth noting the quantitative aspects of such possible bias. For example, to find the expected inverse relation of LTPA at the 2nd survey and later obesity among men with BMI between 25 and 29 kg/m^2, then (based on the crude percentages) more than 29.0% of the future nonattendants reporting inactivity should develop obesity compared to the 11.7% who did so among attendants. Likewise, among women with medium activity in 2nd survey, more than 42.8% nonattendants should have developed obesity compared to the 23.1% who did so among attendants. Thus, the selection bias should be severe to actually reverse the association observed. Furthermore, the distribution of variables at the 2nd survey among future nonattendants and those attending the 3rd survey did not support a strong selection bias. We found no major differences in BMI and in LTPA and no differences in the cross-sectional relation between activity and obesity between future non-attendants and future attendants at the 3rd survey (data not shown).

It is an important question as to whether our inability to find the expected inverse relationship between physical activity and later obesity is due to too crude a measure of physical activity without distinction between no and very low activity. The statistical tools used in the present study, that is, logistic regression models, do not take into account measurement errors on the explanatory variable. This is obviously less of a problem when BMI is the explanatory variable than when LTPA is so, although BMI may also be considered a proxy measure of obesity. Therefore, we may have less confidence in the results of LTPA not being able to predict later obesity, than for the results on BMI strongly predicting future inactivity. Measurement errors of explanatory variables in a complex model may cause unpredictable bias in the estimated effects, but in the crude tabulations, where LTPA is the only explanatory variable, measurement error will weaken associations, and not reverse them.[25] Therefore, the finding that the crude percentages of those developing obesity before the 3rd survey for different levels of LTPA at 2nd survey shows the same tendency as the complex models, suggesting that measurement errors did not produce our results. Furthermore, as seen in the cross-sectional analyses, the measure of LTPA used here was able to demonstrate the strong inverse correlation with concurrent BMI, and the detailed study of Pima Indians showed the same.[24] The finding of the plausible crude as well as multivariate adjusted relation between BMI at the 2nd survey and physical inactivity at the 3rd survey also attests to the usefulness of the measure of the LTPA, as does the observation in the same study population of a strong predictive value of LTPA at 1st and 2nd survey with regard to long-term total mortality.[26,27]

The third potential limitation of this study is the long time interval between the measurements, which could have dilution effects on the results in the same way as measurement errors. In this population, more than 50% have the same level of LTPA in two successive surveys, about 20% increased their level and about 25% reduced their level of LTPA, with the same pattern in women as in men. It is conceivable that reduced physical activity in individuals with a former high LTPA level could be a cause of obesity. However, reducing activity from high LTPA during the intervening years between 1st and 2nd survey was not associated with increased odds of obesity at the 3rd survey. The above-mentioned long-term predictive effects of LTPA

on mortality[26,27] is also in this context a support of the contention that LTPA as assessed at the 2nd survey should show a relation to later development of obesity, had such relationship existed. A closer examination of short-term effects would require follow-up of a cohort at shorter intervals. On the other hand, in a public health perspective it would be essential to demonstrate the long-term relation in a free-living population as also emphasized in the recent reviews.[7,8]

Our results do not support the fact that medium or high physical activity at baseline prevents obesity in the long term. This contradicts the intuitive ideas derived from the energy balance equation and from the observed cross-sectional association. On the other hand, it may be questioned whether it is compatible with the thermodynamic law underlying the energy balance equation to put the question whether physical activity levels at a given point in time are related to later risk of developing obesity. The energy balance equation tells us that changes in physical activity can result in weight change if energy intake does not counterbalance the changes in energy expenditure, which may occur in short-term studies even in free-living subjects.[28] However, it cannot be inferred from the law of the energy balance equation and such short-term experiments as to how the cumulative regulation of the energy balance works in the long term in large populations of freely living individuals. The conclusions drawn in the recent reviews, although not fully adequate from the point of view of the problems discussed here, attest to the relevance of investigating whether various levels of physical activity are related to later risk of development of obesity, irrespective of the lacking control of the energy intake.[7,8] Furthermore, development of obesity corresponds to a very small positive energy balance, usually less than 1% of the total energy turnover, which by itself suggests that factors other than the great differences in LTPA between individuals should be considered.

When obesity has developed, the energy turnover increases and the energy requirement supporting a given physical activity increases as well.[1] In addition, it seems likely that a given level of physical activity elicits on average more discomfort, for example as musculoskeletal complaints, dyspnoea, exhaustion and sweating, the greater the overweight. This may reduce the motivation for physical activity and eventually reduce the actual physical activity.

Although our study does not exclude a short-term effect of LTPA on accumulation of fat in the adipose tissue, our results do not support a long-term effect of physical activity on the risk of later development of obesity. Our study indicates, on the other hand, that the opposite causal direction is operating, namely that obesity leads to less physical activity, a finding that strongly encourages future studies of the prospective relation between physical activity and obesity to integrate the possible temporality of the phenomena in the design and analysis.

Acknowledgements

The study was supported by the Danish Medical Research Council, The Danish Heart Foundation and the Danish National Science Foundation. We thank the staff at the Copenhagen City Heart Study for the skilful examination of the subjects.

References

1 World Health Organistion. *Obesity, preventing and managing the global epidemic. Report of a WHO consultation on obesity*, WHO Tecnical Report Series (no. 894) Geneva; 2000.
2 Fogelholm M, Kukkonen-Hajula K. Does physical activity prevent weight gain—a systematic review. *Obes Rev* 2000; **1**: 95–111.
3 Williamson DF, Madans J, Anda RF, Kleinman JC, Kahn HS, Byers T. Recreational physical activity and ten-year weight change in a US national cohort. *Int J Obes* 1993; **17**: 279–286.
4 Paffenbarger RS, Hyde RT, Wing AL, Hsieh C. Physical activity, all-cause mortality, and longevity of college alumni. *N Engl J Med* 1986; **314** (10): 605–613.
5 Erlichman J, Kerbey AL, James WPT. Physical activity and its impact on health outcomes. Paper 1: the impact of physical activity on cardiovascular disease and all-cause mortality: an historical perspective. *Obes Rev* 2002; **3**: 257–271.
6 Andersen LB, Schnohr P, Schroll M, Hein HO. All-cause mortality associated with physical activity during leisure time, work, sport, and cycling to work. *Arch Intern Med* 2000; **160**: 1621–1628.
7 Erlichman J, Kerbey AL, James WPT. Physical activity and its impact on health outcomes. Paper 2: prevention of unhealthy weight gain and obesity by physical activity: an analysis of the evidence. *Obes Rev* 2002; **3**: 273–287.
8 Saris WHM, Blair SN, van Baak MA, Eaton SB, Davies PSW, Di Pietro L, Fogelholm M, Rissanen A, Schoeller D, Swinburn B, Trembley A, Westerterp KR, Wyatt H. How much physical activity is enough to prevent unhealthy weight gain? Outcome of the IASO 1st Stick Conference and consensus statement. *Obes Rev* 2003; **4**: 101–114.
9 Coakley EH, Rimm EB, Colditz G, Kawachi I, Willett W. Predictors of weight change in men: results from The Health Professionals Follow-up Study. *Int J Obes* 1998; **22**: 89–96.
10 Bild DE, Sholinsky P, Smith DE, Lewis CE, Hardin JM, Burke GL. Correlates and predictors of weight loss in young adults: the CARDIA study. *Int J Obes* 1996; **20**: 47–55.
11 Klesges RC, Klesged LM, Haddock CK, Eck LH. A longitudinal analysis of the impact of dietary intake and physical activity on weight change in adults. *Am J Clin Nutr* 1992; **55**: 818–822.
12 Colditz GA, Willett WC, Stampfer MJ, London SJ, Segal MR, Speizer FE. Patterns of weight change and their relation to diet in a cohort of healthy women. *Am J Clin Nutr* 1990; **51**: 1100–1105.
13 Schnohr P, Jensen G, Lange P, Scharling H, Appleyard M. The Copenhagen City Heart Study. *Eur Heart J* 2001; **3**: H1–H83.
14 Rissanen AM, Heliövaara M, Knekt P, Reunanen A, Aromaa A. Determinants of weight gain and overweight in adult Finns. *Eur J Clin Nutr* 1991; **45**: 419–430.
15 Sørensen TIA, Holst C, Stunkard AJ, Skovgaard LT. Correlations of body mass index of adult adoptees and their biological and adoptive relatives. *Int J Obes* 1992; **16**: 227–236.
16 Saltin B, Grimby G. Physiological analysis of middle-aged and old former athletes. Comparison with still active athletes of the same ages. *Circulation* 1968; **38**: 1104–1115.
17 Saltin B. Physiological effects of physical conditioning. In: Hansen AT, Schnohr P, Rose G (eds). *Ischaemic Heart Disease. The Strategy of Postponement*. Year Book Medical Publishers and Copenhagen, FADL's Forlag: Chicago; 1977. pp 104–115.

18 Mikkelsen KL, Heitmann BL, Keiding N, Sørensen TIA. Independent effects of stable and chancing body weight on total mortality. *Epidemiology* 1999; **10**: 671–678.
19 Clayton D, Hills M. *Statistical Models in Epidemiology*. Oxford University Press: Oxford; 1993.
20 Parker DR, Gonzalez S, Derby CA, Gans KM, Lasater TM, Carleton RA. Dietary factors in relation to weight change among men and women from two southeastern New England communities. *Int J Obes* 1997; **21**: 103–109.
21 Crawford DA, Jeffery RW, French SA. Television viewing, physical activity and obesity. *Int J Obes* 1999; **23**: 437–440.
22 Owens JF, Matthews KA, Wing RR, Kuller LH. Can physical activity mitigate the effect of aging in middle-aged women? *Circulation* 1992; **85**: 1265–1270.
23 Haapanen N, Miilunpalo S, Pasanen M, Oja P, Vuori I. Association between leisure time physical activity and 10-year body mass change among working-aged men and women. *Int J Obes* 1997; **21**: 288–296.
24 Tataranni PA, Harper I, DelParigi, Snitker S, Ravussin E. Physical activity and obesity: findings from cross-sectional and prospective studies. *Int J Obes* 2001; **25**: O45.
25 Carroll RJ, Rupert D, Stefanski LA. *Measurement Error in Nonlinear Models*. Chapman & Hall: London; 1995.
26 Andersen LB, Schnohr P, Schroll M, Hein HO. All-cause mortality associated with physical activity during leisure time, work, sports and cycling to work. *Arch Intern Med* 2000; **160**: 1621–1628.
27 Schnohr P, Scharling H, Jensen JS. Change in Leisure-time Physical Activity and Risk of Death: An Observational Study of 7,000 Men and Women. *Am J Epidemiol* 2003; **158**: 639–644.
28 Stubbs RJ, Sepp A, Hughes DA, Johnstone AM, King N, Horgan G, Blundell JE. The effect of graded levels of exercise on energy intake and balance in free-living women. *Int J Obes* 2002; **26**: 866–869.

論文名	Longitudinal study of the long-term relation between physical activity and obesity in adults.						
著 者	Petersen L, Schnohr P, Sorensen TI						
雑誌名	Int J Obes Relat Metab Disord						
巻・号・頁	28(1) 105-12						
発行年	2004						
PubMedリンク	http://www.ncbi.nlm.nih.gov/pubmed/14647181						
対象の内訳		ヒト	動物	地 域	欧米	研究の種類	縦断研究
	対象	一般健常者	空白		( )		コホート研究
	性別	男女混合	( )		( )		( )
	年齢	20－78歳			( )		前向き研究
	対象数	5000～10000	空白		( )		( )
調査の方法	実測	( )					
アウトカム	予 防	な し	肥満予防	な し	な し	( )	( )
	維持・改善	な し	な し	な し	な し	( )	( )
図 表							
図表掲載箇所							
概 要 (800字まで)	理論的背景:身体活動と肥満との関係に関するこれまでの観察研究結果には矛盾があり、あいまいであった。 目的:適切な交互作用や時間的な流れを考慮して、肥満のその後の進展に対するレジャーにおける身体的活動(LTPA)の長期間の影響と、体重が後の身体不活動へ及ぼす縦断的な影響を調べる. デザイン:Copenhagen City Heart Studyのために、コペンハーゲンの一般住民から20－78歳の3653人の女性と2626人の男性を性・年齢階層別に無作為に選んだ。1回目の調査は1976～1978年で、5年後に行われた2回目の調査時点で肥満でなかった者について、更に5年後の3回目の調査まで観察した。測定項目はLTPA、BMI、いくつかの交絡因子である。肥満(BMIが30kg/m2以上)とLTPAは、10年後の3回目の調査で評価された。 最後の2つの調査の間に増加した肥満のオッズ比(95% CI)は、BMIとLTPAのベースライン値とその前の変化量を考慮して、ロジスティクス回帰分析で推定された。3回目の調査の結果としての身体不活動のオッズ比も、同様の分析で行われた。 結果:低活動レベル群と比較して、中活動レベル群と高活動レベル群が肥満(BMI > 30 kg/m2)になるオッズ比は、女性で0.81(0.53-1.25)と1.16(0.73-1.84)、女性で1.28(0.71-2.33)と1.65(0.91-2.99)だった。一方、BMI中群と比較してBMI高群における低活動レベルのオッズ比は、女性で1.91(1.39-2.61)、男性で1.50(1.01-2.22)であった。						
結 論 (200字まで)	この研究では、長期間にわたる成人の日常生活において、身体活動量が少ないことは肥満の進展につながっていなかったが、逆に肥満は不活発な身体活動に通じるかもしれない。						
エキスパートによるコメント (200字まで)	身体活動が肥満の発現と関連しているのではなく、肥満が不活動をもたらすことが示唆され、興味深い。						

担当者 呉泰雄, 田中茂穂

Factors Associated With 5-Year Risk of Hip Fracture in Postmenopausal Women

John Robbins, MD
Aaron K. Aragaki, MS
Charles Kooperberg, PhD
Nelson Watts, MD
Jean Wactawski-Wende, MD
Rebecca D. Jackson, MD
Meryl S. LeBoff, MD
Cora E. Lewis, MD
Zhao Chen, PhD
Marcia L. Stefanick, PhD
Jane Cauley, DrPH

Context The 329 000 hip fractures that annually occur in the United States are associated with high morbidity, mortality, and cost. Identification of those at high risk is a step toward prevention.

Objective To develop an algorithm to predict the 5-year risk of hip fracture in postmenopausal women.

Design, Setting, and Participants A total of 93 676 women who participated in the observational component of the Women's Health Initiative (WHI), a multiethnic longitudinal study, were used to develop a predictive algorithm based on commonly available clinical features. Selected factors that predicted hip fracture were then validated by 68 132 women who participated in the clinical trial. The model was tested in a subset of 10 750 women who had undergone dual-energy x-ray absorptiometry (DXA) scans for bone mass density assessment.

Main Outcome Measure The prediction of centrally adjudicated hip fracture, measured by the area under the receiver operator characteristic (ROC) curves.

Results During a mean (SD) follow-up of 7.6 (1.7) years, 1132 hip fractures were identified among women participating in the observational study (annualized rate, 0.16%), whereas during a mean follow-up of 8.0 (1.7) years, 791 hip fractures occurred among women participating in the clinical trial (annualized rate, 0.14%). Eleven factors predicted hip fracture within 5 years: age, self-reported health, weight, height, race/ethnicity, self-reported physical activity, history of fracture after age 54 years, parental hip fracture, current smoking, current corticosteroid use, and treated diabetes. Receiver operating characteristic curves showed that the algorithm had an area under the curve of 80% (95% confidence interval [CI], 0.77%-0.82%) when tested in the cohort of different women who were in the clinical trial. A simplified point score was developed for the probability of hip fracture. Receiver operating characteristic curves comparing DXA-scan prediction based on a 10% subset of the cohort and the algorithm among those who participated the clinical trial were similar, with an area under the curve of 79% (95% CI, 73%-85%) vs 71% (95% CI, 66%-76%).

Conclusion This algorithm, based on 11 clinical factors, may be useful to predict the 5-year risk of hip fracture among postmenopausal women of various ethnic backgrounds. Further studies are needed to assess the clinical implication of the algorithm in general and specifically to identify treatment benefits.

JAMA. 2007;298(20):2389-2398 www.jama.com

THE ESTIMATED 329 000 HIP fractures that occur annually in the United States[1] are associated with high morbidity, mortality, and cost.[2] Prevention of hip fracture is a high priority for patients, physicians, and public health. Several studies and consensus opinions have investigated the risk factors for hip fractures.[3-8] The Study of Osteoporotic Fractures (SOF),[3] which included 7782 women over 5 years, set the benchmark for establishing risk of hip fracture to date. The number of women included in the Women's Health Initiative (WHI) is an order of magnitude larger than SOF, and WHI includes nearly 20% minority women.

Although dual-energy x-ray absorptiometry (DXA) scan can precisely predict risk of hip fractures, as it did for a small subset of women participating in WHI, by assessing bone mineral density (BMD), clinicians and patients would benefit from assessing risk by

See also Patient Page.

Author Affiliations: Department of Internal Medicine, University of California at Davis School of Medicine, Sacramento (Dr Robbins); Fred Hutchinson Cancer Research Center, Seattle, Washington (Mr Aragaki and Dr Kooperberg); University of Cincinnati College of Medicine, Cincinnati, Ohio (Dr Watts); Division of Endocrinology, Diabetes and Metabolism, University at Buffalo, Buffalo, New York (Dr Wactawski-Wende); Department of Internal Medicine and Physical Medicine, Ohio State University, Columbus (Dr Jackson); Department of Medicine, Harvard Medical School, Boston, Massachusetts (Dr LeBoff); Division of Preventive Medicine, University of Alabama at Birmingham, (Dr Lewis); Division of Epidemiology and Biostatistics, University of Arizona, Tucson (Dr Chen); Stanford Prevention Research Center, Stanford, California (Dr Stefanick); Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania (Dr Cauley).
Corresponding Author: John Robbins, MD, Department of Internal Medicine, University of California Davis School of Medicine, 4150 V St, Ste 2400, Sacramento, CA 95817 (jarobbins@ucdavis.edu).

other means. Most hip fractures occur in women who are not osteoporotic by BMD testing.[9] Furthermore, it has been suggested by Black et al[4] that an algorithm without BMD is nearly as predictive as one with BMD.

The purpose of our study was to evaluate clinical risk factors for hip fracture in a multiethnic cohort of more than 100 000 postmenopausal women. Our goal was to create and test a predictive model for hip fracture using the WHI cohorts. It is important to investigate the combined effects of risk factors for hip fracture. There is the potential problem of interpreting factors independently of each other. For example, prior studies that had associated the risk of hip fracture with specific ethnic groups may have identified a marker of risk not a cause because they failed to adjust for such factors as weight, smoking status, and other risks.[10] Only hip fracture risk was evaluated. By studying hip fractures, we were able to use data from medical records to clearly identify those with fractures. Had we included other fractures, such as spine fractures, we would have had to rely on self-report. Although spine factures result in significant morbidity and mortally, hip fractures are clearly more detrimental to a woman's health.

METHODS

Study Population

The WHI has multiple components that can be used to build and test a predictive algorithm by taking advantage of an overlapping multicomponent design. Thus, some women were in multiple intervention components of the study. The WHI recruited postmenopausal women aged 50 to 79 years from 40 clinical centers and assigned them to multiple clinical trial components and to an observational study. The dietary modification component included 48 835 eligible women who were randomly assigned to either a sustained low-fat eating pattern (40%) or to eat as they pleased (60%).[11] The hormone therapy clinical trial randomized 27 347 women to trials assessing estrogen plus progestin or estrogen alone compared with placebo; women who still had a uterus received 0.625 mg of conjugated equine estrogen and 2.5 mg of medroxyprogesterone acetate or placebo daily while women without a uterus received estrogen alone or placebo.[12] Approximately 1 year after randomization into 1 of the above components, 36 282 women in the hormone therapy and/or dietary modification trial were randomly assigned to receive 1 g of calcium plus 400 IU of cholecalciferol (vitamin D) or placebo daily.[13]

All of the participants, including those who agreed to being followed up after dropping out of the interventions, are used in this analysis. Mean follow-up of the participants varies by component. The study treatments in the 2 components of the hormone trial were stopped prematurely; however, women continued to be followed up for events until study close-out. Women in the estrogen plus progestin group discontinued intervention after a mean of 5.6 years. Women in the estrogen-only group were followed up while taking the study drugs for 6.8 years. The dietary modification intervention lasted a mean of 8.1 years. Follow-up in the calcium vitamin D trial was a mean of 7 years. Information on the study design, methods, and results of these trials has been previously reported.[3,14-19] The mean (SD) follow-up time for women in the clinical trial was 8.0 (1.7) years (median, 8.0 years; interquartile range, 7.4-9.0 years). The participants in WHI were generally healthier and had more education than the general US population of women in the same age range.[20]

Postmenopausal women who were screened for the clinical trial but were ineligible or unwilling to participate in randomization were asked to enroll in an observational study. Women were ineligible if they did not want to discontinue taking hormone therapy upon study entry, or had a history of breast cancer; they were ineligible for the dietary component if they already followed a low-fat diet or too frequently ate away from home; and they were ineligible for the calcium and vitamin D component if they had a history of kidney stones or were unwilling to limit vitamin D intake.[21] A total of 93 676 women who enrolled in the observational study, were evaluated for multiple risk factors and followed up for a mean (SD) of 7.6 (1.7) years (median, 7.9 years; interquartile range, 6.9-8.9 years). Similar questionnaires and methods were used to determine baseline characteristics for both the clinical trial and the observations study groups. A subset of WHI participants from 3 of 40 clinical sites underwent DXA scans.

Incidence of hip fracture was collected using a standardized medical update questionnaire completed by all participants. These were collected every 6 months for those in the clinical trial and annually for those in the observational study until the study closed between October 2004 and March 2005. Hip fractures were self-reported and then confirmed both locally and centrally by review of medical records including x-ray and surgical reports. Agreement rate between self-reported hip fracture and adjudicated results based on medical records review was good, 78%,[22] but not perfect, and substantiates the need for individual review of outcomes, not just self-report as has been used in a number of other studies. All of the protocols were approved by the appropriate institutional review boards and participants signed informed consents.

Variables

Most of the variables are self-explanatory. (For a complete list of procedures, see http://www.whiscience.org/about/about_collection.php) Height and weight were measured in the clinics with calibrated scales and stadiometers. Two blood pressure and pulse measurements were manually obtained by trained technicians after 5 minutes of rest at 30 seconds apart. Prevalent medical conditions and medications, eg, diabetes, corticosteroid use, were based on self-report. Physical activity was self-reported and measured

as metabolic equivalent tasks (METs), using values derived from the literature and standardized questionnaires, which were validated for reproducibility in this population.[23] Similar questions have been validated against exercise diaries.[24] A MET is the ratio of work metabolic rate to a standard resting metabolic rate of 4.184 kJ/kg per hour.[25] For example, activity intensity were coded as 7 METS for strenuous, 4 for moderate, and 3 for low. Mean walking speed was classed as 3 METS for a 2 to 3 mph, 4 for 3 to 4 mph, and 4.5 for 4 mph or faster. METs per week were calculated as MET-h/wk.

Risk for depression was obtained from the Centers for Epidemiologic Studies–Depression 6-item questionnaire.[26] (This is unrelated to medication or physician diagnosis.) Dietary data were collected via self-report using food frequency questionnaire.[27] Dietary quality was identified using the method described by Neuhouser et al.[28] In brief, dietary intakes of fat, saturated fat, cholesterol, fruit and vegetables, sodium, calcium, protein, and fiber were coded as a 0 if achieved dietary recommendation, 1 if achieved within 30% of dietary recommendation, and 2 for everything else. The 8 scores were then summed. Lower scores indicate a better diet. Race and ethnicity were self-identified by the participants.

Statistical Methods

A prediction model was developed from the WHI observational study dataset and validated by the WHI clinical trial dataset. The observational study population was much larger than the clinical trial and more heterogeneous, thus offered more power for the development of the algorithm.

Model Development

Potential risk factors were identified from the literature and fit 1 at a time in a Cox proportional hazards model, adjusting for age and race/ethnicity. Variables that achieved a modest level of statistical significance ($P<.25$), based on the score test, were included

Table 1. Baseline Characteristics by Hip Fracture During Follow-up in the Observational Study Cohort

Baseline Characteristic[a]	Incident Hip Fracture, No. (%)		*P* Value[b]
	No	Yes	
Age group at screening, y			
50-59	29 603 (32.0)	102 (9.0)	<.001
60-69	40 838 (44.1)	359 (31.7)	
70-79	22 103 (23.9)	671 (59.3)	
Race/ethnicity			
White	76 949 (83.1)	1064 (94.0)	<.001
Black	7612 (8.2)	27 (2.4)	
Hispanic	3612 (3.9)	11 (1.0)	
American Indian	417 (0.5)	5 (0.4)	
Asian/Pacific Islander	2660 (2.9)	11 (1.0)	
Unknown	1294 (1.4)	14 (1.2)	
Marital status			
Never married	4322 (4.7)	68 (6.0)	.04
Divorced/separated	14 593 (15.8)	134 (11.9)	
Widowed	15 964 (17.3)	326 (28.8)	
Presently married/living as married	57 203 (62.1)	602 (53.3)	
Has medical insurance	89 011 (97.2)	1110 (98.8)	.17
Physical activity (METs/wk)			
0, Inactive	12 456 (13.6)	181 (16.3)	<.001
<5	17 522 (19.1)	241 (21.7)	
5-12	21 559 (23.6)	292 (26.3)	
≥12	39 983 (43.7)	395 (35.6)	
Smoking status			
Never smoked	46 458 (50.9)	565 (50.6)	<.001
Past smoker	39 058 (42.8)	456 (40.8)	
Current smoker	5695 (6.2)	96 (8.6)	
Parent broke hip after age 40	12 403 (13.4)	240 (21.2)	<.001
Fracture on or after age 55 y			
No	60 728 (71.2)	655 (65.6)	<.001
Yes	12 228 (14.3)	313 (31.4)	
Not available	12 356 (14.5)	30 (3.0)	
Alcohol consumption, drinks/d			
Nondrinker	38 707 (41.9)	535 (47.3)	<.001
≤1	42 111 (45.6)	459 (40.6)	
>1	11 573 (12.5)	136 (12.0)	
Medication			
Supplemental calcium	55 264 (59.7)	670 (59.2)	.07
Antianxiety or antidepressant	9389 (10.1)	127 (11.2)	.04
Bisphosphonate	1989 (2.1)	45 (4.0)	.008
Oral daily corticosteroid	1162 (1.3)	41 (3.6)	<.001
Thyroid hormone	13 349 (14.4)	213 (18.8)	.08
Hormone therapy			
Never used	37 466 (40.5)	559 (49.4)	.005
Past user	13 721 (14.8)	199 (17.6)	
Current user	41 273 (44.6)	373 (33.0)	
Prior bilateral oophorectomy	18 699 (20.7)	192 (17.6)	.003
Age at menarche, y			
<12	20 328 (22.1)	197 (17.6)	.04
12-13	50 780 (55.1)	610 (54.5)	
≥14	21 041 (22.8)	313 (27.9)	
No. of term pregnancies			
Never pregnant or never had term pregnancy	11 775 (12.8)	163 (14.6)	.06
1-2	32 529 (35.4)	389 (34.9)	
>3	47 619 (51.8)	563 (50.5)	

(continued)

in the pool of variables used to select a final prediction model. Ten-fold cross-validation was used to determine the optimal number of predictors that minimizes an estimate of prediction error.[27,29] Specifically, we divided the training data into 10 parts. Nine-tenths of the data was used to select the best model with k predictors by fitting a hazard regression model, which uses stepwise addition and deletion and considers interactions and nonparametric (spline) terms. For each model, we then evaluated the prediction log-likelihood on the remaining one-tenth of the data that was not used to select the model. For each k, we added these predicted log likelihoods to obtain a prediction score. The value of k that minimizes the cross-validated prediction score is taken to be the optimal number of predictors. A hazard regression model with K* predictors was then selected from the entire WHI observational study data.

The probability of a hip fracture within 5 years was then calculated using a multivariate logistic regression model fit on the WHI observational study dataset, using the K* variables selected above. The Hosmer-Lemeshow statistic was used to ascertain lack-of-fit (calibration) of this model. Participants with missing data in their predictor variables, and 5.5% (n=5161) of the participants who did not have a hip fracture within 5 years or did not have 5 years of follow-up were excluded from the logistic regression model.

Model Validation

To avoid an overly optimistic evaluation of model validity, we use the WHI clinical trial participants as our validation dataset. The women in the clinical trial were different in a multiple ways from the women in the observational study. The women in the clinical trial had volunteered to participate, were taking trial-required medications, and were following diet plans. These differences work to improve the usefulness of the validation; it is important that the algorithm work for women with different characteristics. The probability of a hip fracture within 5 years for the validation data was based on the multivariate logistic regression coefficients calculated exclusively on the WHI observational study data. Receiver operator characteristic (ROC) curves and the corresponding area under the curve (AUC) were used to evaluate how the prediction model preformed on the test data. The AUC was also calculated independently for the factors in the final model to demonstrate the additional value gained from the addition of each factor to the model. ROC curves plot the true-positive rate (sensitivity) vs the false-positive rate (1-specificity) at a continuum of thresholds; a participant is classified as having a hip fracture if her estimated probability of fracture exceeds a particular threshold. The ROC curve is a graphical representation of test characteristics, with sensitivity on the y-axis and 1-specificity on the x-axis, over all possible cut points for defining a positive and a negative test result. For our study, a positive result—predicting that an individual would have a hip fracture—occurs when the probability of fracture lies above a cut point.[30]

Because of the limited number of hip fractures in the DXA subset of women, a 10-fold cross-validation technique was used to compute the ROC curves and

Table 1. Baseline Characteristics by Hip Fracture During Follow-up in the Observational Study Cohort (cont)

Baseline Characteristic[a]	Incident Hip Fracture, No. (%)		*P* Value[b]
	No	Yes	
>10 lb intentional weight loss in last 20 y	49 475 (53.9)	487 (43.4)	.003
Depressive symptom[c]			
0	23 679 (26.1)	273 (24.8)	.04
1-2	33 516 (36.9)	387 (35.1)	
3-4	19 038 (21.0)	259 (23.5)	
>5	14 580 (16.1)	(16.7)	
Baseline general			
Excellent	16 437 (17.9)	139 (12.4)	<.001
Very good	37 303 (40.6)	382 (34.1)	
Good	29 255 (31.8)	414 (37.0)	
Fair	8036 (8.7)	174 (15.5)	
Poor	872 (0.9)	10 (0.9)	
Treated diabetes	38 423 (4.1)	79 (7.0)	<.001
Diet quality index, quartile[d]			
1st	14 387 (16.2)	170 (15.8)	.17
2nd	23 284 (26.2)	285 (26.5)	
3rd	28 818 (32.4)	357 (33.2)	
4th	22 350 (25.2)	263 (24.5)	
	No. Mean (SD)	No. Mean (SD)	
Height, cm	91797 161.7 (6.8)	1123 161.8 (7.1)	<.001
Weight, kg	92077 71.7 (16.9)	1127 67.7 (15.5)	<.001
Dietary calcium, mg	88 839 778.8 (435.3)	1075 765.9 (445.4)	.06
Dietary vitamin D, μg	88 839 5.0 (3.2)	1075 5.0 (3.2)	.10
Change in height from age 18, %	89 612 −1.0 (3.3)	1097 −1.9 (3.6)	<.001
Change in weight from age 35, %	90 951 19.8 (22.3)	1120 12.6 (20.8)	<.001

[a] For brevity, baseline characteristics that did not have a modest marginal association hip fracture ($P > .25$), after adjusting for age and ethnicity, are not shown. These include: use of supplements containing cholecalciferol (vitamin D), multivitamins, thiazides and thiazidelike diuretic, hypnotic medication, benzodiazepines, antiestrogens, oral contraceptive use, age at menopause, resting pulse, education, cups of regular coffee, calcitonin use, age at first birth, and currently following lactose-free diet.

[b] *P* value corresponds to the marginal association of baseline characteristic with hip fracture. *P* value is from a Cox proportional hazards model adjusting for age and ethnicity. *P* values for age and ethnicity correspond to unadjusted marginal associations.

[c] Sum of Center for Epidemiologic Studies–Depression score. A higher score indicates greater depression.

[d] Dietary intakes of fat, saturated fat, cholesterol, fruit and vegetables, sodium, calcium, protein, and fiber were coded as a 0 if achieved dietary recommendation, 1 if achieved within 30% of dietary recommendation, and 2 otherwise. The 8 scores are then summed. Lower scores indicate a better diet.

AUC. The 95% confidence intervals (CIs) were obtained by bootstrapping.

Cox proportional hazards models, logistic regression models, and their corresponding statistics were computed using SAS version 9.1 (SAS Institute Inc, Cary, North Carolina). The hazard regression model fits, stepwise selection, cross-validation, and ROC/AUC were computed using R version 2.1 and R libraries polspline and ROCR (R Development Core Team, http://www.R-project.org).[31-33] $P<.05$ was considered statistically significant.

RESULTS

Over a mean (SD) follow up of 7.6 (1.7) years, women in the observational study experienced 1132 hip fractures, an annual rate of 0.16%, whereas during a mean follow-up of 8.0 (1.7) years 791 women in the clinical trial experienced hip fractures at an annual rate of 0.14%. The 10 750 women with BMD measurements were followed up for 5 years or until they fractured their hip. Eighty hip fractures occurred in the combined groups over a mean (SD) of 8.7 (1.2) years of follow-up. The variables considered for inclusion in the model are shown in TABLE 1. Variables that did not meet the nominal threshold ($P<.25$) for consideration were education; cups of regular coffee; age at menopause; age at first birth; maintaining a lactose-free diet; pulse pressure; intentional weight loss (≥4.5 kg [≥10 lbs]); and use of vitamin D supplements, multivitamins thiazides and thiazidelike diuretics, antihypnotics, benzodiazepines, antiestrogens, calcitonins, and oral contraceptives. The independent frequency or mean after adjustment for age and race/ethnicity in those with and without hip fracture and significance are included.

Development of Algorithm

Cross-validation and stepwise selection of hazard regression models identified 12 variables from Table 1 that were independently predictive of hip fracture. These variables were age, self-reported health, height, change in height since the age of 18 years, change in weight since the age of 35 years, history of fracture after the age of 55 years, race/ethnicity, physical activity, smoking, history of parental fracture after the age of 40 years, diabetes treated with medications, and corticosteroid use. We did not find any pairwise interactions or nonlinear terms that were predictive of hip fracture. The variables change in height since the age of 18 years and change in weight since the age of 35 years were not available for the WHI clinical trial test set that we had planned to use. We therefore chose to use weight as a surrogate for change in weight, this less-than-perfect substitution, errs on the conservative side (TABLE 2).

The participants who were excluded from the logistic regression model (who did not have a hip fracture within 5 years and who did not have 5 years of follow-up) tended to be minorities (28% vs 16%) and older age (66 vs 63 years). More than half of these women died before 5 years of follow-up (n=2768). The Hosmer-Lemeshow statistic indicated no sign of lack of fit ($P=.20$).

An interactive model is available at http://hipcalculator.fhcrc.org.

As a second step, we approximated the additive logistic regression model by multiplying the coefficients by an arbitrary constant (4, selected to yield approximately integer-valued additive factors) and rounded to the nearest integer. This yielded a simple additive score. The 5-year risk of hip fracture can be calculated by totaling the point score. A point total of 9 yields a probability of fracture of 0.1%, a point total of 18 yields a probability of fracture of 1%, and a point total of 24 yields a probability of fracture of 5%.

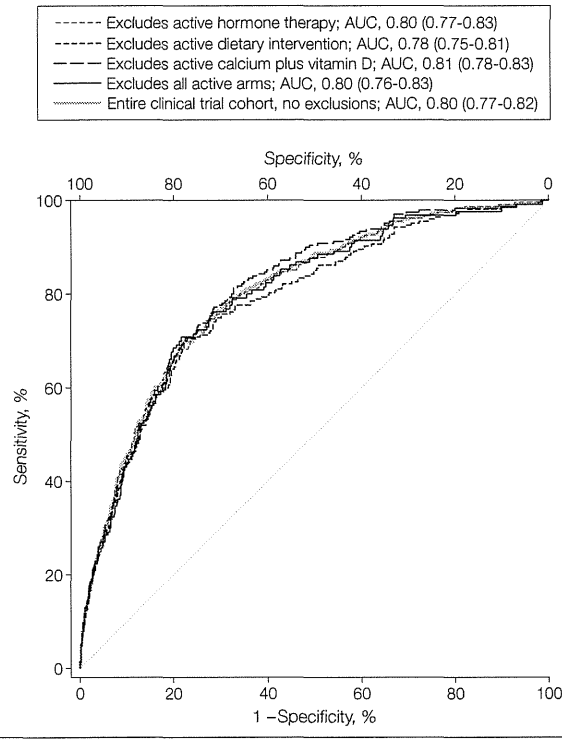
Validation

We tested the ability of the model to identify the 5-year probability of hip

Table 2. Multivariate Logistic Regression Model: Risk Factors for Hip Fracture in the Observational Study

Risk Factors	Odds Ratio (95% Confidence Interval)	*P* Value	Point Score
Age per each year	1.13 (1.11-1.15)	<.001	½ per year >50
Self-reported health			
Fair or poor vs excellent	2.38 (1.66-3.40)	<.001	3
Good vs excellent	1.22 (0.90-1.66)		1
Very good vs excellent	1.11 (0.83-1.49)		0
Height per each inch	1.11 (1.07-1.16)	<.001	½ per inch >64
Weight per each pound	0.99 (0.98-0.99)	<.001	1 per 25 lb <200
Fracture on or after age 55 y			
Not applicable vs no	1.01 (0.51-2.02)	<.001	0
Yes vs no	1.72 (1.41-2.10)		2
Race/ethnicity			White, 3
Unknown vs white	1.00 (0.47-2.14)	<.001	
Asian/Pacific Islander vs white	0.26 (0.10-0.70)		
American Indian vs white	1.60 (0.50-5.10)		
Hispanic vs white	0.32 (0.12-0.86)		
Black vs white	0.41 (0.24-0.70)		
Physical activity, METs			1
5-12 vs ≤12	1.32 (1.04-1.67)	.004	
<5 vs ≤12	1.26 (0.97-1.64)		
Inactive 0 vs ≤12	1.64 (1.24-2.17)		
Smoking status			
Current vs never	2.33 (1.71-3.18)	<.001	3
Past vs never	0.96 (0.79-1.17)		0
Parent broke hip, yes vs no	1.50 (1.20-1.87)	<.001	1
Corticosteroid use, yes vs no	1.94 (1.16-3.25)	.01	3
Use of hypoglycemic agent, yes vs no	1.74 (1.17-2.60)	.006	2

Figure 1. Women's Health Initiative Clinical Trial Test Set Receiver Operating Characteristic Curve



AUC indicates area under the curve. Blue curves in Figure 1 and Figure 2 are the same and are derived from the entire clinical trial cohort.

Table 3. Contributions of Individual Predictors

Variable	AUC%[a]
General health	56
Height	56
Weight	57
Fracture after age 55 y	56
Race/ethnicity	54
Physical activity	53
Currently smoking	53
Parent broke hip	51
Corticosteroid use	50
Diabetes	51
All predictors except age	67
Age	76[b]
Age plus all predictors	80[c]

Abbreviation: AUC, area under the curve.
[a] For predictor variables, other than age, weighted AUCs $\frac{\sum_i w_i AUC_i}{\sum_i w_i}$ are calculated where w_i is the number of hip fractures for ith age group and i goes from 50 to 79. The AUC_i are the AUCs for the ith age group. These are based on logistic regression model that contain the predictor of interest and age (categorical); trained on the observational study and tested on the clinical trial.
[b] Based on a logistic regression model containing age as a single variable; trained on the observational study and tested on the clinical trial.
[c] Based on our full logistic regression model; trained on the observational study and tested on the clinical trial.

fracture in women included in the hormone treatment, dietary, and calcium and vitamin D components of the WHI clinical trial. It should be noted that the women in the observational study cohort had different characteristics than those in clinical trial cohorts. Participants in the clinical trial tended to be younger (mean, 62.7 years), taller (161.1 cm [63.42 in]), heavier (76.1 kg [169.1 lb]), less likely to be white (81.5% were white), with a lower proportion of the clinical trial reporting fair to poor health (8.3%), history of fracture after age 55 years (13.1%), either parent breaking a hip (11.8%), and corticosteroid use (0.1%). A higher proportion of the clinical trial participants reported being physically inactive (19.2%), currently smoking (7.9%), and taking treatment for diabetes (4.8%). These differences between the clinical trial and observational study participants were all statistically significant ($P<.001$).

Using adjudicated hip fractures for women in the clinical trial, ROC curves were developed to test how well the algorithm that was developed from the observational study cohort performed in validation populations. The AUC was tested against the WHI clinical trial. We examined various groups participating in the clinical trial and found similar results in cases in which the AUC ranged from 78% to 81%. The AUC was 80% for all WHI clinical trial participants, all WHI participants receiving placebos, and those who received no active HT intervention (FIGURE 1).

Although there are potentially other variables that are statistically significant in a logistic regression model, they would not appreciably improve prediction and consequently were not included in the model. For example, alcohol consumption was a statistically significant variable when added to the multivariate logistic regression model ($P=.01$) but has little effect on the AUC.

We also tested the various components of the algorithm individually and in combinations that included or excluded age. These results are shown in TABLE 3. This demonstrates that age alone is clearly the best predictor of hip fracture, but added value is gained by the addition of other factors.

The ROC curve in FIGURE 2 shows the accuracy at different estimations of risk tested in all WHI trial participants. This shows the sensitivity and 1-specificity of the prediction of 5-year hip fracture risk for women at different levels of predicted risk. By application of this information, thresholds for further screening can be set based on acceptable risk and desire for certainty. For example, identifying women at risk using a threshold of a 1% 5-year risk would yield a true-positive rate (sensitivity) of about 50%, half of women who would have hip fractures within 5 years, but there would be a false-positive rate (1-specificity) of 15%. Half of the women who would have hip fractures in the next 5 years would be in this group, and 15% who were predicted to have hip fractures would not. A less stringent risk threshold of 0.5% would identify