



Practice of Epidemiology

Usual Physical Activity and Hip Fracture in Older Men: An Application of Semiparametric Methods to Observational Data

Dawn C. Mackey*, Alan E. Hubbard, Peggy M. Cawthon, Jane A. Cauley, Steven R. Cummings, and Ira B. Tager for the Osteoporotic Fractures in Men Research Group

* Correspondence to Dr. Dawn C. Mackey, San Francisco Coordinating Center, California Pacific Medical Center Research Institute, 185 Berry Street, Lobby 5, Suite 5700, San Francisco, CA 94107 (e-mail: dmackey@sfcc-cpmc.net).

Initially submitted August 2, 2010; accepted for publication October 26, 2010.

Few studies have examined the relation between usual physical activity level and rate of hip fracture in older men or applied semiparametric methods from the causal inference literature that estimate associations without assuming a particular parametric model. Using the Physical Activity Scale for the Elderly, the authors measured usual physical activity level at baseline (2000–2002) in 5,682 US men ≥ 65 years of age who were enrolled in the Osteoporotic Fractures in Men Study. Physical activity levels were classified as low (bottom quartile of Physical Activity Scale for the Elderly score), moderate (middle quartiles), or high (top quartile). Hip fractures were confirmed by central review. Marginal associations between physical activity and hip fracture were estimated with 3 estimation methods: inverse probability-of-treatment weighting, G-computation, and doubly robust targeted maximum likelihood estimation. During 6.5 years of follow-up, 95 men (1.7%) experienced a hip fracture. The unadjusted risk of hip fracture was lower in men with a high physical activity level versus those with a low physical activity level (relative risk = 0.51, 95% confidence interval: 0.28, 0.92). In semiparametric analyses that controlled confounding, hip fracture risk was not lower with moderate (e.g., targeted maximum likelihood estimation relative risk = 0.92, 95% confidence interval: 0.62, 1.44) or high (e.g., targeted maximum likelihood estimation relative risk = 0.88, 95% confidence interval: 0.53, 2.03) physical activity relative to low. This study does not support a protective effect of usual physical activity on hip fracture in older men.

aged; confounding factors (epidemiology); exercise; hip fractures; men; motor activity; prospective studies

Abbreviations: CI, confidence interval; ETA, experimental treatment assignment; IPTW, inverse probability-of-treatment weighting; PASE, Physical Activity Scale for the Elderly; T-MLE, targeted maximum likelihood estimation.

Hip fracture is the most serious consequence of osteoporosis. It substantially reduces physical function and increases the likelihood of nursing home admission (1). Approximately 30% of hip fractures occur in men (2), and risk of death in the year after hip fracture is greater in men ($\approx 33\%$) than in women ($\approx 20\%$) (1–3). As the population ages, effective hip fracture prevention strategies are needed.

Physical activity is a potentially attractive intervention for hip fracture prevention; it is inexpensive, is generally suitable for all members of the population, and can be advocated on the basis of prevention and management of several chronic diseases. Structured physical activity interventions in older adults have improved bone mineral density at some skeletal locations (4, 5); improved fall-related risk factors,

such as balance, walking speed, and muscle strength (6–9); and reduced fall risk (10–13). Thus, it has been postulated that physical activity may reduce the risk of hip fracture by maintaining or improving bone strength and reducing fall frequency or severity. On the other hand, physical activity may increase the risk of hip fracture by increasing the incidence of injury, or physical activity may have no effect on hip fracture if the activity is not sufficiently intense or targeted to modify bone strength or fall risk.

To the best of our knowledge, no one has yet conducted a randomized controlled trial to test the effects of physical activity on rates of hip fracture, likely because of the large sample-size requirements for such a trial (14). Results from prospective observational studies of older women have

generally suggested that greater amounts of physical activity were associated with lower hip fracture risk (15–17). The extent to which physical activity influences hip fracture risk in older men remains uncertain. In the present study, we examined the association between usual physical activity level and risk of hip fracture in a prospective study of older men.

MATERIALS AND METHODS

Study population

The men studied were participants in the Osteoporotic Fractures in Men Study. During the baseline examination from 2000 to 2002, a total of 5,995 community-dwelling men ≥ 65 years of age were enrolled at 6 clinical centers in the United States: Birmingham, Alabama; Minneapolis, Minnesota; Palo Alto, California; Monongahela Valley near Pittsburgh, Pennsylvania; Portland, Oregon; and San Diego, California. Men were not eligible to participate if they reported bilateral hip replacement or required assistance with ambulation. Details of the cohort study design and recruitment have been published elsewhere (18, 19). The protocol and consent forms for the study were approved by the institutional review boards of all of the participating institutions. All participants provided written informed consent.

To be included in the analysis data set for this report, participants had to have had nonmissing values for the Physical Activity Scale for the Elderly (PASE) score and all potential confounding variables; 5,851 men had adequate data for inclusion. Data were missing because of a participant's refusal to complete a question or examination, equipment failure, or incorrect protocol administration.

Physical activity

Usual physical activity level was assessed by using PASE (20), a 12-item questionnaire used to assess participation in leisure-time, household, and occupational physical activities during the prior 7 days. Each item was scored by multiplying the activity frequency (hours per day for leisure-time and occupational activities; 0/1 for household activities) by the corresponding activity intensity weight. Item scores were summed to compute the overall PASE score, a unitless relative score that represented the total overall amount of physical activity. Quartiles of the PASE score were used to define low (bottom quartile), moderate (middle quartiles), and high (top quartile) levels of physical activity.

Hip fracture

We contacted participants every 4 months by mailed questionnaire to determine whether they had sustained a hip fracture in the prior 4-month period. Responses to the questionnaires were 99% complete during follow-up. All fractures were validated by centralized physician review of radiology reports.

Potential confounders

We identified an extensive list of risk factors for hip fracture on the basis of published literature and considered these as potential confounders. Risk factor data were obtained at baseline by self-administered questionnaire or through interview or clinical examination by trained, certified staff.

Recorded demographic characteristics included age, self-reported race (white/nonwhite), and clinical site. Height was measured with a wall-mounted Harpenden stadiometer (Holtain Ltd, DyFed, United Kingdom), and body weight was measured with a balance beam scale (except in Portland, where investigators used a digital scale). Body mass index was calculated as weight in kilograms divided by height in meters squared. Average weight at age 25 years was obtained using the questionnaire, and weight change since age 25 years was calculated. Proximal femur and lumbar spine bone mineral density were measured with dual-energy x-ray absorptiometry using Hologic QDR 4500W densitometers (Waltham, Massachusetts) (19). Dietary calcium was calculated from data obtained by using a modified Block food frequency questionnaire (21), and cigarette smoking history was obtained from questionnaire. Overall self-rated health (excellent or good versus fair, poor, or very poor), history of fractures after 50 years of age, and history of falling in the past year were ascertained through questionnaire. Cognitive function was assessed with the Teng Modified Mini-Mental State Examination. Participants reported whether a clinician had ever told them they had certain medical conditions. Osteoporosis was defined as a baseline bone mineral density *T* score (versus a young male reference group) < -2.5 at the femoral neck or total hip (22).

Four tests of physical performance were administered during the clinic examination (23): a 6-m usual-pace walk; a 6-m narrow-balance walk; 5 repeated chair stands; and grip strength tests of the left and right hands using Jamar dynamometers (Sammons Preston Rolyan, Bolingbrook, Illinois). Participants reported degree of difficulty (none = 0, some = 1, much = 2, unable = 3) in accomplishing 5 daily activities: walking 2–3 blocks, climbing up 10 steps without resting, preparing meals, doing heavy housework, and doing their own shopping. Total difficulty was summed across the 5 daily activities.

Statistical analysis

Overview. Our goal was to estimate marginal adjusted associations between baseline physical activity and hip fracture risk. We begin with notation necessary to define the various approaches and parameters of interest. The observed data for each participant consisted of a 3-level treatment (exposure) variable, *A* (low, moderate, or high baseline physical activity level), a binary outcome, *Y* (hip fracture during follow-up), and a set of baseline covariates, *W*, that we assumed preceded treatment and were potential confounders.

First, we will define the optimal data one could collect to infer causation from measured associations by introducing the counterfactual framework (24, 25); later, we will restrict

the analyses to include only our parameters of interest as parameters of the actual data-generating distribution. A counterfactual outcome, Y_a , is defined as the outcome (1 = yes, 0 = no) an individual would experience if the treatment variable A took on a particular value a . The 3 counterfactual outcomes for each participant in this study, denoted as Y_{low} , Y_{mod} , and Y_{high} , were the hip fracture outcomes that would have been observed if the participant's level of physical activity was low, moderate, or high, respectively. Therefore, a theoretical "full" data set would be $X = (Y_a, a \in (\text{low, med, high}), W) \sim P_{X,0}$, where $P_{X,0}$ is the true distribution of these full data. Only 1 of these outcomes, corresponding to the observed level of physical activity, was actually observed, and the other 2 remained unobserved, thus the term counterfactual. The marginal or population-level relative risk for moderate physical activity was defined as the ratio of population means, $E_{X,0}(Y_{mod})/E_{X,0}(Y_{low})$, that is, as the ratio of the mean hip fracture risk if all men in the study population were moderately physically active to the mean hip fracture risk if their levels of physical activity were uniformly low. The marginal relative risk for high physical activity was defined analogously as $E_{X,0}(Y_{high})/E_{X,0}(Y_{low})$.

To estimate these parameters from the observed data, we had to make identifiability assumptions. We defined the observed data as $O = (Y, A, W) \sim P_0$ and noted that under identifiability assumptions (26),

$$EY(a) \equiv E_{W,0}\{E_0(Y|A=a,W)\} \stackrel{\text{assumptions}}{=} E_{X,0}\{Y_a\}.$$

The identifiability assumptions are as follows: 1) The randomization assumption, $A \perp Y_a|W$, requires that treatment be randomized with respect to the set of all counterfactual outcomes, given the covariates; 2) the experimental treatment assignment (ETA) or positivity assumption (27–30) requires that each participant has a non-zero probability of all possible levels of treatment, given his observed covariates; and 3) the consistency assumption (26), which is related to the interdependence of the counterfactual outcomes in a population, sometimes called Rubin's Stable Unit Treatment Value Assumption (25), requires that each individual's observed outcome, under his observed treatment history, be his potential counterfactual outcome under that same treatment history. We concentrated on estimating the parameters as a function of the data-generating distribution (P_0) and noted that the interpretation of these would depend on the validity of these assumptions.

To estimate the associations of interest, we estimated marginal (adjusted) hip fracture probabilities for each of the 3 levels of physical activity and compared these to derive relative risk estimates by implementing 3 estimation techniques: inverse probability of treatment weighting (IPTW), G-computation, and doubly robust targeted maximum likelihood estimation (T-MLE). Each estimation technique relied on different assumptions about estimation of components of P_0 , so the 3 methods allowed us to assess the robustness of association estimates.

All analyses were conducted using R, version 2.5.1 (R Foundation for Statistical Computing, Vienna, Austria), and SAS, version 9.1 (SAS Institute, Inc., Cary, North Carolina).

IPTW. For all of our techniques, to define our measures of association, we needed to estimate the marginally adjusted mean. IPTW can be used to derive the marginal mean by reweighting the data set so that treatment assignment appears to be randomized (i.e., the relation between W and A is broken), and thus the joint distribution of measured covariates is equal across physical activity groups. Larger weights were assigned to participants whose observed physical activity treatment was less probable given their covariates, and smaller weights were assigned to participants whose observed physical activity treatment was more probable given their covariates. Stabilized weights, defined as $\hat{P}(A=a)/\hat{P}(A=a|W)$, were calculated after fitting a multinomial logistic regression model of the physical activity treatment mechanism given potential confounders, $g_0(a|W) \equiv P_0(A=a|W)$. The parameters of this treatment model were not of direct interest, but consistent estimation of the association between physical activity and hip fracture relies on correct specification of the treatment model. The importance of this and the other techniques is they can allow the derivation of these simple measures of association without the assumption of a specific parametric model and thus can "learn" from the data while still providing trustworthy inference. Therefore, we fitted the treatment model with the data-adaptive Deletion/Substitution/Addition algorithm, which combines a flexible and aggressive data-adaptive search and cross-validation to avoid overfitting (31). We allowed for up to 2-way interactions, cubic terms, and 10 total terms in addition to the intercept. Percentile-based 95% confidence intervals for the associations of interest were derived from 10,000 bootstrap iterations created by resampling participants from the study population with replacement. Equivalent Deletion/Substitution/Addition and bootstrap approaches were used to fit all models and derive 95% confidence intervals for G-computation and T-MLE.

We evaluated potential practical violations of the ETA assumption after fitting the treatment model by plotting the distribution of predicted physical activity probabilities for each level of observed physical activity (Figure 1). Empirically, the ETA assumption appeared to be met for low and moderate physical activity levels, as all participants, regardless of their observed physical activity level, had a positive predicted probability (5%–95%) of low and moderate physical activity, given their covariates. However, 169 participants had a predicted probability <5% for high physical activity level given their covariates. To avoid any violation of the ETA assumption, we wanted to include only participants who could reasonably have a high physical activity level, so we excluded these 169 individuals (hip fracture was observed in 15 (8.9%)). The final data set for all analyses consisted of 5,682 participants.

After the treatment weights were assigned, we fitted a standard unadjusted weighted logistic regression model of hip fracture on physical activity (participants were weighted by their IPTW weights), and we estimated hip fracture risks for each level of physical activity from the fitted probabilities. Indicator variables were used for moderate and high

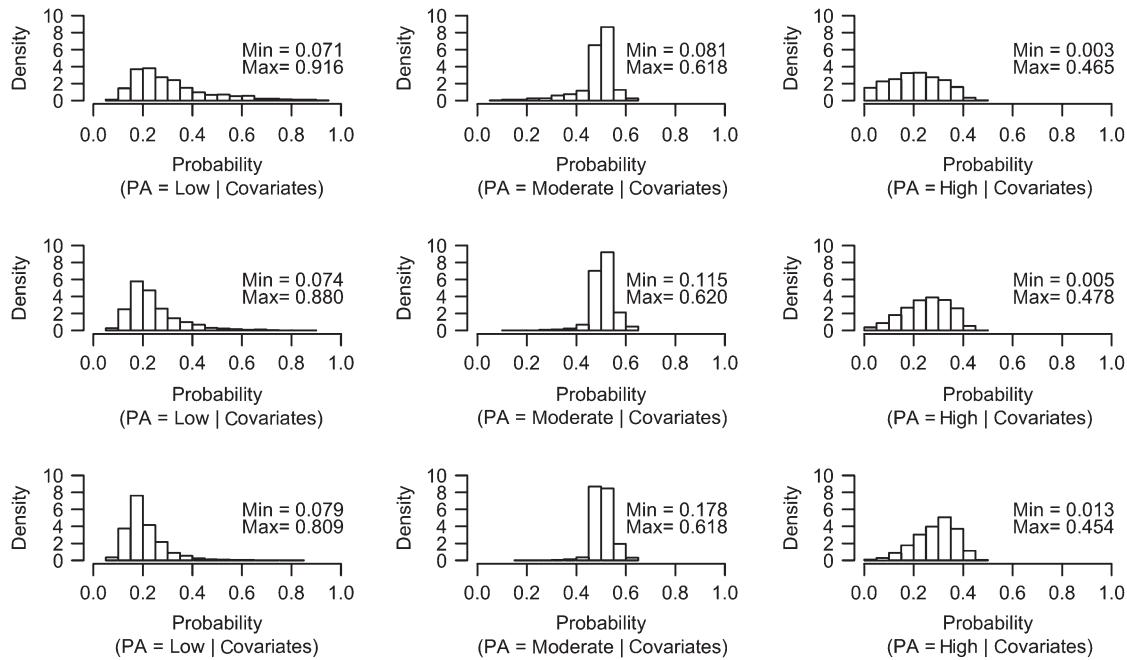


Figure 1. Graphical evaluation of the experimental treatment assignment assumption in the Osteoporotic Fractures in Men Study, 2000–2002. Predicted probabilities of low, moderate, and high levels of physical activity (PA) given covariates among men observed to have low (top panel), moderate (middle panel), and high (bottom panel) PA are shown. Minimum (min) and maximum (max) predicted probabilities are shown on each plot.

physical activity levels; low physical activity level was the reference.

G-computation. In contrast to IPTW, the G-computation technique in essence imputed the hip fracture outcomes for each participant on the basis of a multivariable logistic regression model of physical activity on hip fracture, $Q_0(A, W) \equiv E_0(Y | A, W)$. This was achieved by using the logistic model fitted with the Deletion/Substitution/Addition algorithm to predict the probability of hip fracture for each participant at each of the 3 levels of physical activity while holding his covariates, W , constant at their observed level, and then averaging these outcomes across all participants to estimate the adjusted marginal means, $\hat{E}(Y(a)) = \frac{1}{n} \sum_{i=1}^n \hat{E}(Y | A = a, W_i)$, for $a = (\text{low, moderate, and high})$. We then calculated corresponding relative risks of interest. In contrast to IPTW, which required that the model for the treatment weights be specified correctly, G-computation is a maximum likelihood method and thus required that the logistic model used for estimating $Q_0(A, W)$ be specified correctly.

T-MLE. T-MLE uses an augmentation of the model for $Q_0(A, W)$, such that applying plug-in estimation based on G-computation and maximum likelihood estimation of Q_0 results in a doubly robust (consistent if the model for either Q_0 or g_0 is correctly specified) estimator of $E(Y(a))$ (32, 33). One can also use an estimation equation approach to derive an estimator with the same asymptotic properties, but the virtue of T-MLE estimation is that it guarantees a proper model for the parameter of interest in finite samples and can be often implemented with standard software. It also provides a mechanism (loss-function) for choosing

models for g_0 for estimation of the parameter of interest, $E(Y(a))$ (32).

RESULTS

Participant characteristics

Men who reported higher levels of physical activity were younger, were more highly educated, were more likely to be married, had gained less weight since 25 years of age, had greater total hip and femoral neck bone mineral density, had greater dietary calcium intake, were less likely to be current smokers, were more likely to report excellent or good health, had slightly better cognitive function, and were less likely to have a history of a variety of medical conditions (Table 1). They also had a faster usual walking speed, were more likely to complete the narrow balance walk and 5 repeated chair stands, had greater grip strength, and reported less total difficulty with 5 daily activities.

Physical activity

Among men with a low level of physical activity, the overall PASE score ranged from 0 to 100, with a mean of 67.2 (standard deviation = 24.8); among men with a moderate level of physical activity, it ranged from 101 to 186, with a mean of 143.3 (standard deviation = 23.8); and among men with a high level of physical activity, it ranged from 187 to 486, with a mean of 236.7 (standard deviation = 45.8) (Table 2). Household activities accounted for the largest proportion of the overall PASE score, followed by

Table 1. Baseline Characteristics, by Physical Activity Level, of Men in the Osteoporotic Fractures in Men Study, 2000–2002

Characteristic	Physical Activity Level									P Value ^a
	Low (n = 1,351)			Moderate (n = 2,875)			High (n = 1,456)			
	Mean (SD)	No.	%	Mean (SD)	No.	%	Mean (SD)	No.	%	
Age, years	74.6 (6.2)			73.5 (5.6)			71.9 (5.2)			<0.01
White, non-Hispanic		1,192	88.2		2,599	90.4		1,303	89.5	0.10
Less than high school education		102	7.5		182	6.3		70	4.8	<0.01
Married		1,086	80.4		2,403	83.6		1,220	83.8	0.02
Weight, kg	83.4 (14.2)			83.2 (12.9)			83 (12.7)			0.65
Weight change since age 25 years, kg	10.9 (12.1)			10.5 (11)			9.7 (10.7)			0.01
Height, cm	174.0 (7.0)			174.3 (6.9)			174.4 (6.5)			0.32
Body mass index, kg/m ²	27.5 (4.0)			27.4 (3.7)			27.3 (3.7)			0.21
Total hip BMD, g/cm ²	0.95 (0.14)			0.96 (0.14)			0.97 (0.14)			<0.01
Femoral neck BMD, g/cm ²	0.78 (0.13)			0.79 (0.13)			0.80 (0.13)			<0.01
Lumbar spine BMD, g/cm ²	1.07 (0.20)			1.07 (0.18)			1.07 (0.18)			0.82
Dietary calcium, mg/day	780 (391)			795 (373)			819 (405)			0.03
Smoking history										0.01
Never smoker		488	36.1		1,118	38.9		537	36.9	
Past smoker		799	59.1		1,678	58.4		866	59.5	
Current smoker		64	4.7		79	2.7		53	3.6	
Excellent or good self-rated health		1,089	80.6		2,548	88.6		1,366	93.8	<0.01
Physical Activity Scale for the Elderly score	67.2 (24.8)			143.3 (23.8)			236.7 (45.8)			<0.01
Ever fractured a hip		25	3.4		38	2.4		24	3.0	0.38
Nontraumatic fracture after 50 years of age		251	18.6		456	15.9		241	16.6	0.09
≥1 falls in past year		295	21.8		554	19.3		296	20.3	0.15
Teng Modified Mini-Mental State Examination Score	93.1 (6.2)			93.6 (5.5)			93.8 (5.1)			<0.01
Medical history										
Myocardial infarction		215	15.9		368	12.8		174	12.0	<0.01
Angina		235	17.4		370	12.9		165	11.3	<0.01
Congestive heart failure		83	6.1		136	4.7		51	3.5	<0.01
Stroke		105	7.8		149	5.2		45	3.1	<0.01
Hypertension		641	47.4		1,215	42.3		555	38.1	<0.01
Diabetes		169	12.5		308	10.7		108	7.4	<0.01
High thyroid hormone levels		19	1.4		40	1.4		30	2.1	0.21
Low thyroid hormone levels		100	7.4		204	7.1		80	5.5	0.08
Parkinson's disease		15	1.1		28	1.0		1	0.1	<0.01
Osteoarthritis		281	20.8		552	19.2		239	16.4	0.01
Osteoporosis		56	4.1		85	3.0		35	2.4	0.02
Chronic obstructive pulmonary disease		180	13.3		280	9.7		117	8.0	<0.01
Cancer		426	31.5		808	28.1		407	28.0	0.05
Nonskin cancer		278	20.6		494	17.2		248	17.0	0.02
Usual walking speed, m/second	1.16 (0.23)			1.22 (0.21)			1.26 (0.21)			<0.01
Unable to complete narrow walk		158	11.7		168	5.8		61	4.2	<0.01
Unable to complete 5 chair stands		37	2.7		30	1.0		18	1.2	<0.01
Average grip strength, kg	36.8 (8.0)			39 (7.9)			40.5 (8.0)			<0.01
Total difficulty with 5 daily activities	0.62 (1.27)			0.31 (0.85)			0.18 (0.62)			<0.01

Abbreviations: BMD, bone mineral density; SD: standard deviation.

^a From χ^2 tests for categorical variables and from analysis of variance for continuous variables.

Table 2. Physical Activity Patterns, by Baseline Physical Activity, of Men in the Osteoporotic Fractures in Men Study, 2000–2002

Characteristic	Physical Activity Level					
	Low (n = 1,351)		Moderate (n = 2,875)		High (n = 1,456)	
	Mean (SD)	% (SD)	Mean (SD)	% (SD)	Mean (SD)	% (SD)
Overall Physical Activity Scale for the Elderly score	67.2 (24.8)		143.3 (23.8)		236.7 (45.8)	
Leisure-time subscore	19.2 (16.9)		31.3 (24.1)		64.0 (44.8)	
Household subscore	45.8 (26.5)		103.0 (31.3)		127.7 (34.3)	
Occupational subscore	2.3 (8.4)		9.0 (21.5)		44.9 (52.7)	
Leisure-time activities, hours/day						
Walking	0.44 (0.47)		0.72 (0.70)		1.38 (1.21)	
Light sports	0.12 (0.37)		0.23 (0.52)		0.49 (0.85)	
Moderate sports	0.08 (0.27)		0.16 (0.42)		0.40 (0.76)	
Strenuous sports	0.10 (0.29)		0.15 (0.35)		0.34 (0.67)	
Muscle strength and endurance	0.13 (0.23)		0.16 (0.29)		0.31 (0.53)	
Household activities ^a						
Light housework	63.0 (48.3)		83.8 (36.9)		89.3 (30.9)	
Heavy housework	31.8 (46.6)		75.9 (42.8)		87.1 (33.5)	
Home repairs	14.8 (35.5)		50.5 (50.0)		73.5 (44.2)	
Lawn work/yard care	31.2 (46.4)		81.3 (39.0)		90.3 (29.6)	
Outdoor gardening	19.9 (39.9)		59.0 (49.2)		73.4 (44.2)	
Caring for another person	6.8 (25.2)		19.8 (39.8)		41.1 (49.2)	
Occupational activities, hours/day ^b	0.10 (0.40)		0.43 (1.02)		2.14 (2.51)	

Abbreviation: SD; standard deviation.

^a Values represent percent of participants who reported engaging in the activity in the past 7 days.

^b Work for pay or volunteer.

leisure-time and occupational activities. Light and heavy housework and lawn work/yard care were the most common household activities; walking was the most common leisure-time activity.

Hip fracture

During a mean 6.5 years (standard deviation = 1.4) of follow-up, 95 men (1.7%) experienced a hip fracture, corresponding to an incidence of 2.6 per 1,000 person-years (95% confidence interval (CI): 2.1, 3.1). Most hip fractures ($n = 66$, 69.5%) involved a fall from standing height.

Hip fracture was observed in 31 men (2.3%) with a low level of physical activity (3.7 per 1,000 person-years; 95% CI: 2.4, 5.0), in 47 men (1.6%) with a moderate level of physical activity (2.5 per 1,000 person-years; 95% CI: 1.8, 3.2), and 17 men (1.2%) with a high level of physical activity (1.2 per 1,000 person-years; 95% CI: 0.9, 2.6). Unadjusted relative risk estimates were 0.71 (95% CI: 0.45, 1.12) for moderate and 0.51 (95% CI: 0.28, 0.92) for high physical activity.

The logistic and multinomial regression models selected by the Deletion/Substitution/Addition algorithm are presented in Web Table 1 (available at: <http://aje.oxfordjournal->

[s.org/](http://aje.oxfordjournal.org/)). The resulting stabilized IPTW weights ranged from 0.38 to 5.04, with a mean of 1.00.

Estimated absolute hip fracture risks if all participants had a given level of physical activity (low, moderate, and high) were generally consistent across the 3 estimation methods (Table 3). There was a slightly graded association between amount of physical activity and hip fracture risk, suggesting that greater amounts of physical activity reduced risk of hip fracture. However, confidence intervals were wide and overlapping. Indeed, relative risk estimates using IPTW showed that risk of hip fracture was not significantly reduced by moderate or high levels of physical activity relative to low levels of physical activity. Relative risk estimates from G-computation were very similar to IPTW estimates. Use of T-MLE resulted in somewhat attenuated relative risk estimates compared with the other estimators.

DISCUSSION

In the present large study of older men, unadjusted analyses showed that men with high physical activity levels had a 49% lower risk of hip fracture than did men with low physical activity levels. However, physical activity levels were strongly associated with other risk factors for hip

Table 3. Estimated Absolute Hip Fracture Risks and Relative Risks, by Baseline Physical Activity Level, for Men in the Osteoporotic Fractures in Men Study, 2000–2002

	Estimated Absolute Hip Fracture Risk, %						Relative Risk ^a			
	Low PA Level		Moderate PA Level		High PA Level		Moderate PA Level		High PA Level	
	Risk	95% CI	Risk	95% CI	Risk	95% CI	RR	95% CI	RR	95% CI
Unadjusted							0.71	0.45, 1.12	0.51	0.28, 0.92
Inverse probability of treatment weighting	2.34	1.88, 3.62	1.98	1.56, 2.89	1.84	1.14, 4.31	0.85	0.52, 1.28	0.79	0.40, 1.76
G-computation	2.64	2.37, 4.27	2.31	1.89, 3.24	2.08	1.39, 3.49	0.87	0.53, 1.16	0.79	0.40, 1.19
Targeted maximum likelihood estimation	2.52	2.08, 3.92	2.32	1.99, 3.59	2.22	1.62, 5.40	0.92	0.62, 1.44	0.88	0.53, 2.03

Abbreviations: CI, confidence interval; PA, physical activity; RR, relative risk.

^a Low physical activity level served as the reference level.

fracture, and semiparametric analyses that controlled for confounding did not support an association between physical activity level and hip fracture. Although all of the relative risk point estimates for moderate and high physical activity (vs. low) were <1 and each estimator suggested a similar trend of decreasing hip fracture risk with increasing physical activity, the confidence intervals were wide and spanned the null. T-MLE analyses, which we regarded to be the least biased because the estimation procedure was doubly robust, indicated little difference in hip fracture risk for men with moderate or high physical activity levels relative to those with low physical activity level.

Previous studies of physical activity and rates of hip fracture in older men have had inconsistent results. As in the current study, researchers in some previous studies found no significant reduction in hip fracture risk across physical activity levels (17, 34–36). Others found lower hip fracture risk among more physically active men (34, 35, 37–39), but some did not account for important sources of confounding other than age (35, 39).

The current results do not preclude the possibility that an exercise program targeted to maintenance of bone health and fall risk reduction could decrease the risk of hip fracture in older men, but they suggest that typical forms of physical activity in older men have a limited ability to prevent hip fracture. A possible explanation is that the physical activity usually undertaken by older men does not provide sufficient magnitudes or rates of loading to enhance bone strength. Although the ideal physical activity prescription to optimize bone health in older adults remains unclear, animal studies indicate that short, novel, vigorous loading bouts separated by rest periods enhance the osteogenic response (40, 41). Exercise interventions in older adults that reduce bone fragility involve progressive and high-intensity resistance and agility training, dynamic movements, multidirectional acceleration/deceleration, and moderate-to-high impact loads (4, 5). It is unlikely that household activities or walking, the most common activities in the present study, provided sufficient stimuli to maintain or increase bone strength. Indeed, most men experienced some amount of bone loss at the hip during follow-up (42).

We chose to analyze the relation between baseline physical activity and hip fracture by using semiparametric

methods to control for confounding without assuming an arbitrary parametric model. A conventional analysis approach might use a multivariable logistic regression model with only main terms. The results of conventional analyses (Web Table 2) were qualitatively consistent with the main results presented in this article. Despite this similarity, there are many advantages to the use of semiparametric methods. First, one never knows whether a traditional analysis will yield similar results unless one estimates adjusted associations without making the arbitrary assumptions that are typical of standard analyses. In addition, even if one did such an analysis, it is erroneous to report inference under that model, because in truth one does not know the form of Q . Therefore, the apparent advantages of traditional parametric regression approaches are misleading because these approaches include bias but do not account for either the bias or the variance in the correct model for the data-generating distribution (and that is one where almost nothing is known). Regardless, all methods should have diagnostics that inform the amount of information a data set contains about the parameter of interest, and the ETA assumption is particularly relevant in this regard. By testing this assumption, we ensured there was experimentation in our data set, that is, that all men in the study population could have achieved each possible level of physical activity, conditional on his covariate distribution. Without sufficient experimentation, causal interpretations and deductive inference are not valid.

This study has certain limitations. There was an unquantifiable amount of measurement error in the PASE scores. We do not know the error structure, but we expect any bias to be toward the null. Our results are susceptible to bias from unmeasured confounding. For example, low serum 25-hydroxyvitamin D levels are associated with increased hip fracture risk (43, 44), but we did not include serum 25-hydroxyvitamin D in the set of potential confounders because these measures were available only in a subset of the cohort ($n = 1,608$). Relative risks were estimated rather imprecisely because of the relatively small number of hip fractures. In all analyses, we modeled risk over the entire duration of follow-up. We could have modeled the hazard over discrete periods of time, such as 1 year. Because hip fracture was rare, it is unlikely that modeling the hazard would substantially change the results.

Despite these limitations, the Osteoporotic Fractures in Men Study cohort is one of the largest cohorts of well-characterized older men with adjudicated hip fractures. Moreover, for reasons of prohibitive size and cost, it is unlikely that a randomized controlled trial will be conducted to test the effects of physical activity on hip fracture. Thus, associations from observational studies, such as this one, represent the best available evidence about the potential impacts of physical activity on hip fracture.

In conclusion, the results of this study, based on an approach that was completely agnostic to the data-generating distribution but that used optimal estimation techniques within that model, do not support a protective effect of overall amount of usual physical activity on hip fracture risk in older men.

ACKNOWLEDGMENTS

Author affiliations: Division of Epidemiology, School of Public Health, University of California, Berkeley, California (Dawn C. Mackey, Ira B. Tager); San Francisco Coordinating Center, California Pacific Medical Center Research Institute, San Francisco, California (Dawn C. Mackey, Peggy M. Cawthon, Steven R. Cummings); Division of Biostatistics, School of Public Health, University of California, Berkeley, California (Alan E. Hubbard); and Department of Epidemiology, Graduate School of Public Health, University of Pittsburgh, Pittsburgh, Pennsylvania (Jane A. Cauley).

The Osteoporotic Fractures in Men Study is supported by National Institutes of Health funding. The following institutes provide support: the National Institute of Arthritis and Musculoskeletal and Skin Diseases, the National Institute on Aging, the National Center for Research Resources, and the National Institutes of Health Roadmap for Medical Research under the following grant numbers: U01 AR45580, U01 AR45614, U01 AR45632, U01 AR45647, U01 AR45654, U01 AR45583, U01 AG18197, U01-AG027810, and UL1 RR024140.

Results contained within this article were presented at the Society for Epidemiologic Research Annual Meeting, June 23–26, 2009, Anaheim, California, and published in abstract form (45).

Conflict of interest: none declared.

REFERENCES

- Bentler SE, Liu L, Ombrian M, et al. The aftermath of hip fracture: discharge placement, functional status change, and mortality. *Am J Epidemiol*. 2009;170(10):1290–1299.
- Brauer CA, Coca-Perrillon M, Cutler DM, et al. Incidence and mortality of hip fractures in the United States. *JAMA*. 2009;302(14):1573–1579.
- Center JR, Nguyen TV, Schneider D, et al. Mortality after all major types of osteoporotic fracture in men and women: an observational study. *Lancet*. 1999;353(9156):878–882.
- Liu-Ambrose TY, Khan KM, Eng JJ, et al. Both resistance and agility training increase cortical bone density in 75- to 85-year-old women with low bone mass: a 6-month randomized controlled trial. *J Clin Densitom*. 2004;7(4):390–398.
- Karinkanta S, Heinonen A, Sievänen H, et al. A multi-component exercise regimen to prevent functional decline and bone fragility in home-dwelling elderly women: randomized, controlled trial. *Osteoporos Int*. 2007;18(4):453–462.
- Fiatarone MA, O'Neill EF, Ryan ND, et al. Exercise training and nutritional supplementation for physical frailty in very elderly people. *N Engl J Med*. 1994;330(25):1769–1775.
- Lord SR, Ward JA, Williams P, et al. The effect of a 12-month exercise trial on balance, strength, and falls in older women: a randomized controlled trial. *J Am Geriatr Soc*. 1995;43(11):1198–1206.
- Schlicht J, Camaione DN, Owen SV. Effect of intense strength training on standing balance, walking speed, and sit-to-stand performance in older adults. *J Gerontol A Biol Sci Med Sci*. 2001;56(5):M281–M286.
- Wolf SL, Barnhart HX, Ellison GL, et al. The effect of Tai Chi Quan and computerized balance training on postural stability in older subjects. *Phys Ther*. 1997;77(4):371–381.
- Liu-Ambrose T, Khan KM, Eng JJ, et al. Resistance and agility training reduce fall risk in women aged 75 to 85 with low bone mass: a 6-month randomized, controlled trial. *J Am Geriatr Soc*. 2004;52(5):657–665.
- Lord SR, Castell S, Corcoran J, et al. The effect of group exercise on physical functioning and falls in frail older people living in retirement villages: a randomized, controlled trial. *J Am Geriatr Soc*. 2003;51(12):1685–1692.
- Province MA, Hadley EC, Hornbrook MC, et al. The effects of exercise on falls in elderly patients. A preplanned meta-analysis of the FICSIT Trials. Frailty and Injuries: Cooperative Studies of Intervention Techniques. *JAMA*. 1995;273(17):1341–1347.
- Campbell AJ, Robertson MC, Gardner MM, et al. Randomised controlled trial of a general practice programme of home based exercise to prevent falls in elderly women. *BMJ*. 1997;315(7115):1065–1069.
- Moayyeri A. The association between physical activity and osteoporotic fractures: a review of the evidence and implications for future research. *Ann Epidemiol*. 2008;18(11):827–835.
- Feskanich D, Willett W, Colditz G. Walking and leisure-time activity and risk of hip fracture in postmenopausal women. *JAMA*. 2002;288(18):2300–2306.
- Gregg EW, Cauley JA, Seeley DG, et al. Physical activity and osteoporotic fracture risk in older women. Study of Osteoporotic Fractures Research Group. *Ann Intern Med*. 1998;129(2):81–88.
- Høidrup S, Sørensen TI, Strøger U, et al. Leisure-time physical activity levels and changes in relation to risk of hip fracture in men and women. *Am J Epidemiol*. 2001;154(1):60–68.
- Blank JB, Cawthon PM, Carrion-Petersen ML, et al. Overview of recruitment for the Osteoporotic Fractures in Men Study (MrOS). *Contemp Clin Trials*. 2005;26(5):557–568.
- Orwoll E, Blank JB, Barrett-Connor E, et al. Design and baseline characteristics of the Osteoporotic Fractures in Men (MrOS) Study—a large observational study of the determinants of fracture in older men. *Contemp Clin Trials*. 2005;26(5):569–585.
- Washburn RA, Smith KW, Jette AM, et al. The Physical Activity Scale for the Elderly (PASE): development and evaluation. *J Clin Epidemiol*. 1993;46(2):153–162.
- Block G, Woods M, Potosky A, et al. Validation of a self-administered diet history questionnaire using multiple diet records. *J Clin Epidemiol*. 1990;43(12):1327–1335.

22. Cummings SR, Cawthon PM, Ensrud KE, et al. BMD and risk of hip and nonvertebral fractures in older men: a prospective study and comparison with older women. *J Bone Miner Res.* 2006;21(10):1550–1556.
23. Cawthon PM, Fullman RL, Marshall L, et al. Physical performance and risk of hip fractures in older men. *J Bone Miner Res.* 2008;23(7):1037–1044.
24. Neyman J. On the application of probability theory to agricultural experiments. *Stat Sci.* 1990;5(4):465–480.
25. Rubin DB. Bayesian inference for causal effects: the role of randomization. *Ann Statist.* 1978;6(1):34–58.
26. van der Laan MJ, Robins JM. *Unified Methods for Censored Longitudinal Data and Causality.* New York, NY: Springer-Verlag; 2003.
27. Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol.* 2008;168(6):656–664.
28. Mortimer KM, Neugebauer R, van der Laan M, et al. An application of model-fitting procedures for marginal structural models. *Am J Epidemiol.* 2005;162(4):382–388.
29. Petersen ML, Wang Y, van der Laan MJ, et al. Assessing the effectiveness of antiretroviral adherence interventions. Using marginal structural models to replicate the findings of randomized controlled trials. *J Acquir Immune Defic Syndr.* 2006;43(suppl 1):S96–S103.
30. Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology.* 2000;11(5):550–560.
31. Sinisi SE, van der Laan MJ. Deletion/substitution/addition algorithm in learning with applications in genomics. *Stat Appl Genet Mol Biol.* 2004;3(1):article 18. (doi: 10.2202/1544-6115.1069).
32. van der Laan MJ, Gruber S. Collaborative double robust targeted maximum likelihood estimation. *Int J Biostat.* 2010;6(1):article 17. (doi: 10.2202/1557-4679.1181).
33. van der Laan MJ, Rubin D. Targeted maximum likelihood learning. *Int J Biostat.* 2006;2(1):article 11. (doi: 10.2202/1557-4679.1043).
34. Kujala UM, Kaprio J, Kannus P, et al. Physical activity and osteoporotic hip fracture risk in men. *Arch Intern Med.* 2000;160(5):705–708.
35. Meyer HE, Tverdal A, Falch JA. Risk factors for hip fracture in middle-aged Norwegian women and men. *Am J Epidemiol.* 1993;137(11):1203–1211.
36. Mussolino ME, Looker AC, Madans JH, et al. Risk factors for hip fracture in white men: the NHANES I Epidemiologic Follow-up Study. *J Bone Miner Res.* 1998;13(6):918–924.
37. Grisso JA, Kelsey JL, O'Brien LA, et al. Risk factors for hip fracture in men. Hip Fracture Study Group. *Am J Epidemiol.* 1997;145(9):786–793.
38. Michaëlsson K, Olofsson H, Jensevik K, et al. Leisure physical activity and the risk of fracture in men. *PLoS Med.* 2007;4(6):e199. (doi: 10.1371/journal.pmed.0040199).
39. Paganini-Hill A, Chao A, Ross RK, et al. Exercise and other factors in the prevention of hip fracture: the Leisure World Study. *Epidemiology.* 1991;2(1):16–25.
40. Robling AG, Burr DB, Turner CH. Partitioning a daily mechanical stimulus into discrete loading bouts improves the osteogenic response to loading. *J Bone Miner Res.* 2000;15(8):1596–1602.
41. Robling AG, Burr DB, Turner CH. Recovery periods restore mechanosensitivity to dynamically loaded bone. *J Exp Biol.* 2001;204(pt 19):3389–3399.
42. Cawthon PM, Ewing SK, McCulloch CE, et al. Loss of hip BMD in older men: the Osteoporotic Fractures in Men (MrOS) Study. *J Bone Miner Res.* 2009;24(10):1728–1735.
43. Cauley JA, Lacroix AZ, Wu L, et al. Serum 25-hydroxyvitamin D concentrations and risk for hip fractures. *Ann Intern Med.* 2008;149(4):242–250.
44. Looker AC, Mussolino ME. Serum 25-hydroxyvitamin D and hip fracture risk in older U.S. white adults. *J Bone Miner Res.* 2008;23(1):143–150.
45. Mackey DC, Hubbard A, Tager I, et al. Physical activity effects on hip fracture in older men [abstract]. *Am J Epidemiol.* 2009;169(suppl):S94.