



Original article

Long-term weight changes and risk of rheumatoid arthritis among women in a prospective cohort: a marginal structural model approach

Nathalie E. Marchand ¹, Jeffrey A. Sparks ¹, Susan Malspeis¹, Kazuki Yoshida¹, Lauren Prisco¹, Xuehong Zhang^{2,3}, Karen Costenbader¹, Frank Hu^{2,3,4}, Elizabeth W. Karlson¹ and Bing Lu¹

Abstract

Objective. To examine the association of long-term weight change with RA risk in a large prospective cohort study.

Methods. The Nurses' Health Study II started in 1989 (baseline); after exclusions, we studied 108 505 women 25–42 years old without RA. Incident RA was reported by participants and confirmed by medical record review. Body weight was reported biennially through 2015. We investigated two time-varying exposures: weight changes from baseline and from age 18; change was divided into five categories. We used a marginal structural model approach to account for time-varying weight change and covariates.

Results. Over 2 583 266 person-years, with a median follow-up time of 25.3 years, 541 women developed RA. Compared with women with stable weight from baseline, weight change was significantly associated with increased RA risk [weight gain 2–<10 kg: RR = 1.98 (95% CI 1.38, 2.85); 10–<20 kg: RR = 3.28 (95% CI 2.20, 4.89); ≥20 kg: RR = 3.81 (95% CI 2.39, 6.07); and weight loss >2 kg: RR = 2.05 (95% CI 1.28, 3.28)]. Weight gain of 10 kg or more from age 18 compared with stable weight was also associated with increased RA risk [10–< 20 kg: RR = 2.12 (95% CI 1.37, 3.27), ≥20 kg: RR = 2.31 (95% CI 1.50, 3.56)]. Consistent findings were observed for seropositive and seronegative RA.

Conclusion. Long-term weight gain was strongly associated with increased RA risk in women, with weight gain of ≥20 kg associated with more than a three-fold increased RA risk. Maintenance of healthy weight may be a strategy to prevent or delay RA.

Key words: rheumatoid arthritis, epidemiology, statistics, study design, adipose obesity

Rheumatology key messages

- Long-term weight gain during adult life may nearly quadruple rheumatoid arthritis risk in women.
- Rheumatoid arthritis risk increased starting with a weight gain of 2–10 kilograms from study baseline.

Introduction

Several lifestyle factors, including body composition [1], diet [2–7], smoking [8] and physical activity [9] have been associated with RA risk, while being overweight or obese [1] is among the strongest of these risk factors. However, while previous studies have examined the association of prevalent obesity with RA risk, the effect of long-term weight change, an indicator of incident obesity, on RA risk has received limited attention [10]. In

¹Division of Rheumatology, Inflammation, and Immunity, Brigham & Women's Hospital and Harvard Medical School, ²Channing Division of Network Medicine, Department of Medicine, Brigham & Women's Hospital, ³Department of Nutrition and ⁴Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA

Submitted 8 March 2021; accepted 2 July 2021

Correspondence to: Bing Lu, Division of Rheumatology, Inflammation and Immunity, Brigham & Women's Hospital, 60 Fenwood Rd—0616v, Boston, MA 02115, USA. E-mail: blu1@bwh.harvard.edu

general, long-term weight change, and weight gain in particular, can bring with it myriad negative consequences including an increased risk of diabetes [11], some cancers [12] and mortality [13, 14]. Thus, the association between long-term weight change and incident RA risk is of interest, particularly weight changes leading to incident obesity during adult life.

An analysis of weight change and RA risk in prospective cohort studies may be limited by time-varying confounders, which may themselves be affected by previous weight change, that lie on the causal pathway between weight change and RA [15] (Fig. 1), limiting the ability of traditional statistical methods to account for these mediators and detect a possible true association. Marginal structural models (MSM) address these factors and allow us to obtain unbiased estimates [15]. Because being overweight and obesity has been associated with increased RA risk [1], we hypothesized that long-term weight change, particularly weight gain, might be associated with RA risk independently of baseline weight. We studied weight change from study baseline, and also from age 18, to isolate the effect of adult weight gain on the risk of RA. Onset of RA has already been associated with weight loss, and rheumatoid cachexia in particular [16]; therefore, we also conducted lagged analyses to address potential reverse causation.

Materials and methods

Study population

The NHSII is a prospective cohort study established in 1989, that enrolled 116 429 female nurses 25–42 years old living in 14 US states. Participants have provided data regarding health, weight, lifestyle and medical history on extensive biennial questionnaires. We studied women, followed from 1989 to 2015, excluding those who did not provide data on weight ($n=5581$), and those with prevalent RA or other self-reported CTD diagnosed before 1989 ($n=615$), leaving 108 505 women for

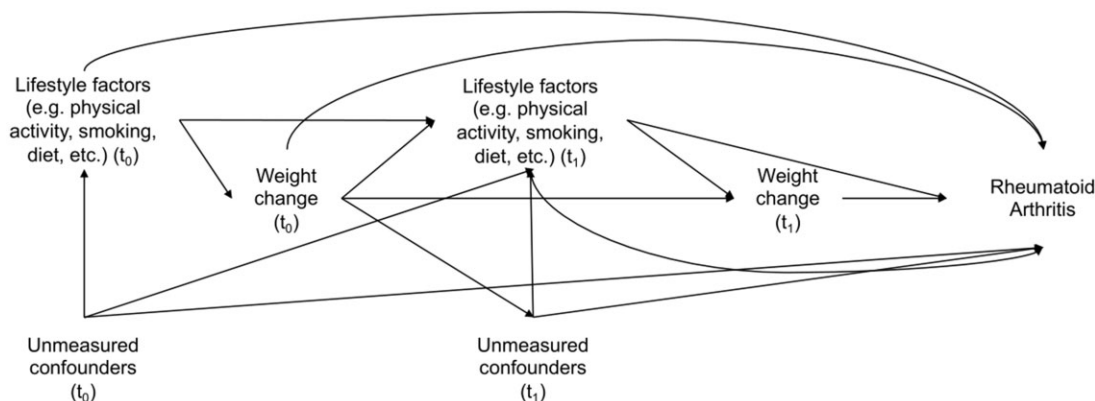
analysis at baseline. Study participants provided written informed consent. The study protocol was approved by the Institutional Review Boards of the Brigham and Women’s Hospital and the Harvard T.H. Chan School of Public Health; the Partners HealthCare Institutional Review Board also approved (approval #2011P001730) the study.

Weight assessment

Information on self-reported body weight was collected every 2 years from 1989 continuing through to the end of the study, using a previously validated instrument [17]. We truncated values above 350 pounds (any value >350 was equal to 350; $n=200$). Self-reported and technician measured weight were highly correlated, with a Pearson correlation coefficient of 0.97 [17, 18]. If women were pregnant when body weight was measured, their previous weight was utilized when available (when not available, they were excluded from the study). Weight change from baseline was calculated for every 2-year interval of study cycles as weight reported in 1989 subtracted from current reported weight. The calculated weight change was divided into five categories (as has been done similarly in other NHSII studies [4]): weight loss >2 kg from baseline; stable weight (reference group) defined as remaining within 2 kg above or below baseline weight; weight gain of 2 to <10 kg; weight gain of 10 to <20 kg; and weight gain of 20 kg or more from baseline. Similar cut-off points for weight change have been used in Nurses’ Health Study analyses examining other endpoints [14].

Weight at age 18 was self-reported on the 1989 questionnaire; extreme values were truncated at 350 pounds ($n=44$). A validation study conducted in NHSII comparing recalled weight to records reporting weight taken at physical examinations found a high correlation of 0.87 between the two measures, with recalled weight tending to be slightly lower (therefore the mean difference in BMI was 0.5 kg/m²) [18, 19]. Weight change from age 18

Fig. 1 Directed acyclic graph showing the relationships between weight change and rheumatoid arthritis in the presence of time-varying confounding



was calculated for every 2-year study cycle as the difference between current reported weight and weight at age 18, and divided into five categories: weight loss >2 kg; stable weight defined as remaining within 2 kg of reported weight at age 18; weight gain of 2 to <10 kg; weight gain of 10 to <20 kg; and weight gain of 20 kg or more.

Identification of incident RA cases

Women were asked about having a self-reported physician diagnosis of RA on the biennial questionnaires. Women who reported a new RA diagnosis were sent the validated CTD screening questionnaire [20]. Two rheumatologists then independently reviewed their medical records to confirm RA meeting the 1987 ACR [21] or 2010 ACR/EULAR criteria [22]. The date of incident RA was defined as the date of clinical diagnosis. Serostatus was determined by collecting clinical data from medical records on tests done for RF or ACPA at the time of diagnosis for cases diagnosed after ACPA assays became available clinically. RA cases were classified as having either seropositive (RF and/or ACPA positive) or seronegative (RF and ACPA negative) RA.

Covariates

Covariate data were collected using biennial questionnaires. Time-varying covariates, associated with both weight change (Table 1) and RA, included age (years), smoking (0, 0–20 and 20+ pack years) [23], parity and breastfeeding (nulliparous, parous/no breastfeeding, parous/1–12 months breastfeeding, parous/>12 months breastfeeding) [24], census tract median family income (quartiles) [25], Alternative Healthy Eating Index (AHEI) score (quartiles) [2] (based on food frequency questionnaires collected every 4 years [26]), physical activity [0–<3, 3–<9, 9–<18, ≥18 metabolic equivalents (METs) per week] [9], menopausal status and postmenopausal hormone use (pre-menopausal, post-menopausal with never use, current use and past use) [27]. Methodology for the AHEI score has been described previously [28]. Briefly, it is a score of diet quality that includes 11 components including vegetables, fruit, whole grains, sugar-sweetened beverages and fruit juice, nuts and legumes, red/processed meats, trans fat, long-chain omega-3 fatty acids (eicosapentaenoic acid and docosahexaenoic acid), poly-unsaturated fatty acids, sodium and alcohol [28]. All components are scored from 0 (worst) to 10

TABLE 1 Age-standardized characteristics in Nurses' Health Study II at baseline in 1989

	Categories of weight change in 1991 from 1989 ^a				
	Loss (>2 kg) (n = 14 510)	Stable (–2 to <2 kg) (n = 51 202)	Gain (2 to <10 kg) (n = 37 516)	Gain (10 to <20 kg) (n = 4558)	Gain (≥20 kg) (n = 719)
Age (years) ^b	34.5 (4.6)	34.6 (4.6)	34.5 (4.7)	33.7 (4.7)	34.1 (4.8)
Weight (kg) in 1989	72.7 (17.0)	62.1 (12.9)	66.0 (13.8)	74.2 (16.7)	80.4 (16.9)
Census tract median family income (\$1000)	42.5 (14.8)	43.7 (15.5)	42.5 (14.9)	41.9 (14.5)	40.4 (14.4)
Alternative Healthy Eating Index score ^c	47.3 (10.7)	48.5 (11.15)	48.9 (11.0)	47.9 (10.7)	47.2 (10.6)
Physical activity (MET-h/wk)	22.5 (34.4)	26.2 (38.8)	24.3 (35.6)	24.1 (32.9)	28.5 (42.7)
Smoking status					
0 pack years, %	62.5	66.1	65.3	63.9	60.4
0–20 pack years, %	30.7	28.8	28.9	29.6	30.9
20+ pack years, %	6.8	5.1	5.8	6.5	8.7
Parity and breastfeeding					
Nulliparous, %	31.4	33.5	32.8	38.4	51.1
Parous/breastfeeding 0 mo, %	15.0	13.7	14.5	14.6	14.7
Parous/breastfeeding ≤12 mo, %	26.1	24.4	25.7	24.6	19.1
Parous/breastfeeding >12 mo, %	27.5	28.4	27.0	22.4	15.2
Menopausal status and postmenopausal hormone use					
Premenopausal, %	97.3	97.9	97.5	97.3	96.4
Postmenopausal-never use, %	0.1	0.1	0.1	0.0	0.2
ostmenopausal-current use, %	2.3	1.9	2.2	2.4	3.1
Postmenopausal-past use, %	0.2	0.1	0.2	0.3	0.4

^aWeight change categories for this table defined using weight in 1991 (weight in 1989 subtracted from weight in 1991).

^bValue is not age adjusted. ^cThe Alternative Healthy Eating Index score has a potential of range of 0 (no adherence and therefore poorest diet quality) to 110 (perfect adherence and therefore highest diet quality) points. Values are means (standard deviation) or medians (Q25, Q75) for continuous variables; percentages or ns or both for categorical variables, and are standardized to the age distribution of the study population. Values of polytomous variables may not sum to 100% due to rounding.

(best), and the total AHEI score can range from 0 to 110. A higher AHEI score indicates better diet quality. Previous studies have revealed a protective association between moderate alcohol intake and RA risk [5]. Women with moderate alcohol intake defined as 0.5–1.5 drinks/day received a maximum score of 10 but non-drinkers or heavy drinkers were assigned lower scores [28]. Physical activity was measured using validated questionnaires where participants were asked about average amount of time spent engaged in several recreational physical activities per week [29]. A total recreational physical activity estimate was calculated for each woman using the average time spent on 11 common activities using a published MET score [30]. Hours per week spent on each activity was multiplied by the MET score, and these were summed together to create a total estimate of physical activity.

Statistical analyses

Censoring events included date of RA diagnosis, death, loss to follow-up (censoring), or 1 June 2015 (end of study period), whichever occurred first. We calculated age-adjusted descriptive statistics in order to summarize participant characteristics. We used MSMs to estimate the effect of weight changes (multinomial exposure variables) both from baseline and from age 18 on RA risk [15, 31, 32]. MSMs allow for exposure switching (switching between weight change categories over the course of the study) and time-varying confounders [15, 32].

To develop our MSMs, we first calculated the stabilized weight for the exposure variable as a ratio of two conditional probabilities: the numerator was the probability of adherence to a particular weight change category (in these analyses we had five categories, ranging from weight loss to weight gain) given weight change in the previous cycle [33] and the denominator was the probability of adherence to a certain weight change category, given weight change in the previous cycle, time-varying and baseline confounders. We used a multinomial logistic regression model with our categorical weight change variable as the response variable where, as described above, the numerator included weight change in previous cycle as the explanatory variable, and the model for the denominator included weight at baseline (for the weight change from baseline analysis) or at age 18 (for the weight change from age 18 analysis), weight change in the previous cycle, and time-varying confounders including: age smoking, parity and breastfeeding, census tract median family income, alcohol intake, physical activity, and menopausal status and postmenopausal hormone use. Similarly, we calculated the stabilized weight for censoring using a binomial logistic regression model where censoring was the response variable. The overall weight was the product of stabilized weights for exposure and censoring. The final weight was the cumulative product of all previous weights for each study period, thereby accounting for all information from previous periods at each timepoint, truncated at 99.5% and 0.5% [34, 35]. We applied the final weight to our study sample to create

a pseudo-population and developed pooled logistic regression models to estimate the odds ratio (OR) and robust 95% CI using a Generalized Estimating Equation approach [32, 33]. With rare outcomes such as incident RA, the OR can be approximated as a risk ratio (RR).

We then performed secondary analyses to investigate the association between weight change and RA risk by serostatus (seropositive RA or seronegative RA) as separate outcomes using the same methods as the primary analysis. Separate datasets were created to exclude those with seropositive RA for the seronegative RA analyses, and likewise to exclude those with seronegative RA for the seropositive RA analyses. In addition, to investigate whether reverse causation of weight change as a manifestation of early RA may have explained the findings, we conducted lagged analyses, where onset of RA was separated by one additional study cycle (2–4 years) from the assessment of weight changes but not other covariates. In sensitivity analyses, adjustment for smoking as a never/past/current variable, rather than as pack years of smoking, did not change results. We did not conduct analyses stratified by age – for example, dichotomized at age 50 or 55 – as has been done in previous Nurses' Health Study analyses, because most participants in NHSII were <50 years of age at baseline.

All statistical tests were two-sided at a statistical significance level of 0.05, performed using SAS 9.4 (SAS Institute, Cary, NC, USA).

Results

Baseline characteristics

Of the 108 505 women included in this study over a maximum of 26 years, 541 developed RA; 357 of these cases were seropositive. Table 1 shows the age-standardized baseline characteristics according to categories of weight change. Women in the highest category of weight gain (≥ 20 kg) had a higher weight at study baseline compared with women in all other categories of weight change. Women with the greatest weight gain also participated in a greater number of total hours of physical activity and had a higher pack years of smoking at baseline. Further, more were postmenopausal and more were current or past users of post-menopausal hormones. Fewer of these women had children, and when they had children, fewer breastfed their babies for 12 or more months. They also consumed less alcohol and lived in areas with a lower median census tract family income. These preliminary results at baseline may provide insights for multivariable analyses. Additionally, seropositive and seronegative RA cases had similar baseline characteristics (Supplementary Table S1, available at *Rheumatology* online).

Weight change from baseline and RA risk

Weight gain from study baseline was associated with a significantly increased risk of all RA, seropositive RA and seronegative RA. Compared with stable weight, defined as remaining within a range of 2 kg above or

below baseline weight, weight gain from baseline of 2 to <10, 10 to <20 and ≥20 kg was associated with increased RA risk (RR=1.98, 95% CI 1.38, 2.85; RR=3.28, 95% CI 2.20, 4.89; RR=3.81, 95% CI 2.39, 6.07, respectively) (Table 2). Weight gain was significantly associated with increased seropositive RA risk (RR for weight gain of 2 to <10 kg=1.66, 95% CI 1.08, 2.56; RR for weight gain of 10 to <20 kg=3.17, 95% CI 1.98, 5.08; RR for weight gain of ≥20 kg=3.54, 95% CI 2.04, 6.12) and seronegative RA (RR for weight gain of 2 to <10 kg=2.94, 95% CI 1.49, 5.80; RR for weight gain of 10 to <20 kg=3.62, 95% CI 1.71, 7.67; RR for weight gain of ≥20 kg=4.64, 95% CI 1.92, 11.21) (Table 2). Weight loss of >2 kg was associated with increased RA risk (RR=2.05, 95% CI 1.28, 3.28) and seropositive RA risk (RR=2.08, 95% CI 1.21, 3.58) but not increased seronegative RA risk (RR=1.95, 95% CI 0.76, 5.05) (Table 2).

Weight change from age 18 and RA risk

Weight gain from age 18 of 10 kg or more was associated with significantly increased RA risk, as well as seropositive and seronegative RA risk. Weight gain of 10 to <20 and ≥20 kg from age 18 years was associated with increased RA risk (RR=2.12, 95% CI 1.37, 3.27; RR=2.31, 95% CI 1.50, 3.56, respectively), seropositive RA risk (RR=1.84, 95% CI 1.09, 3.10; RR=2.16, 95% CI 1.29, 3.61, respectively) and seronegative RA risk (RR=2.93, 95% CI 1.34, 6.40; RR=2.77, 95% CI 1.26, 6.08, respectively) (Table 2).

Lagged analyses

Results for lagged analyses, in which weight change from baseline was assessed 2–4 years prior to RA incidence, in order to protect against potential reverse causation, were similar but attenuated for all RA and

TABLE 2 Associations between weight change from baseline and age 18 with RA

Categories of time-updated weight change from baseline					
	Loss (>2 kg)	Stable (-2 to <2 kg)	Gain (2 to <10 kg)	Gain (10 to <20 kg)	Gain (≥20 kg)
All RA					
# of cases/person-years	48/197 984	51/450 712	193/974 788	131/453 192	69/201 566
RR (95% CI) Unadjusted	2.10 (1.42, 3.12)	Ref	1.71 (1.26, 2.33)	2.46 (1.78, 3.40)	2.88 (2.01, 4.14)
RR (95% CI) MSM ^a	2.05 (1.28, 3.28)	Ref	1.98 (1.38, 2.85)	3.28 (2.20, 4.89)	3.81 (2.39, 6.07)
Seropositive RA					
# of cases/person-years	35/197 778	38/450 190	116/973 710	91/452 792	48/201 404
RR (95% CI) Unadjusted	2.06 (1.30, 3.26)	Ref	1.38 (0.96, 1.99)	2.29 (1.57, 3.35)	2.69 (1.76, 4.12)
RR (95% CI) MSM ^a	2.08 (1.21, 3.58)	Ref	1.66 (1.08, 2.56)	3.17 (1.98, 5.08)	3.54 (2.04, 6.12)
Seronegative RA					
# of cases/person-years	13/197 425	13/449 678	77/972 561	40/452 145	21/201 117
RR (95% CI) Unadjusted	2.24 (1.04, 4.82)	Ref	2.68 (1.49, 4.82)	2.95 (1.58, 5.51)	3.44 (1.72, 6.87)
RR (95% CI) MSM ^a	1.95 (0.76, 5.05)	Ref	2.94 (1.49, 5.80)	3.62 (1.71, 7.67)	4.64 (1.92, 11.21)
Categories of time-updated weight change since age 18					
	Loss (>2 kg)	Stable (-2 to <2 kg)	Gain (2 to <10 kg)	Gain (10 to <20 kg)	Gain (≥20 kg)
All RA					
# of cases/person-years	32/174 225	27/210 913	111/801 307	168/676 627	201/698 871
RR (95% CI) Unadjusted	1.43 (0.85, 2.38)	Reference	1.08 (0.71, 1.65)	1.94 (1.29, 2.91)	2.23 (1.49, 3.33)
RR (95% CI) MSM ^a	1.23 (0.72, 2.11)	Reference	1.16 (0.74, 1.81)	2.12 (1.37, 3.27)	2.31 (1.50, 3.56)
Seropositive RA					
# of cases/person-years	26/174 065	19/210 693	68/800 407	104/675 826	140/698 123
RR (95% CI) Unadjusted	1.65 (0.91, 2.98)	Reference	0.94 (0.57, 1.57)	1.70 (1.05, 2.78)	2.21 (1.37, 3.56)
RR (95% CI) MSM ^a	1.30 (0.70, 2.42)	Reference	0.95 (0.56, 1.62)	1.84 (1.09, 3.10)	2.16 (1.29, 3.61)
Seronegative RA					
# of cases/person-years	6/173 867	8/210 494	43/799 612	64/674 884	61/696 922
RR (95% CI) Unadjusted	0.90 (0.31, 2.60)	Reference	1.41 (0.66, 3.01)	2.49 (1.19, 5.20)	2.29 (1.09, 4.78)
RR (95% CI) MSM ^a	1.03 (0.34, 3.06)	Reference	1.77 (0.79, 3.99)	2.93 (1.34, 6.40)	2.77 (1.26, 6.08)

Study conducted in the Nurses' Health Study II between 1989 and 2015. ^aAdjusted for previous questionnaire cycle weight change, baseline weight in 1989, age (years, continuous), smoking (never smoker, <20 pack years of smoking, 20+ pack years of smoking), parity and breastfeeding status (nulliparous, parous/no breastfeeding, parous/1–12 months breastfeeding, parous/>12 months breastfeeding), menopausal status and hormone use (pre-menopausal, post-menopausal with never use, current use and past use), census tract median family income (quartiles), physical activity (0–2.9, 3–8.9, 9–17.9, ≥18 METs/week), Alternative Healthy Eating Index (quartiles). kg, kilograms; MSM, marginal structural model.

seropositive RA (RR for weight gain of ≥ 20 kg = 2.24, 95% CI 1.37, 3.65 and 2.28, 95% CI 1.28, 4.06, respectively), but there was not a significantly increased risk of RA with weight loss for all RA, seropositive RA or seronegative RA (Table 3). In lagged analyses, weight loss since age 18 was likewise not significantly associated with RA risk, and results for weight gain since age 18 were similar but attenuated for all RA (RR for weight gain of 10 to < 20 kg = 1.81, 95% CI 1.17, 2.78; RR for weight gain of ≥ 20 kg = 1.72, 95% CI 1.12, 2.65) and seronegative RA (RR for weight gain of 10 to < 20 kg = 2.29, 95% CI 1.11, 4.74; RR for weight gain of ≥ 20 kg = 2.32, 95% CI 1.13, 4.79). Increased risk was not statistically significant for seropositive RA (RR for weight gain of 10 to < 20 kg = 1.65, 95% CI 0.98, 2.79; RR for weight gain of ≥ 20 kg = 1.53, 95% CI 0.90, 2.59) (Table 3).

Discussion

In this large prospective cohort of women, we found that long-term weight gain, both from study baseline (covering a period of up to 26 years; median age at baseline was 35 years) and from the age of 18, was associated with significantly increased RA risk. For those with the greatest amount of weight gain (≥ 20 kg), weight accumulated during adult life was associated with a nearly 4-fold increased RA risk, and a > 2 -fold increased risk when weight gain was calculated from age 18. This association was present for both seropositive and seronegative RA and was slightly stronger for seronegative RA, which is consistent with results from a Danish case control study that found that obesity was associated with seronegative RA and all RA [36]. Perhaps this is because the seronegative RA group includes a potentially heterogeneous population that may include those with different types of inflammatory arthritis, of which inflammatory osteoarthritis has previously been positively associated with BMI [36]. The increased RA risk we observed in association with increasing weight gain remained even in lagged analyses of 2–4 years before diagnosis, further supporting our findings.

Weight loss was also associated with an increased risk of all RA and seropositive RA in our main analysis, but this association did not remain in lagged analyses. This suggests that reverse causation due to RA cachexia (RA related weight loss 2–4 years prior to diagnosis) in the non-lagged analyses may have been responsible for these findings, and weight loss may, thus, be an indicator of prodromal disease, rather than a risk factor. Increased levels of inflammatory cytokines, in particular TNF- α , are observed in those with RA, which may serve to shift the body towards a catabolic state as it stimulates body protein degradation [37] and these increased cytokine levels have been observed up to 12 years before the onset of RA [38], indicating that rheumatoid cachexia may also begin before RA is diagnosed.

Being overweight or obese is an established risk factor for RA [39], but these studies have been studying the effect of prevalent obesity on disease risk. To our knowledge, the effect of weight change, particularly during adult life, on incident RA in the general population is under-researched. This includes the novel approach of including an evaluation of the association between incident obesity and RA risk, which is important given that the increased risk of incident RA we found was in younger and middle-aged women, and women reach their peak BMI between the ages of 50 and 59 years of age [40]. A previous study of 92 participants found that weight change 2 years after bariatric surgery was not significantly associated with later incident RA risk [10] even though weight loss following bariatric surgery has been associated with reduced disease activity among those with RA [41]. Long-term weight gain may play a role in increasing RA risk via systemic inflammation that is a known underlying feature of both RA [38] and weight gain [42]. Adipose tissue is a recognized source of inflammatory biomarkers including TNF- α and IL-6 [43], and adipokines including adiponectin, leptin, progranulin and lipocalin-2 [44]. Weight gain has been associated with changes in CRP, IL-6, leptin and adiponectin levels [42, 45]. Furthermore, those with RA have been shown to have significantly higher levels of soluble TNF receptor 2 (sTNFR2) (a proxy for TNF- α) [38] and modestly higher levels of IL-6 [38] many years in advance of disease onset, and leptin has been associated with increased disease duration [46]. We did not assess whether or not the weight gain found in this cohort was attributable predominantly to a gain in fat mass, rather than fat-free mass; however, body composition change during the menopause transition, which many of the women in the NHSII cohort would have been going through during the time of our study (occurring around 50 years of age), is typically characterized by fat mass change [47]. Therefore, for those already at an increased risk of RA via genetic background [48] or lifestyle choices (e.g. smoking, reduced physical activity), limiting weight gain and its associated inflammation may be a beneficial preventive action. However, as a single risk factor of RA, increased adiposity might not significantly increase clinical diagnosis of RA, compared with those who maintained a stable weight.

Using an MSM approach in our analyses allowed us to deal with the time-varying confounding. In addition, by conducting our analyses in the ‘pseudo-population’ we were able to statistically approximate the study conditions of a randomized controlled trial (RCT) in which we could specify hypothetical weight-change interventions of interest [49]. An RCT studying the effect of weight change on RA risk would face several prohibitory limitations: participants might not adhere to dietary or lifestyle interventions leading to a weight change of up to 20 or more kg [50]; ethical concerns associated with randomizing participants to long-term weight gain; and the large number of participants required to observe a sufficient number of RA cases over a potentially long

TABLE 3 Associations between weight change from baseline and age 18 with rheumatoid arthritis (2-year lag)

	Categories of time-updated weight change since baseline				
	Loss (>2 kg)	Stable (-2 to <2 kg)	Gain (2 to <10 kg)	Gain (10 to <20 kg)	Gain (≥20 kg)
All RA					
# of cases/person-years	45/177 960	66/411 716	188/891 110	119/396 235	52/170 106
RR (95% CI) Unadjusted	1.58 (1.08, 2.31)	Ref	1.32 (1.00, 1.75)	1.86 (1.38, 2.52)	1.88 (1.31, 2.71)
RR (95% CI) by MSM ^a	1.25 (0.77, 2.02)	Ref	1.57 (1.12, 2.22)	2.11 (1.45, 3.07)	2.24 (1.37, 3.65)
Seropositive RA					
# of cases/person-years	28/177 779	40/411 264	123/890 186	85/395 914	36/169 981
RR (95% CI) Unadjusted	1.62 (1.00, 2.63)	Ref	1.43 (1.00, 2.04)	2.20 (1.51, 3.20)	2.15 (1.37, 3.38)
RR (95% CI) by MSM ^a	1.18 (0.65, 2.16)	Ref	1.54 (1.00, 2.37)	2.19 (1.36, 3.51)	2.28 (1.28, 4.06)
Seronegative RA					
# of cases/person-years	17/177 482	26/410 809	65/889 135	34/395 381	16/169 768
RR (95% CI) Unadjusted	1.52 (0.82, 2.79)	Ref	1.16 (0.74, 1.83)	1.35 (0.81, 2.25)	1.47 (0.79, 2.74)
RR (95% CI) by MSM ^a	1.38 (0.63, 3.05)	Ref	1.65 (0.95, 2.86)	1.94 (1.07, 3.53)	2.15 (0.86, 5.39)
Categories of time-updated weight change since age 18					
	Loss (>2 kg)	Stable (-2 to <2 kg)	Gain (2 to <10 kg)	Gain (10 to <20 kg)	Gain (≥20 kg)
All RA					
# of cases/person-years	26/165 325	31/200 139	112/757 758	167/618 377	181/620 722
RR (95% CI) Unadjusted	1.02 (0.60, 1.71)	Ref	0.95 (0.64, 1.42)	1.73 (1.18, 2.54)	1.86 (1.27, 2.72)
RR (95% CI) by MSM ^a	0.80 (0.45, 1.42)	Ref	0.96 (0.62, 1.50)	1.81 (1.17, 2.78)	1.72 (1.12, 2.65)
Seropositive RA					
# of cases/person-years	19/165 176	21/199 929	68/756 943	112/617 692	123/620 086
RR (95% CI) Unadjusted	1.10 (0.59, 2.04)	Ref	0.85 (0.52, 1.39)	1.71 (1.07, 2.73)	1.86 (1.17, 2.96)
RR (95% CI) by MSM ^a	0.77 (0.39, 1.53)	Ref	0.76 (0.44, 1.30)	1.65 (0.98, 2.79)	1.53 (0.90, 2.59)
Seronegative RA					
No. of cases/person-yr	7/164 997	10/199 758	44/756 177	55/616 809	58/619 081
RR (95% CI) Unadjusted	0.85 (0.32, 2.23)	Ref	1.16 (0.58, 2.31)	1.77 (0.90, 3.47)	1.85 (0.94, 3.61)
RR (95% CI) MSM ^a	0.89 (0.32, 2.46)	Ref	1.62 (0.77, 3.41)	2.29 (1.11, 4.74)	2.32 (1.13, 4.79)

Study conducted in the Nurses' Health Study II between 1989 and 2015. ^aAdjusted for previous questionnaire cycle weight change, baseline weight, age (years, continuous), smoking (never smoker, <20 pack years of smoking, 20+ pack years of smoking), parity and breastfeeding status (nulliparous, parous/no breastfeeding, parous/1-12 months breastfeeding, parous/>12 months breastfeeding), menopausal status and hormone use (pre-menopausal, post-menopausal with never use, current use and past use), census tract median family income (quartiles), physical activity (0-2.9, 3-8.9, 9-17.9, ≥18 METs/week), Alternative Healthy Eating Index (quartiles), kg, kilograms; MSM, marginal structural model.

study period. To our knowledge, there have been no RCTs studying the effect of lifestyle interventions on RA risk, likely for the above-mentioned reasons. Therefore, using a statistical approach to approximate an RCT in an already existing longitudinal cohort was warranted.

To obtain unbiased estimates using an MSM approach, the following conditions are required: no measurement error, no unmeasured confounding and no model misspecification [15, 33]. In regard to the first, validated questionnaires were used to obtain all data on exposure, covariates [26, 29] and RA outcome [20], which was additionally confirmed by two independent rheumatologists. The presence of unmeasured confounding is possible, though the homogeneous nature of the NHSII cohort may have limited this. Model misspecification was probably also limited, as our results were consistent in both the main and secondary analyses, and robust when various cut-points were chosen for the exposure variable.

Notable strengths of this study include the long-term nature of this prospective cohort study with repeated measures of exposure, covariates and RA confirmed by ACR criteria [20, 26, 29] in which a relatively small percentage of women were lost to follow-up (90% active follow-up rate).

This study also has limitations of note. There may have been some measurement error for current weight; however, self-reported weight measurement was validated in this cohort [17, 18] with a high correlation (0.97) having been reported between technician and self-reported weight. Reported weight at age 18 has also been validated [18, 19] and still had a relatively high correlation (0.87) with medical records despite the recall nature of this question. Further, measurement error for weight would be non-differential (similar among cases and non-cases) and attenuate results [51]. There may also have been differential misclassification for diagnosed RA, instead of other arthritis types or chronic pain conditions. However, all study participants had multiple questionnaire years in which to report RA, all were trained nurses, and self-reported RA was verified by two independent rheumatologists following ACR criteria, thereby limiting the possibility of such misclassification. Due to the common concern for residual confounding, or confounders also being mediators, in cohort analyses, we used an MSM approach that effectively deals with time-varying confounding and mediators in the context of treatment switching (i.e. movement between categories of weight change over the study period) [15, 32]. Further, the incidence of RA in our cohort was slightly lower than that found in the United States (US) generally [52]. In addition, the NHSII cohort is predominantly composed of mostly white and well-educated health professionals, with initial age of 25–42 years. The total US population, with a much wider age distribution, may contain more people closer to the peak age of RA onset. Therefore, results may not be generalizable to other populations; replication in more diverse cohorts is recommended.

In conclusion, long-term weight gain during adult life is an important risk factor for incident RA among women.

This finding has considerable clinical implications for counselling of individuals at high risk for RA, in particular those with a family history of the disease, and should be confirmed in other populations including men and non-Whites.

Acknowledgements

We thank the participants of the NHSII for their dedicated participation in this longitudinal study as well as the staff members at the Channing Division of Network Medicine (Department of Medicine, Brigham and Women's Hospital and Harvard Medical School).

All authors were involved in drafting the article or revising it critically for important intellectual content, and all authors approved the final version to be submitted for publication. B.L. and N.E.M. had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Study conception and design: B.L., N.E.M., J.A.S., X.Z., K.C., K.Y., E.W.K., X.Z., F.H. Analysis and interpretation of data: B.L., N.E.M., J.A.S., S.M., X.Z., K.C., K.Y., E.W.K., F.H.

Funding: This work was supported by the National Institutes of Health [AR071326, AR061362, AR047782, AR049880, AR059073, AR069688, UM1 CA186107, U01 CA176726, U01 HL145386]. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Disclosure statement: N.E.M. reports personal fees from Pritikin Longevity Center; J.A.S. and K.C. report grants from National Institutes of Health; J.A.S. reports grants from Rheumatology Research Foundation, grants and personal fees from Bristol-Myers Squibb, and personal fees from Gilead, Inova Diagnostics, Optum and Pfizer outside the submitted work. No other disclosures relevant to this article were reported.

Data availability statement

Further information including the procedures to obtain and access data from the Nurses' Health Studies and Health Professionals Follow-up Study is described at <https://www.nurseshealthstudy.org/researchers> (contact email: nhsaccess@channing.harvard.edu) and <https://sites.sph.harvard.edu/hpfs/for-collaborators/>.

Supplementary data

Supplementary data are available at *Rheumatology* online.

References

- 1 Lu B, Hiraki LT, Sparks JA *et al*. Being overweight or obese and risk of developing rheumatoid arthritis among women: a prospective cohort study. *Ann Rheum Dis* 2014;73:1914–22.

- 2 Hu Y, Sparks JA, Malspeis S *et al.* Long-term dietary quality and risk of developing rheumatoid arthritis in women. *Ann Rheum Dis* 2017;76:1357–64.
- 3 Hu Y, Costenbader KH, Gao X *et al.* Mediterranean diet and incidence of rheumatoid arthritis in women. *Arthritis Care Res* 2015;67:597–606.
- 4 Hu Y, Costenbader KH, Gao X *et al.* Sugar-sweetened soda consumption and risk of developing rheumatoid arthritis in women. *Am J Clin Nutr* 2014;100:959–67.
- 5 Lu B, Solomon DH, Costenbader KH, Karlson EW. Alcohol consumption and risk of incident rheumatoid arthritis in women: a prospective study. *Arthritis Rheumatol* 2014;66:1998–2005.
- 6 Sparks JA, O'Reilly ÉJ, Barbhuiya M *et al.* Association of fish intake and smoking with risk of rheumatoid arthritis and age of onset: a prospective cohort study. *BMC Musculoskelet Disord* 2019;20:2.
- 7 Sparks JA, Barbhuiya M, Tedeschi SK *et al.* Inflammatory dietary pattern and risk of developing rheumatoid arthritis in women. *Clin Rheumatol* 2019;38:243–50.
- 8 Costenbader KH, Feskanich D, Mandl LA, Karlson EW. Smoking intensity, duration, and cessation, and the risk of rheumatoid arthritis in women. *Am J Med* 2006;119:503.e1–e9.
- 9 Liu X, Tedeschi SK, Lu B *et al.* Long-term physical activity and subsequent risk for rheumatoid arthritis among women: a prospective cohort study. *Arthritis Rheumatol* 2019;71:1460–71.
- 10 Maglio C, Zhang Y, Peltonen M *et al.* Bariatric surgery and the incidence of rheumatoid arthritis – a Swedish Obese Subjects study. *Rheumatology* 2020;59:303–9.
- 11 Resnick HE, Valsania P, Halter JB, Lin X. Relation of weight gain and weight loss on subsequent diabetes risk in overweight adults. *J Epidemiol Community Health* 2000;54:596–602.
- 12 Keum N, Greenwood DC, Lee DH *et al.* Adult weight gain and adiposity-related cancers: a dose-response meta-analysis of prospective observational studies. *J Natl Cancer Inst* 2015;107:djv088.
- 13 Chen C, Ye Y, Zhang Y, Pan X-F, Pan A. Weight change across adulthood in relation to all cause and cause specific mortality: prospective cohort study. *BMJ* 2019;367:l5584.
- 14 Sun Q, Townsend MK, Okereke OI *et al.* Adiposity and weight change in mid-life in relation to healthy survival after age 70 in women: prospective cohort study. *BMJ* 2009;339:b3796.
- 15 Fewell Z, Hernán MA, Wolfe F *et al.* Controlling for time-dependent confounding using marginal structural models. *Stata J* 2004;4:402–20.
- 16 Santo RCE, Fernandes KZ, Lora PS, Filippin LI, Xavier RM. Prevalence of rheumatoid cachexia in rheumatoid arthritis: a systematic review and meta-analysis: systematic Review of RA cachexia prevalence. *J Cachexia Sarcopenia Muscle* 2018;9:816–25.
- 17 Rimm EB, Stampfer MJ, Colditz GA *et al.* Validity of self-reported waist and hip circumferences in men and women. *Epidemiology* 1990;1:466–73.
- 18 Colditz GA, Coakley E. Weight, weight gain, activity, and major illnesses: the Nurses' Health Study. *Int J Sports Med* 1997;18(Suppl 3):S162–70.
- 19 Troy LM, Hunter DJ, Manson JE *et al.* The validity of recalled weight among younger women. *Int J Obes Relat Metab Disord* 1995;19:570–2.
- 20 Karlson EW, Sanchez-Guerrero J, Wright EA *et al.* A connective tissue disease screening questionnaire for population studies. *Ann Epidemiol* 1995;5:297–302.
- 21 Arnett FC, Edworthy SM, Bloch DA *et al.* The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis. *Arthritis Rheum* 1988;31:315–24.
- 22 Aletaha D, Neogi T, Silman AJ *et al.* 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Arthritis Rheum* 2010;62:2569–81.
- 23 Hernández Avila M, Liang MH, Willett WC *et al.* Reproductive factors, smoking, and the risk for rheumatoid arthritis. *Epidemiology* 1990;1:285–91.
- 24 Karlson EW, Mandl LA, Hankinson SE, Grodstein F. Do breast-feeding and other reproductive factors influence future risk of rheumatoid arthritis?: results from the Nurses' Health Study. *Arthritis Rheum* 2004;50:3458–67.
- 25 Bengtsson C, Nordmark B, Klareskog L, Lundberg I, Alfredsson L, EIRA Study Group. Socioeconomic status and the risk of developing rheumatoid arthritis: results from the Swedish EIRA study. *Ann Rheum Dis* 2005;64:1588–94.
- 26 Willett WC, Sampson L, Browne ML *et al.* The use of a self-administered questionnaire to assess diet four years in the past. *Am J Epidemiol* 1988;127:188–99.
- 27 Bengtsson C, Malspeis S, Orellana C *et al.* Association between menopausal factors and the risk of seronegative and seropositive rheumatoid arthritis: results from the nurses' health studies. *Arthritis Care Res* 2017;69:1676–84.
- 28 Chiuvè SE, Fung TT, Rimm EB *et al.* Alternative dietary indices both strongly predict risk of chronic disease. *J Nutr* 2012;142:1009–18.
- 29 Wolf AM, Hunter DJ, Colditz GA, Manson JE *et al.* Reproducibility and validity of a self-administered physical activity questionnaire. *Int J Epidemiol* 1994;23:991–9.
- 30 Ainsworth BE, Haskell WL, Leon AS *et al.* Compendium of physical activities: classification of energy costs of human physical activities. *Med Sci Sports Exerc* 1993;25:71–80.
- 31 Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;11:550–60.
- 32 SAS Institute. *Analysis of Observational Health Care Data Using SAS*. Cary, NC: SAS Publishing, 2010.
- 33 Hernán MA, Brumback B, Robins JM. Marginal structural models to estimate the causal effect of zidovudine on the survival of HIV-positive men. *Epidemiology* 2000;11:561–70.
- 34 Cole SR, Hernán MA. Constructing inverse probability weights for marginal structural models. *Am J Epidemiol* 2008;168:656–64.

- 35 Williamson T, Ravani P. Marginal structural models in clinical research: when and how to use them? *Nephrol Dial Transplant* 2017;32:ii84–90.
- 36 Pedersen M, Jacobsen S, Klarlund M *et al.* Environmental risk factors differ between rheumatoid arthritis with and without auto-antibodies against cyclic citrullinated peptides. *Arthritis Res Ther* 2006;8:R133.
- 37 Rall LC, Roubenoff R. Rheumatoid cachexia: metabolic abnormalities, mechanisms and interventions. *Rheumatology* 2004;43:1219–23.
- 38 Karlson EW, Chibnik LB, Tworoger SS *et al.* Biomarkers of inflammation and development of rheumatoid arthritis in women from two prospective cohort studies. *Arthritis Rheum* 2009;60:641–52.
- 39 Zhou Y, Sun M. A meta-analysis of the relationship between body mass index and risk of rheumatoid arthritis. *EXCLI J* 2018;17:1079–89.
- 40 Chumlea WC, Guo SS, Kuczmarski RJ *et al.* Body composition estimates from NHANES III bioelectrical impedance data. *Int J Obes Relat Metab Disord* 2002; 26:1596–609.
- 41 Sparks JA, Halperin F, Karlson JC, Karlson EW, Bermas BL. Impact of bariatric surgery on patients with rheumatoid arthritis: bariatric surgery and RA. *Arthritis Care Res* 2015;67:1619–26.
- 42 Fransson EI, Batty GD, Tabák AG *et al.* Association between change in body composition and change in inflammatory markers: an 11-year follow-up in the Whitehall II Study. *J Clin Endocrinol Metab* 2010;95:5370–4.
- 43 Galic S, Oakhill JS, Steinberg GR. Adipose tissue as an endocrine organ. *Mol Cellular Endocrinol* 2010;316: 129–39.
- 44 Francisco V, Pino J, Gonzalez-Gay MA *et al.* Adipokines and inflammation: is it a question of weight?: obesity and inflammatory diseases. *Br J Pharmacol* 2018;175: 1569–79.
- 45 Kimura Y, Pham NM, Yasuda K *et al.* Association of adulthood weight gain with circulating adipokine and insulin resistance in the Japanese population. *Eur J Clin Nutr* 2015;69:462–6.
- 46 Olama SM, Senna MK, Elarman M. Synovial/serum leptin ratio in rheumatoid arthritis: the association with activity and erosion. *Rheumatol Int* 2012;32: 683–90.
- 47 Greendale GA, Sternfeld B, Huang M *et al.* Changes in body composition and weight during the menopause transition. *JCI Insight* 2019;4:e124865.
- 48 Karlson EW, Chibnik LB, Cui J *et al.* Associations between human leukocyte antigen, PTPN22, CTLA4 genotypes and rheumatoid arthritis phenotypes of autoantibody status, age at diagnosis and erosions in a large cohort study. *Ann Rheum Dis* 2007;67:358–63.
- 49 Lajous M, Willett WC, Robins J *et al.* Changes in fish consumption in midlife and the risk of coronary heart disease in men and women. *Am J Epidemiol* 2013;178: 382–91.
- 50 Hernán MA, Alonso A, Logan R *et al.* Observational studies analyzed like randomized experiments: an application to postmenopausal hormone therapy and coronary heart disease. *Epidemiology* 2008;19:766–79.
- 51 Hart JE, Liao X, Hong B *et al.* The association of long-term exposure to PM2.5 on all-cause mortality in the Nurses' Health Study and the impact of measurement-error correction. *Environ Health* 2015;14:38.
- 52 Hunter TM, Boytsov NN, Zhang X *et al.* Prevalence of rheumatoid arthritis in the United States adult population in healthcare claims databases, 2004-2014. *Rheumatol Int* 2017;37:1551–7.