Osteoarthritis and Cartilage 28 (2020) 932–940

Osteoarthritis and Cartilage

Development and validation of a prediction model for incident hand osteoarthritis in the HUNT study

M.B. Johnsen † ‡ § * a, K. Magnusson ‖ ¶ a, S. Børte † ‡ §, M.E. Gabrielsen §, B.S. Winsvold ‡ §, A.H. Skogholt §, L. Thomas § #, K. Storheim ††, K. Hveem §, J.-A. Zwart † ‡ §

† *Institute of Clinical Medicine, University of Oslo, Oslo, Norway*
‡ *Department of Research, Innovation and Education, Division of Clinical Neuroscience, Oslo University Hospital, Oslo, Norway*
§ *K. G. Jebsen Center for Genetic Epidemiology, Department of Public Health and Nursing, Faculty of Medicine and Health, Norwegian University of Science and Technology, Trondheim, Norway*
‖ *Lund University, Faculty of Medicine, Department of Clinical Sciences, Clinical Epidemiology Unit, Lund, Orthopaedics, Lund, Sweden*
¶ *National Advisory Unit on Rehabilitation in Rheumatology, Department of Rheumatology, Diakonhjemmet Hospital, Oslo, Norway*
# *Department of Clinical and Molecular Medicine, Norwegian University of Science and Technology, Trondheim, Norway*
†† *Research and Communication Unit for Musculoskeletal Health, Oslo University Hospital, Oslo, Norway*

ARTICLE INFO

Article history:
Received 13 December 2019
Accepted 21 April 2020

Keywords:
Osteoarthritis
Hand
Prediction
Risk
Discrimination

SUMMARY

Objective: To develop and externally validate prediction models for incident hand osteoarthritis (OA) in a large population-based cohort of middle aged and older men and women.
Design: We included 17,153 men and 18,682 women from a population-based cohort, aged 35–70 years at baseline (1995–1997). Incident hand OA were obtained from diagnostic codes in the Norwegian National Patient Register (1995–2018). We studied whether a range of self-reported and clinically measured predictors could predict hand OA, using the Area Under the receiver-operating Curve (AUC) from logistic regression. External validation of an existing prediction model for male hand OA was tested on discrimination in a sample of men. Bootstrapping was used to avoid overfitting.
Results: The model for men showed modest discriminatory ability (AUC = 0.67, 95% CI 0.62–0.71). Adding a genetic risk score did not improve prediction. Similar discrimination was observed in the model for women (AUC = 0.62, 95% CI 0.59–0.64). Prediction was not improved by adding a genetic risk score or hormonal and reproductive factors. Applying external validation, similar results were observed among men in HUNT (The Nord-Trøndelag Health Study) as in the developmental sample (AUC = 0.62, 95% CI 0.57–0.65).
Conclusion: We developed prediction models for incident hand OA in men and women. For women, the model included body mass index (BMI), heavy physical work, high physical activity and perceived poor health. The model showed moderate discrimination. For men, we have shown that a prediction model including BMI, education and information on sleep can predict incident hand OA in several populations with moderate discriminative ability.

Introduction

The disease burden of osteoarthritis (OA) of the hands is regarded to be comparable to, or higher than that experienced with rheumatoid arthritis[1]. Since treatment options are limited, identifying factors for prevention and management of hand OA is important. Prediction models allow for the estimation of individual risk, and may be utilized for risk reduction and improved disease outcomes. Predicting the OA risk from a combination of risk factors has been attempted in several study samples for hip and knee OA,

* Address correspondence and reprint requests to: M.B. Johnsen, Institute of Clinical Medicine, University of Oslo, PO Box 1072 Blindern, N-0316, Oslo, Norway. Tel.: 47-23-01-63-24; Fax: 47-23-01-61-50.
E-mail addresses: m.b.johnsen@medisin.uio.no (M.B. Johnsen), mangusson_karin@outlook.com (K. Magnusson), sigrid.borte@gmail.com (S. Børte), maiken.e.gabrielsen@ntnu.no (M.E. Gabrielsen), bendik.s.winsvold@gmail.com (B.S. Winsvold), anne.heidi.skogholt@ntnu.no (A.H. Skogholt), laurent.thomas@ntnu.no (L. Thomas), kjersti.storheim@medisin.uio.no (K. Storheim), kristian.hveem@ntnu.no (K. Hveem), j.a.zwart@medisin.uio.no (J.-A. Zwart).
[a] shared 1. Authorship.

https://doi.org/10.1016/j.joca.2020.04.005

demonstrating moderate ability to discriminate between those who later do, or do not, develop hip and knee OA[2–4]. However, only one prediction model exists for hand OA[5]. This model had a moderate ability to predict hand OA (Area Under Curve [AUC] = 0.6)[5]. The study sample, in which the hand OA prediction model was developed, consisted of young men only, and its performance in other samples is unexplored. Thus, no prediction model exist for hand OA in middle aged or older men, or in women. Given the high prevalence of hand OA in these groups[6,7], the development, validation and refinement of prediction models is important for improved disease prevention. Potential relevant predictors for women, other than those relevant for men, may be hormonal and reproductive factors like parity, age at menarche and use of hormone replacement therapy[8–10]. Further, genetic factors seem to play an important role also in hand OA[11]. To date, two genome-wide association studies (GWAS) on hand OA have been performed with a few genetic variants being successfully identified[12,13]. Thus, the aim of this study was to develop new prediction models and externally validate the existing one, in a large population-based cohort of middle aged and older men and women.

Method

Study population

The Nord-Trøndelag Health Study (HUNT) is a large population-based cohort in Norway where all residents in the county of Nord-Trøndelag, aged 20 years and older, have been invited to participate. Data was collected through three cross-sectional surveys, HUNT1 (1984–1986), HUNT2 (1995–1997) and HUNT3 (2006–2008), and has been described in detail elsewhere[14,15], with the fourth survey recently completed (HUNT4 2017–2019). For the present study, we included participants in HUNT2 (1995–1997) with genotype information, who were free of hand OA, and between 35 and 70 years of age at baseline (Fig. 1). The HUNT2 baseline data were linked to medical diagnoses obtained from the Norwegian National Patient Register (NPR) (individual level data). Persons having hand OA prior to 1995 (baseline) were excluded. The NPR is a national administrative health register that contains information of all admissions to hospital, both public and private. All participants in HUNT2 have signed a written informed consent regarding the use of data from questionnaires, biological samples and linkage to other registries for research purposes[14,15]. The current study was approved by the Regional Committee for Medical and Health Research Ethics (REK) 2015/573. The study conforms to the recommendation of Transparent Reporting of a multivariable prediction model for Individual Prognosis Or Diagnosis (TRIPOD)[16].

Definition of hand OA

We defined incident hand OA as the first record in secondary care between baseline in HUNT2 (1995–1997) and December 31, 2018, with relevant ICD-9 and ICD-10 codes from the NPR (either one of the following; 715, M15, M18 or M19) (Supplementary material 1).

Self-reported predictors

We selected predictors based on the previous prediction model for hand OA[5], in addition to other putative predictors available from the baseline questionnaire in HUNT2 that could contribute to a better predictive performance (Fig. 2). Self-reported predictors included physically demanding work, leisure time physical activity, education, smoking status, alcohol consumption, diabetes, anxiety and depression [The Hospital and Anxiety Depression Scale (HADS)], health status and sleep problems. For women, we also

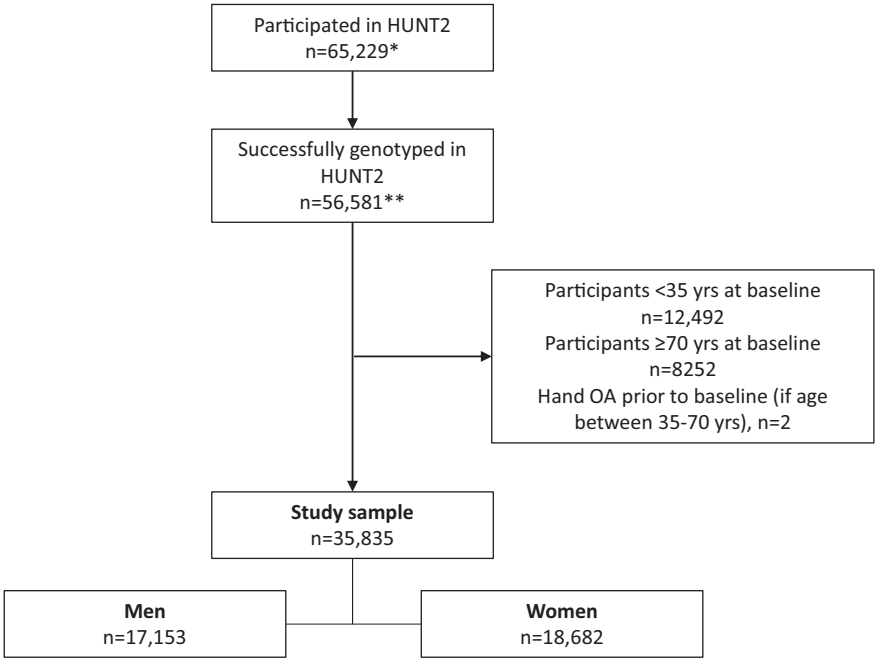

Fig. 1 Flowchart of the study sample included from the Nord-Trøndelag Health Study (HUNT). * Numbers of participants are available at https://hunt-db.medisin.ntnu.no/hunt-db/. ** Passed quality control (QC) and have information on hand OA (ICD codes).

Osteoarthritis and Cartilage

Self-reported predictors	Heavy physical work (yes/no) Education (primary, secondary, college/university) Leisure time physical activity (inactive/low, medium, high) Smoking status (never, former, current) Alcohol consumption (abstainers, <3, ≥3 units per week) Diabetes (yes/no) Anxiety and depression (mean HADS score) Sleep problems (never, seldom, often) Perceived health (good/poor) Parity (0, 1-2, ≥3)* Age of menarche (≤11, 12, 13, 14, ≥15 years)* Use of hormonal replacement therapy (never, past, current)*
Measured predictors	Systolic blood pressure (mmHg) Diastolic blood pressure (mmHg) Body mass index (kg/m^2) Genetic risk score**

Fig. 2 Baseline candidate predictors. *women only, ** DNA extracted from blood samples, including genotyping of genetic variants associated with hand OA to calculate the genetic risk score. DNA, deoxyribonucleic acid.

included parity, age at menarche and hormonal replacement therapy (HRT) as potential predictors that have previously been associated with OA[8–10]. A more detailed description of the measurement units, categorization and collection of the predictors are given in Supplementary material 2.

Clinically measured predictors

Body mass index (BMI) was calculated from height and weight that were measured at attendance to the clinical examination at baseline in HUNT2. Systolic and diastolic blood pressure (BP) levels were measured with three consecutive automatic oscillometric BP-measurements (Dinamap 845XT; Critikon, Tampa, FL, USA), recorded at 1-min intervals. The mean of the second and the third readings were calculated. BP was registered to the nearest 2 mm Hg (Supplementary material 2).

DNA was extracted from blood samples collected at and stored at the HUNT Biobank. Details of the calling, quality control and imputation of the nuclear genotypes in HUNT have been described elsewhere[17]. We used two previously published genome-wide association (GWA) studies on hand OA to select single-nucleotide polymorphisms (SNPs) for calculating a genetic risk score[12,13]. Three independent SNPs were included in the current genetic risk score (Supplementary material 3). Similar to previous hip and knee OA studies[2,3], we selected SNPs with *P*-value below $1x\ 10^{-7}$ and calculated the genetic risk score as the dosage of the effect allele for each SNP, weighted by the effect estimated reported in the GWA studies on hand OA[12,13]; Σ risk allele (non − weighted) × (log) odds ratio from GWAS *GWAS* (weighted).

Statistical analyses

All analyses were performed for men and women separately due to difference in incidence and prevalence of hand OA between genders[7]. First, we externally validated the existing model, in addition to an updated and simplified version of the existing prediction model for male hand OA from the Swedish study[5]. We used a sample of men (aged 35–70 years at baseline) from the general population who had participated in HUNT2. With the logit of the probability of hand OA as outcome [logit (p) = log(p /1 − p)], the existing prediction model included:

$$\text{Logit (hand OA)} = -6.84 + 1.03\text{*upper secondary school} > 3\ \text{yrs} + (-5.90\text{*college / university} > 3\ \text{yrs}) + 0.08\text{*BMI} + 0.18\text{*sleep problems}) + 0.22\text{*BMI*college / university} > 3\ \text{yrs}$$

The existing model was then updated to reflect differences in case mix and different measurement categories of predictors between our sample and the sample used to develop the existing model[5], e.g., to include age and a slightly different categorization of educational level. Moreover, we did not constrain the regression coefficients or the model intercept, but updated the intercept to reflect baseline risk of incident hand OA in the current study sample.

For women, all models were developed from the material available in HUNT2. For both men and women, the model development building procedure was as follows: First, we performed univariate analyses to assess the relationship between each candidate predictor and incident hand OA. Second, we included all the predictors from the univariate analysis in a saturated multivariable model, independent of statistical significance, since all the variables were selected based on the rationale of being a risk factor for hand OA. We performed logistic regression analysis with backward elimination of predictors with $P > 0.20$. In each subsequent step, we removed the predictor with the highest *P*-value

(and/or based on quality of the predictor, e.g., missing data and the information the variable comprised) if the removal of the predictor in question did not lead to a drop in the model's discriminative ability (AUC). Third, nonlinear effects of continuous variables were assessed using restricted cubic spline plots with the recommendations of Harrell[18] regarding the number and location of knots. If deviation from linearity were observed, we created linear splines with knots based on the plot to estimate linear effects. For categorical variables, we reduced or recoded the number categories if certain categories were found not to be of statistical significance ($P > 0.20$), or contribute to discrimination. We examined the fit of the models using the likelihood ratio test and Akaike information criterion (AIC). After identifying the most optimal parameterization and categorization, we included the predictors in a final model. Lastly, we included the genetic risk score and reproductive and hormonal factors (the latter for women only) in the final models to determine the independent and added predictive value of these variables[16,19], as performed in former studies predicting hip and knee OA[2,3]. Our models were fitted using samples without any missing data (complete-case analysis). A comparison of characteristics between those included vs excluded is depicted in Supplementary material 4.

The discrimination of the models was compared using Harrell's C statistic (AUC). To find the most universal model that can be applied across different samples in future populations and studies, we did not study interactions. However, as an alternative to interaction testing, we developed separate models for men and women[16], based on the hypothesis that the effect of the predictor(s) on hand OA might differ by sex.

For analysis of the reproductive and hormonal factors, we excluded women who had undergone a hysterectomy and/or oophorectomy as both these procedures can induce premature menopause[20,21], thus affect the use of HRT.

To correct for any potential overfitting and internally validate the final models, we performed a bootstrap resampling procedure with 1,000 iterations[22]. The AUCs and 95% CIs of the final models were based on this bootstrap resampling. We studied calibration plots of the final models, showing the agreement between observed and predicted values by sample deciles, where perfect predictions align along the 45° line[23]. All analyses were performed in STATA MP (general-purpose statistical software package for multiprocessor computers) v. 15.1.

	Men (n = 17,153)	Women (n = 18,682)
Age, years, mean (SD)	51.0 (9.6)	50.9 (9.7)
BMI, kg/m^2, mean (SD)	26.8 (3.4)	26.4 (4.5)
N incident hand OA cases	206	732
Education, n (%)		
Primary	5,482 (32.0)	7,757 (41.5)
Secondary	7,525 (43.9)	6,803 (36.4)
College/university	3,678 (21.4)	3,500 (18.7)
Missing	468 (2.7)	622 (3.3)
Smoking status, n (%)		
Never	5,683 (33.1)	7,460 (39.9)
Former	6,034 (35.2)	4,744 (25.4)
Current	5,261 (30.7)	6,191 (33.1)
Missing	175 (1.0)	287 (1.5)
Poor health, n (%)	4,223 (24.6)	5,412 (29.0)
Missing	109 (0.6)	189 (1.0)

SD = standard deviation, BMI = body mass index, OA = osteoarthritis.

Table I Baseline characteristics of the HUNT study sample

Osteoarthritis and Cartilage

Results

The study sample included in total 35,835 participants without hand OA at baseline, 17,153 men and 18,682 women. Mean age in the sample was 51.0 (SD 9.7), BMI 26.6 (SD 4.0) and mean follow-up of 20.6 years (SD 4.5). Incident hand OA was seen in 1.2% of the men and 3.9% of the women. Baseline characteristics of the sample are depicted in Table I. The results from the univariate analysis and the saturated multivariable models for men and women are shown in Supplementary material 5 and 6, respectively.

In the model including reproductive and hormonal factors for women, we excluded 1,432 women who reported at baseline to have undergone a hysterectomy and/or oophorectomy. This resulted in 17,250 women eligible for analysis, where 665 had incident hand OA (Supplementary material 6).

Validation of the existing risk prediction model for men (developed within the Swedish conscription register)

$$\text{Logit (hand OA)} = -6.84 + 1.03 * \text{upper secondary school} > 2\ \text{yrs} + (-5.90 * \text{college / university} \geq 4\ \text{yrs} + 0.18 * \text{sleep problems} + 0.22 * \text{BMI} * \text{upper secondary school} > 2\ \text{yrs}$$

External validation of the existing prediction model on hand OA amongst the men in HUNT showed a discriminatory ability of Area Under the receiver-operating Curve (AUC) 0.60 (0.56–0.64). This is comparable to the findings in the original Swedish sample (AUC 0.62, 95% CI 0.58–0.64).

Validation of the simplified model

$$\text{Logit (hand OA)} = -7.53 + 0.02 * \text{age} + 0.07 * \text{BMI} + 0.05 * \text{upper secondary school} > 2\ \text{yrs} + (-0.39 * \text{college / university} \geq 4\ \text{yrs}) + 0.55 * \text{sleep problems}$$

The validation of a simplified version of the existing prediction model on hand OA showed similar discriminatory ability amongst men in HUNT2, AUC 0.62 (95% CI 0.57–0.65), as in the original Swedish sample, AUC = 0.62 (95% CI 0.58–0.64)[5].

Development of risk prediction models for men in HUNT

For men, age was included as linear splines (knots at 50 and 60 years) and BMI were kept as continuous. Educational level (primary vs higher education) and sleep problems (never/seldom vs weekly) were collapsed from three to two categories based on the *P*-values and effect on the model discrimination. The discriminatory ability (AUC) of the final model was 0.67 (95% CI 0.62–0.71) for men (Table II). The model showed good agreement between observed and predicted values of incident hand OA (Fig. 3). AUC of the genetic risk score alone was 0.53 (95% CI 0.49–0.57). Including the genetic risk score in the final model did not change the discriminatory ability (Table II).

1. Final model for incident hand OA for men.

$$\text{Logit (hand OA)} = -10.82 + 0.07 * \text{age spline1} + (-0.05 * \text{age spline2}) + 0.06 * \text{age spline3} + 0.07*\text{BMI} + 0.39 * \text{higher education} + 0.77 * \text{heavy work} + 0.46 * \text{sleep problems}$$

2. Final model + genetic risk score for men.

$$\text{Logit (hand OA)} = -10.57 + 0.07 * \text{age spline1} + (-0.05 * \text{age spline2}) + 0.06 * \text{age spline3} + 0.07 * \text{BMI} + 0.39 * \text{higher education} + 0.77*\text{heavy work} + 0.46*\text{sleep problems} + (-0.40*\text{genetic risk score})$$

Development of risk prediction models in for women in HUNT

For women, age was included as linear splines (knots at 48, 51 and 53 years) and BMI (kg/m^2) was kept as continuous. We collapsed the physical activity categories into low physical activity (inactive/low) and high physical activity (medium/high), based on the *P*-values and effect on the model discrimination. AUC of the final prediction model was 0.62 (95% CI 0.59–0.64) for women (Table II). The model showed poor agreement between observed and predicted values of incident hand OA (Fig. 4). AUC of the genetic risk score alone was 0.53 (95% CI 0.51–0.55). Parity, age at menarche and HRT together had an AUC of 0.55 (95% CI 0.53–0.58). Neither the genetic risk score nor the reproductive and hormonal factors significantly changed the discriminatory ability when added to the final model (Table II).

	Men		Women		
	Final model†	Final model + genetic risk score†	Final model‡	Final model + genetic risk score‡	Final model + reproductive and hormonal factors§
AUC (95% CI)*	0.67 (0.62–0.71)	0.67 (0.62–0.71)	0.62 (0.59–0.64)	0.62 (0.60–0.64)	0.62 (0.60–0.65)
Sensitivity	65.5%	65.5%	56.7%	56.5%	56.1%
Specificity	61.1%	61.3%	59.2%	59.2%	61.6%
Proportion correctly classified	61.2%	61.1%	59.0%	59.1%	61.4%

* Bias corrected 95% CI, AUC = area under the curve, OA = osteoarthritis, the cut-offs for classification were the observed proportions with hand OA in men (0.012) and women (0.039). Included in the models.
† $n = 13,580$.
‡ $n = 14,858$.
§ $n = 12,888$.

Table II Discriminatory performance of the prediction models for hand OA in men and women

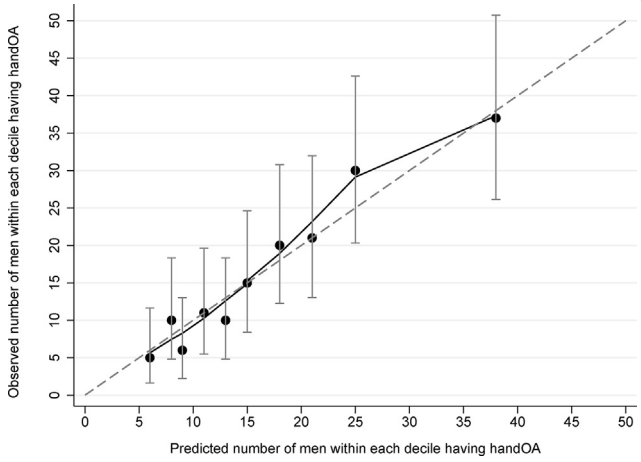

Fig. 3 Calibration plot for the final prediction model for incident hand OA for men. Black dots represent the agreement between observed and predicted values of hand OA with 95% confidence intervals. The 45° line represents perfect predictions. The solid black line is a smoothed calibration line depicting the average agreement between the observed and predicted number of men with hand OA.

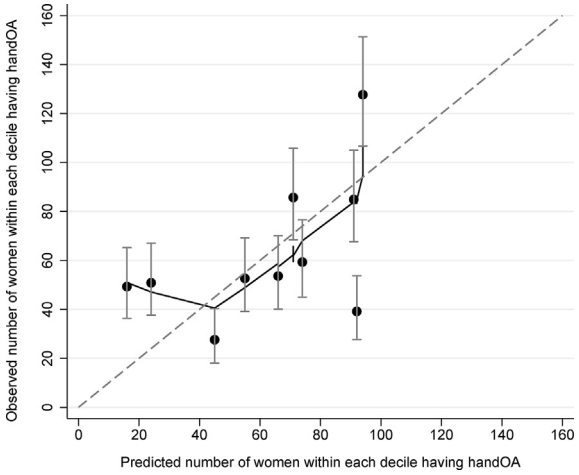

Fig. 4 Calibration plot for the final prediction model for incident hand OA for women. Black dots represent the agreement between observed and predicted values of hand OA with 95% confidence intervals. The 45° line represents perfect predictions. The solid black line is a smoothed calibration line depicting the average agreement between the observed and predicted number of women with hand OA.

1. Final model for incident hand OA for women.

$$\text{Logit}\,(\text{hand OA}) = -6.20 + 0.05 * \text{age spline1} + (-0.06 * \text{age spline2}) + (0.13 * \text{age spline3}) + (-0.05 * \text{age spline4}) + 0.02 * \text{BMI} + 0.31 * \text{heavy work} + (-0.13 * \text{high physical activity}) + 0.54 * \text{poor health}.$$

2. Final model including the genetic risk score for women.

$$\text{Logit}\,(\text{hand OA}) = -6.33 + 0.05 * \text{age spline1} + (-0.06 * \text{age spline2}) + 0.13 * \text{age spline3} + (-0.05 * \text{age spline4}) + 0.02 * \text{BMI} + 0.31 * \text{heavy work} + (-0.13 * \text{high physical activity}) + 0.54 * \text{poor health} + 0.27 * \text{genetic risk score}.$$

3. Final model including reproductive and hormonal factors for women.

$$\text{Logit(hand OA)} = -5.99 + 0.05\text{* age spline1} + (-0.11\text{ * age spline2}) + 0.16\text{ * age spline3} + (-0.05\text{ * age spline4}) + 0.02\text{ * BMI} + 0.34\text{ * heavy work} + (-0.11\text{ * high physical activity}) + 0.51\text{ * poor health} + (-0.23\text{ * parity, } 1-2) + (-0.25\text{ * parity} \geq 3) + 0.18\text{ * age menarche 12 yrs} + 0.18\text{* age menarche 13 yrs} + 0.07\text{ * age menarche 14 yrs} + 0.24\text{* age menarche} \geq 15\text{ yrs} + 0.31\text{ * HRT former use} + 0.28\text{ * HRT curent use.}$$

Discussion

In a population-based sample of more than 35,000 persons, we found that an existing prediction model on hand OA for young men performed equally well in our sample of men as in the development sample. We also found that incident hand OA could be equally well predicted by using a novel prediction model for women. Including a genetic risk score or information on female reproductive and hormonal history (women only) did not improve prediction. Except for the existing prediction model for men[5], we are not aware of other studies predicting hand OA to compare of our findings.

The discriminative ability of the prediction models for hand OA in the current study was modest, but comparable to what has been previously reported for hip and knee OA[2,3,24]. To date, a good AUC (>0.75) in prediction models of OA has only been reached in studies that included structural changes visual on X-ray both as a predictor *and* as outcome[2,3]. However, this is expected and may be of small clinical use because two similar constructs are likely to predict each other perfectly. Kerkhof *et al.*[2] reported a discrimination (AUC) of 0.66 in a model developed for incident knee OA in a sample of 2,628 individuals from the Rotterdam Study (RS-I). Adding a genetic risk score did not improve discrimination, however, inclusion of baseline Kellgren & Lawrence (KL) score increased AUC to 0.79[2]. Prediction of incident hip OA was performed also in the RS-I cohort (AUC 0.67) with external validation in the RS-II (AUC 0.60) and the Cohort Hip and Cohort Knee Study (CHECK) (AUC 0.54)[3]. Again, only inclusion of imaging variables increased discrimination (AUC 0.78)[3]. Radiographic information was not available in our study. However, as previously acknowledged[2,3], these minor radiographic changes might be interpreted as an early state of OA and should be considered perhaps more as predictors of disease progression than of future risk of OA.

The population in the current study were slightly younger at baseline (mean age 51 years) than in the former prediction studies on OA[2,3]. Our study sample is perhaps most comparable to the RS-I cohort, except for the older age, comprising a large sample from the general population. While the external validation cohorts consisted either of a high-risk population with early signs of symptomatic hip or knee OA (CHECK Study)[25], or women only, derived from a general practitioner's register in the Chingford Study[26]. In contrast to the former studies[2,3], we developed separate models for men and women and defined incident hand OA from diagnostic codes, i.e., a clinically relevant end point as a proxy for more severe hand OA (only those referred to specialist). In comparison to the former OA prediction studies, we also had longer follow-up time. The first diagnosis of incident hand OA occurred after 1.6 years and the average time to incident hand OA was 11.7 years (SD 4.6). We acknowledge that also differences between study populations, including difference in presence and distributions of predictors, may affect risk prediction and explain the somewhat lower discriminative ability in our study than what were observed in studies of knee OA using radiographic changes as predictors and as definitions of incident disease.

In an attempt to explore predictors that can predict future hand OA equally well as smoking can for lung cancer[27], we included a genetic risk score. However, the genetic risk score itself showed poor prediction, both in the models for men and women (AUC 0.53), and did not add any discriminative ability to the models. Only a few genetic variants for hand OA have been identified so far[12,13]. We used strict criteria when selecting SNPs for the genetic risk score, both regarding the *P*-value threshold and in that we did not include SNPs in high linkage disequilibrium. Thus, the genetic risk score comprised three genetic variants only; one from the ALDH1A2 gene[13], the Matrix Gla Protein (MGP) gene[12], and one rare variant at the 1p31[13]. The poor prediction could be due to the few genetic markers, and that these explain only a small fraction of the disease[28]. Moreover, we added reproductive and hormonal factors to the model to investigate if these factors, despite being controversial risk factors for OA[29], could improve prediction of hand OA in women. However, introducing these factors only slightly affected the performance of the model in the current study.

The final prediction model for men and women demonstrated modest ability to discriminate between those with and without incident hand OA, based on the general suggestion that an AUC of 0.60–0.75 refers to poor to potentially helpful discrimination[30,31]. However, the final model for men showed good agreement between observed frequencies and predicted probabilities of hand OA, as depicted in the calibration plot. The final model for women showed lower ability to correctly detect individuals with hand OA (i.e., sensitivity) than the model for men. Furthermore, poorer agreement between observed and predicted values of hand OA for women, for those in the low and high risk groups, indicated that low risks were underestimated and high risk were overestimated. The poor calibration for women shows that prediction of hand OA is complex, perhaps due to the unclear and heterogeneous etiology of the disease[7]. Future prediction studies of hand OA might explore more sophisticated statistical methods, like penalized regression methods[32], in order to improve the calibration and subsequent validation.

Our study had several limitations. First, we had a relatively large percentage missing data for some of the potential predictors, in particular heavy physical work (missing 14% and 12% for men and women, respectively) and sleep problems (missing for 19% of the men). We did not performed multiple imputation of data since the aim of the study was to develop a prediction model. Imputed data would, similar to the prediction model itself, be based on predictions from a model developed from the same sample (risk of overfitting etc.). However, including only participants with complete data may lead to biased results if these participants are not representative of the whole study sample[16]. We observed some differences in baseline characteristics between those included vs those excluded due to missing data. Thus, some predictors might have been under- or overestimated in the included sample, which could bias the results. A second limitation is that we do not have available hand OA diagnoses from primary care, therefore we

cannot conclude on a prediction model for milder hand OA cases based on our data. However, we believe the inclusion of data from specialist care might have reduced the risk of misclassification, since rheumatologists and orthopaedic surgeons working in specialist care will have more experience in making a correct hand OA diagnoses than general practitioners who see a broad range of patients. Therefore, to better predict milder hand OA disease, we suggest that future prediction models use a combination of clinical data and register data.

Third, we did not externally validate our results. In particular for the model for women this would have been valuable since there are no other existing prediction models to compare with. Important strengths of our study were the prospective design and inclusion of a large population-based sample representative for the entire Norwegian population. We included less conventional predictors than in the previous prediction model for hand OA[5], more comparable to the existing prediction studies on hip and knee OA[2,3]. We also used updated statistical methods, e.g., to assess nonlinear effects, using restricted cubic splines. When nonlinearity is present, splines represents a better and more flexible way to model a continuous predictor, i.e., due to reducing the loss of information from categorizing and potentially reducing the risk of over- or underestimation of the model[16,33]. We acknowledge that splines may not be as easily interpretable as categorization. Therefore, age was also included as a categorical variable in the models (Supplementary material 7). However, this was performed only to enhance the interpretation and comparison to other studies. A model including splines could still be used to obtain individual predictions[16]. Thus, our studies may have direct clinical implications, by allowing for the calculation of the total individual risk, at least for men, since prediction models are more or less similar across two large population-based cohorts in Sweden and Norway. As an example, using the final incident model, a man aged 60–70 years who has all of the predictors and a BMI of 30, would have a 4.2 times higher risk of hand OA than the mean aged man with mean BMI and none of the predictors (predicted risks 5.1% vs 1.2%). Hence, as long as no disease-modifying treatments exist for (hand) OA, the information obtained through prediction models may be utilized to potentially reduce the individual risk of hand OA by intervening on modifiable risk factors like BMI and sleep problems. However, causality cannot be inferred directly from a prediction study. The effect of the predictors on an outcome, and the effect of intervening on them, can only be investigated in comparative, preferably randomized, studies[16]. For women, future studies should focus on the validation and refinement of the prediction models from the current study in order to improve the generalizability and clinical usefulness of the models.

In conclusion, we have developed the first prediction model for hand OA in women. The models' performance should be validated in other samples. For men, there is increasing evidence that BMI, education and information of sleep can predict incident hand OA in several populations with moderate discriminative ability.

Contributions

Conception and design: Johnsen, Magnusson.

Analysis and interpretation of data: Johnsen, Magnusson, Børte, Winsvold.

Collection and assembly of data: Børte, Gabrielsen, Winsvold, Skogholt, Thomas.

Obtaining funding: Storheim, Hveem, Zwart.

Drafting of the article and critical revision of the article for intellectual content: Johnsen, Magnusson, Børte, Gabrielsen, Winsvold, Skogholt, Thomas, Storheim, Hveem, Zwart.

Final approvement of the article: Johnsen, Magnusson, Børte, Gabrielsen, Winsvold, Skogholt, Thomas, Storheim, Hveem, Zwart.

Conflict of interest

We declare no conflict of interests.

Funding sources

Dr. Johnsen is currently funded by the Research Council of Norway (grant number 248817). The funding source had no inovelvement in any aspects of this manscript.

Acknowledgments

We would like to thank the participants of the Nord-Trøndelag Health Study (The HUNT Study) and the HUNT Research Centre for their cooperation. The HUNT Study is a collaboration between HUNT Research Centre (Faculty of Medicine and Health Sciences, NTNU, Norwegian University of Science and Technology), Trøndelag County Council, Central Norway Regional Health Authority, and the Norwegian Institute of Public Health.

Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.joca.2020.04.005.

References

1. Slatkowsky-Christensen B, Mowinckel P, Kvien TK. Health status and perception of pain: a comparative study between female patients with hand osteoarthritis and rheumatoid arthritis. Scand J Rheumatol 2009;38:342–8, https://doi.org/10.1080/03009740902913496.
2. Kerkhof HJ, Bierma-Zeinstra SM, Arden NK, Metrustry S, Castano-Betancourt M, Hart DJ, *et al*. Prediction model for knee osteoarthritis incidence, including clinical, genetic and biochemical risk factors. Ann Rheum Dis 2014;73:2116–21, https://doi.org/10.1136/annrheumdis-2013-203620.
3. Saberi Hosnijeh F, Kavousi M, Boer CG, Uitterlinden AG, Hofman A, Reijman M, *et al*. Development of a prediction model for future risk of radiographic hip osteoarthritis. Osteoarthritis Cartilage 2018;26:540–6, https://doi.org/10.1016/j.joca.2018.01.015.
4. Zhang W, McWilliams DF, Ingham SL, Doherty SA, Muthuri S, Muir KR, *et al*. Nottingham knee osteoarthritis risk prediction models. Ann Rheum Dis 2011;70:1599–604, https://doi.org/10.1136/ard.2011.149807.
5. Magnusson K, Turkiewicz A, Timpka S, Englund M. Prediction of midlife hand osteoarthritis in young men. Osteoarthritis Cartilage 2018;26:1027–32, https://doi.org/10.1016/j.joca.2018.05.010.
6. Srikanth VK, Fryer JL, Zhai G, Winzenberg TM, Hosmer D, Jones G. A meta-analysis of sex differences prevalence, incidence and severity of osteoarthritis. Osteoarthritis Cartilage 2005;13:769–81, https://doi.org/10.1016/j.joca.2005.04.014.
7. Haugen IK, Englund M, Aliabadi P, Niu J, Clancy M, Kvien TK, *et al*. Prevalence, incidence and progression of hand osteoarthritis in the general population: the Framingham Osteoarthritis Study. Ann Rheum Dis 2011;70:1581–6, https://doi.org/10.1136/ard.2011.150078.
8. Liu B, Balkwill A, Cooper C, Roddam A, Brown A, Beral V. Reproductive history, hormonal factors and the incidence of hip and knee replacement for osteoarthritis in middle-aged women. AnnRheumDis 2009;68:1165–70, https://doi.org/10.1136/ard.2008.095653. doi: ard.2008.095653 [pii].
9. Wise BL, Niu J, Zhang Y, Felson DT, Bradley LA, Segal N, *et al*. The association of parity with osteoarthritis and knee replacement in the multicenter osteoarthritis study.

Osteoarthritis Cartilage 2013;21:1849–54, https://doi.org/10.1016/j.joca.2013.08.025.
10. Hellevik AI, Nordsletten L, Johnsen MB, Fenstad AM, Furnes O, Storheim K, *et al*. Age of menarche is associated with knee joint replacement due to primary osteoarthritis (The HUNT Study and the Norwegian Arthroplasty Register). Osteoarthritis Cartilage 2017;25:1654–62, https://doi.org/10.1016/j.joca.2017.06.010.
11. Valdes AM, Spector TD. The contribution of genes to osteoarthritis. Med Clin 2009;93:45–66, https://doi.org/10.1016/j.mcna.2008.08.007. doi: S0025-7125(08)00124-7.
12. den Hollander W, Boer CG, Hart DJ, Yau MS, Ramos YFM, Metrustry S, *et al*. Genome-wide association and functional studies identify a role for matrix Gla protein in osteoarthritis of the hand. Ann Rheum Dis 2017;76:2046–53, https://doi.org/10.1136/annrheumdis-2017-211214.
13. Styrkarsdottir U, Thorleifsson G, Helgadottir HT, Bomer N, Metrustry S, Bierma-Zeinstra S, *et al*. Severe osteoarthritis of the hand associates with common variants within the ALDH1A2 gene and with rare variants at 1p31. Nat Genet 2014;46:498–502, https://doi.org/10.1038/ng.2957.
14. Holmen J, Midthjell K, Krüger Ø, Langhammer A, Holmen TL, Bratberg G, *et al*. The nord-trøndelag health study 1995-97 (HUNT 2): objectives, contents, methods and participation. Norsk Epidemiologi 2003;13:19–32.
15. Krokstad S, Langhammer A, Hveem K, Holmen TL, Midthjell K, Stene TR, *et al*. Cohort profile: the HUNT study, Norway. Int J Epidemiol 2013;42:968–77, https://doi.org/10.1093/ije/dys095.
16. Moons KG, Altman DG, Reitsma JB, Ioannidis JP, Macaskill P, Steyerberg EW, *et al*. Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. Ann Intern Med 2015;162:W1–W73, https://doi.org/10.7326/m14-0698.
17. Nielsen JB, Fritsche LG, Zhou W, Teslovich TM, Holmen OL, Gustafsson S, *et al*. Genome-wide study of atrial fibrillation identifies seven risk loci and highlights biological pathways and regulatory elements involved in cardiac development. Am J Hum Genet 2018;102:103–15, https://doi.org/10.1016/j.ajhg.2017.12.003.
18. Harrell JFE, Bickel P, Diggle P, Fienberg SE, Gather U, Olkin I, *et al*. Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. 2nd edn. Cham, Cham: Springer International Publishing; 2015. Edition, 2015.
19. Steyerberg EW, Pencina MJ, Lingsma HF, Kattan MW, Vickers AJ, Van Calster B. Assessing the incremental value of diagnostic and prognostic markers: a review and illustration. Eur J Clin Invest 2012;42:216–28, https://doi.org/10.1111/j.1365-2362.2011.02562.x.
20. Moorman PG, Myers ER, Schildkraut JM, Iversen ES, Wang F, Warren N. Effect of hysterectomy with ovarian preservation on ovarian function. Obstet Gynecol 2011;118:1271–9, https://doi.org/10.1097/AOG.0b013e318236fd12.
21. Shuster LT, Rhodes DJ, Gostout BS, Grossardt BR, Rocca WA. Premature menopause or early menopause: long-term health consequences. Maturitas 2010;65:161–6, https://doi.org/10.1016/j.maturitas.2009.08.003.
22. Steyerberg EW, Harrell Jr FE, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: efficiency of some procedures for logistic regression analysis. J Clin Epidemiol 2001;54:774–81.
23. Steyerberg EW. Clinical Prediction Models : A Practical Approach to Development, Validation, and Updating. New York: Springer; 2009.
24. Magnusson K, Turkiewicz A, Timpka S, Englund M. A prediction model for the 40-year risk of knee osteoarthritis in adolescent men. Arthritis Care Res 2019;71:558–62, https://doi.org/10.1002/acr.23685.
25. Wesseling J, Boers M, Viergever MA, Hilberdink WK, Lafeber FP, Dekker J, *et al*. Cohort profile: cohort hip and cohort knee (CHECK) study. Int J Epidemiol 2016;45:36–44, https://doi.org/10.1093/ije/dyu177.
26. Hart DJ, Spector TD. Cigarette smoking and risk of osteoarthritis in women in the general population: the Chingford study. Ann Rheum Dis 1993;52:93–6, https://doi.org/10.1136/ard.52.2.93.
27. Tammemagi CM, Pinsky PF, Caporaso NE, Kvale PA, Hocking WG, Church TR, *et al*. Lung cancer risk prediction: prostate, lung, colorectal and ovarian cancer screening trial models and validation. J Natl Cancer Inst 2011;103:1058–68, https://doi.org/10.1093/jnci/djr173.
28. Zengini E, Finan C, Wilkinson JM. The genetic epidemiological landscape of hip and knee osteoarthritis: where are we now and where are we going? J Rheumatol 2016;43:260–6, https://doi.org/10.3899/jrheum.150710.
29. de Klerk BM, Schiphof D, Groeneveld FP, Koes BW, van Osch GJ, van Meurs JB, *et al*. No clear association between female hormonal aspects and osteoarthritis of the hand, hip and knee: a systematic review. Rheumatology 2009;48:1160–5, https://doi.org/10.1093/rheumatology/kep194.
30. Alba AC, Agoritsas T, Walsh M, Hanna S, Iorio A, Devereaux PJ, *et al*. Discrimination and calibration of clinical prediction models: users' guides to the medical literature. Jama 2017;318:1377–84, https://doi.org/10.1001/jama.2017.12126.
31. Hosmer J, David W, Lemenshow S, Sturdivant RX. Assessing the fit of the model, Applied Logistic Regression. Hoboken, NJ, USA: John Wiley & Sons, Inc.; 2013153–225.
32. Pavlou M, Ambler G, Seaman S, De Iorio M, Omar RZ. Review and evaluation of penalised regression methods for risk prediction in low-dimensional data with few events. Stat Med 2016;35:1159–77, https://doi.org/10.1002/sim.6782.
33. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. Stat Med 2006;25:127–41, https://doi.org/10.1002/sim.2331.