Development of a prediction model for future risk of radiographic hip osteoarthritis

Fatemeh Saberi Hosnijeh, Maryam Kavousi, Cindy G. Boer, Andre G. Uitterlinden, Albert Hofman, Max Reijman, Edwin H.G. Oei, Sita M. Bierma-Zeinstra*, Joyce B.J. van Meurs*

Fatemeh Saberi Hosnijeh: Department of Internal Medicine and Department of Immunology, Erasmus University Medical Center, Rotterdam, Netherlands, E-mail: f.saberihosnijeh@erasmusmc.nl

Maryam Kavousi, Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands, E-mail: m.kavousi@erasmusmc.nl

Cindy G. Boer, Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, Netherlands, E-mail: c.boer@erasmusmc.nl

Andre G. Uitterlinden, Department of Internal Medicine and Department of Epidemiology, Erasmus University Medical Center, Rotterdam, Netherlands, E-mail: a.g.uitterlinden@erasmusmc.nl

Albert Hofman, Department of Epidemiology, Erasmus University Medical Center, Rotterdam, The Netherlands, E-mail: a.hofman@erasmusmc.nl

Max Reijman, Department of Orthopedics, Erasmus University Medical Center, Rotterdam, Netherlands, E-mail: m.reijman@erasmusmc.nl

Edwin H.G. Oei, Department of Radiology, Erasmus University Medical Center, Rotterdam, The Netherlands

Sita M. Bierma-Zeinstra, Department of General Practice, Erasmus University Medical Center, Rotterdam, Netherlands, E-mail: s.bierma-zeinstra@erasmusmc.nl

Joyce B.J. van Meurs: Department of Internal Medicine, Erasmus University Medical Center, Rotterdam, Netherlands, E-mail: j.vanmeurs@erasmusmc.nl

* Co-last authors

corresponding author:

Fatemeh Saberi Hosnijeh, MD PhD

Department of Immunology, Erasmus University Medical Center

P.O. Box 2040, 3000 CA Rotterdam, The Netherlands

E-mail: f.saberihosnijeh@erasmusmc.nl

Tel: +31 10 703 84 21

Running head: Risk assessment model for hip osteoarthritis

Word counts: Abstract (249), main text (4000)

ABSTRACT

Objective: To develop and validate a prognostic model for incident radiologic hip osteoarthritis (HOA) and determine the value of previously identified predictive factors.

Design: We first validated previously reported predictive factors for HOA by performing univariate and multivariate analyses for all predictors in 3 large prospective cohorts (total sample size of 4548 with 653 incident cases). The prognostic model was developed in 2327 individuals followed for 10 years from the Rotterdam Study-I (RS-I) cohort. External validation of the model was tested on discrimination in two other cohorts: RS-II (n=1435) and the Cohort Hip and Cohort Knee (CHECK) study (n=786).

Results: From the total number of 28 previously reported predictive factors, we were able to replicate 13 factors, while 15 factors were not significantly predictive in a meta-analysis of the 3 cohorts. The basic model including the demographic, questionnaire, and clinical examination variables (area under the curve (AUC)=0.67) or genetic markers (AUC=0.55) or urinary C-terminal cross-linked telopeptide of type II collagen (*u*CTX-II) levels (AUC=0.67) alone were poor predictors of HOA in all cohorts. Imaging factors showed the highest predictive value for the development of HOA (AUC=0.74). Addition of imaging variables to the basic model led to substantial improvement in the discriminative ability of the model (AUC=0.78) compared with *u*CTX-II (AUC=0.74) or genetic markers (AUC=0.68). Applying external validation, similar results were observed in the RS-II and the CHECK cohort.

Conclusions: The developed prediction model included demographic, a limited number of questionnaire, and imaging risk factors seems promising for prediction of HOA.

Keywords: osteoarthritis, prediction, Hip, risk

INTRODUCTION

The number of people affected with osteoarthritis (OA) is likely to increase due to the high prevalence of obesity and ageing of the population whilst, to date, there are no curative treatment options which allow regeneration of damaged cartilage (1,2). The current focus in OA research and clinical practice is on persons with radiographic symptomatic disease (3). Despite extensive researches, modern therapies are largely palliative and only modestly effective. There is presently no disease-modifying osteoarthritis drug with a consistent, documented effect despite several clinical attempts in late-stage phases. This reinforces the need for interventions in early OA or even before development of symptoms to develop preventive strategies (3). Therefore, clinicians have the challenging task to identify those individuals that will develop OA as early as possible. Identification of high-risk individuals will be beneficial for efficient screening of novel therapeutic options, since follow-up time and number of individuals included in the studies can be reduced when only subjects are included which are at a high risk of developing OA.

Conventional risk prediction models, although well established in other disease areas, such as cardiovascular disease, have not been developed in OA and in particular not for hip OA. Two recent studies reported on prediction models for incidence and progression of knee OA (4,5). In the first study (4), three different models were evaluated for incident and progressive radiographic knee OA and symptomatic knee OA among a small high-risk cohort of 99 cases and 179 controls with 12 years of follow-up using conventional risk factors such as age, gender, body mass index (BMI), family history of OA, occupational risk, and joint injury. However, external validation of these models in the Osteoarthritis Initiative showed poor discrimination. In another study among Rotterdam Study (RS) participants (5), different risk prediction models were developed using basic risk factors (i.e., age, gender and BMI) and were compared with less conventional risk factors such as radiographic features, genetic risk score, and biochemical marker of urinary C-terminal cross-linked telopeptide of type II collagen (*u*CTX-II) levels. They showed that all risk factor groups by themselves had limited and rather similar predictive value, while the full model had useable predictive value and showed good external validation.

In the current study, we focus on prediction of incident hip OA. A recent review highlighted the fact that hip OA requires specific attention separate from other OA phenotypes, since its etiology is different from knee OA (6). Surprisingly very little prospective data on hip OA is available, since the majority of risk factor assessment has been done on cross-sectional data. In addition, no study has tried to make a prediction model for hip OA using multiple risk factors, something that was recently highlighted to be essential for future monitoring and managing of hip OA patients (6). In the current manuscript, we therefore aimed to determine the value of various sets of risk factors, which were selected based on previous literature. These included anthropometric /demographic characteristics, routine questionnaire and clinical examination parameters, imaging risk markers, biochemical marker of uCTX-II, and genetic markers in hip OA risk prediction. We attempted to validate the previously suggested risk factors for hip OA in 3 large prospective cohorts, and then used that information to develop and validate a prognostic model for incident hip OA. We used the first cohort of the large prospective population-based Rotterdam Study (RS-I) to develop the risk prediction model. The model was externally validated in the second cohort of the Rotterdam Study (RS-II), which is an independent cohort with similar population characteristics, and the Cohort Hip and Cohort Knee (CHECK) study, which is a study of people who for the first time consult the general practitioner for their joint complaints. Our study is the first to compare the value of different risk factor groups including the clinical, imaging, genetic, and biochemical markers in prediction of hip OA and provides the first risk prediction model for incident hip OA.

METHODS

Study populations

The Rotterdam Study is a large prospective population-based cohort study of men and women aged 55 years and older in the municipality of Rotterdam, the Netherlands. The study design and rationale are described elsewhere in detail (7). The first cohort (RS-I) was initiated in 1989 and included 7983 individuals. The second cohort (RS-II) was initiated in 2000 and included 3011 inhabitants who reached the age of 55 years after the baseline examination and persons aged 55 years or older who migrated into the research area. Written informed consent was

obtained from all participants and the study was approved by the Medical Ethics Committee of the Erasmus Medical Center (7). Baseline measurements were obtained through a home interview and visits to the research center for physical examinations and imaging and laboratory assessments. The present study includes the cohort's participants for whom hip radiographs at baseline and 10 years follow-up were present and scored. Patients with ankylosing spondylitis, rheumatoid arthritis (RA), and subjects with a total hip replacement (THR) due to fracture were excluded from the study. After these exclusions, 2327 participants from RS-I were used to develop the risk prediction model and 1435 participants from RS-II were used for external validation of the model (Figure 1).

The CHECK study included 1002 participants aged 45-65 years living in The Netherlands, with early symptomatic OA characterized by pain of knee and/or hip, entered the cohort in the period October 2002 to September 2005. They were included at or within 6 months of their first visit to the general practitioner for these symptoms (8). Participants with a pathological condition that could explain the hip symptoms were excluded including intra-articular fractures, RA, congenital dysplasia, Perthes disease, subluxation, osteochondritis dissecans, septic arthritis, previous hip joint replacement, previous hip surgery, and individuals having only symptoms of bursitis or tendinitis. This left 786 participants with 8 years follow-up from the CHECK study for external validation of the risk prediction model (Figure 1).

Outcome assessment

Weight-bearing antero-posterior radiograph of the pelvis was obtained at baseline and follow-ups and scored for the presence of a THR and OA according to the K&L score (9). Radiographic hip OA was defined as a K&L score≥2 of one or both joints or a THR. The incidence of hip OA was defined as a K&L<2 at baseline and THR or K&L≥2 at follow-up. All subjects were free of hip OA in both side at baseline.

Risk factors

Previously suggested hip OA risk factors or predictors that were available in RS were included in the study. These include age, obesity (BMI and waist to hip ratio (WHR)), gender, height, family

history of OA, occupation, smoking, education, alcohol intake, diabetes, high blood pressure, joint pain (hip, knee, low back), hip joint morning stiffness, lower limb disability index, and baseline measurement of total cholesterol, high-density lipoprotein cholesterol level (HDL), and C reactive protein (CRP), hip baseline K&L score (0 or 1), hand OA divided as finger OA and thumb OA, subtle acetabular dysplasia, cam morphology, femoral neck bone marrow density at baseline, urinary CTX-II, and a genetic risk score (Supplementary Table 1). The Supplementary Material describes the methods for measuring all of the risk predictors.

Statistical analysis

Age and sex adjusted generalized estimating equation (GEE) model (a logit model) was used to evaluate the association between the predictive factors and hip OA incidence within each cohort , followed by a meta-analyses of the results.

Development of the Risk Prediction Model

GEE model was used to fit the models for correlations between the right and left extremity in each individual using RS-I cohort. Using a backward elimination, we made a basic model that included the demographic, questionnaire, and clinical and routine laboratory examination variables. The variable that was least significant was removed and the model was refitted. Each subsequent step removed the least significant variable in the model until all remaining variables had individual P-values smaller than 0.10. To determine the added prediction value of the predictors which generally are not used in the routine clinical examination, the basic model was extended by the addition of the imaging risk factors, *u*CTX-II, and genetic risk score, separately and combined. We, therefore, constructed 5 models; (1) the basic model, (2) the basic model + imaging variables, (3) the basic model + *u*CTX-II levels, (4) the basic model+ genetic markers, and (5) the basic model + imaging variables + *u*CTX-II levels. Assessment of the predictive discrimination of the various models was made using ROC curve by plotting the sensitivity against the corresponding false-positive rate. The AUC was used as a measure of how well a group of variables predicted the development of hip OA. AUC was corrected for optimism by 100 bootstrap repetitions (Supplementary Material). The models were developed for 10-year

hip OA risk prediction. As some risk factors might be more predictive for the Hip OA occurring in a shorter time period, we further repeated all models using 5 years follow-up.

External Validation

Calibration and discrimination abilities of the models were externally examined in two cohorts: (1) RS-II which is an independent cohort with similar population characteristics; (2) CHECK cohort which is a high risk population (details in Supplementary Material). All analyses were performed using SPSS (IBM SPSS Statistics 21) and R, version 3.3.3.

RESULTS

The baseline characteristics of all three studies are shown in Table 1. In total, 2327 individuals from RS-I with both radiographic examinations at baseline and follow-up were included in the model development. The RS-I participants had mean age of 64 years (SD=5.8), 56.3% women, mean BMI of 26.2 (SD=3.5), and mean follow-up of 10.8 years (SD=1.1). Incident hip OA was seen in 7% of the hips. Participants of the RS-II and CHECK cohorts were younger than the RS-I cohort. The proportion of diabetics and smokers were higher in RS-II. The CHECK cohort included the highest percentage of female participants and OA incidence, while RS-I subjects had the highest frequency of K&L score of one at baseline. Hip pain, cam morphology and dysplasia were more frequent in CHECK participants. Around half of the RS-I and RS-II subjects dropped out, because no follow up radiograph was available (Figure 1). These subjects were older, less educated, more frequently smoking and disabled, and had higher level of CRP and more hip pain compared to the subjects with follow-up radiograph. Compared to the total cohort of RS-I and RS-II, included study subjects were younger, less diabetic, more educated, and had lower CRP levels and less hip pain (data not shown). The genetic risk scores were normally distributed (Supplementary Figure 1) in all cohorts and was higher in CHECK cohort compared with RS-I and RS-II.

Since many of the previously identified risk factors have not been validated to a large extent in longitudinal prospective cohorts, we first set out to determine the relationship between these factors and incident hip OA; these results are shown in Supplementary Table 3-6. We found a

significant negative association between physically demanding work and hip OA incidence which could be due to a health-based selection bias (Supplementary Table 6). Therefore this variable was excluded from our analyses. We confirmed a number of previously reported risk factors (age, height, WHR, hip pain, LLD, diabetes, CRP and *u*CTX-II level, BMD, presence of thumb OA, hip dysplasia and cam morphology, and doubtful mild degenerative changes in hip joint (K&L=1)), but did not observe a significant association with incident hip OA for the following reported factors: female gender, BMI, alcohol intake, education, smoking status, pain in knee and low back, hip stiffness, presence of HBP, hypercholesterolemia and finger OA, positive family history of osteoarthritis, and genetic markers (Supplementary Table 6).

For comparison of the discriminative ability of different sets of risk factors for hip OA risk prediction, we used the AUC. A total of five different groups of risk factors were created; (1) age, sex, and BMI, (2) all questionnaire, clinical examination variables, and routine lab test, (3) *u*CTX-II level, (4) imaging variables, (5) genetic risk score. The results are shown in Supplementary Table 7. The first three groups showed a poor prediction ability for incident hip OA (AUC range 0.60-0.67) while imaging variables showed the highest AUC (0.74). The genetic risk score had no discriminative ability for prediction of 10-year incidence of hip OA.

We then constructed the prediction models. We developed 5 prediction models. The 'basic model' included demographic, information gathered by questionnaire, and routine lab tests: gender, age, BMI, hip pain, education, smoking status, diabetes, family history of OA, and CRP level (Table 2). The AUC of the basic model was 0.67 for 10-year hip OA risk prediction. Subsequently imaging features were added to the basic model. The imaging variables that remained significant in the model included cam morphology, subtle hip dysplasia, presence of thumb OA, and baseline hip K&L grade of one. Compared with *u*CTX-II and genetic markers, addition of the imaging markers to the base model showed the highest increase in model discrimination (AUC=0.78) (Table 2). Adding the biochemical markers to the model increases the predictive value of the model to an AUC of 0.82. However, as the sample size for the last model (model 5) was limited (82 cases and 1073 controls, Table 2), the results should be interpreted with caution.

We checked the interaction of continuous variables with sex and age. None of them was significant except a borderline association between sex and *u*CTX-II. Therefore, no interaction term was included in the models.

The AUC of the different models in external validation cohorts are depicted in Table 3. The basic model showed a moderate discriminative ability in RS-II cohort (AUC=0.60, 95% CI=0.56-0.64) while in CHECK cohort it showed no discriminative ability (AUC=0.54, 95% CI=0.50-0.58). In both RS-II and CHECK cohorts, addition of imaging variables to the basic model provided the highest discrimination (AUC=0.75 (0.72-0.79) and 0.71 (0.66-0.75) in RS-II and CHECK cohorts respectively).

Calibration plots of the basic model and the basic model together with *u*CTX-II in RS-II showed reasonable agreement between predicted and actual probabilities while the models including imaging variables underestimated the actual probabilities (Figure 2). In the CHECK cohort, all models showed a poor calibration. As the CHECK cohort includes a high risk population, this systematic underestimation of the risk by the models is expected.

Analyses for 5 years follow-up showed essentially similar results as 8-10 years follow-up for all three cohorts (Supplementary Table 8 and supplementary Figure 2) except that the imaging markers showed to be better predictors for 5 years follow-up.

DISCUSSION

Osteoarthritis of the hip affects 7%-25% of Caucasian people over the age of 55 years (11). In addition to the related pain and discomfort, hip OA has substantial economic consequences, and with the current aging of the population in western societies this problem will increase (11). An individual risk assessment tool for future OA is needed to develop preventive strategies. could be applied. Furthermore, clinical decision making of general practitioners will be affected. In addition, a good prediction model could be used in selection of participants for trials. We here present the first study to evaluate the predictive values of different risk factor groups and to develop a risk prediction model for radiographic hip OA based on conventional risk factors, imaging features, biochemical -, and genetic markers. The models were developed

in a population-based cohort of elderly people followed for 10 years. The basic model including gender, age, BMI, hip pain, education, smoking status, diabetes, family history of OA, and CRP level showed limited discriminative value in all cohorts. Among the different groups of predictors, imaging factors showed the best predictive value for the development of OA. Addition of these variables to the basic model led to substantial improvement in the discriminative ability of the model.

The model with best predictive performance included demographic variables (age, sex, BMI) and a limited set of questionnaire/routine lab test variables (diabetes, smoking status, educational level, hip pain, positive family history of OA, and CRP level) and imaging risk factors, and seems promising for prediction of hip OA. Among the imaging features, doubtful mild degenerative changes (defined as a K&L score of one) are important predictors of future incident hip OA. As the predictive measure is similar to the outcome measure, a strong correlation between this predictive variable and the outcome is expected. In a study using incident knee OA data from RS, K&L score of one at baseline also was a strong predictor for knee OA development (5). Although radiologists do not always report a K&L score of one, this finding may be used by the practitioner and the patient for making decisions regarding preventative measures, enrolment in clinical trials for early OA, and monitoring of disease progression. Indeed, these minor radiographic changes represent an early stage of OA in which some structural damage has already occurred and therefore, the probability of progression to a definite OA is higher compared with people without any signs of OA on a radiograph. By excluding the baseline K&L score variable from the model, AUC reduced to 0.71. Moreover, our results indicated that individuals who have cam-type morphology are at high risk of developing osteoarthritis. A cam-type morphology might be a modifiable risk factor that can be diagnosed before severe hip damage is present, providing an opportunity to prevent hip osteoarthritis (12).

While our developed risk prediction models provided a reasonable calibration in the second cohort of the Rotterdam Study, the calibrations were poor for the CHECK cohort which is a clinical cohort of people coming to the general practitioners for the first time with joint

complaints. Several differences between populations may affect the results of the validation. We found that models including imaging variables showed poor calibration. To explore the effect estimates of imaging variables, we applied the same GEE model as discovery in the validation cohorts. The results suggest that these variables are stronger predictors for participants of the validation cohorts than the RS-I population (the discovery). We observed higher estimates for K&L baseline and cam morphology in the RS-II and CHECK cohorts compared with RS-I when adjusted for all other variables in the model. Moreover, distribution of several variables differed between the development and validation cohorts. We previously showed that cam morphology and hip dysplasia increased the risk of hip OA in subjects younger than 65 years (13). Therefore, the obtained estimates in RS-I could have resulted in an underestimation of the risks in RS-II and CHECK cohorts, which consist of younger participants. The CHECK cohort includes a high risk population who at recruitment had pain or stiffness of knee or hip joints. So it was expected that a prediction model developed in a general elderly population might underestimate the risk in this population.

Heritability of hip OA has been estimated to be around 40-60% (14,15). However, to date only few genetic variants have been successfully identified (14,16,17). Current SNPs used in this study are among common variants which explain only a small part of the disease variation. Our study did not show large predictive ability for this set of markers for incident of hip OA, which might be explained by the limited genetic variants available up to now. In addition, a number of genetic variants included in the overall genetic risk score, have previously been found to have effects in only a part of the population (for example only in women or men). In a study using incident knee OA data from Rotterdam Study, a moderate predictive value of a set of genetic markers chosen form literature was seen (AUC=0.62) (5).

Besides the prediction model, we present prospective data on a total of 28 previously suggested risk factors for hip OA. Previous reports on risk modeling has focused exclusively on knee OA, while hip OA is also a major form of OA disabling many individuals around the world. A recent review specifically highlighted the need to examine risk factors for hip OA separately from knee OA (7). We here presented data on a total of 4548 individuals from 3 prospective

cohorts on hip OA, and therefore present valuable information for definite validation of these factors. We found similar risk estimates for age within each cohort suggesting that aging is the robust identified risk factor for the development of hip OA. Our data did not show the relationship between sex and hip OA, which was consistent with the results of a recent prospective study including more than 3 million people from Spain (18), and was inconsistent with the results of a meta-analysis of two old studies showing an increased incidence of hip OA in females (19). The different results may relate to different inclusion/exclusion criteria and definition of hip OA.

We could not confirm the association between BMI and hip OA incidence, while WHR that shows abdominal obesity was associated with increased risk of hip OA. The latter finding is in contrast with the results of two studies (a case-control study and a population based prospective cohort study) reported no association between WHR and hip OA arthroplasty (20,21). The mechanism of the association between obesity and hip OA traditionally was thought to be purely biomechanical. However, recent advances in adipose biology have suggested that the adipocytokine, leptin has effects on chondrocytes (20). After adjustment for BMI, the risk estimate for WHR although non-significantly, was still high (OR=2.42, 95% CI=0.51-11.4, p=0.26), even higher than BMI estimate (OR=1.04, 95% CI=1.00-1.07, p=0.05).

Consistent with our finding, there is evidence that persons with hip or knee OA are more likely to have greater BMD than age-matched persons without OA at these sites (2). The contribution of high bone density to the pathogenesis of OA might be explained by the fact that OA is a process characterized by increased subchondral thickness and bony sclerosis. In contrast, progressive OA is often characterized by bony attrition. In the Framingham cohort, risk for progressive knee OA decreased from 34.4% to 22.0%, 20.3%, and 18.9% as BMD increased (22). Thus, it appears plausible that the bone effects may differ at different stages of the disease (2).

Strengths of the current study include its large sample size, comparison of multiple markers that were all measured with standardized methods, and use of hard end points to avoid misclassification bias. This study has several limitations as well. There are a few potentially important risk factors for hip OA for which data were not available in our study including hip

injury, physical activity, and physical examination of the hips (RS cohort). The studies among CHECK participants (23,24) have shown that painful internal rotation of the hip is a good predictor of 6 years THR. Moreover, our validation cohorts lacked some of the model variables such as family history of OA (RS-II and CHECK), and hand OA (CHECK) which could result in underestimation of the predicted probabilities. Second, our study subjects were 55 years or older at baseline. At this age, some people already have osteoarthritis and were excluded from the study. In addition, we used a subset of RS participants who were able to visit the research center at baseline and follow-up. These subjects were younger, more educated and mobile (less disability) and survived in the follow-up period. These factors could have led to the relatively healthier and younger study population and therefore this may limit the model's generalizability. Moreover, the effect of some risk predictors might be underestimated in our prediction models because of the excluded individuals. As the same condition was present for included subjects of RS-II, this may partly explain the modest calibration of the models in RS-II (lower predicted probabilities).

In conclusion, a basic model including the demographic, questionnaire, and clinical examination variables or model containing a genetic markers or *u*CTX-II levels alone were not good predictors of incident radiographic hip OA. In contrast, a model including the basic model with imaging features reached a fair predictive value and might be applicable in clinical practice when validated in other studies. The reported models could be seen as pilots to lead further research in this area. The models may be applied at the individual level to predict the risk, and to encourage risk reduction.

ACKNOWLEDGMENTS

The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists. This study was funded by the Dutch Arthritis foundation. We would like to thank all participants of the CHECK cohort and all collaborators of the different sites for their efforts. CHECK is funded by the Dutch Arthritis Association on the lead of a steering committee comprising 16 members with expertise in different fields of OA.

Author contributions: FSH analyzed the data and wrote the first draft of the manuscript. All authors contributed to the study design, data interpretation, and were involved in revising the article critically for important intellectual content and approved the final version to be published.

Funding: The Rotterdam Study is supported by the Netherlands Organization of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012), the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) (project nr. 050-060-810) and the Erasmus Medical Center and Erasmus University, Rotterdam. This study is funded by the Dutch Arthritis foundation (project nr 13-1-201).

Competing interest statement: Dr. Bierma-Zeinstra reports grants from Dutch Arthritis Foundation, during the conduct of the study; grants from Dutch Arthritis Foundation, the Netherlands Organisation for Health Research and Development, EU Horizon 2020, Stichting Coolsingel, Nuts-Ohra, and EU Fp7, other from Regeneron, Infirst healthcare, personal fees from Osteoarthritis & Cartilage, personal fees from OARSI, EULAR, Regeneron, and Infirst health care, outside the submitted work. The other authors declare no competing financial interests.

REFERENCES

1. Felson DT, Lawrence RC, Dieppe PA, Hirsch R, Helmick CG, Jordan JM, et al. Osteoarthritis: new insights. Part 1: the disease and its risk factors. Annals of Internal Medicine 2000;133:635-46.

2. Issa SN, Sharma L. Epidemiology of osteoarthritis: an update. Current Rheumatology reports 2006;8:7-15.

3. Dam EB, Loog M, Christiansen C, Byrjalsen I, Folkesson J, Nielsen M, et al. Identification of progressors in osteoarthritis by combining biochemical and MRI-based markers. Arthritis Research & Therapy 2009;11:1-11.

4. Zhang W, McWilliams DF, Ingham SL, Doherty SA, Muthuri S, Muir KR, et al. Nottingham knee osteoarthritis risk prediction models. Annals of the Rheumatic Diseases 2011;70:1599-604.

5. Kerkhof HJM, Bierma-Zeinstra SMA, Arden NK, Metrustry S, Castano-Betancourt M, Hart DJ, et al. Prediction model for knee osteoarthritis incidence, including clinical, genetic and biochemical risk factors. Annals of the Rheumatic Diseases 2014;73:2116-21.

6. Murphy NJ, Eyles JP, Hunter DJ. Hip Osteoarthritis: Etiopathogenesis and Implications for Management. Advances in Therapy 2016;33:1921-46.

7. Hofman A, Brusselle GO, Murad S, van Duijn C, Franco O, Goedegebure A, et al. The Rotterdam Study: 2016 objectives and design update. European Journal of Epidemiology 2015;30:661-708.

8. Wesseling J, Boers M, Viergever MA, Hilberdink WKHA, Lafeber FPJG, Dekker J, et al. Cohort Profile: Cohort Hip and Cohort Knee (CHECK) study. International Journal of Epidemiology 2016;45:36-44.

9. Kellgren JH, Lawrence JS. Radiological Assessment of Osteo-Arthrosis. Annals of the Rheumatic Diseases 1957;16:494-502.

10. Steyerberg EW, Harrell FE Jr, Borsboom GJ, Eijkemans MJ, Vergouwe Y, Habbema JD. Internal validation of predictive models: Efficiency of some procedures for logistic regression analysis. Journal of Clinical Epidemiology 2001; 54: 774–781.

11. Lievense AM, Bierma-Zeinstra SMA, Verhagen AP, Verhaar JAN, Koes BW. Prognostic factors of progress of hip osteoarthritis: A systematic review. Arthritis Care & Research 2002;47:556-62.
12. Agricola R, Heijboer MP, Bierma-Zeinstra SMA, Verhaar JAN, Weinans H, Waarsing JH. Cam impingement causes osteoarthritis of the hip: a nationwide prospective cohort study (CHECK). Annals of the Rheumatic Diseases 2013;72:918-23.
13. Saberi Hosnijeh F, Zuiderwijk ME, Versteeg M, Smeele HTW, Hofman A, Uitterlinden AG, et al. Cam deformity and acetabular dysplasia as risk factors for hip osteoarthritis. Arthritis & Rheumatology 2017;69:86-93.
14. Castaño-Betancourt MC, Evans Daniel S., Ramos Y.F.M., Boer C.G., Metrustry S., Y. L, et al. Novel genetic variants for Cartilage Thickness and Hip Osteoarthritis. PLOS Genetics. 2016;12:e1006260.
15. Spector TD, Cicuttini F, Baker J, Loughlin J, Hart D. Genetic influences on osteoarthritis in women: a twin study. Bmj. 1996;312(7036):940-3.
16. Castaño Betancourt MC, Cailotto F, Kerkhof HJ, Cornelis FMF, Doherty SA, Hart DJ, et al. Genome-wide association and functional studies identify the DOT1L gene to be involved in cartilage thickness and hip osteoarthritis. Proceedings of the National Academy of Sciences 2012;109:8218-23.
17. arc OC. Identification of new susceptibility loci for osteoarthritis (arcOGEN): a genome-wide association study. The Lancet 2012;380(9844):815-23.
18. Prieto-Alhambra D, Judge A, Javaid MK, Cooper C, Diez-Perez A, Arden NK. Incidence and risk factors for clinically diagnosed knee, hip and hand osteoarthritis: influences of age, gender and osteoarthritis affecting other joints. Annals of the Rheumatic Diseases 2014;73:1659-64.
19. 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.
20. Holliday KL, McWilliams DF, Maciewicz RA, Muir KR, Zhang W, Doherty M. Lifetime body mass index, other anthropometric measures of obesity and risk of knee or hip osteoarthritis in the GOAL case-control study. Osteoarthritis and Cartilage 2011;19:37-43.

21. Lohmander LS, de Verdier MG, Rollof J, Nilsson PM, Engström G. Incidence of severe knee and hip osteoarthritis in relation to different measures of body mass: a population-based prospective cohort study. Annals of the Rheumatic Diseases 2009;68:490-6.
22. Zhang Y, Hannan MT, Chaisson CE, et al. Bone mineral density and risk of incident and progressive radiographic knee osteoarthritis in women: the Framingham Study. The Journal of Rheumatology 2000;27:1032-7.
23. Bastick AN, Verkleij SPJ, Damen J, Wesseling J, Hilberdink WKHA, Bindels PJE, et al. Defining hip pain trajectories in early symptomatic hip osteoarthritis; 5 year results from a nationwide prospective cohort study (CHECK). Osteoarthritis and Cartilage 2016;24:768-75.
24 Lievense AM, Koes BW, Verhaar JAN, Bohnen AM, Bierma-Zeinstra SMA. Prognosis of hip pain in general practice: A prospective followup study. Arthritis Care & Research 2007;57:1368-74.

Figure 1. Flowchart of participants of RS-I (10 years follow-up), RS-II (10 years follow-up), and CHECK cohorts (8 years follow-up) included in the study; rheumatoid arthritis (RA); ankylosing spondylitis (AS); osteoarthritis (OA); total hip replacement (THR)

Figure 2. Calibration plots for different models in RS-II and CHECK cohorts; mean predicted probability of outcome and mean observed outcome for the basic model, basic + imaging variables, basic + CTX-II, and basic + imaging variables + CTX-II were 0.07/0.08, 0.05/0.08, 0.07/0.08, 0.04/0.08 for RS-II and 0.06/0.19, 0.04/0.19, 0.06/0.19, 0.04/0.19 for CHECK cohort respectively.

Table 1. Descriptive statistics of the included cohorts

	Model development	Model Validation	
	RS-I	**RS-II**	**CHECK**
Sample size, N	2327	1435	786
Age*	64.0 (5.80)	61.8 (5.08)	55.8 (5.21)
Women, %	56.3	55.2	78.6
BMI, kg/m^2 *	26.2 (3.48)	27.2 (3.78)	26.2 (4.02)
Hip pain, %	7	13.9	39.4
Diabetes, %	6.1	9	2.9
Smoking, %			
Never	31.9	67.3	31.3
Former	46.7	10.5	55
Current	21.4	22.2	13.7
Education, %			
Low	43.9	49.3	19.7
Mediate	44.1	31	52.2
High	12	19.7	28.1
Family history of OA, %	21.8	NA	NA
CRP, mg/l*	2.60 (4.88)	2.17 (3.67)	3.13 (7.58)
Cam morphology†, %	6.5	5.6	8.8
Acetabular dysplasia†, %	4	3.8	10.9
Thumb OA, %	28	17.2	NA
uCTX-II, ng/mmol*	2.24 (0.21)	2.28 (0.23)	2.26 (0.27)
K&L score 0 at baseline, N (%)†	2943 (63.2)	2520 (87.8)	1198 (76.2)
K&L score 1 at baseline, N (%)†	1711 (36.8)	350 (12.2)	374 (23.8)
Hips with OA incidence, N (%)†	321 (6.9)	222 (7.7)	292 (18.6)
Hips without OA incidence, N (%)†	4254 (91.4)	2628 (91.6)	1271 (80.9)
Follow-up time*	10.8 (1.13)	10.3 (0.98)	7.7 (1.16)

Body mass index (BMI); osteoarthritis (OA); Kellgren and Lawrence (K&L); C reactive protein (CRP);* mean (SD); † numbers are based on hips; †Acetabular dysplasia and cam morphology were defined as the presence of a center-edge angle <20° and an alpha angle of >60°, respectively.

Table 2. Multivariate models including different risk factors groups in prediction of incident hip OA in RS-I

	Basic model (B)		**(B)+ Imaging variables**		**(B)+ uCTX-II levels**		**(B)+ genetic markers**		**(B)+ Imaging + uCTX-II**	
	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P	OR (95% CI)	P
Age	1.08 (1.05-1.10)	<0.0001	1.06 (1.03-1.08)	<0.0001	1.09 (1.05-1.14)	<0.0001	1.08 (1.06-1.11)	<0.0001	1.07 (1.02-1.12)	0.01
women	1.96 (1.44-2.69)	<0.0001	2.41 (1.70-3.41)	<0.0001	2.12 (1.20-3.73)	0.01	1.90 (1.36-2.62)	0.0001	3.37 (1.73-6.59)	0.0004
BMI	1.04 (1.01-1.08)	0.02	1.06 (1.02-1.10)	0.01	0.99 (0.93-1.05)	0.71	1.04 (1.00-1.08)	0.04	1.02 (0.95-1.10)	0.52
CRP	1.22 (1.00-1.50)	0.05	1.27 (1.01-1.58)	0.04	1.58 (1.12-2.22)	0.01	1.30 (1.06-1.61)	0.01	1.48 (0.96-2.30)	0.08
Smoking		0.07		0.18		0.23		0.11		0.52
Former	1.04 (0.76-1.43)		0.98 (0.69-1.39)		1.34 (0.75-2.39)		1.12 (0.81-1.58)		1.39 (0.70-2.75)	
Current	1.52 (1.04-2.23)		1.40 (0.92-2.14)		1.86 (0.91-3.81)		1.56 (1.04-2.35)		1.58 (0.66-3.76)	
OA in family	1.32 (0.97-1.78)	0.08	1.27 (0.90-1.80)	0.18	1.58 (0.85-2.92)	0.15	1.36 (0.99-1.86)	0.06	1.82 (0.86-3.85)	0.12
Diabetes	1.61 (0.98-2.64)	0.06	0.93 (0.49-1.77)	0.84	2.02 (0.93-4.40)	0.08	1.53 (0.90-2.60)	0.11	0.60 (016-2.25)	0.45
Hip pain	1.52 (1.03-2.23)	0.04	1.48 (0.97-2.25)	0.07	1.83 (0.89-3.78)	0.10	1.65 (1.11-2.47)	0.01	1.47 (0.63-3.43)	0.37
Education		0.06		0.25		0.47		0.27		0.66
Mediate	1.30 (0.98-1.73)		1.08 (0.79-1.49)		1.07 (0.65-1.76)		1.21 (0.89-1.63)		0.88 (0.47-1.59)	
High	1.61 (1.03-2.50)		1.51 (0.93-2.46)		0.60 (0.24-1.49)		1.42 (0.88-2.31)		0.62 (0.20-1.94)	
Thumb OA			1.78 (1.30-2.44)	0.0003					2.61 (1.44-4.75)	0.002
Baseline KL score			4.43 (3.41-5.77)	<0.0001					3.96 (2.40-6.55)	<0.0001
Hip dysplasia			2.36 (1.44-3.84)	0.001					2.97 (1.06-8.33)	0.04
Cam morphology			1.84 (1.18-2.87)	0.01					1.96 (0.80-4.79)	0.14
*u*CTX-II levels					9.83 (2.76-35.0)	0.0004			5.54 (1.08-28.4)	0.04
Genetic risk score							1.11 (1.04-1.18)	0.002		
AUC of the model	**0.67 (0.64-0.70)**		**0.78 (0.75-0.81)**		**0.74 (0.69-0.79)**		**0.68 (0.64-0.71)**		**0.82 (0.77-0.86)**	
OA/no-OA	**321/4254**		**258/3574**		**113/1212**		**293/3904**		**82/1073**	

Urinary C-terminal cross-linked telopeptide of type II collagen (*u*CTX-II); osteoarthritis (OA); area under the receiver-operating characteristic curve (AUC); body mass index (BMI); C reactive protein (CRP); Distribution of most variables was the same for subjects with and without uCTX-II data and the drop out did not affect the results. All models were repeated using the same sample size (69 cases and 996 controls) in which the data were available for all variables and the conclusions remained same. Model performance was classified according to AUC scores: 0.50-0.60 = fail, 0.60-0.70 = poor, 0.70-0.80 = fair, 0.80-0.90 = good, and 0.90-1.0 = excellent.

Table 3. Discrimination of the risk prediction models in external validation cohorts

Model	Discrimination: AUC (95% CI)			
	OA/no-OA	RS-II 10 y Follow-up	OA/no-OA	CHECK 8 y follow-up
Basic model	222/2628	0.60 (0.56-0.64)	292/1271	0.54 (0.50-0.58)
Basic model + imaging variables	191/2086	0.75 (0.72-0.79)	230/1003	0.71 (0.66-0.75)
Basic model + uCTX-II levels	146/1705	0.65 (0.60-0.70)	292/1271	0.61 (0.57-0.65)
Basic model + genetic markers	189/2037	0.62 (0.58-0.66)	271/1143	0.57 (0.53-0.61)
Basic model + imaging variables + uCTX-II	135/1566	0.75 (0.71-0.80)	230/1003	0.69 (0.65-0.73)

AUC, area under the curve; in order to get more robust estimates, the largest sample size for each model was used.