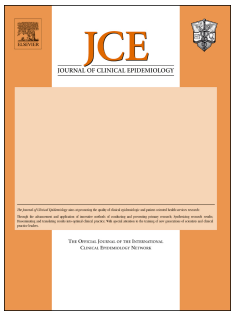
Accepted Manuscript



Weighted Cumulative Exposure (WCE) models helped identify an association between early knee pain consultations and future knee OA diagnosis

Dahai Yu, George Peat, John Bedson, John J. Edwards, Aleksandra Turkiewicz, Kelvin P. Jordan

PII: S0895-4356(16)30003-8

DOI: 10.1016/j.jclinepi.2016.02.025

Reference: JCE 9119

To appear in: *Journal of Clinical Epidemiology*

Received Date: 15 May 2015

Revised Date: 23 February 2016

Accepted Date: 26 February 2016

Please cite this article as: Yu D, Peat G, Bedson J, Edwards JJ, Turkiewicz A, Jordan KP, Weighted Cumulative Exposure (WCE) models helped identify an association between early knee pain consultations and future knee OA diagnosis, *Journal of Clinical Epidemiology* (2016), doi: 10.1016/j.jclinepi.2016.02.025.

TITLE PAGE

Weighted Cumulative Exposure (WCE) models helped identify an association between early knee pain consultations and future knee OA diagnosis

Dahai Yu[1*], George Peat[1], John Bedson[1], John J Edwards[1], Aleksandra Turkiewicz[2], Kelvin P Jordan[1]

[1] Research Institute for Primary Care & Health Sciences, Keele University, Staffordshire, England, ST5 5BG

[2] Clinical Epidemiology Unit, Orthopaedics, Department of Clinical Sciences Lund, Lund University, Lund, Sweden

*Corresponding author: Dahai Yu, Research Associate, Research Institute for Primary Care & Health Sciences, Keele University, Keele, Staffordshire, England, ST5 5BG. Tel: 00 44 1782 734891; Fax: 00 44 1782 734719. Email: d.yu@keele.ac.uk

Short title: *Weighted cumulative exposure models for early diagnosis*

Words in main text: 4,029
Words in abstract: 200
Tables: 3
Figures: 3
References: 39
Online supplementary data files: 3

ABSTRACT

Objective

To establish the association between prior knee pain consultations and early diagnosis of knee osteoarthritis (OA) by weighted cumulative exposure (WCE) models.

Study Design and Setting

Data were from an electronic healthcare record (EHR) database (Consultations in Primary Care Archive [CiPCA]). WCE functions for modelling the cumulative effect of time-varying knee pain consultations weighted by recency were derived as a predictive tool in a population-based case-control sample and validated in a prospective cohort sample.

Two WCE functions ((i) weighting of the importance of past consultations determined *a priori*; (ii) flexible spline-based estimation) were comprehensively compared with two simpler models ((iii) time since most recent consultation; total number of past consultations) on model goodness of fit, discrimination, and calibration both in derivation and validation phases.

Results

People with the most recent and most frequent knee pain consultations were more likely to have high WCE scores that were associated with increased risk of knee OA diagnosis both in derivation and validation phases. Better model goodness of fit, discrimination, and calibration were observed for flexible spline-based WCE models.

Conclusion

WCE functions can be used to model pre-diagnostic symptoms within routine EHR data and provide novel low-cost predictive tools contributing to early diagnosis.

Keywords

Weighted cumulative exposure; flexible modelling; risk prediction; knee osteoarthritis; knee pain

What is new

Key findings

- The application of weighted cumulative exposure (WCE) functions to patterns of pre-diagnostic symptoms / consultations (codes) in the routine electronic healthcare record may yield novel, low-cost prediction tools for early diagnosis.
- A flexible approach to weighting using regression cubic splines appeared to perform better than a weighting function assigning highest weights to the most recent consultations that was defined *a priori*.

What this adds to what is known

- WCE approaches have been used extensively for modelling time-varying causal exposures but their application to time-varying patterns of consultations recorded in the routine EHR is novel.
- Knee osteoarthritis diagnosis in primary care is typically preceded by multiple symptom-coded consultations over a period of months and sometimes years, offering the potential for earlier diagnosis.

What is the implication, what should change now

- The pattern of primary healthcare contacts and presented symptoms recorded in the routine EHR may be usefully modelled for early diagnosis using WCE methods.
- Flexible approaches to modelling these patterns should be investigated further in the context of early diagnosis of diseases other than OA.

INTRODUCTION

Weighted cumulative exposure (WCE) models have been developed in aetiologic research to provide an appropriate 'exposure metric' [1] to represent the complex cumulative effects of duration, intensity, and timing of time-varying exposures on health outcomes. These models, originating from concepts of time-weighted cumulative exposures introduced by Breslow et al [2] and Thomas [3], and extended notably by Vacek [4] and Abrahamowicz et al [5], have already found a wide range of applications including modelling the effects of lifestyle behaviours [6], environmental hazards [7], and prescription drugs [8, 9]. WCE approach estimates the total effect of past exposures by using a weighted sum of these exposures, with the weights dependent on time since exposure. A review of WCE models and their development can be found in Sylvestre and Abrahamowicz 2009[10].

In this study we consider a novel application of WCE models intended to serve the purpose of improving early diagnosis and identifying 'pre-diagnostic' cases for recruitment into research studies (e.g. early diagnosis, trials of early intervention). This application uses primary care electronic health record data and arose from the observation that for some conditions (e.g. breast cancer[11] and asthma[12]), disease diagnosis in primary care occurs after a sequence of one or more consultations for symptoms. These consultations may be unevenly interspersed over a variable period of time ranging from a matter of days to several years. In this study we aimed to assess the feasibility of using WCE models within primary care electronic health data to determine the association between presenting symptoms and diagnosis in primary care settings. We take as our example, knee osteoarthritis (OA). The majority of cases present to primary care with pain and the diagnosis is made on clinical grounds without routine use of imaging [13]. However, the diagnosis often occurs relatively late in the disease process [14], is strongly determined by (older) age of the patient [13], and is often preceded by consultations assigned non-specific symptom codes (e.g. 'knee pain' or 'knee arthralgia') [15]. It was our hypothesis that these non-specific consultations - discrete, time-varying binary events - could be modelled using WCE approaches and that the resultant weighted cumulative scores would be superior to simple counts of knee-related consultations within a particular time-window in predicting future knee OA diagnosis.

Since the exact nature of the exposure-outcome relationship is seldom confidently known in advance, a range of approaches to choosing an appropriate weighting function have been investigated, informed by knowledge of underlying pathophysiologic mechanisms (including pharmacokinetic properties of drugs when these are the exposures of interest). Specific weight functions may be specified in advance [16] or several may be compared in head-to-

head comparisons using model fit statistics [4]. However, modelling methods that require selecting the parametric form of the weight function in the absence of any prior knowledge about its shape might lead to invalid results if the function is incorrectly specified [17, 18].

Alternatively, the functional form of the weight function can be estimated from the data, using flexible nonparametric or quasi-parametric methods [10]. In our study we contrasted two WCE approaches: a weight function assigning highest weights to the most recent consultations that was defined *a priori*, and a weight function estimated by cubic regression splines.

METHODS

Data source and study design

We conducted a study of a primary care population included in the Consultations in Primary Care Archive (CiPCA), using all recorded consultation data by GPs and practice nurses from 11 general practices in North Staffordshire, England who contributed data continuously between 2000 and 2010. The total practice population consisted of 94,565 people in 2010 [19].

Ethics approval for CiPCA was given by the North Staffordshire Local Research Ethics Committee to download, store and analyse anonymised medical record information for research use from participating general practices (REC Reference: 03/04). Patients are informed by a poster at their practice and by leaflet that the practice is a Keele research practice and that their anonymised records (with identifiable information removed) may be used for research, and that they can opt out if they wish by informing the practice staff. Therefore no separate ethical approval was required for our study.

Our study was conducted in two phases. In the first phase, we conducted a population-based case-control study (outcome = incident knee OA diagnosis) in which we derived weighted cumulative exposure (WCE) scores using: (1) a weight function assigning highest weights to the most recent consultations defined *a priori* [5]; (2) a weight function using restricted cubic regression splines fitted to the case-control data [10]. In conditional logistic regression models we then compared how well each of the two WCE scores discriminated and validated between cases and controls. In the second phase, we conducted a prospective validation study, following controls forward in time to evaluate how well each of the two WCE scores predicted future knee OA diagnosis in these individuals. In both derivation and validation phases, we compared the models with the two weight functions against simpler, unweighted models which defined the exposure by categorizing, respectively, either (1) time since the most recent consultation, or (2) total number of past consultations. The models were compared in terms of model goodness of fit, discrimination and calibration.

Phase 1: Derivation and comparative discriminative ability of weighted cumulative exposure scores

Incident diagnosis of knee OA

The outcome of interest was incident diagnosis of knee OA. A consultation of OA was defined as a Read code starting N05 (‘Osteoarthritis and allied disorders’; equivalent to ICD9 codes beginning 715). As the main outcome, knee OA was defined by Read code lists drawn up through a consensus process involving local GPs (codelists are available on request from the authors [20] and also through the website: www.keele.ac.uk/mrr). Instead of marking the onset of disease, date of first Read-coded OA diagnosis is a clinical milestone denoting the expressed need for primary healthcare for a painful or disabling joint problem that is attributed by the doctor to a chronic, incurable condition for the first time [21, 22].

We used the maximum available run-in period within the CiPCA database (10 years) to identify new diagnoses of knee OA. An incident knee OA case was defined as one with a relevant knee OA code recorded in 2010 with no prior recorded OA code and complete registration in the previous 10 years. Hence, prevalent knee OA cases who were diagnosed with OA during the 10-year run-in period from January 1 2000 until December 31 2009 were not eligible to be defined as incident cases. We considered the date of the first diagnosis of knee OA in 2010 to be the index date for cases. We confirmed in our previous work [23] that a 10-year run-in period is likely to be sufficient to define an incident case of osteoarthritis within CiPCA in all adult age groups.

Selection of controls

Risk-set sampling [24] was used to select 10 practice-, age-, and gender-matched controls per case from patients continuously registered in CiPCA between 1 January 2000 and 31 December 2010. Controls were assigned an index date identical to that of their matched case. Eligible controls had no knee OA diagnosis by the age of the matched cases.

Exposure definition

Knee pain consultations recorded in general practice between January 1 2000 and the index date were defined as exposures based on Read Codes identified through consensus of local GPs [25, 26]. Six GPs, all with an interest in musculoskeletal research, were independently asked to decide whether or not individual codes relating to knee joint pain may relate to clinical OA. Codes were excluded if the majority of GPs decided they would

be unlikely to relate to OA. The final list of Read codes for knee pain included consultations recorded as ‘knee pain’ ,’knee joint pain’, and ‘arthralgia of knee’[26]. The codelist for knee pain is available through our website: http://www.keele.ac.uk/mrr/.

Derivation of WCE scores

Weight function assigning highest weights to the most recent consultations defined a priori

The cumulative exposure was estimated by a recency-weighted cumulative function, with the weight function defined *a priori* as proposed by *Abrahamowicz* [5]. This takes into account exposure duration (time interval between knee pain consultation and knee OA diagnosis) and cumulative dose (total number of knee pain consultations within a given time period). The *a priori* weight function was specified (i.e., the function that determines how the weights change with increasing time since the exposure) and used to assign appropriate numerical weights to past exposure. *Abrahamowicz* [5] proposed the positive (decreasing) half of the Normal `Gaussian’ density function that assigns highest weights to most recent exposures, and in which the cumulative exposure is calculated as a weighted sum of the past exposure. The rationale was that in common pharmacoepidemiological applications, the current risk may be affected by a recent increase of the exposure [5, 27]. The implementation of the WCE model with the weight function determined *a priori* is described in the Supplemental Technical Appendix.

We considered several different `half-Normal’ decreasing weight functions, assuming that the effect of past consultations may last between 1 and 10 years and compared their model goodness of fit. Based on the minimum Akaike information criterion (AIC) & Bayesian information criterion (BIC), we selected the function in which the weight assigned to a consultation that occurred 2 years before was reduced to one half of the weight assigned to a current consultation (**Supplemental Table 1**).

Weighted cumulative function estimated by restricted cubic regression spline

In contrast to the above approach which assumes a specific shape of the weight function, we also used weight functions fitted to the data using flexible restricted cubic regression splines [28]. Briefly, the weight function was built up by placing internal knots at equal time intervals, and modelled in conditional logistic regression models [29]. The number of knots and the length of time window prior to index date were identified from the model with optimal model-fit statistics. The number of knots determines the flexibility of the estimated spline function and the model's degrees of freedom. *Sylvestre and Abrahamowicz* [10] suggested that no more than five knots are enough to model a smooth spline model while reducing the risk of major over-fitting bias. In our analysis, model containing three to five knots over with the length of time window increasing at 0.5 year interval from1 to 10 years were tested and compared (**Supplemental Table 2**). The spline functions were generated and modelled in conditional logistic regression models. We then relied on the Bayesian Information Criteria (BIC) incorporating a penalizing term corresponding to the number of cases to select the optimal-fitting model [30].

The knots with their coefficients forming the optimal-fitting model were used to define the weight function [10, 29]. The linear combination of the estimates from the weight function was used to calculate a WCE score for each individual (see the Supplemental Technical Appendix). Z-Scores were calculated to allow better comparability between the two WCE function scores.

Comparison of discriminative ability of WCE scores

The discriminative abilities of the weight function determined *a priori* and the flexible spline-based weight function were compared in conditional logistic regression models fitted to the case-control data. In addition, we compared their discriminative abilities against two simpler unweighted

exposure models. The first model defined the exposure simply as categorised time since most recent knee pain consultation (none, 1-6 months, 7-12 months, 13-24 months, 25-36 months, > 36 months), whereas the second model categorized only the total number of previous knee consultations (0, 1, 2, 3, 4-5, ≥6). The results of the four models were compared with respect to discriminative ability, estimated by the C-statistic with bootstrap resampling (1000 samples) [31], and calibration evaluated for each decile of predicted probability of knee OA by plotting observed proportions versus predicted probability [31].

Phase 2: Prospective validation of WCE scores

To evaluate the predictive validity of scores from the two WCE functions, the same WCE scores derived from knee pain consultations up to the index date in Phase 1 (i.e. without including knee pain consultations after the index date) were used to predict the prospective risk of being diagnosed with knee OA from index date in 2010 to 31 December 2011.

The associations between unweighted recency of consultation, unweighted consultation counts and WCE scores derived from knee pain consultations up to the index date in Phase 1 and the outcome of diagnosis of OA between the index date and 31 December 2011 among controls were investigated using Cox proportional hazard models with adjustment for age, gender and practice. To allow easier interpretation of the model, the WCE scores in the Cox models were categorised into three groups for each of the two WCE functions: score equals zero (no consultation for knee pain), scores below the median of the non-zero scores, and scores above the median of the non-zero scores. The follow-up time was stratified into 3 levels: less than 6 months, 6-12 months and more than 12 months.

The proportional hazards assumption was tested using Schoenfeld residuals [32]. The four different exposure models outlined above were compared with respect to model discrimination quantified by C-statistics, goodness of fit measured by BIC, and calibration assessed by calculating the predicted risk and the observed risk at 2 years (maximally available follow-up period) and comparing these by decile of predicted risk [33].

RESULTS

Characteristics of study participants

Among the 203 newly diagnosed cases of knee OA in 2010 (mean age: 68.7 years, range 39.1-97.5 years; 52.7% female), 163 (80%) had at least 1 knee pain consultation in the previous 10 years, with 93 (46%) having 3 or more (**Table 1**). The most recent knee pain consultation occurred most commonly within 6 months prior to knee OA diagnosis. Among the 1964 controls, 709 (36%) had at least 1 knee pain consultation, 283 (14%) had 3 or more, and 107 (6%) had their most recent knee pain consultation within 6 months of index date.

Associations between the risk of diagnosed knee OA and simple categorization of time since most recent knee pain consultation and/or total number of previous knee pain consultations are presented in **Table 1**. The timing of most recent prior knee pain consultation was strongly associated with OA diagnosis, even after adjustment for total number of previous consultations. Irrespective of recency, having 1 or more prior knee pain consultations was strongly associated with OA diagnosis but there was no evidence of a further increase in the risk of OA diagnosis with increasing number of previous knee pain consultations.

Derivation of WCE scores

The Z-Scores by weight function among cases and controls are shown in **Table 2**. The spline function based on 5 knots equally spaced over 4.5 years before index date emerged as the best-fitting flexible spline-based function (**Supplemental Table 2**). The weights from both sets of WCE function are presented visually by time in **Supplemental Figure 1**. **Table 2** illustrates the higher Z-Scores from both methods for individuals with a more recent knee pain consultation. Higher scores in cases indicate that for the same number of consultations cases had, on average, more recent (i.e. higher-weighted) consultations, and for the same recency, cases had a higher total number of consultations.

Of the four exposure models, the best model fit statistics (AIC=800.80; BIC=806.11) and C-statistics (0.68 (95% CI: 0.65, 0.72)) were found for the model with the flexible spline-based function (**Supplemental Tables 3 & 4**).

Comparison of predictive performance of WCE scores

From the case-control analyses, both WCE scores showed a similar pattern of overall dose-response relationship with the outcome of incident knee OA diagnosis (**Figure 1**).

Figure 2 displays calibration plots for both WCE models (**Figure 2-a; Figure 2-b**) and both simple unweighted models (**Figure 2-g; Figure 2-h**). The agreement ratio was calculated as the predicted probability of OA diagnosis divided by the observed proportion (**Figure 2c** for WCE models; **Figure 2-i** for unweighted models). An agreement ratio closer to 1 represents better agreement between predicted probability and observed proportion receiving an OA diagnosis. Closer agreement between the predicted probability and observed proportion of knee OA diagnosis and flexible spline-based function was found across most levels of predicted risk. For example, the mean difference between 1 and the agreement ratio was 0.28 for the WCE model with

the flexible spline-based function, compared with 0.78, 1.02 and 1.89 for the WCE model with weight function determined *a priori*, simple unweighted model categorizing time since last knee pain consultation, and simple unweighted model categorizing total number of past knee pain consultations, respectively.

Prospective validation analysis in controls

The incidence rate of 30.8 per 1000 person-years was obtained from 72 newly diagnosed knee OA cases in 1,778 controls followed for median 1.32 years (IRQ: 1.11-1.64) (2337.9 person-years of observation). In the final Cox model of OA diagnosis, residual checks indicated no violation of the assumption of proportional hazards. A higher risk of diagnosed knee OA was more likely to be found among individuals with most recent consultations, individuals with more consultations, and individuals with higher scores on both WCE functions in comparison with individuals without any prior consultation (**Table 3**).The risk of diagnosed knee OA increased with the increase in weight function scores (both *a priori* determined function and flexible spline-based function) (**Figure 3**).

Model goodness of fit was better for the flexible spline-based function than for the unweighted models and the weight function determined *a priori* (**Supplemental Table 3**). Discriminative abilities for all models were higher than observed in the derivation phase (C statistics for weight function determined *a priori* 0.65 (95%CI: 0.59 to 0.72); flexible spline-based function 0.69 (0.63 to 0.75)) (**Supplemental Table 4**). **Figure 2** compares the predicted and observed risks of knee OA diagnosis at 2 years across each decile of predicted risk for the WCE models (**Figure 2-d; Figure 2-e**) and the unweighted models (**Figure 2-j; Figure2-k**). The agreement ratio between predicted risk and observed risk is presented in **Figure 2-f** for the WCE models and in **Figure 2-l** for the unweighted models. Closer agreement between the predicted probability and observed proportion of knee OA

diagnosis was found across most levels of predicted risk when using flexible spline-based function. The mean difference between 1 and agreement ratio was 0.42, 0.95, 1.64 and 2.92 for WCE model with flexible spline-based function, WCE model with the weight function determined a priori, simple unweighted model categorizing time since last knee pain consultation and simple unweighted model categorizing total number of past knee pain consultations, respectively.

DISCUSSION

Our study findings provide empirical support for the use of weighted cumulative exposure approaches to model pre-diagnostic symptom consultations in primary care. In derivation and validation studies of knee osteoarthritis diagnosis in general practice a flexible non-parametric function (weights estimated by regression cubic splines) was found to discriminate better than a recency-weighted function (weights based on a weight function determined *a priori* that assigned highest weights to the most recent consultations). Prospective validation revealed better model calibration from the model with a flexible spline-based function in predicting the 2-year risk of diagnosed knee OA by pattern of knee pain consultations in the prior 4.5 years. Both weight functions were superior to simpler unweighted models of recency and number of prior consultations in terms of model goodness of fit, discrimination and calibration, both in derivation and validation phases.

Scores from both the weight function determined *a priori* and the flexible spline-based function suggest that people with more recent knee pain consultations and more consultations are more likely to receive a diagnosis of knee OA in the future. High scores might nonetheless reflect different diagnostic processes and pathways for knee OA in primary care settings. At one extreme, ‘fast diagnosis’ may be achieved through a combination of patient-profile (e.g. classical risk factors, advanced signs and symptoms of osteoarthritis present), and clinician diagnostic preference (e.g. comfortable making the diagnosis of osteoarthritis on clinical grounds alone). The prognostic utility of WCE approaches will be limited in cases who already achieve a ‘fast diagnosis’. In addition these instances may be particularly susceptible to 'reverse causation' where knee pain consultations are erroneously believed to precede OA diagnosis due to misclassification of the date of OA diagnosis. This might happen, for example, if there were delays in administrative staff entering into the primary care record a diagnosis that was made in secondary care and reported in a letter to the practice. In the current application, where few diagnoses of OA are made in secondary care, this is unlikely to have been a major concern although in other settings it remains an important consideration. At the other extreme, ‘slow diagnosis’ may reflect the absence of one or more of these elements. In our data, other symptom records or referrals were not available but we would advocate such investigation of the underlying diagnostic process to accompany applications of weighted cumulative exposure models for clinical diagnosis.

An important feature of weight functions for early diagnosis is that they are a relatively inexpensive prediction tool derived from record coded symptoms or consultations available in primary care electronic health records. For descriptive prediction models, information that is easy to obtain would normally be considered before information that is more difficult to obtain. Use of readily available information from the routine electronic health record, such as the weight function reported in this study, may be incorporated in basic diagnostic models (e.g. with age, sex and other well-

known risk factors). The incremental value of more costly or difficult-to-obtain imaging or biochemical markers could be evaluated when added to such basic models [34].

A weight function, calculated at prognostic zero-time (e.g. the point in time of the patient presenting with eligible symptoms/symptom codes to primary care or alternatively a point in calendar time chosen by an investigator to identify patients at high/low risk of future diagnosis), and based on modelling patterns of consultations prior to prognostic zero-time, represents a novel prediction tool. This tool used alone, or more likely in combination with other recorded prognostic factors (e.g. patient age, gender, risk factors for the disease of interest), may have potential applications in clinical practice and in enriching patient recruitment to clinical research studies by identifying patients early who are at an increased risk of having undiagnosed disease and might benefit from further clinical investigations/interventions. The selection of relevant prior consultations, the optimal weighting approach, the relevant time interval before diagnosis, and the consequences of misclassification, however, are all likely to vary from one disease to the next and between healthcare systems.

There were several limitations in our study. Firstly the WCE functions in our study were only applied to knee pain consultations to predict the knee OA diagnosis. The performance of WCE functions to predict the early diagnosis of other diseases by symptoms / consultations recorded in primary care settings should be further investigated.

Secondly, some important covariables (e.g. body mass index, comorbidity) [35] were not adjusted for both in our derivation and calibration models. The incorporation of those variables might potentially improve the discrimination of our prediction models. However, the discriminative ability of our models was comparable or higher to that seen in other prognostic models for incident knee OA [36].

Thirdly, although prospective validation was undertaken in our study, we would encourage further external validation. A relatively small sample size was available for the current study. Under such circumstances it is well recognised that model uncertainty may be large and reliable predictions may not be derived [37]. Over-fitting is also a concern due to the limited sample size, especially in the case of more flexible models that require additional degrees-of-freedom [38]. Both the validation sample and the derivation sample in the current study were from the same database (CiPCA) rather than two independent databases. Moreover the validation sample comprised controls from the derivation sample, which made the relatedness between derivation and validation samples high [39]. Different models were used in the derivation phase (conditional logistic regression model) and the validation phase (Cox regression model), which made comparisons (model goodness of fit, discrimination and calibration) difficult between the derivation and validation phases.

WCE models [5, 10, 29] can be applied to patterns of pre-diagnostic symptoms and consultations within the routine primary care electronic health record. These relatively simple, low-cost predictors may have the potential to contribute to improving early diagnosis and warrant further investigation.

Acknowledgements

The Keele GP Research Partnership and the Informatics team at the Arthritis Research UK Primary Care Centre.

CiPCA database is funded by the North Staffordshire Primary Care Research Consortium and Keele University Research Institute for Primary Care and Health Sciences.

Conflicts of interest: None.

References

1. Abrahamowicz M, Beauchamp ME, Sylvestre MP. Comparison of alternative models for linking drug exposure with adverse effects. Stat Med 2012; 31(11-12):1014-30.

2. Breslow NE, Lubin JH, Marek P, Langholz B. Multiplicative models and cohort analysis. Journal of the American Statistical Society 1983; 78(381):1-12.

3. Thomas DC. Models for exposure-time-response relationships with applications to cancer epidemiology. Annu Rev Public Health 1988; 9:451-82.

4. Vacek PM. Assessing the effect of intensity when exposure varies over time. Stat Med 1997; 16(5):505-13.

5. Abrahamowicz M, Bartlett G, Tamblyn R, du Berger R. Modeling cumulative dose and exposure duration provided insights regarding the associations between benzodiazepines and injuries. J Clin Epidemiol 2006; 59(4):393-403.

6. Bjerkaas E, Parajuli R, Engeland A, Maskarinec G, Weiderpass E, Gram IT. The association between lifetime smoking exposure and breast cancer mortality - results from a Norwegian cohort. Cancer Med 2014; 3(5):1448-57.

7. Atkinson RW, Carey IM, Kent AJ, van Staa TP, Anderson HR, Cook DG. Long-term exposure to outdoor air pollution and the incidence of chronic obstructive pulmonary disease in a national English cohort. Occup Environ Med 2015; 72(1):42-8.

8. Alonso A, Bengtson LG, MacLehose RF, Lutsey PL, Chen LY, Lakshminarayan K. Intracranial hemorrhage mortality in atrial fibrillation patients treated with dabigatran or warfarin. Stroke 2014; 45(8):2286-91.

9. Dixon WG, Abrahamowicz M, Beauchamp ME, *et al*. Immediate and delayed impact of oral glucocorticoid therapy on risk of serious infection in older patients with rheumatoid arthritis: a nested case-control analysis. Ann Rheum Dis 2012; 71(7):1128-33.

10. Sylvestre MP, Abrahamowicz M. Flexible modeling of the cumulative effects of time-dependent exposures on the hazard. Stat Med 2009; 28(27):3437-53.

11. Shen N, Mayo NE, Scott SC, *et al*. Factors associated with pattern of care before surgery for breast cancer in Quebec between 1992 and 1997. Med Care 2003; 41(12):1353-66.

12. Kawasumi Y, Abrahamowicz M, Ernst P, Tamblyn R. Development and validation of a predictive algorithm to identify adult asthmatics from medical services and pharmacy claims databases. Health Serv Res 2011; 46(3):939-63.

13. Zhang W, Doherty M, Peat G, *et al*. EULAR evidence-based recommendations for the diagnosis of knee osteoarthritis. Ann Rheum Dis 2010; 69(3):483-9.

14. Bijlsma JW, Berenbaum F, Lafeber FP. Osteoarthritis: an update with relevance for clinical practice. Lancet 2011; 377(9783):2115-26.

15. Heintjes EM, Berger MY, Koes BW, Bierma-Zeinstra SM. Knee disorders in primary care: design and patient selection of the HONEUR knee cohort. BMC Musculoskelet Disord 2005; 6:45.

16. Berry G. Models for mesothelioma incidence following exposure to fibers in terms of timing and duration of exposure and the biopersistence of the fibers. Inhal Toxicol 1999; 11(2):111-30.

17. Hastie T, Tibshirani R. Generalized additive models for medical research. Stat Methods Med Res 1995; 4(3):187-96.

18. Seixas NS, Robins TG, Becker M. A novel approach to the characterization of cumulative exposure for the study of chronic occupational disease. Am J Epidemiol 1993; 137(4):463-71.

19. Jordan KP, Joud A, Bergknut C, *et al*. International comparisons of the consultation prevalence of musculoskeletal conditions using population-based healthcare data from England and Sweden. Ann Rheum Dis 2014; 73(1):212-8.

20. Jordan KP, Kadam UT, Hayward R, Porcheret M, Young C, Croft P. Annual consultation prevalence of regional musculoskeletal problems in primary care: an observational study. BMC Musculoskelet Disord 2010; 11:144,2474-11-144.

21. Bedson J, McCarney R, Croft P. Labelling chronic illness in primary care: a good or a bad thing? Br J Gen Pract. 2004;54(509):932-8.

22. Bedson J, Croft PR. The discordance between clinical and radiographic knee osteoarthritis: a systematic search and summary of the literature. BMC Musculoskelet Disord. 2008;9:116.

23. Yu D, Peat G, Bedson J, Jordan KP. Annual consultation incidence of osteoarthritis estimated from population-based health care data in England. Rheumatology (Oxford). 2015;54(11):2051-60.

24. Wacholder S, Silverman DT, McLaughlin JK, Mandel JS. Selection of controls in case-control studies. II. Types of controls. Am J Epidemiol 1992; 135(9):1029-41.

25. Jordan KP, Jöud A, Bergknut C, Croft P, Edwards JJ, Peat G, Petersson IF, Turkiewicz A, Wilkie R, Englund M. International comparisons of the consultation prevalence of musculoskeletal conditions using population-based healthcare data from England and Sweden. Ann Rheum Dis 2014; 73(1):212-8.

26. Jordan KP, Kadam UT, Hayward R, Porcheret M, Young C, Croft P. Annual consultation prevalence of regional musculoskeletal problems in primary care: an observational study. BMC Musculoskelet Disord 2010;11:144.

27. Bespalov A, Jongen-Relo AL, van Gaalen M, Harich S, Schoemaker H, Gross G. Habituation deficits induced by metabotropic glutamate receptors 2/3 receptor blockade in mice: reversal by antipsychotic drugs. J Pharmacol Exp Ther 2007; 320(2):944-50.

28. Harrell, Jr., Frank E. Regression Modeling Strategies: With Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis. Springer. 2015

29. Hauptmann M, Wellmann J, Lubin JH, Rosenberg PS, Kreienbrock L. Analysis of exposure-time-response relationships using a spline weight function. Biometrics 2000; 56(4):1105-8.

30. Schwarz G. Estimating the dimension of a model. Annals of Statistics 1978; 6(2): 461-4.

31. Steyerberg EW, Vickers AJ, Cook NR, *et al*. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010; 21(1):128-38.

32. Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. Stat Med 2002; 21(15):2175-97.

33. Collins GS, Altman DG. Identifying patients with undetected colorectal cancer: an independent validation of QCancer (Colorectal). Br J Cancer 2012; 107(2):260-5.

34. Moons KG, de Groot JA, Linnet K, Reitsma JB, Bossuyt PM. Quantifying the added value of a diagnostic test or marker. Clin Chem 2012; 58(10):1408-17.

35. Glyn-Jones S, Palmer AJ, Agricola R, Price AJ, Vincent TL, Weinans H, Carr AJ. Osteoarthritis. Lancet 2015. pii: S0140-6736(14)60802-3.

36. Zhang W, McWilliams DF, Ingham SL, Doherty SA, Muthuri S, Muir KR, Doherty M. Nottingham knee osteoarthritis risk prediction models. Ann Rheum Dis 2011;70(9):1599-604.

37. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating. Springer. 2009

38. Abrahamowicz M, MacKenzie T, Esdaile JM. Time-Dependent Hazard Ratio: Modeling and Hypothesis Testing With Application in Lupus Nephritis. Journal of the American Statistical Association 1996; 91(436):1432-39.

39. Debray TP, Vergouwe Y, Koffijberg H, Nieboer D, Steyerberg EW, Moons KG. A new framework to enhance the interpretation of external validation studies of clinical prediction models. J Clin Epidemiol 2015; 68(3):279-89.

Figure Legends

Figure-1. Association between weighted cumulative exposure (WCE) scores (*a priori* determined the weight function that assigned highest weights to most recent consultations & flexible spline-based function) of knee pain consultation and risk of diagnosed knee osteoarthritis

The incidence rate ratio was presented as thick short dash dot line with 95% confidence interval as thin dash lines.

Figure-2. Validation plots of prediction models

Grey square indicates predicted probability, black dot indicated observed proportion; black hollow triangle indicated agreement ratio from model-1 (Figure-2(c) and Figure-2(f)) or model-3 (Figure-2(i) and Figure-2(l)); grey hollow diamond indicates agreement ratio from model-2 (Figure-2(c) and Figure-2(f)) or model-4 (Figure-2(i) and Figure-2(l)).

Model-1, WCE model with spline-based function; Model-2, WCE model with the weight function determined a priori; Model-3, simple unweighted model categorizing time since last knee pain consultation; Model-4, simple unweighted model categorizing total number of past knee pain consultations.

Figure-3. Associations between weighted cumulative exposure (WCE) scores of knee-pain consultations and relative risk of diagnosed knee osteoarthritis in validation cohort, followed up from index dates

The incidence rate ratio was presented as thick short dash dot line with 95% confidence interval as thin dash lines.

Tables

Table 1. Recency and total number of previous knee-pain consultations among cases and controls

	Knee-OA cases (N=203)	Controls (N=1964)	Crude IRR	Adjusted† IRR
	n (%)	n (%)	**IRR (95%CI)**	**IRR (95%CI)**
Timing of most recent knee pain consultation				
No previous consultation	40 (20)	1255 (64)	1	1
1-6 months	99 (49)	107 (6)	31.2(19.7, 49.4)	24.83 (13.57, 45.43)
7-12 months	15 (7)	68 (4)	7.03(3.7, 13.60)	5.61 (2.61, 12.07)
13-24 months	16 (8)	67 (3)	8.25(4.26, 16.0)	6.67 (3.13, 14.23)
25-36 months	9 (4)	88 (5)	3.63(1.67, 7.92)	3.02 (1.30, 7.01)
≥37 months	24 (13)	379 (19)	2.10(1.23, 3.56)	1.77 (0.96, 3.26)
Total number of knee pain consultations				
0	40 (20)	1255 (64)	1	1
1	37 (18)	277 (14)	4.66(2.89, 7.54)	13.37 (8.09, 22.09)
2	33 (16)	149 (8)	7.32(4.44, 12.2)	14.46 (8.75, 23.90)
3	31 (15)	80 (4)	13.46(7.92, 23.1)	22.58 (13.14, 38.78)
4-5	26 (13)	103 (5)	8.65 (4.97, 15.09)	14.59 (8.61, 24.75)
≥6	36 (18)	100 (5)	12.82 (7.61, 21.59)	16.26 (9.89, 26.74)

IRR, incidence rates ratio; †IRR for timing of most recent knee pain consultation was adjusted for the total number of knee pain consultations; IRR for total number of knee pain consultations was adjusted for the timing of most recent knee pain consultation.

Table 2. Z-Scores of Weighted cumulative exposure (WCE) function assigning highest weights to the most recent consultations that was defined a priori and flexible spline-based function in cases and controls, overall and by recency and total number of knee pain consultations

	***A priori* determined function**		**Flexible spline-based function**	
	Knee-OA	**Controls**	**Knee-OA**	**Controls**
Overall	1.05 (0.03, 2.63)	0 (0, 0)	2.82 (0, 8.49)	0 (0, 0)
Timing of most recent knee pain consultation				
1-6 month	0.36 (-1.03, 0.67)	0.08 (-0.75, 0.84)	0.48 (-0.20, 0.60)	0.26 (-0.23, 0.52)
7-12 months	0.34 (-0.38, 0.69)	0.00 (-0.14, 0.16)	0.04 (-0.53, 0.41)	-0.09 (-0.69, 0.52)
13-24 months	-0.25 (-0.64, 0.43)	-0.36 (-0.84, 0.71)	-0.06 (-0.54, -0.06)	-0.11 (-0.78, 0.58)
25-36 months	-0.97 (-0.97, 1.01)	-0.66 (-0.97, 1.01)	-0.15 (-0.81, 0.31)	-0.25 (-0.57, 0.14)
≥37 months	0 (0, 0)	0 (0, 0)	-0.45 (-0.54, -0.17)	-0.36 (-0.39, -0.12)
No previous consultation	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)
Total number of previous knee pain consultations				
0	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)
1	0.66 (-0.88, 0.73)	-0.66 (-0.76, 0.77)	0.49 (-0.60, 0.59)	-0.07 (-0.84, 1.13)
2	0.56 (-0.03, 0.63)	-0.61 (-0.90, 1.22)	0.31 (-0.29, 0.71)	0.08 (-0.94, 1.09)
3	0.66 (-0.55, 0.68)	-0.45 (-1.00, 1.16)	0.51 (-0.42, 0.67)	0.20 (-1.13, 0.97)
4-5	0.61 (-1.06, 0.63)	-0.43 (1.00, 1.15)	0.46 (-0.31, 0.79)	-0.15 (-0.53, 0.93)
≥6	0.64 (-0.69, 0.66)	0.55 (-1.12, 0.98)	0.21 (-1.18, 0.97)	-0.29 (-1.04, 1.34)

Scores are presented as median (interquartile range) Z-score.

Table 3. Association between unweighted recency, unweighted counts, and both weight function scores of prior knee pain consultations and knee osteoarthritis diagnosis, by follow-up time period

	Follow-up period	Person-years	Diagnosed OA cases	OA diagnoses per 1000 person years	HR (95% CI)	
					Unadjusted	Adjusted‡
Simple unweighted model categorizing time since last knee pain consultation						
	0-6 months	689.9	11	15.9 (8.8, 28.8)	Reference	Reference
No previous consultation	**7-12 months**	632.7	13	20.5 (11.9, 35.4)	Reference	Reference
	>12 months	577.9	19	32.9 (19.8, 51.3)	Reference	Reference
Timing of most recent knee pain	**0-6 months**	21.0	6	285.7 (104.9, 621.9)	18.0 (11.9, 21.6)	18.8 (5.4, 65.5)
consultation by index date:	**7-12 months**	19.5	0	0.0	0.0	0.0
1-6 month	**>12 months**	14.3	2	139.9 (16.9, 505.2)	4.3 (0.9, 9.8)	4.5 (1.0, 19.5)
Timing of most recent knee pain	**0-6 months**	14.7	1	68.0 (1.7, 379.0)	4.3 (0.2, 13.2)	2.0 (0.2, 20.0)
consultation by index date:	**7-12 months**	12.7	1	78.7 (2.0, 438.7)	3.8 (0.2, 12.4)	2.8 (0.3, 23.9)
7-12 months	**>12 months**	10.0	0	0.0	0.0	0.0
Timing of most recent knee pain	**0-6 months**	13.4	1	74.6 (1.9, 415.8)	4.7 (0.2, 14.4)	3.5 (0.4, 28.7)
consultation by index date:	**7-12 months**	12.5	0	0.0	0.0	0.0
13-24 months	**>12 months**	11.2	1	89.3 (2.3, 497.5)	2.7 (0.1, 9.7)	2.6 (0.3, 19.6)
Timing of most recent knee pain	**0-6 months**	22.5	0	0.0	0.0	0.0
consultation by index date:	**7-12 months**	21.2	1	47.2 (1.2, 262.8)	2.3 (0.1, 7.4)	3.1 (0.4, 26.0)
25-36 months	**>12 months**	19.5	3	153.8 (31.7, 449.6)	4.7 (1.6, 8.8)	3.7 (1.1, 12.7)
Timing of most recent knee pain	**0-6 months**	92.3	1	10.8 (0.3, 60.4)	0.7 (0.0, 2.1)	0.4 (0.06, 3.5)
consultation by index date:	**7-12 months**	82.0	7	85.4 (34.3, 175.9)	4.2 (2.9, 5.0)	4.7 (1.8, 12.3)
≥37 months	**>12 months**	70.6	5	70.8 (23.0, 165.3)	2.2 (1.2, 3.2)	2.1 (0.8, 5.8)
Simple unweighted model categorizing total number of past knee pain consultations						
	0-6 months	689.9	11	15.9 (8.8, 28.8)	Reference	Reference
No previous consultation	**7-12 months**	632.7	13	20.5 (11.9, 35.4)	Reference	Reference
	>12 months	577.9	19	32.9 (19.8, 51.3)	Reference	Reference
Total number of knee pain	**0-6 months**	76.673	2	26.1 (3.2, 94.2)	1.6 (0.4, 3.3)	1.3 (0.3, 6.0)
consultations by index date:	**7-12 months**	69.124	2	69.1 (3.5, 104.5)	3.4 (0.3, 3.0)	2.2 (0.5, 10.2)
1	**>12 months**	56.687	6	56.7 (38.8, 230.4)	1.7 (2.0, 4.5)	1.5 (1.4, 1.7)

Total number of knee pain consultations by index date: 2	**0-6 months**	32.599	4	122.7 (33.4, 314.2)	7.7 (3.8, 10.9)	7.2 (4.8, 10.9)
	7-12 months	30.146	1	33.2 (0.8, 184.8)	1.6 (0.1, 5.2)	1.4 (0.5, 4.0)
	>12 months	27.404	1	36.5 (0.9, 203.3)	1.1 (0.05, 4.0)	1.1 (0.2, 8.6)
Total number of knee pain consultations by index date: 3	**0-6 months**	19.463	1	51.4 (1.3, 286.3)	3.2 (0.1, 9.9)	3.1 (1.5, 6.4)
	7-12 months	18.734	1	53.4 (1.4, 297.4)	2.6 (0.1, 8.4)	2.8 (0.9, 8.9)
	>12 months	20.213	0	0.0	0.0	0.0
Total number of knee pain consultations by index date: 4-5	**0-6 months**	19.805	1	50.5 (1.3, 281.3)	3.2 (0.1, 9.8)	2.0 (0.1, 32.0)
	7-12 months	15.873	4	252.0 (68.7, 645.2)	12.3 (5.8, 18.2)	9.0 (2.8, 29.6)
	>12 months	11.947	2	167.4 (20.3, 604.7)	5.1 (1.0, 11.8)	5.3 (1.2, 23.0)
Total number of knee pain consultations by index date: ≥6	**0-6 months**	15.445	1	64.7 (1.6, 360.7)	4.1 (0.2, 12.5)	1.6 (0.2, 13.8)
	7-12 months	14.091	1	71.0 (1.8, 395.4)	3.5 (0.2, 11.2)	3.1 (0.4, 25.9)
	>12 months	9.257	2	216.1 (26.2, 780.5)	6.6 (1.3, 15.2)	6.9 (2.0, 23.2)
Weighted cumulative exposure model with weight function assigning highest weights to the most recent consultations that was defined a priori						
***A priori* determined function score=0**	**0-6 months**	689.9	11	15.9 (8.8, 28.8)	Reference	Reference
	7-12 months	632.7	13	20.5 (11.9, 35.4)	Reference	Reference
	>12 months	577.9	19	32.9 (19.8, 51.3)	Reference	Reference
***A priori* determined function score 0-0.001**	**0-6 months**	94.7	1	10.6 (1.5, 74.9)	0.7 (0.1, 5.1)	0.5 (0.07, 3.9)
	7-12 months	86.2	0	0.0	0.0	0.0
	>12 months	78.9	2	25.4 (3.1, 91.6)	0.8 (0.2, 3.3)	0.8 (0.2, 3.2)
***A priori* determined function score >0.001**	**0-6 months**	68.8	11	159.8 (88.5, 288.6)	10.0 (4.3, 23.1)	5.7 (2.3, 14.2)
	7-12 months	61.6	7	113.7 (54.2, 238.5)	5.5 (2.2, 13.9)	4.4 (1.7, 11.5)
	>12 months	47.2	8	169.6 (73.2, 334.3)	5.2 (2.3, 11.8)	5.2 (2.3, 11.8)
Weighted cumulative exposure model with the flexible spline-based function						
Flexible spline-based function score=0	**0-6 months**	689.9	11	15.9 (8.8, 28.8)	Reference	Reference
	7-12 months	632.7	13	20.5 (11.9, 35.4)	Reference	Reference
	>12 months	577.9	19	32.9 (19.8, 51.3)	Reference	Reference
	0-6 months	139.5	3	21.5 (6.9, 66.7)	1.4 (0.4, 4.8)	1.1 (0.3, 4.0)

Flexible spline-based function score 0-2.6	**7-12 months**	128.8	0	0.0	0.0	0.0
	>12 months	114.8	1	8.7 (0.2, 48.5)	0.3 (0.04, 2.0)	0.3 (0.03, 1.9)
	0-6 months	24.1	9	374.0 (194.6, 718.7)	23.5 (9.7, 56.6)	7.6 (2.8, 21.1)
Flexible spline-based function score>2.6	**7-12 months**	18.9	7	370.4 (176.6, 776.9)	18.0 (7.2, 45.2)	5.8 (2.1, 16.0)
	>12 months	11.2	9	801.7 (366.3, 1520.7)	24.4 (11.0, 53.9)	27.9 (12.4, 62.5)

‡, adjusted for age gender, and practice. HR, hazard ratio. CI, confidence interval.