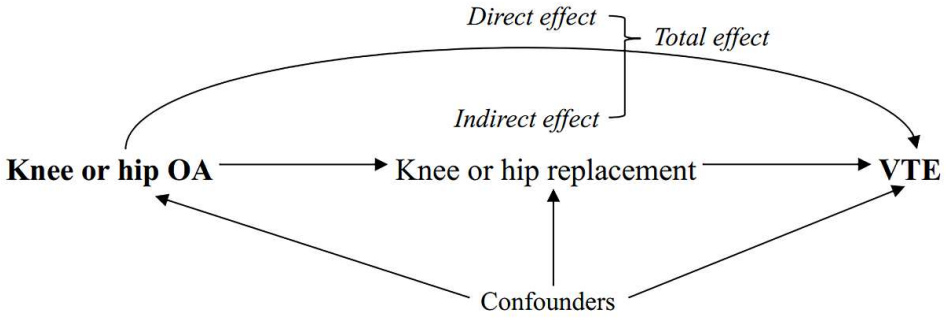
Supplemental Table S1. READ codes for knee osteoarthritis, hip osteoarthritis, hand osteoarthritis, pulmonary embolism and deep vein thrombosis

Knee osteoarthritis	N05z611, N05zL00, N05zM00, N053611, N051B00
Hip osteoarthritis	N053512, N05z511, N05zJ00, N051900, N051A00, N052511, N053500, N051500, N052500, N054500, N05zK00, N05z500
Hand osteoarthritis	N05z411, N053400, N051400, N051D00, N052400, N054400, N05zH00, N05zD00, N05zF00, N05zG00, N05z400, N05zE00, N05z412, N05z311, N051C00, N053900, N050600
Pulmonary embolism	G401.00, G401.12, G401000, G401100
Deep vein thrombosis	G801.00, G801.11, G801.12, G801.13, G801500, G801B00, G801C00, G801D00, G801E00, G801F00, G801z00

Zeng C, *et al*. *Ann Rheum Dis* 2020; 79:1616–1624. doi: 10.1136/annrheumdis-2020-217782



Supplemental Figure S1. A directed acyclic graph to describe the mediation effect between knee or hip OA and VTE. OA, osteoarthritis; VTE, venous thromboembolism.
Indirect effect: effect of knee or hip OA on the risk of VTE mediated through knee or hip replacement.
Direct effect: effect of knee or hip OA on the risk of VTE not mediated through knee or hip replacement.
Total effect: sum of indirect and direct effects.
Confounders: confounders of relations of knee or hip OA to knee or hip replacement and to VTE (i.e., age, sex, body mass index, region, smoking status, alcohol drinking, comorbidities, medication use, trauma, fracture, surgery, hospitalization and number of general practitioner visits).

Zeng C, *et al. Ann Rheum Dis* 2020; 79:1616–1624. doi: 10.1136/annrheumdis-2020-217782

Supplemental Text

Regarding missing values for smoking status and alcohol consumption, there were respectively 204 and 1,091 individuals in the knee OA cohort (883 and 4,768 in its matched non-OA cohort), 115 and 531 individuals in the hip OA cohort (453 and 2,381 in its matched non-OA cohort), and 52 and 300 individuals in the hand OA cohort (219 and 1,367 in its matched non-OA cohort).

For all analyses, missing values for smoking status and alcohol consumption were imputed using IVEware for SAS V.9.4 (SAS Institute, Cary, North Carolina, USA), which is based on chained equations, assuming that smoking status and alcohol consumption were missing at random and conditioning on a set of covariates. To minimize random error, we imputed 20 datasets and then combined estimates from these datasets.

The observed covariates included in the imputation model were age (continuous), sex (dichotomous), region (categorical), BMI (continuous), number of GP visits (ordinal), socioeconomic deprivation index (ordinal), venous thromboembolism (dichotomous), knee/hip/hand osteoarthritis (dichotomous), referral (dichotomous), aspirin (dichotomous), glucocorticoids (dichotomous), hormone replacement therapy (dichotomous), anticoagulation (dichotomous), cancer (dichotomous), congestive heart failure (dichotomous), stroke (dichotomous), myocardial infarction (dichotomous), atrial fibrillation (dichotomous), hypertension (dichotomous), diabetes

Zeng C, *et al*. *Ann Rheum Dis* 2020; 79:1616–1624. doi: 10.1136/annrheumdis-2020-217782

mellitus (dichotomous), chronic obstructive pulmonary disease (dichotomous), systemic lupus erythematosus (dichotomous), rheumatoid arthritis (dichotomous), chronic kidney disease (dichotomous), inflammatory bowel disease (dichotomous), varicose veins (dichotomous), coagulation disturbances (dichotomous), trauma (dichotomous), fracture (dichotomous), surgery (dichotomous), hospitalization (dichotomous), and survival time (continuous, square root transformed). For non-normally distributed variables, a square root approach was used to transform such variables to approximate normality. Including all the variables included in the analysis models plus possibly other predictors makes the missing at random assumption and the conclusions from these results more plausible.

Results from analyses restricted to complete cases are shown in **Supplemental Table 2**.

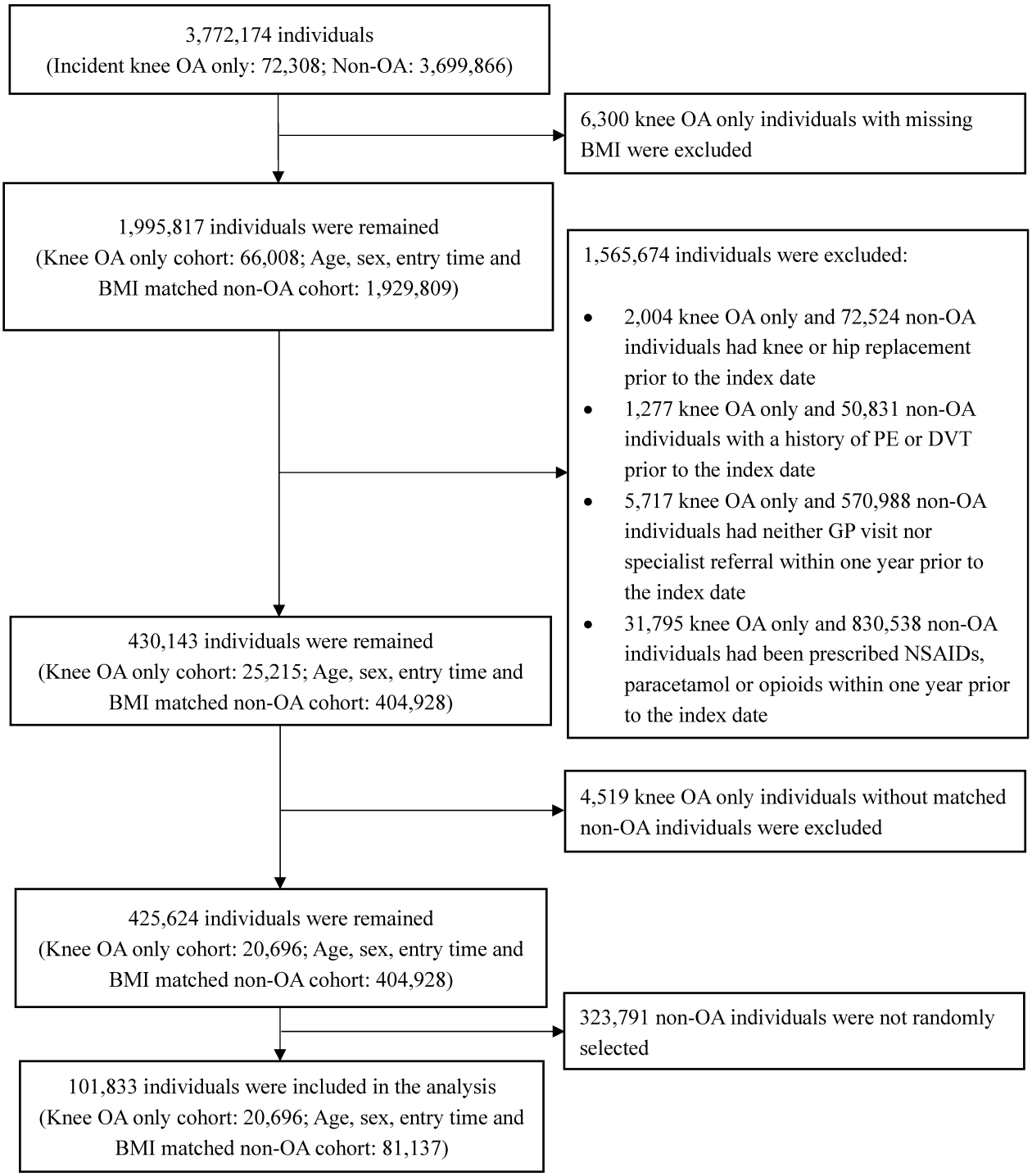
Supplemental Table S2. Association between incident OA and risk of venous thromboembolism (complete case analyses)

	Knee OA versus Non-OA		Hip OA versus Non-OA		Hand OA versus Non-OA	
	(n=19,345)	**(n=71,632)**	**(n=9,763)**	**(n=36,686)**	**(n=5,954)**	**(n=22,446)**
Venous thromboembolism						
Events (n)	313	821	181	395	54	207
Mean follow-up (years)	5.81	5.91	5.77	5.85	5.81	5.85
Incidence rate, per 1000 person-years	2.8	1.9	3.2	1.8	1.6	1.6
RD (95% CI), per 1000 person-years	0.9 (0.6, 1.2)	0.0 (reference)	1.4 (0.9, 1.9)	0.0 (reference)	0.0 (-0.5, 0.5)	0.0 (reference)
Age-, sex-, BMI-, and entry year-matched HR (95% CI)	1.48 (1.32, 1.67)	1.00 (reference)	1.74 (1.48, 2.03)	1.00 (reference)	0.97 (0.73, 1.29)	1.00 (reference)
Multivariable-adjusted HR (95% CI) *	1.46 (1.30, 1.65)	1.00 (reference)	1.72 (1.46, 2.03)	1.00 (reference)	0.94 (0.70, 1.26)	1.00 (reference)
Pulmonary embolism						
Events (n)	164	375	93	203	25	100
Mean follow-up (years)	5.84	5.93	5.82	5.87	5.83	5.86
Incidence rate, per 1000 person-years	1.5	0.9	1.6	0.9	0.7	0.7
RD (95% CI), per 1000 person-years	0.6 (0.4, 0.8)	0.0 (reference)	0.7 (0.3, 1.1)	0.0 (reference)	0.0 (-0.3, 0.3)	0.0 (reference)
Age-, sex-, BMI-, and entry year-matched HR (95% CI)	1.79 (1.52, 2.11)	1.00 (reference)	1.74 (1.39, 2.17)	1.00 (reference)	0.95 (0.63, 1.43)	1.00 (reference)
Multivariable-adjusted HR (95% CI) *	1.83 (1.54, 2.18)	1.00 (reference)	1.82 (1.44, 2.29)	1.00 (reference)	0.85 (0.53, 1.34)	1.00 (reference)

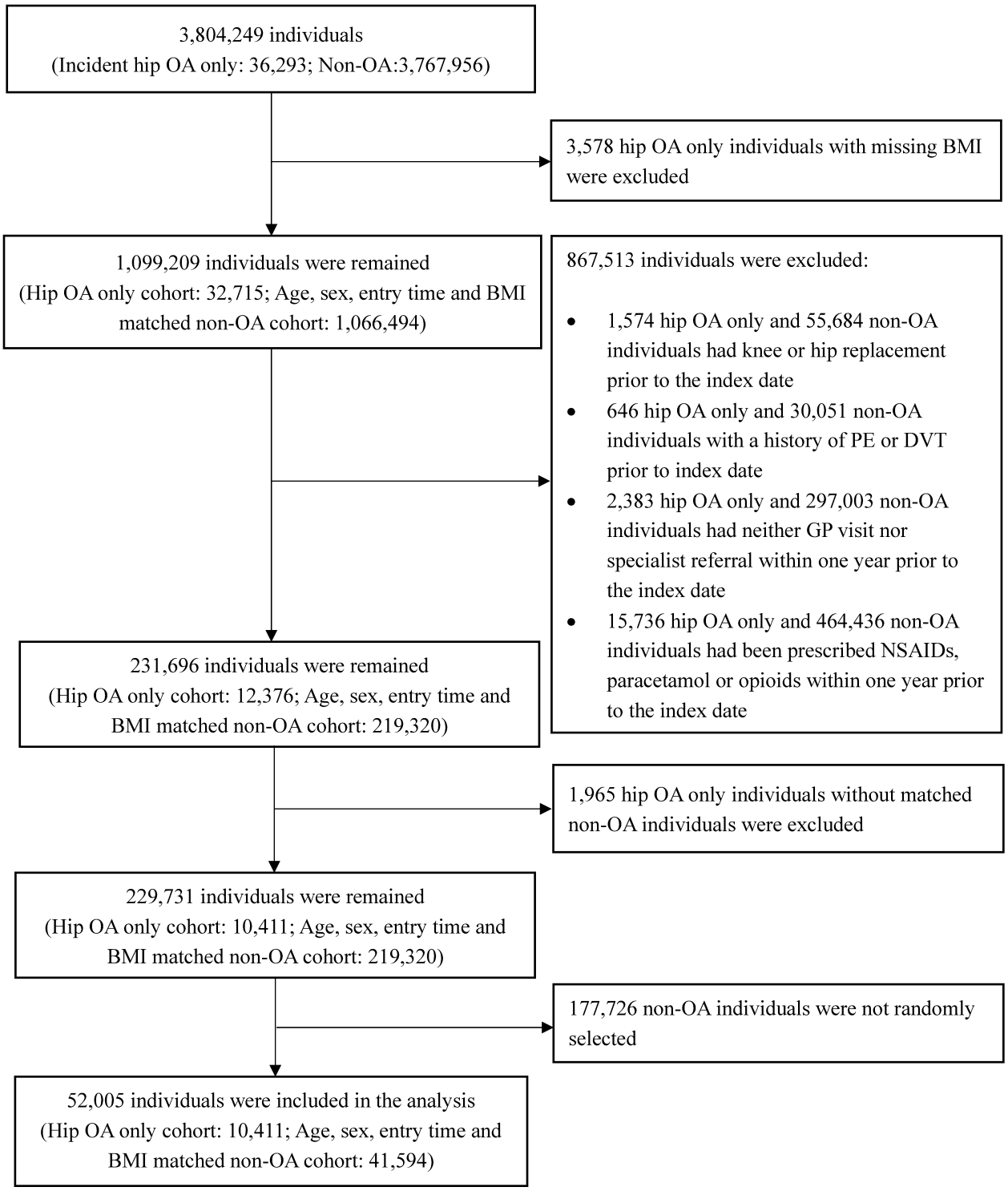
Deep vein thrombosis						
Events (n)	163	497	98	214	33	112
Mean follow-up (years)	5.84	5.92	5.80	5.86	5.81	5.86
Incidence rate, per 1000 person-years	1.5	1.2	1.7	1.0	1.0	0.9
RD (95% CI), per 1000 person-years	0.3 (0.1, 0.5)	0.0 (reference)	0.7 (0.3, 1.1)	0.0 (reference)	0.1 (-0.3, 0.5)	0.0 (reference)
Age-, sex-, BMI-, and entry year-matched HR (95% CI)	1.24 (1.06, 1.46)	1.00 (reference)	1.74 (1.40, 2.16)	1.00 (reference)	1.12 (0.77, 1.62)	1.00 (reference)
Multivariable-adjusted HR (95% CI) *	1.20 (1.02, 1.42)	1.00 (reference)	1.71 (1.36, 2.14)	1.00 (reference)	1.06 (0.73, 1.55)	1.00 (reference)

OA, osteoarthritis; n, number; HR, hazard ratio, RD, rate difference, BMI, body mass index.

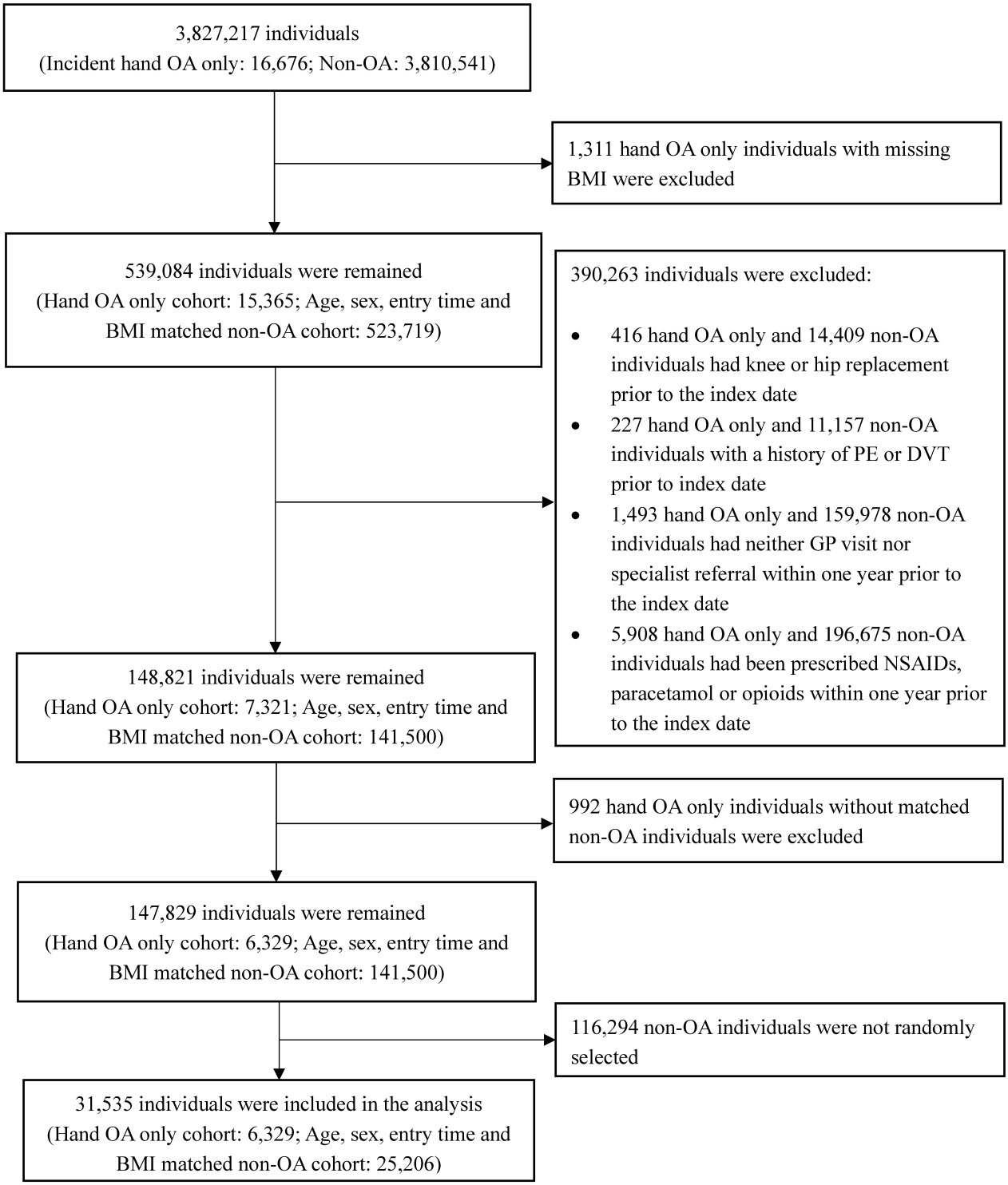
*In addition to the matching variables, multivariable models were adjusted for region, smoking status (i.e., non-smokers, ex-smokers and current smokers), alcohol drinking (i.e., non-drinkers, ex-drinkers and current drinkers), number of GP visits, comorbidities, medication use, trauma, fracture, surgery and hospitalization prior to the index date.



Supplemental Figure S2. Selection process of included individuals for the comparison between knee OA and non-OA. OA, osteoarthritis; THIN, The Health Improvement Network; BMI, body mass index; GP, general practitioner; PE, pulmonary embolism; DVT, deep vein thrombosis; NSAIDs, non-steroidal anti-inflammatory drugs.



Supplemental Figure S3. Selection process of included individuals for the comparison between hip OA and non-OA. OA, osteoarthritis; BMI, body mass index; GP, general practitioner; PE, pulmonary embolism; DVT, deep vein thrombosis; NSAIDs, non-steroidal anti-inflammatory drugs.



Supplemental Figure S4. Selection process of included individuals for the comparison between hand OA and non-OA. OA, osteoarthritis; BMI, body mass index; GP, general practitioner; PE, pulmonary embolism; DVT, deep vein thrombosis; NSAIDs, non-steroidal anti-inflammatory drugs.