Diabetologia (2012) 55:2173–2182
DOI 10.1007/s00125-012-2544-y

ARTICLE

Circulating 25-hydroxyvitamin D concentration and the risk of type 2 diabetes: results from the European Prospective Investigation into Cancer (EPIC)-Norfolk cohort and updated meta-analysis of prospective studies

N. G. Forouhi · Z. Ye · A. P. Rickard · K. T. Khaw · R. Luben · C. Langenberg · N. J. Wareham

Received: 15 December 2011 / Accepted: 8 March 2012 / Published online: 15 April 2012

Abstract
Aims/hypothesis Epidemiological evidence is suggestive, but limited, for an association between circulating 25-hydroxyvitamin D (25[OH]D) and risk of type 2 diabetes. We conducted a systematic review and meta-analysis that included new data from previously unpublished studies.
Methods Using a nested case–cohort design in the European Prospective Investigation into Cancer (EPIC)-Norfolk study, we identified a random subcohort and incident type 2 diabetes cases occurring between baseline (1993–1997) and 2006. In the Ely prospective study we identified incident type 2 diabetes cases between 1990 and 2003. We conducted a systematic review of prospective studies on 25(OH)D and type 2 diabetes published in MEDLINE or EMBASE until 31 January 2012, and performed a random-effects meta-analysis combining available evidence with results from the EPIC-Norfolk and Ely studies.
Results In EPIC-Norfolk, baseline 25(OH)D was lower among incident type 2 diabetes cases (mean [SD] 61.6 [22.4] nmol/l; n=621) vs non-case subcohort participants (mean 65.3 [23.9] nmol/l; n=826). There was an inverse association between baseline 25(OH)D and incident type 2 diabetes in multivariable-adjusted analyses: HR (95% CI) 0.66 (0.45, 0.97), 0.53 (0.34, 0.82), 0.50 (0.32, 0.76), p trend <0.001, comparing consecutive increasing 25(OH)D quartiles with the lowest. In Ely, 37 incident type 2 diabetes cases were identified among 777 participants. In meta-analysis, the combined RR of type 2 diabetes comparing the highest with lowest quartile of 25(OH)D was 0.59 (0.52, 0.67), with little heterogeneity (I^2=2.7%, p=0.42) between the 11 studies included (3,612 cases and 55,713 non-cases).
Conclusions/interpretation These findings demonstrate an inverse association between circulating 25(OH)D and incident type 2 diabetes. However, causal inference should be addressed through adequately dosed randomised trials of vitamin D supplementation or genetic Mendelian randomisation experiments.

Keywords Causal inference · 25-Hydroxyvitamin D · Meta-analysis · Prospective · Systematic review · Type 2 diabetes · Vitamin D

Abbreviations

EPIC	European Prospective Investigation of Cancer
25(OH)D	Circulating 25-hydroxyvitamin D
MRC	Medical Research Council
PAF	Population-attributable fraction

Introduction

The potential impact of feasible approaches to reduce the increasing global burden [1] of diabetes could be vast. While primary prevention of type 2 diabetes through lifestyle modification represents a promising strategy among high-risk individuals [2], it is acknowledged that many key

N. G. Forouhi (✉) · Z. Ye · A. P. Rickard · C. Langenberg · N. J. Wareham
Medical Research Council Epidemiology Unit,
Institute of Metabolic Science, Addenbrooke's Hospital,
Box 285, Hills Road,
Cambridge CB2 0QQ, UK
e-mail: nita.forouhi@mrc-epid.cam.ac.uk

K. T. Khaw · R. Luben
Department of Public Health and Primary Care,
University of Cambridge,
Cambridge, UK

lifestyle changes are difficult to implement or sustain. It has been suggested that increasing the concentration of circulating serum 25-hydroxyvitamin D (25[OH]D), an indicator of vitamin D status, could provide a practical complementary approach.

Epidemiological studies suggest an overall tendency towards an inverse association between vitamin D status and risk of type 2 diabetes [3–5]. However, some studies assessed only dietary vitamin D intake [6, 7], which does not include the major non-dietary component of vitamin D from sun exposure. Other studies measured circulating 25(OH)D, the best indicator of vitamin D status [8], but had small sample sizes or a limited number of cases [9–11]. A recent study reported no association between 25(OH)D and risk of diabetes among older women [12]. A systematic review by Pittas et al [13] identified three studies of the association between vitamin D status and diabetes risk. Two assessed dietary vitamin D intake and one included 25(OH)D; meta-analysis was not possible. Further work by the same group included a meta-analysis of data from four studies, which showed an inverse association, but 25(OH)D status reflected measured level as well as the use of a predicted 25(OH)D score [14]. Vitamin D trial evidence has shown no significant effect of supplementation on outcomes of glycaemia or incident diabetes, but individual trials have had limitations, including their post-hoc nature, small sample size, inadequate dose of supplement, or use of combination supplementation with calcium such that individual effects could not be teased out easily [13–16]. While further evidence from trials would help to address the issue of a causal association, currently the epidemiological evidence for a possible association between circulating 25(OH)D and type 2 diabetes remains inconclusive, and the question is important to resolve.

Therefore, our aim was to conduct a systematic review and updated meta-analysis of the longitudinal association between circulating 25(OH)D and incident type 2 diabetes, including the contribution of previously unpublished data.

Methods

European Prospective Investigation into Cancer-Norfolk study

Described in detail previously [17], the European Prospective Investigation into Cancer (EPIC)-Norfolk study included 25,639 men and women from Norfolk general practices, aged 40–75 years at the baseline study visit (1993–1997). We designed a nested case–cohort (n=1,852), including 892 incident type 2 diabetes cases occurring before 31 July 2006 and 989 subcohort participants (including 29 incident diabetes cases), selected at random from the entire cohort and representative of the whole cohort. Prevalent cases of diabetes (n=32) were excluded from analysis. Clinically incident type 2 diabetes cases were ascertained using multiple sources and verified by record linkage with general practice, hospital and death registries. Complete data were available for n=1,447 for current analysis, with n=621 type 2 diabetes cases and n=849 subcohort participants (including 23 incident diabetes cases and 826 non-case participants in the subcohort, as per the design of the case–cohort study). Ultra-performance liquid chromatography–tandem mass spectrometry was used to measure 25(OH) vitamin D_2 and 25(OH) vitamin D_3, which were summed to provide a measure of total 25(OH) vitamin D. This assay was standardised against the National Institute of Standards and Technology (NIST). Written informed consent was obtained from all participants, and the study was approved by the Norwich District Ethics Committee. In addition to including the EPIC-Norfolk study in the meta-analysis, we used this cohort to investigate the effect of adjusting for potential confounding factors that did not occur consistently across other studies.

Ely study

The Medical Research Council (MRC) Ely study [18] is a population-based prospective study of European-origin adults in which 1,122 participants were randomly selected from a sampling frame of all adults aged 40–69 years and without diabetes, registered at a single general practice in Ely. Baseline measurements were performed between 1990 and 1992, with follow-up visits at phase 2 (1994–1996) and phase 3 (2000–2003) among individuals who were non-diabetic at previous visits. By a median of 10 years of follow-up, 61 cases of incident diabetes were diagnosed based on WHO criteria [18]. A subset of 740 non-diabetic cohort participants and 37 incident type 2 diabetes cases for whom data on baseline serum 25(OH)D concentration, measured by radioimmunoassay, was available were included in the current analysis. All participants gave written informed consent and the study was approved by the local research ethics committee. A prior analysis examined the association of serum 25(OH)D with continuous metabolic traits [19], but did not examine diabetes incidence.

Systematic review: literature search and data extraction

Long-term prospective cohort studies published before 31 January 2012 that reported on the association between 25(OH)D and type 2 diabetes were sought using MEDLINE and EMBASE, and by scanning the reference lists of articles identified for relevant studies and reviews. The search terms

were related to vitamin D levels ('25-hydroxyvitamin D' or '25(OH)D' or 'vitamin D'), and diabetes risk ('diabetes' or 'glucose' or 'metabolic syndrome' or 'hyperglycaemia') without limits on publication date or language. One author (Z. Ye) reviewed all identified titles (n=2,265) and subsequently the abstracts and full articles (Fig. 1). Studies were included if they had a prospective study design, included adult men or women (or constituted a nested case–control or nested case–cohort within a prospective study), measured circulating (serum or plasma) 25(OH)D at baseline, ascertained incident type 2 diabetes based on self-report of physician diagnosis or by biochemical criteria defined by WHO or the American Diabetes Association, and provided an RR and corresponding 95% CI for type 2 diabetes in relation to plasma 25(OH)D levels. We excluded studies examining only dietary vitamin D intake or supplement use, or those with predicted 25(OH)D score, including only those that measured circulating 25(OH)D. For each contributing study, information was extracted according to a pre-specified protocol, and corresponding authors were contacted to submit data where appropriate (n=2) (Table 1). We used data from the maximally adjusted analysis presented in each study. The results variously reported as ORs, RRs and HRs were assumed to approximate the same measure of RR [20]. We included unpublished results from the EPIC-Norfolk study, described in more detail below. Additionally, our previously published analyses for the association between 25(OH)D and continuous metabolic-syndrome-related traits in the Ely Study [19] were extended to the endpoint of incident type 2 diabetes [18].

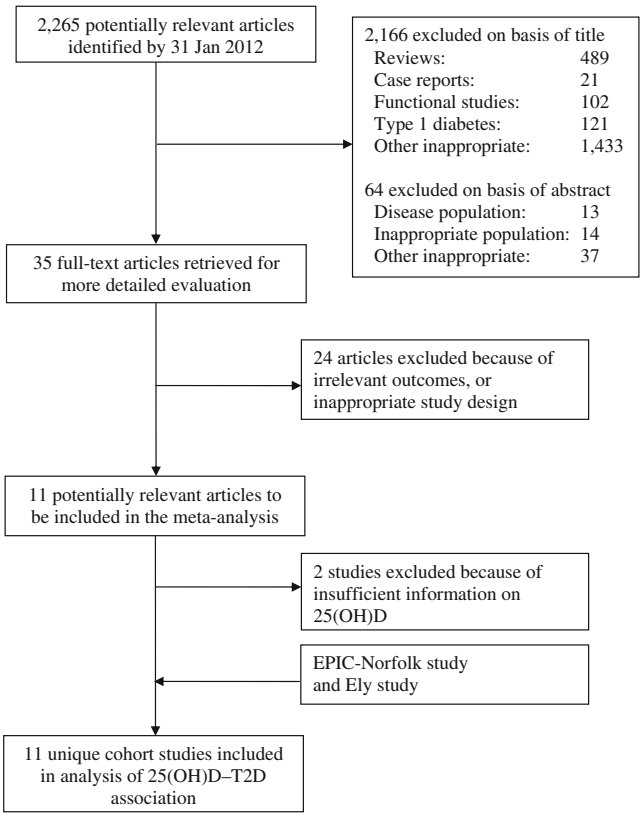

Fig. 1 Flow diagram of studies identified in the literature search for the association between circulating 25(OH) D and the risk of type 2 diabetes. T2D, type 2 diabetes

Statistical analysis

EPIC-Norfolk study Baseline characteristics of the participants were examined across quartiles of 25(OH)D in the subcohort. Differences were examined using mean and SD for continuous variables, with p values from an ANOVA test or using number and per cent for categorical variables, with χ^2 test. Multivariable Prentice-weighted Cox-regression was used, allowing for the nested case–cohort design [21, 22], to examine associations between total 25(OH)D (quartiles) and incident diabetes. Models 1 to 4 examined the association including age (continuous) as the underlying timescale, sex (1 = men, 2 = women), season of blood test (1 = January–March, 2 = April–June, 3 = July–September, 4 = October–December), cigarette smoking (1 = current smoker, 2 = former smoker, 3 = never smoker), education (0 = none reported, 1 = O level or equivalent [to age 16 years], 2 = A level or equivalent [to age 18 years], 3 = degree or higher [university/college/equivalent level]), family history of diabetes [1 = yes, 2 = no], physical activity index derived from self-reported physical activity [23] [1 = inactive, 2 = moderately inactive, 3 = moderately active, and 4 = active], self-reported alcohol intake [continuous] and supplement use [0 = yes for any vitamins or cod liver oil use, 1 = no]. Model 5 additionally adjusted for BMI (continuous). The population-attributable fraction (PAF) for diabetes associated with 25(OH)D level was calculated using the equation below

$$\mathrm{PAF} = 100 \times \frac{[\mathrm{Px} \times (\mathrm{HR} - 1)]}{1 + [\mathrm{Px} \times (\mathrm{HR} - 1)]}$$

where Px is the prevalence of the exposure (defined as vitamin D insufficiency, with 25[OH]D <50 nmol/l [24]), and HR is the hazard ratio of diabetes when comparing the exposed (25[OH]D <50 nmol/l) vs the unexposed (25[OH]D ≥50 nmol/l) groups in a multivariable-adjusted model that included the covariates specified in model 5 above. Assuming causality, PAF is the estimated proportion by which type 2 diabetes incidence would be reduced in the entire population if the exposure (i.e. low 25-hydroxyvitamin D, <50 nmol/l) was eliminated.

Ely study We examined the association between baseline 25 (OH)D (comparing quartiles) and incident type 2 diabetes using logistic regression, and constructed an adjusted model as specified in Table 1.

Table 1 Prospective studies of the association between circulating 25(OH)D and type 2 diabetes included in the meta-analysis

Study	Study design/ follow-up (years)	Location	Demographics of sample: age (mean ± SD [or range], years); % women; ethnicity	Sample size	Sample selection		25(OH)D measurement method/mean 25(OH)D (SD)	Adjustments
					Ascertainment of cases	Controls		
Mattila et al, (2007) [9], Mini-Finland	Cohort/17	Finland	53 (range 40–69 years); 53%; NA (majority European)	Case 187; non-case 3,910	Mini-Finland Health Survey with linkage to national re-imbursement registry for costs of diabetes medication	Non-case cohort	Radioimmunoassay/ 43.6 (19.5) nmol/l	Age, sex, month, BMI, leisure time exercise, smoking, education, blood pressure
Knekt et al, (2008) [10], FMC	Nested case–control/22	Finland	40–74; 54%; NA (majority European)	Case 230; control 452	Finnish Mobile Clinic Health Examination Survey & Mini-Finland Health Survey with linkage to national re imbursement registry for costs of diabetes medication	Matched controls from cohorts	Radioimmunoassay/ 47.3 nmol/l	Age, BMI, physical activity, smoking, education
Grimnes et al, (2010) [11], Tromso	Cohort/11	Norway	58.9±10.2; 60%; majority European	Case 247; non-case 5,872	Self-report, verified by medical records, or elevated HbA_{1c}, or linkage to local hospital records in Tromso cohort study	Non-case cohort	Electrochemiluminescent immunoassay/59.3 nmol/l	Age, sex, month, physical activity, BMI
Anderson et al, (2010) [5], Intermountain Healthcare study	Prospective/1.3 (maximum 9.3)	USA	55±21; 74.8%; 86% white	Case 724; non-case 31,877	Electronic medical record database of the integrated healthcare system ICD-9 codes	All patients in a healthcare system with a vitamin D measure available	Chemiluminescent immunoassay/NA	Age, sex, history of hypertension or hyperlipidaemia or heart failure or infection or depression or renal failure
Pittas et al, (2010) [4], Nurses' Health Study	Nested case–control/14	USA	56.4 (SD NA); 100%; NA	Case 608; control 559	Self-report and confirmation by supplementary questionnaire in the Nurses Health Study	Matched controls from cohort	Radioimmunoassay/ 56.7 nmol/l	Age, race, fasting status, month, laboratory batch, latitude, history of hypercholesterolaemia or hypertension, family history of diabetes, smoking, physical activity, alcohol, multivitamin use, dietary variables, BMI
Robinson et al, (2011) [12], WHI	Nested case–control/7.3	USA	66.3±7.1;100%; white, black, Hispanic	Case 317; control: 4,823	Nested in Women's Health Initiative—self-report of physician diagnosis or use of oral hypoglycaemic agent or insulin on a medication inventory	Matched controls from cohort	Chemiluminescent immunoassay/NA	Age, ethnicity, latitude of clinical centre, month, WHI study indicators, BMI, hypertension, fibre intake, magnesium intake and physical activity
Gagnon et al, (2011) [35], AusDiab	Cohort/5	Australia	50.8±12.5; 54.7%; 92.2% European	Case 199; non-case 5,001	Treatment with insulin or oral hypoglycaemic agents, fasting plasma glucose ≥7 mmol/l or 2 h plasma glucose post-OGTT ≥11.1 mmol/l in AusDiab cohort	Non-case cohort	Chemiluminescent immunoassay/65 (25) nmol/l	Age, ethnicity, waist circumference, family history of diabetes, smoking, physical activity, season, latitude, hypertension, serum triacylglycerols, fasting plasma glucose

Table 1 (continued)

Study	Study design/ follow-up (years)	Location	Demographics of sample: age (mean ± SD [or range], years); % women; ethnicity	Sample size	Sample selection		25(OH)D measurement method/mean 25(OH)D (SD)	Adjustments
					Ascertainment of cases	Controls		
Thorand et al, (2011) [36], MONICA/KORA	Case–cohort/11	Germany	52.7±7.1; 47.2%; 100% European	Case 416; non-case 1,267	Within cohort, self-report, validated by contacting physi cian or medical chart review	Randomly selected, representative of MONICA/ KORA	Immunossay/37.8 nmol/l in cases/42.0 nmol/l in non-cases[a]	Age, sex, survey, season, BMI, smoking, alcohol consumption, systolic blood pressure, total cholesterol/ HDL-cholesterol, parental history of diabetes, C-reactive protein, IL-6, solu ble intercellular adhesion molecule-1, and inducible protein-10/CXCL10
Gonzalez-Molero et al, (2012) [37]	Cohort/5	Spain	50.3±14.4;57.0%; NA (majority European)	Case 26; non-case 386	OGTT and glycosylated haemoglobin in cohort	Non-case cohort	Electrochemiluminescent immunoassay/51.8 (14.5) nmol/l in cases/58.3(15.5) nmol/l in non-cases	Age, sex, obesity, smoking, outdoor activity, alcohol and month of extraction
EPIC-Norfolk study (current analysis)	Case–cohort/10	UK	60±9.1; 58%; >99% European	Case 621; non-case 826	In the EPIC-Norfolk cohort, self-report verified by record linkage with medical records	Randomly selected, representative of EPIC-Norfolk	Liquid chromatography–tandem-mass spectrometry/ 65.3 (23.9)[a] nmol/l	Age, sex, season, family history of diabetes, smoking, physical activity, education level, alcohol, supplement use, BMI
Ely study (analysis from Forouhi et al [18, 19])	Cohort/10	UK	63.5±7.7/57.7%; >99% European	Case 37; non case 740	OGTT in the Ely study	Non-case cohort	Radioimmunoassay/ 58.8 (25.4) nmol/l	Age, sex, season, family history of diabetes, cholesterol, alcohol intake, smoking status, socioeconomic status, BMI, physical activity

[a] In non-case subcohort only

AusDiab, Australian Diabetes, Obesity and Lifestyle Study; CXCL10, chemokine (C-X-C motif) ligand 10; FMC, Finnish Mobile Clinic Health Examination Survey; ICD-9, International Classification of Diseases, 9th Revision (www.icd9data.com/2007/Volume1/240-279/250-259/250/default.htm); NA, not available; WHI, Women's Health Initiative

Meta-analysis Pooled estimates across studies were obtained by random-effect summary measures of the reported log-risk ratios weighted by the inverse of the variance. Heterogeneity was assessed using the Q statistic test [25] and the I^2 statistic [26]. Evidence of publication bias was assessed using funnel plots and Begg's test [27].

All analyses were performed using Stata/SE10.1 (Stata, College Station, TX, USA). All statistical tests were two-sided and used a significance level of $p<0.05$.

Results

EPIC-Norfolk study

The mean age at baseline was 58.0 (SD 9.4) years and 58% of the subcohort were women. The mean (SD) total 25(OH)D concentration was lower in cases (61.6 [22.4] nmol/l, $n=621$) than among the non-case subcohort (65.3 [23.9] nmol/l, $n=826$), $p<0.001$. Circulating 25(OH)D levels were highest over July to September (82.2 nmol/l [SD 25.1]), and lowest over January to March (53.8 nmol/l [SD 20.8]). Table 2 shows the thresholds and ranges for the quartile distribution of 25(OH)D in the subcohort. The cut-off for the lowest quartile in this population was 48.8 nmol/l, which is close to the accepted definition of vitamin D insufficiency (levels <50 nmol/l). Supplement use (vitamins or cod liver oil) was significantly greater across increasing quartiles of 25(OH)D, and there were more current smokers in the lower quartiles of the 25(OH)D distribution, while mean BMI was slightly higher in the two middle quartiles (quartiles two and three), than in quartiles one and four (Table 2). Table 3 shows that 25(OH)D was inversely associated with incident type 2 diabetes. The proportional hazards assumption was valid ($p=0.389$). Comparing the highest and lowest quartiles of 25(OH)D, the HR for diabetes was 0.64 (95% CI 0.47, 0.88) in age–sex adjusted analysis, reducing to 0.47 (0.33, 0.67) on adjustment for season of blood test. There was no interaction between 25(OH)D and season on diabetes risk ($p=0.644$). Further adjustment made negligible difference, with an overall 50% lower hazard of developing type 2 diabetes (HR 0.50, 95% CI 0.32, 0.76) in model 5. Additional adjustment for blood pressure or lipids did not change our findings. The PAF for incident type 2 diabetes was 17.6%, with a prevalence of exposure of 26.5% in the subcohort, defined by 25(OH)D <50 nmol/l.

Ely study

Mean age at baseline was 63.5 (SD 7.7) years and 57.7% were women. The mean (SD) 25(OH)D concentration was slightly lower in cases (57.5[20.6] nmol/l, $n=37$) than in non-case participants (58.7 [25.3] nmol/l, $n=740$). The OR of type 2 diabetes was 0.69 (95% CI 0.17, 2.91), $p=0.69$ in adjusted analysis (Table 1), comparing the highest with the lowest quartile of 25(OH)D.

Systematic review and meta-analysis

Figure 1 shows that the initial searches identified 2,265 articles (PubMed 1,732 articles, EMBASE 682 articles, duplicates removed). After exclusions, 11 relevant studies were initially identified, but two [28, 29] were ineligible because of limited information on 25(OH)D; thus, nine unique prospective studies were retained [4, 5, 9–12, 30–32]. Our meta-analysis included 11 studies, nine published plus new, previously unpublished, data from the EPIC-Norfolk and Ely studies [19] (Table 1); this gave a total of 3,612 type 2 diabetes cases and 55,713 non-cases. A comparison of individuals in the top quartile with those in the bottom quartile of baseline 25(OH)D yielded a combined RR of 0.59 (95% CI 0.52, 0.67) (Fig. 2). There was little evidence of heterogeneity ($I^2=2.7\%$ [95% CI 0%, 61%], $p=0.42$). No publication bias was observed when either using Begg's test ($p=0.82$) or visually inspecting the funnel plot (not shown). As the Intermountain Healthcare study [5], the largest contributing study in this meta-analysis, had a somewhat different design compared with the other studies included (Table 1), despite it meeting our inclusion criteria, we also repeated the analysis with its exclusion. The pooled RR of type 2 diabetes after the exclusion of this study was 0.64 (95% CI 0.54, 0.76).

Discussion

The present meta-analysis of 11 prospective studies involving a total of 3,612 cases and 55,713 non-case participants provides the largest and most comprehensive assessment thus far of the association between circulating 25(OH)D levels and type 2 diabetes. It suggests a strong inverse association between serum 25(OH)D concentration and incident type 2 diabetes. The combined RR of 0.59 suggests that the risk of future diabetes may be reduced by 41% (95% CI 33%, 48%) by being in the top rather than the bottom quartile of 25(OH)D at baseline. If 25(OH)D levels are causally related to diabetes risk, then this finding could have a substantial public health impact, as we estimated a PAF of 17.6% associated with vitamin D insufficiency (levels <50 nmol/l) in the EPIC-Norfolk study.

The strengths of this study are that this meta-analysis included only prospective studies with data on measured circulating 25(OH)D, the best indicator of vitamin D status, and the analysis of EPIC-Norfolk data enabled us to examine the impact on the estimated association of a range of important potential confounders. The meta-

Table 2 Baseline characteristics of study participants by quartiles of serum 25(OH)D concentration in the subcohort: EPIC-Norfolk Study

Characteristic	25(OH)D (nmol/l)				*p* value
	Quartile 1 <48.8	Quartile 2 49.0-63.5	Quartile 3 63.6-80.0	Quartile 4 >80.0	
n	215	210	213	211	
Age (years)	57.9 (9.2)	58.7 (9.8)	58.6 (9.1)	58.6 (9.5)	0.82
Women (%)	133 (61.9)	123 (58.6)	117 (54.9)	121 (57.4)	0.53
Serum calcium (mmol/l)	2.43 (0.23)	2.46 (0.18)	2.46 (0.19)	2.44 (0.25)	0.47
Using supplement or cod liver oil (%)	84 (39.1)	93 (44.3)	110 (51.6)	126 (59.7)	<0.001
Alcohol (units/week)	6.1 (8.5)	6.6 (7.9)	6.9 (7.5)	6.6 (8.5)	0.81
BMI (kg/m^2)	25.7 (3.7)	26.8 (3.8)	26.0 (3.6)	25.4 (3.1)	<0.001
Systolic blood pressure (mmHg)	134.7 (18.5)	134.6 (19.3)	134.9 (17.7)	132.7 (18.1)	0.58
Diastolic blood pressure (mmHg)	82.2 (11.5)	81.7 (11.4)	82.1 (10.8)	81.7 (11.5)	0.96
Total cholesterol (mmol/l)	6.06 (1.10)	6.19 (1.15)	6.06 (1.03)	6.22 (1.15)	0.30
HDL-cholesterol (mmol/l)	1.44 (0.45)	1.41 (0.42)	1.41 (0.38)	1.45 (0.41)	0.65
LDL-cholesterol (mmol/l)	3.89 (1.03)	3.97 (0.98)	3.93 (0.95)	4.01 (1.01)	0.67
Triacylglycerols (mmol/l)	1.5 (1.0, 2.0)	1.6 (1.2, 2.3)	1.4 (1.0, 2.1)	1.4 (1.0, 2.2)	0.050
Smoking					0.052
Never (%)	95 (44.2)	99 (47.1)	108 (50.7)	94 (44.6)	
Former (%)	77 (35.8)	82 (39.1)	85 (39.9)	92 (43.6)	
Current (%)	43 (20.0)	29 (13.8)	20 (9.4)	25 (11.9)	
Highest educational achievement					0.40
No reported qualifications (%)	70 (32.6)	75 (35.7)	89 (41.8)	77 (36.5)	
O level or equivalent (to age 16 years) (%)	23 (10.7)	17 (8.1)	28 (13.2)	20 (9.5)	
A level or equivalent (to age 18 years) (%)	88 (40.9)	89 (42.4)	73 (34.3)	85 (40.3)	
Degree or higher (university/college/equivalent) (%)	34 (15.8)	29 (13.8)	23 (10.8)	29 (13.7)	
Physical activity					0.29
Inactive (%)	66 (30.7)	58 (27.6)	54 (25.4)	46 (21.8)	
Moderately inactive (%)	64 (29.8)	66 (31.4)	61 (28.6)	58 (27.5)	
Moderately active (%)	48 (22.3)	53 (25.2)	48 (22.5)	57 (27.0)	
Active (%)	37 (17.2)	33 (15.7)	50 (23.5)	50 (23.7)	
Season					<0.001
Jan–Mar (%)	100 (46.5)	63 (30.0)	47 (22.1)	22 (10.4)	
Apr–Jun (%)	57 (26.5)	61 (29.1)	66 (31.0)	45 (21.3)	
Jul–Sep (%)	13 (6.1)	31 (14.8)	43 (20.2)	89 (42.2)	
Oct–Dec (%)	45 (20.9)	55 (26.2)	57 (26.8)	55 (26.1)	

Results are from the subcohort (n=849), and are either mean (SD) for continuous variables or number (%) for categorical variables

Triacylglycerols had a skewed distribution, so were log transformed for the ANOVA; medians and IQRs are presented

p value is from a χ^2 test for categorical variables and from ANOVA for continuous variables

analysis included a large sample, with over 3,600 cases, substantially greater than that included in previous work [14]. We were able to compute a pooled RR of incident type 2 diabetes comparing the top with the bottom quartile of 25(OH)D across all the included studies, offering a 'harmonisation' of the exposure variable across studies, which is statistically more robust than comparing extreme categories of exposure where the categories differ across studies.

Our study has some limitations. It predominantly included participants of European descent, and our findings cannot therefore be applied to other ethnic populations; specifically designed studies that include other ethnic groups are warranted. A number of the studies included in the meta-analysis were conducted prior to the standardisation of the 25(OH)D assay and consequently 25(OH)D was measured by a number of methods, which may reduce the ability to directly compare studies. Notwithstanding, our inclusion of

Table 3 Association between baseline serum 25(OH)D and incident type 2 diabetes: EPIC-Norfolk study

Model	25(OH)D (nmol/l)				*p* value for linear trend
	Quartile 1 (reference) <48.8 *n*=396	Quartile 2 49.0–63.5 *n*=375	Quartile 3 63.6–80.0 *n*=345	Quartile 4 >80.0 *n*=331	
1	1	0.93 (0.69–1.25)	0.71 (0.52–0.97)	0.64 (0.47–0.88)	0.0015
2	1	0.88 (0.65–1.19)	0.62 (0.45–0.86)	0.47 (0.33–0.67)	<0.001
3	1	0.88 (0.63–1.22)	0.61 (0.43–0.88)	0.46 (0.31–0.67)	<0.001
4	1	0.89 (0.65–1.24)	0.64 (0.44–0.91)	0.48 (0.33–0.71)	<0.001
5	1	0.66 (0.45–0.97)	0.53 (0.34–0.82)	0.50 (0.32–0.76)	<0.001

Results are HRs and 95% CIs using Prentice-weighted Cox regression

p values for linear trends were obtained by including the quartile variable (coded 0, 1, 2, 3) as a linear term in the model, and using a Wald test to test the null hypothesis that the effect of this variable is 0

Model 1: age as underlying timescale, sex

Model 2: model 1 plus season

Model 3: model 2 plus family history of diabetes, cigarette smoking, physical activity and education level

Model 4: model 3 plus alcohol intake and supplement and/or cod liver oil use

Model 5: model 4 plus BMI

only those studies that measured circulating 25(OH)D with a named method, and exclusion of studies in which 25(OH)D was derived or vitamin D status assessed only from dietary questionnaires, offers a more direct comparison than has previously been done. Also, in all the included studies, there was a single baseline measurement of 25(OH)D, which may not reflect vitamin D status over long periods. The ascertainment of the outcome (type 2 diabetes) was by self-report in several of the included studies, and this could include misclassification of participants with undiagnosed diabetes as non-cases. However, the overall effect of such an error would be likely to attenuate the effect estimates towards unity. The level of adjustment for potential confounders varied in the different studies included (as shown in Table 1); thus, the effect estimates used in the meta-analysis might have been over- or underestimates of the true effect size. Future studies could improve the validity of findings if repeat measurements of 25(OH)D become available, in tandem with biochemical assessment of diabetes incidence. The latter goal should now be more feasible, as the logistical constraints of repeat fasting glucose or OGTTs required for the diagnosis of diabetes can now be overcome with the measurement of non-fasted HbA_{1c}, which is approved as a diagnostic test for diabetes [30].

Our meta-analysis demonstrates an inverse association of 25(OH)D with the risk of diabetes, and plausible mechanisms for such association have been proposed, including,

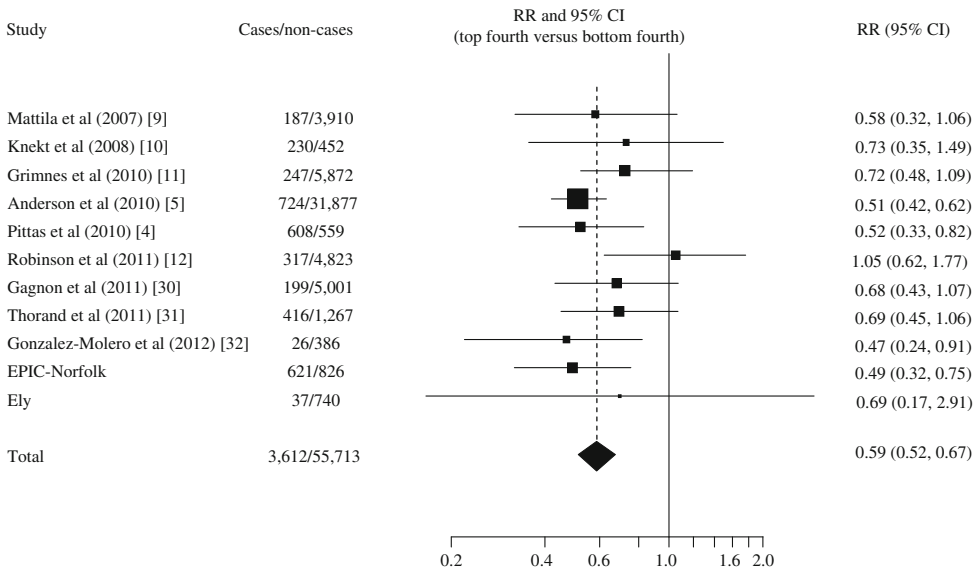

Fig. 2 The association between circulating 25(OH)D and incident type 2 diabetes: meta-analysis of prospective studies. In the forest plot, the sizes of the boxes for individual studies are inversely proportional to the variances of log RRs, and horizontal lines represent 95% CI. There was no significant heterogeneity (I^2=2.7%, *p* value=0.42)

among others, potential effects through the presence of vitamin D receptors in pancreatic beta cells influencing insulin secretion or through the effects of 25(OH)D on calcium metabolism [31, 32]. However, our meta-analysis cannot address the issue of whether the observed association is likely to be causal. Meta-analysis principally deals with the issues of consistency and precision and can provide a better estimate of the measure of association. However, it does not resolve the problem of confounding that is universal in observational studies. Classic adjustment for a wide range of confounding factors in the EPIC-Norfolk study did not materially alter the measure of association, but we cannot exclude confounding by factors we did not consider, or residual confounding by factors for which we did not adjust sufficiently. To address the question of causality, there is a need for specifically designed randomised controlled trials that use adequate doses of supplementation and a pre-specified outcome of diabetes incidence. There are currently no such trials, as summarised in the recent systematic review and meta-analysis by Mitri *et al* [14]. An alternative approach that could also help to address causality would be the use of Mendelian genetic randomisation studies. By examining the association between genetic variants [33], 25(OH)D level and type 2 diabetes, it is possible to investigate consistency between risk estimates obtained from genotype-disease studies and those from phenotype-disease studies using genes as instrumental variables to address the problems of confounding [34].

In conclusion, the findings of our meta-analysis provide evidence for a strong inverse association between circulating 25(OH)D levels and risk of incident type 2 diabetes, an association that remained largely unchanged in a new analysis of the EPIC-Norfolk study that considered a range of relevant potential confounding factors. However, there is as yet no demonstrable evidence of causality, and clinicians should exercise caution when interpreting or acting on the epidemiological evidence alone that currently dominates this field.

Acknowledgements We thank all study participants, general practitioners and the EPIC-Norfolk study team for their helpful input, and S. Sharp for statistical help. We thank J. Anderson and H. May, Cardiovascular Department, Intermountain Medical Center, UT, USA and I. Gonzalez-Molero, Servicio de Endocrinologia y Nutricion, Hospital Universitario Carlos Haya, for providing the additional data. We thank HFL Sport Science (a trading division of LGC) for the measurement of 25(OH)D.

Funding The EPIC-Norfolk study is supported by funding from the MRC, the Food Standards Agency, Cancer Research UK and the British Heart Foundation.

No funding source had any role in the design and conduct of the study, the collection, management, analysis, and interpretation of the data or the preparation, review or approval of the manuscript.

Duality of interest The authors declare that there is no duality of interest associated with this manuscript.

Contribution statement NJW, KTK (principal investigators) and RL collected data for the EPIC-Norfolk study, and NGF and NJW arranged the measurement of vitamin D status in the EPIC-Norfolk study. NGF, NJW and CL conceived this study. ZY, APR and NGF analysed data and wrote early draft manuscripts. NGF wrote the final manuscript. All authors contributed to the interpretation of data and to revising the manuscript critically for important intellectual content. All authors approved the final text.

References

1. Shaw JE, Sicree RA, Zimmet PZ (2010) Global estimates of the prevalence of diabetes for 2010 and 2030. Diabetes Res Clin Pract 87:4–14
2. Gillies CL, Abrams KR, Lambert PC et al (2007) Pharmacological and lifestyle interventions to prevent or delay type 2 diabetes in people with impaired glucose tolerance: systematic review and meta-analysis. Br Med J 334:299
3. Pittas AG, Lau J, Hu FB, Dawson-Hughes B (2007) The role of vitamin D and calcium in type 2 diabetes. A systematic review and meta-analysis. J Clin Endocrinol Metab 92:2017–2029
4. Pittas AG, Sun Q, Manson JE, Dawson-Hughes B, Hu FB (2010) Plasma 25-hydroxyvitamin D concentration and risk of incident type 2 diabetes in women. Diabetes Care 33:2021–2023
5. Anderson JL, May HT, Horne BD et al (2010) Relation of vitamin D deficiency to cardiovascular risk factors, disease status, and incident events in a general healthcare population. Am J Cardiol 106:963–968
6. Pittas AG, Dawson-Hughes B, Li T et al (2006) Vitamin D and calcium intake in relation to type 2 diabetes in women. Diabetes Care 29:650–656
7. Liu S, Song Y, Ford ES, Manson JE, Buring JE, Ridker PM (2005) Dietary calcium, vitamin D, and the prevalence of metabolic syndrome in middle-aged and older US women. Diabetes Care 28:2926–2932
8. Holick MF (2007) Vitamin D deficiency. N Engl J Med 357:266–281
9. Mattila C, Knekt P, Mannisto S et al (2007) Serum 25-hydroxyvitamin D concentration and subsequent risk of type 2 diabetes. Diabetes Care 30:2569–2570
10. Knekt P, Laaksonen M, Mattila C et al (2008) Serum vitamin D and subsequent occurrence of type 2 diabetes. Epidemiology 19:666–671
11. Grimnes G, Emaus N, Joakimsen RM et al (2010) Baseline serum 25-hydroxyvitamin D concentrations in the Tromso Study 1994–95 and risk of developing type 2 diabetes mellitus during 11 years of follow-up. Diabet Med 27:1107–1115
12. Robinson JG, Manson JE, Larson J et al (2011) Lack of association between 25(OH)D levels and incident type 2 diabetes in older women. Diabetes Care 34:628–634
13. Pittas AG, Chung M, Trikalinos T et al (2010) Systematic review: vitamin D and cardiometabolic outcomes. Ann Intern Med 152:307–314
14. Mitri J, Muraru MD, Pittas AG (2011) Vitamin D and type 2 diabetes: a systematic review. Eur J Clin Nutr 65:1005–1015
15. de Boer IH, Tinker LF, Connelly S et al (2008) Calcium plus vitamin D supplementation and the risk of incident diabetes in the Women's Health Initiative. Diabetes Care 31:701–707
16. Avenell A, Cook JA, MacLennan GS, McPherson GC (2009) Vitamin D supplementation and type 2 diabetes: a substudy of a randomised placebo-controlled trial in older people (RECORD trial, ISRCTN 51647438). Age Ageing 38:606–609
17. Day N, Oakes S, Luben R et al (1999) EPIC-Norfolk: study design and characteristics of the cohort European Prospective Investigation of Cancer. Br J Cancer 80(Suppl 1):95–103

18. Forouhi NG, Luan J, Hennings S, Wareham NJ (2007) Incidence of Type 2 diabetes in England and its association with baseline impaired fasting glucose: the Ely study 1990–2000. Diabet Med 24:200–207
19. Forouhi NG, Luan J, Cooper A, Boucher BJ, Wareham NJ (2008) Baseline serum 25-hydroxy vitamin D is predictive of future glycemic status and insulin resistance: the Medical Research Council Ely Prospective Study 1990–2000. Diabetes 57:2619–2625
20. Danesh J, Collins R, Appleby P, Peto R (1998) Association of fibrinogen, C-reactive protein, albumin, or leukocyte count with coronary heart disease: meta-analyses of prospective studies. JAMA 279:1477–1482
21. Onland-Moret NC, van der AD, van der Schouw YT et al (2007) Analysis of case-cohort data: a comparison of different methods. J Clin Epidemiol 60:350–355
22. Barlow WE, Ichikawa L, Rosner D, Izumi S (1999) Analysis of case-cohort designs. J Clin Epidemiol 52:1165–1172
23. Wareham NJ, Jakes RW, Rennie KL et al (2003) Validity and repeatability of a simple index derived from the short physical activity questionnaire used in the European Prospective Investigation into Cancer and Nutrition (EPIC) study. Public Health Nutr 6:407–413
24. Ross AC, Manson JE, Abrams SA et al (2011) The 2011 report on dietary reference intakes for calcium and vitamin D from the institute of medicine: what clinicians need to know. J Clin Endocrinol Metab 96:53–58
25. Dersimonian R, Laird N (1986) Meta-analysis in clinical trials. Control Clin Trials 7:177–188
26. Higgins JP, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. Br Med J 327:557–560
27. Begg CB, Mazumdar M (1994) Operating characteristics of a rank correlation test for publication bias. Biometrics 50:1088–1101
28. Doro P, Grant WB, Benko R, Matuz M, Toth T, Soos G (2008) Vitamin D and the seasonality of type 2 diabetes. Med Hypotheses 71:317–318
29. Liu E, Meigs JB, Pittas AG et al (2010) Predicted 25-hydroxyvitamin D score and incident type 2 diabetes in the Framingham Offspring Study. Am J Clin Nutr 91:1627–1633
30. World Health Organization (2011) Use of glycated haemoglobin (HbA_{1c}) in the diagnosis of diabetes mellitus. WHO, Geneva
31. Tai K, Need AG, Horowitz M, Chapman IM (2008) Vitamin D, glucose, insulin, and insulin sensitivity. Nutrition 24:279–285
32. Palomer X, Gonzalez-Clemente JM, Blanco-Vaca F, Mauricio D (2008) Role of vitamin D in the pathogenesis of type 2 diabetes mellitus. Diabetes Obes Metab 10:185–197
33. Wang TJ, Zhang F, Richards JB et al (2010) Common genetic determinants of vitamin D insufficiency: a genome-wide association study. Lancet 376:180–188
34. Lawlor DA, Harbord RM, Sterne JA, Timpson N, Davey SG (2008) Mendelian randomization: using genes as instruments for making causal inferences in epidemiology. Stat Med 27:1133–1163
35. Gagnon C, Lu ZX, Magliano DJ et al (2011) Serum 25-hydroxyvitamin D, calcium intake, and risk of type 2 diabetes after 5 years: results from a national, population-based prospective study (the Australian Diabetes, Obesity and Lifestyle study). Diabetes Care 34:1133–1138
36. Thorand B, Zierer A, Huth C et al (2011) Effect of serum 25-hydroxyvitamin D on risk for type 2 diabetes may be partially mediated by subclinical inflammation: results from the MONICA/KORA Augsburg study. Diabetes Care 34:2320–2322
37. González-Molero I, Rojo-Martinez G, Morcillo S et al (2012) Vitamin D and incidence of diabetes: a prospective cohort study. Clin Nutr. doi:10.1016/j.clnu.2011.12.001