Vitamin D supplementation
and risk of diabetes
in patients at risk for type 2 diabetes

Protocol for a
Systematic Review and
Meta-analysis of
Individual Participant Data
From Randomized Controlled Trials

PROSPERO registration number: PENDING

A consortium among the principal investigators of the eligible randomized controlled trials that contributed individual-participant data (IPD) in the meta-analysis will be established.

Consortium members

Principal investigators of eligible trials that provided IPD

- To be determined

Staff / collaborators

- Anastassios G. Pittas, MD MS (Principal Investigator)
  apittas@tuftsmedicalcenter.org

- Edith Angellotti, MD (Research Associate)
  eangellotti@tuftsmedicalcenter.org
- Bess Dawson-Hughes, MD (Co-investigator)
  Bess.dawson-hughes@tufts.edu
- Jason Nelson, MPH (Data Analyst)
  Jnelson2@tuftsmedicalcenter.org
- Ellen Vickery, MS (Data Manager)
  evickery@tuftsmedicalcenter.org

Brown University Center for Evidence Synthesis in Health collaborators

- Ethan M. Balk, MD MPH (Co-investigator)
  ethan_balk@brown.edu
- Thomas A. Trikalinos, MD PhD (Co-investigator)
  thomas_trikalinos@brown.edu

INTRODUCTION

Rationale

Observational studies have long supported an association between low blood 25-hydroxyvitamin D levels and risk of type 2 diabetes. However, the question of whether vitamin D supplementation significantly lowers risk of diabetes has only recently been tested in randomized controlled trials (RCTs). Meta-analyses of RCTs have reported no significant effects of vitamin D supplementation on prevention of diabetes;[1,2] however, the included studies were heterogeneous in populations and types of vitamin D supplementation, had inadequate follow-up times for incident diabetes (e.g., <2 years) and small sample sizes, and were of low quality, thus reflecting the controversial application of meta-analyses of aggregate data in nutrition studies that has been suggested by Barnard et al.[3] Most importantly, none of the included trials were specifically designed and conducted to test the effect of vitamin D supplementation on incident diabetes as a primary outcome among people with prediabetes. Recently, RCTs have been published that were designed and conducted to specifically assess whether vitamin D supplementation reduces the rate of progression to diabetes in patients with prediabetes.[4-6] In each of these trials, which recruited participants not selected for vitamin D insufficiency, the risk of diabetes was lower in the vitamin D supplementation group than in the placebo group, but the observed differences were not statistically significant.[7-9] The hazard ratios of the effect of vitamin D on diabetes risk were nearly identical among the largest three trials, but the trials were each powered to detect a larger reduction in risk than was observed. Each individual trial was also underpowered to address additional clinical questions of interest, such as whether the effect of vitamin D supplementation differs based on one's baseline blood 25-hydroxyvitamin D level. Given the potential misleading interpretation of results from meta-analyses of aggregate data, the use of individual participant data (IPD) in meta-analysis has been strongly encouraged.[3] IPD meta-analysis may also allow investigation of secondary outcomes that the original RCTs were not adequately powered to address and may provide additional insights.

Objectives

The aim of this work is to conduct a systematic review and perform an IPD meta-analysis of RCTs of vitamin D supplementation versus placebo among patients with prediabetes to (a) derive a more precise hazard ratio of vitamin D supplementation on diabetes risk in patients with prediabetes and (b) increase the statistical power to allow additional analyses of interest (e.g., effect of vitamin D supplementation in subgroups of patients defined by the baseline 25-hydroxyvitamin D level).

Key Questions

Among patients with prediabetes:
- What is the efficacy of vitamin D supplementation to decrease the risk of new-onset diabetes, compared with placebo treatment?
- Does the efficacy of vitamin D supplementation to decrease the risk of new-onset diabetes differ across pre-specified subgroups of patients defined by baseline characteristics?
- What is the safety of vitamin D supplementation in this population?

METHODS

Overview of data collection and analysis
The study involves collection and analyses of de-identified data that have already been collected for the primary purpose (effect of vitamin D on incident diabetes) and all original trials had ethics approval. Each eligible trial will contribute de-identified data, i.e., staff conducting the analyses will not be able to trace data back to an individual participant. For example, age will be collected rather than date of birth.

Protocol Registration and Adherence to IPD guidelines
The systematic review and IPD meta-analysis has been registered with the PROSPERO International Prospective Register of Systematic Reviews (www.crd.york.ac.uk/PROSPERO; registration number: TBD). The protocol follows the general guidelines provided as part of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses)-IPD statement.[10]

Systematic review to identify eligible trials
For this IPD meta-analysis, a consortium among the principal investigators of the RCTs included will be established.

Eligibility criteria

1) Study Design: Randomized, double-blind, placebo-controlled trials.

2) Population of Interest: Adults (≥18 years old) with prediabetes at enrollment, as defined by each trial based on standard (e.g., ADA, WHO) glycemic criteria based on fasting glucose (FG), hemoglobin A1c (HbA1c) and/or glucose 2 hours after a 75-gram glucose load (2hG). Trials that include broader populations (e.g., adults without diabetes) that report subgroup analysis specific to those with prediabetes are eligible (only data from participants with prediabetes at randomization will be analyzed).

We will exclude trials with children, pregnant or lactating women, hospitalized patients (including long-term care facilities) and patients with end-stage renal disease, known diabetes (any type), or HIV at enrollment.

3) Intervention of Interest: Oral vitamin D supplementation with any formulation (e.g., ergocalciferol [vitamin D_2], cholecalciferol [vitamin D_3], eldecalcitol [a vitamin D analog]).

We will exclude trials using *(a)* non-oral supplementation of vitamin D (e.g., intramuscular); *(b)* vitamin D-fortified food (e.g., yogurt); *(c)* combination supplementation of vitamin D and calcium (or other intervention).

4) Duration of intervention and follow-up Given that progression to diabetes is a slow process, only trials with a planned duration of intervention and a follow-up of 2 years or longer will be included, i.e., trials of intervention *or* follow-up <2 years will be excluded.

5) Comparator: Placebo supplement containing no vitamin D. Trials without a placebo (e.g., comparator is "no supplementation" in a single-blind design) or with a placebo that had any amount of vitamin D will be excluded.

6) Outcomes: The primary outcome is new-onset diabetes as defined by each trial (glycemic criteria or a diagnosis outside of the study). The secondary outcome is safety (adverse events). Additional possible outcomes to be analyzed include regression from prediabetes to normoglycemia, glycemic measures, measures of insulin secretion and insulin resistance, and kidney function.

7) Results status: Study results must be available in manuscripts, as posted results in clinicaltrials.gov, from the study researchers, or in other forms.

Identification of studies and selection process
We will search the online databases Medline (through PubMed) and Embase, and the trial registry ClinicalTrials.gov. No date restriction will be applied, and the date of the last search will be as close to the manuscript submission as possible. We will present the full final electronic search strategies.

Researchers with clinical experience (all MDs) will perform a double, independent screening of abstracts and potentially relevant full text articles. Conflicts will be resolved by a third screener (and/or by group consensus). Detailed information regarding the number of studies identified, abstracts and full-text evaluated, and included/excluded articles will be provided. We will provide reasons for exclusion of articles retrieved in full text.

Data collection process
The principal investigator of each eligible trial willing to participate to the consortium IPD meta-analysis will agree to the terms of collaboration, as specified in a data sharing agreement.

For eligible trials for which we are unable to obtain relevant individual participant data, aggregate data will be extracted by one researcher with review and confirmation by a second researcher. Extracted data will be limited to data of interest (as listed below).

From each trial, we will collect data about population, intervention, comparator and outcomes, and information to assess risk of bias per the revised Cochrane risk of bias tool (RoB 2).[11] We will extract numerical information for the overall effect of supplementation (by intention-to-treat and "per protocol") and any eligible subgroup analyses. Data permitting, and depending on the granularity and type of reported analyses, we will conduct evidence syntheses (via hierarchical meta-analytic models) of: *(a)* unadjusted main treatment effect estimates (via typical meta-analytic approaches, *(b)* treatment-by-subgroup interactions (synthesizing the interaction effects only or, in a multivariate meta-analysis, both the main effect and the interaction effect, jointly). If necessary (i.e., if we are not able to obtain IPD for any eligible RCTs), we will explore the feasibility of combining IPD and aggregate data in the same analysis using hierarchical models with two levels. The first (within-study) level models patient data, and the second level (between-

studies) models the distribution of effects across studies. Trials with IPD will contribute information in both levels. Trials with aggregate data will only contribute information to the second level. The exact specification of the aforementioned analyses will depend on the available data.

Data items

The following **trial-level variables** will be collected:

- Name of trial and principal investigator
- Country in which study was carried out
- Year of trial completion (or publication data if completion date not reported)
- Recruitment setting (e.g., primary care, hospitals, specialist clinics)
- Eligibility criteria
- Details of intervention (e.g., vitamin D formulation, dosage, frequency of administration)
- Number of participants randomized
- Number of participants allocated to intervention (vitamin D) and placebo
- Trial pill adherence, as specified in each trial, overall and by group
- Expected relative risk of reduction in new-onset diabetes, per reported power analysis
- 25-hydroxyvitamin D assay method
- Risk of bias metrics per Cochrane RoB 2 tool[11]

Participant-level variables required for the analysis were agreed upon among the research team. A table with relevant IPD variables will be generated and shared among the trial principal investigators to ensure that data for these variables are available in each trial. Variables of potential interest that are not available in all eligible trials will not be requested (e.g., physical activity - see below).

The following **participant-level variables** will be collected, where available. During the conduct of the meta-analysis, additional variables of interest may be added and requested.

- At baseline (only)
  - Age (years)
  - Sex (female/male)
  - Race (Asian/Black/White/Other)
  - Smoking (yes-current, yes-former, never)
  - Body mass index
- At randomization (only)
  - Date of visit
  - Assigned treatment group (vitamin D or placebo)
- At baseline and follow-up visits
  - Date of visit
  - Vitamin D intake from supplements
  - Calcium intake from supplements
  - Serum or plasma 25-hydroxyvitamin D

- o Fasting (plasma or serum) glucose
- o 30-minute (plasma or serum) glucose
- o 2-hour (plasma or serum) glucose after a 75-gram glucose load (OGTT)
- o Hemoglobin A1c
- o Serum calcium
- o Serum creatinine (and calculation of estimated glomerular filtration rate [eGFR])
- o Fasting (plasma or serum) insulin and C-peptide
- o 30-minute (plasma or serum) insulin and C-peptide
- o 2-hour (plasma or serum) insulin and C-peptide
- During follow-up visits (at ≥2 years of follow-up)
  - o Diabetes (yes/no) and date
    - Method of diagnosis (intra-study with trial-specific glycemic criteria, outside of the study)
  - o Death (yes/no) and date
  - o Withdrawal from trial (yes/no) and date
  - o Date of last follow up encounter
  - o First use of diabetes medication (yes/no) and date
  - o First use of weight-loss medication (yes/no) and date
  - o Permanent discontinuation of trial pills (yes/no) and date
  - o Adverse events: hypercalciuria, hypercalcemia, kidney stones (yes/no)

The following variables will *not* be collected
- o *Ethnicity*
  - This variable is assessed in different ways in each trial (e.g., Hispanic/not Hispanic; Japanese) and is likely not to be available in all trials. It is mostly relevant to US-based trials and it is not relevant to the primary aim of this IPD MA.
- o *Family history of diabetes*
  - This variable may not be available in all trials. It is not relevant to the primary aim of this IPD MA.
- o *Physical activity*
  - This variable may not be available in all trials. If available, we would need to reconcile different assessment methods and different units of measurement (MET hours/week, hours/weeks, hours/day). It is not relevant to the primary aim of this IPD MA.
- o *Vitamin D intake from food*
  - This variable may not be available in all trials. It is not relevant to the primary aim of this IPD MA and the information is better captured by the blood 25-hydroxyvitamin D level.
- o *Calcium intake from food*
  - This variable may not be available in all trials.

IPD integrity

In preparation for receiving the data, data request forms with the relevant variables will be created and shared with the principal investigators of eligible trials. If applicable, efforts will be made to standardize the names and options of each variable before the data transfer (e.g., age in years). Data will be de-identified at source before being transferred to Tufts Medical Center via secure email. After receiving the data in the appropriate statistical format, the data analyst and at least one other member of the research team will assess the integrity of the data by performing internal consistency checks. Trial investigators will be contacted to provide missing data and to respond to queries deriving from this integrity check. Clean data will then be uploaded to the IPD meta-analysis database.

Data Harmonization

The IPD approach ensures an improved quantity and quality of data compared to an aggregate data meta-analysis, and enables standardization of outcomes across trials and detailed data checking.[10] Moreover, the IPD approach allows researchers to derive standardized classifications of participant characteristics or their disease/condition or translate different definitions to a common scale.

Below is one example of how we plan to harmonize data among trials:
- Smoking: this variable is often collected differently across studies. Some studies assess "general" smoking status (e.g. options never, former, current) and others assess "current" smoking status (yes/no). For this IPD meta-analysis, we will collect the raw data as assessed in each study and we will then standardize data to examine the "current" smoking status during the study period with yes/no options in the meta-analysis database.

Quality Assessment (risk of bias). Methodological quality of each study will be assessed using the Cochrane revised tool (RoB 2) for assessing risk of bias in RCTs (adequate sequence generation, allocation concealment, etc.).[11] Risk of bias will be independently assessed by two researchers. Discrepancies will be solved by a third reviewer (and/or by group consensus).

Specification of outcomes
- **Primary outcome:** new-onset diabetes, as defined by each trial (glycemic criteria or a diagnosis outside of the study).
  - Intention-to-treat ("treatment policy estimand").
    - In the entire cohort (including sensitivity analysis that includes only diabetes cases according to trial-specific criteria – see below for definition).
    - In subgroups defined by key baseline criteria - see below for subgroups.
  - Per-protocol ("trial product estimand") analyses - see below for definition.
    - In the entire cohort
- **Secondary outcome: safety**
  - Specific adverse events of interest: new-onset hypercalcemia, hypercalciuria, kidney stones.
    - In the entire cohort

- o **Additional outcomes.** These outcomes will be included as time, resources, and available data permit.
  - o Regression from prediabetes to normoglycemia [time to event analysis].
  - o Change in individual glycemic variables: HbA1c, FG, 2hG.
  - o Change in insulin (and/or C-peptide) based indices of insulin resistance and secretion, e.g. homeostasis model assessment estimates based on fasting values.
  - o To identify subpopulations most likely to benefit, we will conduct heterogeneity of treatment effect analyses for the primary outcome of new-onset diabetes.[12]
  - o Change in kidney function (e.g., eGFR).

Data Synthesis/ Statistical analyses plan

After combining IPD, *intention-to-treat* analyses will compare groups defined by the randomization procedure and include all participants irrespective of adherence to assigned treatment or the protocol (e.g., adherence to treatment assignment, use of diabetes or weight-loss medications, use of out-of-study high-dose vitamin D). Follow-up time for all analyses will be calculated as time from randomization until the occurrence of the primary outcome, death, withdrawal, or last follow-up encounter free of diabetes. Data for the longest available follow-up time for each trial will be used in the analysis.

Kaplan-Meier estimates will be plotted for each group (vitamin D vs. placebo). Cox proportional hazard models will be used to compare the hazard rate of incident diabetes between the two groups. The model will include group assignment as its main predictor variable, and will also include the RCT. Comparisons between the groups for key baseline characteristics, rates of participant withdrawals, permanent discontinuation of trial pills, use of diabetes or weight-loss medications, and supplemental intake above the trial limit will use the Fisher exact test, the chi-square test, the Wilcoxon rank-sum test, or the pooled variance estimate t test.

Sensitivity analyses: In some trials, the diagnosis of diabetes may be established outside of the study. Given that the diagnosis of diabetes by study procedures (e.g., screening for diabetes at regular intervals, use of common laboratory criteria) is a robust, unbiased way compared to diagnosis outside of the study which depends on many random, uncontrolled factors, we will run a sensitivity analysis where the primary outcome is defined as new-onset diabetes according to trial-specific glycemic criteria only, i.e., diagnoses of diabetes made outside of the study will not be included and follow-up of these participants will be censored when diabetes was diagnosed outside of the study.

Variability of response to vitamin D supplementation on new-onset diabetes will be assessed in **pre-specified subgroups** based on the following key baseline variables for the combined cohort (i.e., all participants included):
- Blood 25-hydroxyvitamin D level, ng/mL
  - o <12 vs. 12-19 vs. ≥20 ng/mL
- BMI, kg/m^2
  - o <30 vs. 30-34.9 vs. ≥35

- - < median vs. ≥ median
- Age, years
  - < median vs. ≥ median
- Sex
  - Male vs. Female
- Race
  - White vs. Black vs. Other (non-White, non-Black)
- Glycemic risk
  - Higher risk (met all three [FG, 2hG, HbA1c] glycemic criteria for prediabetes) vs. lower risk (met one or two criteria for prediabetes)
- Total calcium intake (supplements), mg/day
  - <600 vs. ≥600
  - <1200 vs. ≥1200

Each subgroup analysis will include a test for interaction, and effect modification will be claimed if the test for interaction reaches statistical significance. No adjustments will be made for multiple comparisons. Whenever more than one threshold is examined, we will rank the fit of models using different thresholds using informational criteria, namely the AIC. We will consider that a difference of at least 4 points in the AIC indicates better fit.

We will conduct a **_pre-specified per-protocol analysis_** that will censor follow-up when a participant stopped trial pills, started a diabetes or weight-loss medication, or took out-of-study vitamin D from supplements above study limitation (as done in the publication by Pittas et al NEJM 2019[8]). This analysis provides a "trial product estimand" which evaluates the effect of the intervention (vitamin D supplementation) under the assumption that all participants continued taking the trial medication and did not use rescue medication (diabetes/weight-loss medication or high-dose vitamin D).

Rates of adverse events relevant to vitamin D (e.g., hypercalcemia, hypercalciuria, kidney stones) will be compared between the two groups.

Among trials that measured blood 25-hydroxyvitamin D throughout the follow-up, we will also conduct analyses to test the association between cumulative exposure to vitamin D (assessed by mean cumulative 25-hydroxyvitain D level during follow-up, regardless of how vitamin D was obtained) and incident diabetes.

References

1. Seida JC, Mitri J, Colmers IN, et al. Clinical review: Effect of vitamin D3 supplementation on improving glucose homeostasis and preventing diabetes: a systematic review and meta-analysis. J Clin Endocrinol Metab 2014;99:3551-60.
2. Tang H, Li D, Li Y, Zhang X, Song Y, Li X. Effects of Vitamin D Supplementation on Glucose and Insulin Homeostasis and Incident Diabetes among Nondiabetic Adults: A Meta-Analysis of Randomized Controlled Trials. Int J Endocrinol 2018;2018:7908764.
3. Barnard ND, Willett WC, Ding EL. The Misuse of Meta-analysis in Nutrition Research. JAMA 2017;318:1435-6.
4. Sollid ST, Hutchinson MY, Fuskevag OM, et al. No effect of high-dose vitamin D supplementation on glycemic status or cardiovascular risk factors in subjects with prediabetes. Diabetes Care 2014;37:2123-31.
5. Pittas AG, Dawson-Hughes B, Sheehan PR, et al. Rationale and design of the Vitamin D and Type 2 Diabetes (D2d) study: a diabetes prevention trial. Diabetes Care 2014;37:3227-34.
6. Kawahara T, Suzuki G, Inazu T, et al. Rationale and design of Diabetes Prevention with active Vitamin D (DPVD): a randomised, double-blind, placebo-controlled study. BMJ open 2016;6:e011183.
7. Jorde R, Sollid ST, Svartberg J, et al. Vitamin D 20,000 IU per Week for Five Years does not Prevent Progression from Prediabetes to Diabetes. J Clin Endocrinol Metab 2016:jc20154013.
8. Pittas AG, Dawson-Hughes B, Sheehan P, et al. Vitamin D Supplementation and Prevention of Type 2 Diabetes. N Engl J Med 2019.
9. Kawahara T. Eldecalcitol, a Vitamin D analogue, for diabetes prevention in impaired glucose tolerance: DPVD study. 2018.
10. Tierney JF, Vale C, Riley R, et al. Individual Participant Data (IPD) Meta-analyses of Randomised Controlled Trials: Guidance on Their Use. PLoS Med 2015;12:e1001855.
11. Sterne JAC, Savovic J, Page MJ, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ (Clinical research ed 2019;366:l4898.
12. Kent DM, Paulus JK, van Klaveren D, et al. The Predictive Approaches to Treatment effect Heterogeneity (PATH) Statement. Ann Intern Med 2019.