Working Title	The effect of supplementing older adults with vitamin D on all-cause mortality: the D-Health Trial		
Date of plan	17 March 2021		
Person conducting analysis	Mary Waterhouse	Email	mary.waterhouse@qimrberghofer.edu.au
Authors	Rachel E. Neale, Catherine Baxter, Briony Duarte Romero, Donald S. A. McLeod, Dallas R. English, Bruce K. Armstrong, Peter R. Ebeling, Gunter Hartel, Michael G. Kimlin, Rachel L. O'Connell, Jolieke C. van der Pols, Alison J. Venn, Penelope M. Webb, David C. Whiteman, Mary Waterhouse		
Journal Options	NEJM; The Lancet; JAMA; Ann Intern Med; BMJ		
Background and overview			
Background	Many observational studies have reported inverse associations between serum 25-hydroxy vitamin D (25(OH)D) concentration (used as a measure of vitamin D status) and risk of mortality.[1, 2] However, findings from observational studies may be a result of reverse causality or uncontrolled confounding, so cannot be used to reliably infer that increasing 25(OH)D concentrations in the population (through supplementation or food fortification) would be beneficial. Meta- and pooled analyses of trials have found that vitamin D supplementation results in a small decrease in all-cause mortality,[1, 3, 4] but the generalizability of these results is uncertain. Many studies were designed to assess falls and bone health, so included primarily women, and a substantial proportion of participants lived in aged care facilities. Mortality was usually a secondary outcome. The D-Health trial aims to determine if monthly high-dose vitamin D supplementation of the general older population can reduce mortality.[5]		
Aims	The <u>primary aim</u> of this analysis is to determine whether randomisation to monthly supplementation with 60,000 international units (IU) of vitamin D for up to five years reduces all-cause mortality. <u>Secondary aims</u> are to investigate the effect of randomisation to vitamin D supplementation on: 1. All-cause mortality within subgroups of: age (<70, ≥70 years); sex; body mass index (BMI) (<25, ≥25 kg/m^2); and predicted deseasonalised baseline serum 25(OH)D concentration (<50, ≥50 nmol/L); 2. All-cause mortality, taking into account non-compliance (with study tablets) and contamination (with off-trial supplementation) (i.e. per protocol analysis); and 3. Mortality from cancer, cardiovascular disease, and other causes.		

Outcomes and instrument	
Outcomes	The primary outcome is death from any cause. Follow-up will begin at date of randomisation and end at the date of death or censoring, whichever comes first; see Appendix 1. The secondary outcome is death from cancer, cardiovascular disease, and other causes. We will classify the death as being due to: • Cancer if the underlying cause of death (UCOD)[1] is a malignant neoplasm (ICD-10 codes C00-C97); • Cardiovascular disease if the UCOD is any disease of the circulatory system (ICD-10 codes I00-I99); • Other causes if the UCOD is known and is neither a malignant neoplasm nor a disease of the circulatory system.[1] [1] UCOD is defined as "The disease or injury which initiated the train of morbid events leading directly to death." [2] This category includes in situ neoplasms, benign neoplasms, and neoplasms of uncertain or unknown behaviour.
Instrument	**Mortality** Mortality will be ascertained primarily via linkage to state death registries. Since death registries use probabilistic linkage, and do not capture deaths that occur outside Australia, and because we do not have death registry data from the Australian Capital Territory, there will be some participants who died but do not appear in the registry data. In cases where we are confident[1] that a participant died (but they have no registry data[2]) and their death was ascertained prior to the unblinding of participants, we will count their death as an event and use the date of death from our records; there are fewer than 50 (currently n=46) such cases. If the death was reported to us after participants were informed of their tablet allocation, we will censor the participant at the date of death (i.e. the death will not be counted). **Underlying cause of death** We will ascertain UCOD via linkage to state death registries. The UCOD for those deaths not ascertained through linkage will be classified as missing. [1] The project manager and chief investigator will review the source of death information in the D-Health database to reach consensus about the reliability of the information. If consensus is not reached the death will not be included and the date last known alive will be imputed as for all other participants. The project manager and chief investigator will have no knowledge of randomisation allocation when making this decision. [2] Because either the death: (1) was not identified by probabilistic linkage; or (2) occurred outside Australia; or (3) occurred within the Australian Capital Territory; or (4) was not registered; or (5) occurred after the linkage date for the state in which the participant lived.

Documentation	
Analysis packages	SAS 9.4 R version 4.0.3 Stata 15.0
Dataset	*The location of the original dataset provided, and the final dataset containing any constructed variables* **Original datasets:** L:\Lab_RachelN\DHealthDataAnalysis\Projects\Mortality\Data\original data\ **Final datasets:** • L:\Lab_RachelN\DHealthDataAnalysis\Projects\Mortality\Data\final data\ • R:\
Statistical code	*The location of the statistical code used to produce all results. (Please ensure code is adequately documented to enable analyses to be reproduced. In addition to documenting code, please produce a document that includes names of relevant analysis files, along with comprehensive description of their purpose. If on site, please include links).* Code used to prepare original datasets: • L:\Lab_RachelN\DHealthDataAnalysis\SASCode\CoreDataManipulation\ProjectSetup\Mortality\Mortality_prepareData.sas (determines date of event or censoring) • L:\Joint_Projects\DHealth_StatsGroup\mortality\code\ o mortality_createScrambledDatasets.sas: will create "dummy" datasets (i.e. participants have been randomised to two new groups that have no relationship to the true treatment allocation) that are used to develop/debug code. o mortality_createFinalDataset.sas: will create blinded dataset that includes all participants Code used to prepare and analyse data will be saved here: • L:\Lab_RachelN\DHealthDataAnalysis\Projects\Mortality\code\ • R:\
Participants and data	
Participants and eligibility	We will include all randomised participants except the five people who withdrew their consent and requested that their data be deleted from our database (i.e. the number analysed will be n=21,310).
Exposure variable(s)	Randomisation group
Covariates	<u>Randomisation stratification variables:</u> • Age: 60-64; 65-69; 70-74; 75+ • Sex: F; M

	• State: NSW; QLD; SA; TAS; VIC; WA <u>Variables that will be used to derive interaction terms:</u> • Sex (male vs female) • Age at randomisation (<70 years vs ≥70 years) • BMI at randomisation (<25 vs ≥25 kg/m^2) • Predicted baseline 25(OH)D concentration (<50 vs ≥50 nmol/L) Predicted baseline 25(OH)D concentration has been generated for all participants.[6] <u>Variables used to derive inverse probability weights:</u> These will be baseline and post-randomisation prognostic variables that influence the probability of non-adherence to the trial protocol and mortality. Refer to the Supplementary Analysis section for more details.
Data cleaning	<u>Date of death</u> • If a state provides only the month and year of death, we will set the date of death to be the approximate middle of the specified month (i.e. the 15th day for February, April, June, September or November, and the 16th day for all remaining months). <u>Underlying cause of death</u> • If a death was ascertained via linkage, and the ICD-10 code for UCOD is missing, and a non-codified (text) cause(s) of death has been provided, we will use the latter to classify the UCOD as either: (1) cancer (i.e. malignant neoplasm); (2) cardiovascular disease; (3) other – external cause; (4) other – non-external cause; or (5) unable to be coded. o We will use deaths that have both an ICD-10 code and a text description to develop a sense of how text relates to UCOD coding. o In cases where the UCOD is not obvious from the text description, three investigators (RN, MW, DM) will attempt to reach consensus on the UCOD. If this cannot be achieved, the UCOD will be classified as "unable to be coded". o If a death was ascertained via linkage and it has neither a code nor text (for UCOD), or the death was ascertained via our records, the UCOD will be classified as missing.
Handling missing data	We will not impute any data. <u>Intention-to-treat analyses</u> Of the covariates used in the intention-to-treat analysis, BMI is the only one for which there is missing data (n=119 (0.6%)). Participants with missing BMI will be excluded from analyses of interaction between randomisation group and BMI.

	Per protocol analysis Some baseline covariates that will be used to derive inverse probability weights have missing data (BMI (n=119), self-rated overall health (n=345), ever smoker (n=185), highest qualification obtained (n=263)). These covariates will be included as factors in the weight determining models and missing values will be coded as 'unknown'.
Maintaining blinding	1. Code will be written and debugged using data in which people have been randomly assigned to two new groups so that there is no relationship between the new groups and the true treatment allocation. Development and testing of code will include producing all results, including tables and figures, as they will appear in the manuscript. Randomisation group allocation will not be provided until all investigators have approved the tables and figures. We will conduct the per protocol analysis (secondary aim #2) after unblinding.
Analysis details	
Analysis	• Participant flow will be shown using a CONSORT diagram (**Figure 1**). • We will present selected baseline characteristics according to randomisation group (**Table 1**). • We will report the number and percentage of deaths within each randomisation group, displaying how many of these of deaths were not ascertained using probabilistic linkage. • We will plot cumulative incidence of all-cause mortality for each randomisation group, using Kaplan-Meier methods (**Figure 2**). • For our primary intention-to-treat analysis of the effect of vitamin D supplementation on all-cause mortality, we will fit two flexible parametric survival models (FPSMs).[7, 8] In both FPSMs, we will model the baseline log cumulative hazard function using a restricted cubic spline with two internal knots (placed at the 33rd and 67th percentiles of the uncensored log survival times). Both models will include randomisation group and the randomisation stratification variables of age, sex, and state of residence at baseline. Model 1 will not include any time-varying coefficients; it will be used to estimate an "overall" hazard ratio (HR) and 95% confidence interval (CI); the HR (95% CI) will be embedded in **Figure 2**. Model 2 will include an interaction between randomisation group and a restricted cubic spline with one internal knot (placed at the median of uncensored log survival times), thereby allowing the HR for randomisation group to vary with time. • We will report the p-value from the likelihood ratio test comparing Models 1 and 2 (i.e. testing the effect of including the interaction between time and randomisation group). • Using Model 2, we will plot the following as functions of time since randomisation: (a) the estimated HR (**Figure 3a**); and (b) the difference in standardised survival curves (**Figure 3b**). The plots will include 95% CIs. We will also report the estimated HR (95% CI) and difference in

standardised survival (95% CI) at 2, 4 and 6 years post-randomisation (**Table 2**). We will show the standardised survival curves for each randomisation group in the Supplementary Material (**SFigure 1**).

Interactions and subgroup analyses

- We will investigate whether the effect of supplementation is modified by the following baseline characteristics:
    - Age (< 70 years, ≥ 70 years);
    - Sex (men, women);
    - BMI (< 25 kg/m^2, ≥ 25 kg/m^2); and
    - Predicted deseasonalised 25(OH)D concentration (< 50 nmol/L, ≥ 50 nmol/L).

For each baseline characteristic listed above we will do the following:

- Use Kaplan-Meier methods to plot cumulative incidence of all-cause mortality for each randomisation group within each stratum of the characteristic (**SFigures 2-5**).
- We will fit a FPSM that includes randomisation group, the baseline characteristic of interest, an interaction term between randomisation group and the baseline characteristic of interest, age, sex and state of residence at baseline. We will model the baseline log cumulative hazard function using a restricted cubic spline with two internal knots. Using this model, we will report estimated HRs (95% CI) for each level of the baseline characteristic (**Figure 4;** estimates will also be embedded in the relevant supplementary figure of cumulative incidence). We will report the p-value from the likelihood ratio test comparing models with and without the interaction term.
- We will refit the FPSM from the previous step, this time including an interaction between randomisation group and a restricted cubic spline with one internal knot. We will report the p-value from the likelihood ratio test comparing models with and without the interaction term between randomisation group and: the baseline characteristic of interest; and time since randomisation. For each level of the baseline characteristic, we will plot the following as functions of time since randomisation: (a) the estimated HR; and (b) the difference in standardised survival curves (**SFigures 6-9**).

Sensitivity analysis (Predicted deseasonalised 25(OH)D concentration)

- We will repeat the interaction and subgroup analysis for predicted deseasonalised 25(OH)D concentration using a cut-off of 60 nmol/L. The results will be presented in the supplementary material (**SFigures 10** and **11**); they will not appear in Figure 4.
- The motivation for this sensitivity analysis is that the assay that we used to measure 25(OH)D concentrations (in a subset of participants) tends to run a bit high, and these measured 25(OH)D concentrations were used to develop our prediction models.

	Potential exploratory analysis • If we find an interaction between randomisation group and BMI and there is an effect in the subgroup with BMI < 25 kg/m^2, we may conduct an exploratory analysis after unblinding where we exclude people with BMI < 18.5 kg/m^2. Any results would be included in the Supplementary Material. **Supplementary modelling of hazard ratios** • To enable comparisons with other studies, in particular VITAL, we will estimate HRs (95% CIs) using Cox proportional hazards models. We will produce estimates for all participants and according to age, sex, BMI, and predicted 25(OH)D concentration (< 50 nmol/L, ≥ 50 nmol/L) (**SFigure 12**). All models will include randomisation code, and age, sex, and state of residence at baseline. Interactions between randomisation group and each selected baseline characteristic will be assessed using the likelihood ratio test. **Underlying cause of death** • We will describe UCOD (**STable 1**). • We will perform a completing risks analysis of deaths due to: (1) cancer; (2) cardiovascular disease; and (3) other causes (excluding "ill-defined or unknown" and "unable to be coded"). We will use a flexible parametric modelling approach for direct likelihood inference on the cause-specific cumulative incidence function.[9, 10] We will fit a non-proportional log-cumulative subdistribution hazards flexible parametric model. Deaths where the UCOD is "ill-defined or unknown", "unable to be coded" or missing will be censored at the date of death. We will plot the cause-specific cumulative incidence functions (**SFigure 13**). **Retention, compliance, and adverse events** • We will describe retention and compliance with study tablets (**STable 2**), and intake of off-trial supplementary vitamin D (**STable 3**) according to randomisation group. • We will describe the distribution of measured 25(OH)D concentrations and plot the distribution by year of blood draw according to randomisation group (**SFigure 14**). • We will report the number of people who experienced at least one adverse event according to randomisation group. Adverse events will analysed using unadjusted Poisson regression (**Table 3**).
Supplementary analysis	**Per-protocol analysis (conducted after unblinding)** We will use inverse probability weighting to estimate the effect of vitamin D supplementation on mortality over a six year period that would have been observed had all participants adhered to the protocol.[11-13] Adherence will be defined as taking ≥ 80% of the study tablets overall *and* not taking more than 500 IU/day of off-study vitamin D supplementation at any time during the intervention period.

For this analysis, we will use discretised interval data (annual intervals), and we will use 6 years of follow-up time from randomisation. Follow-up will begin at date of randomisation and end at the date of death, censoring, or 6 years after randomisation, whichever comes first.

Notation

- Let i index the i^{th} participant,
- Let j denote years from randomisation (with $j = 0$ corresponding to time of randomisation).
- Let $C_{ij} = 1$ indicate that participant i was censored at or before year j, 0 otherwise.

Censoring

- We will artificially censor a participant if and when they first stopped adhering to the protocol (unless they withdrew for a clinical reason [14]). See Appendix 2 for how we will estimate the date a participant was first non-compliant with study tablets (for those who took < 80% overall).

Determining inverse probability weights

- We will generate time-varying stabilised weights for each individual (W_{ij}):
- $$W_{ij} = \begin{cases} \prod_{k=0}^{j} \frac{P(C_{ik} = 0|\text{baseline covariates only})}{P(C_{ik} = 0|\text{baseline and time-varying covariates})}, & \text{if } C_{ij}=0 \\ 0, & \text{if } C_{ij}=1 \end{cases}$$
-
- We will use pooled logistic regression models to estimate the probabilities in the numerator and denominator. We will fit separate models for each randomisation group.[12]
- Baseline covariates will include:
    - Age (60-64, 65-69, 70-74, ≥ 75 years)
    - Sex (men, women)
    - BMI (< 18.5, 18.5 to < 25, 25 to < 30, ≥ 30 kg/m^2, unknown)
    - Predicted baseline 25(OH)D concentration (<50, ≥50 nmol/L)
    - Self-rated overall health (excellent or very good, good, fair or poor, unknown)
    - Ever smoker (no, yes, unknown)
    - Highest qualification obtained (none, school or intermediate certificate, higher school or leaving certificate, apprenticeship or certificate, university degree or higher, unknown)
- Time-varying covariates will be fixed at their values at the start of each period, and will include:
    - Cumulative number of serious adverse events since randomisation (from ad hoc reporting of adverse events)

	○ Compliance to study tablets in the previous year (none, 1-3, 4-6, 7-9, 10-12) ○ Self-reported hospitalisation for cancer in the previous year (yes, no) ○ Self-reported incidence of any of the following in the previous year: myocardial infarction, stroke, transient ischaemic attack, angioplasty, insertion of a coronary stent, or a blood clot (yes, no) • We will check the distribution of weights for extreme values. If necessary, we will modify the weight-determining model. • Once satisfied with the weight-determining model, we will truncate weights at 99th percentile to prevent outliers having excessive influence. Weighted outcome model • We will use weighted pooled logistic regression to estimate the HR for all-cause mortality, artificially censoring participants who did not adhere to the protocol (as described above). • The model will include randomisation group, and the baseline covariates included in the weight determining model. • We will compute robust standard errors to estimate conservative CIs.
Planned main tables	**See Appendix 3** Table 1. Baseline characteristics according to randomisation group. Table 2. Effect of supplementation with vitamin D on all-cause mortality. Predicted difference in standardised survival and time-varying hazard ratio at 2, 4, and 6 years post-randomisation, and predicted overall hazard ratio. Table 3. Self-reported adverse events according to randomisation group.
Planned supplementary tables	**See Appendix 4** STable 1. Underlying cause of death according to randomisation group. STable 2. Retention, and compliance with study tablets according to randomisation group. STable 3. Intake of vitamin D from off-trial supplements according to randomisation group.
Planned main figures	**Appendix 5** shows selected main figures, created using dummy data. Figure 1. Participant flow for analyses of all-cause mortality (CONSORT flow diagram). Figure 2. Cumulative incidence of all-cause mortality according to time since randomisation in Group 1 (solid line) and Group 0 (dashed line). Figure 3. Effect of vitamin D supplementation on all-cause mortality. Panel A shows the time-varying hazard ratio and panel B shows the difference in standardised survival curves as a function of time since randomisation.

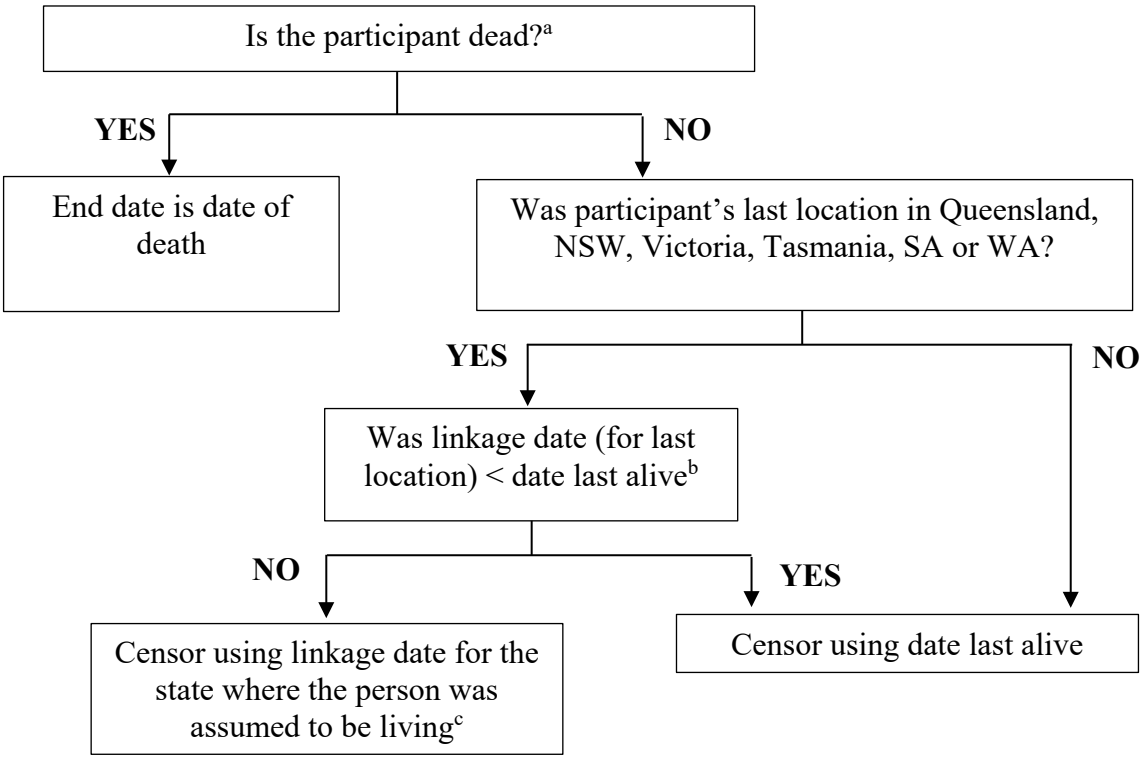
	Figure 4. Effect of vitamin D supplementation on all-cause mortality for all participants and by selected baseline characteristics; forest plot of the average hazard ratios.
Planned supplementary figures	**Appendix 6** shows selected supplementary figures, created using dummy data. SFigure 1. Standardised survival curves according to randomisation group; estimates from a flexible parametric survival model that included age, sex, and state of residence at baseline. Randomisation group was included as a time-varying covariate. SFigure 2. Cumulative incidence of all-cause mortality according to time since randomisation in Group 1 (solid line) and Group 0 (dashed line) within strata of age (< 70 years; ≥ 70 years). SFigure 3. Cumulative incidence of all-cause mortality according to time since randomisation in Group 1 (solid line) and Group 0 (dashed line) within strata of sex. SFigure 4. Cumulative incidence of all-cause mortality according to time since randomisation in Group 1 (solid line) and Group 0 (dashed line) within strata of body mass index (< 25 kg/m^2; ≥ 25 kg/m^2). SFigure 5. Cumulative incidence of all-cause mortality according to time since randomisation in Group 1 (solid line) and Group 0 (dashed line) within strata of predicted deseasonalised serum 25(OH)D concentration at baseline. SFigure 6. Effect of vitamin D supplementation on all-cause mortality according to age group (< 70 years (top row); ≥ 70 years (bottom row)). Plots on the left show the time-varying hazard ratio. Plots on the right show the difference in standardised survival curves as a function of time since randomisation. SFigure 7. Effect of vitamin D supplementation on all-cause mortality according to sex (men (top row); women (bottom row)). Plots on the left show the time-varying hazard ratio. Plots on the right show the difference in standardised survival curves as a function of time since randomisation. SFigure 8. Effect of vitamin D supplementation on all-cause mortality according to body mass index (< 25 kg/m2 (top row); ≥ 25 kg/m2 (bottom row)). Plots on the left show the time-varying hazard ratio. Plots on the right show the difference in standardised survival curves as a function of time since randomisation. SFigure 9. Effect of vitamin D supplementation on all-cause mortality according to predicted deseasonalised baseline serum 25(OH)D concentration (< 50 nmol/L (top row); ≥ 50 nmol/L (bottom row)). Plots on the left show the time-varying hazard ratio. Plots on the right show the difference in standardised survival curves as a function of time since randomisation. SFigure 10. Sensitivity analysis: cumulative incidence of all-cause mortality according to time since randomisation in Group 1 (solid line) and Group 0 (dashed line) within strata of predicted deseasonalised serum 25(OH)D concentration at baseline (< 60 nmol/L; ≥ 60 nmol/L).

	SFigure 11. Sensitivity analysis of the effect of vitamin D supplementation on all-cause mortality according to predicted deseasonalised baseline serum 25(OH)D concentration (< 60 nmol/L (top row); ≥ 60 nmol/L (bottom row)). Plots on the left show the time-varying hazard ratio. Plots on the right show the difference in standardised survival curves as a function of time since randomisation. SFigure 12. Effect of vitamin D supplementation on all-cause mortality for all participants and within strata of pre-specified baseline characteristics; hazard ratios estimated using Cox proportional hazards regression models adjusted for baseline age, sex and state of residence. SFigure 13. Predicted cause-specific cumulative incidence functions showing the effect of vitamin D supplementation on mortality from cancer, cardiovascular disease, and other causes. SFigure 14. Boxplots of serum 25(OH)D concentrations by year of blood draw according to randomisation group.

References

1. Autier, P., et al., *Vitamin D status and ill health: a systematic review.* Lancet Diabetes Endocrinol, 2014. **2**(1): p. 76-89.
2. Gaksch, M., et al., *Vitamin D and mortality: Individual participant data meta-analysis of standardized 25-hydroxyvitamin D in 26916 individuals from a European consortium.* PLoS One, 2017. **12**(2): p. e0170791.
3. Autier, P., et al., *Effect of vitamin D supplementation on non-skeletal disorders: a systematic review of meta-analyses and randomised trials.* Lancet Diabetes Endocrinol, 2017. **5**(12): p. 986-1004.
4. Bjelakovic, G., et al., *Vitamin D supplementation for prevention of mortality in adults.* Cochrane Database Syst Rev, 2014(1): p. Cd007470.
5. Neale, R.E., et al., *The D-Health Trial: A randomized trial of vitamin D for prevention of mortality and cancer.* Contemp Clin Trials, 2016. **48**: p. 83-90.
6. Waterhouse, M., et al., *Predicting deseasonalised serum 25 hydroxy vitamin D concentrations in the D-Health Trial: an analysis using boosted regression trees.* Contemp Clin Trials, 2020. **(in press)**.
7. Royston, P. and M.K. Parmar, *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): p. 2175-97.
8. Lambert, P.C. and P. Royston, *Further development of flexible parametric models for survival analysis.* The Stata Journal, 2009. **9**(2): p. 265-290.
9. Mozumder, S.I., M. Rutherford, and P. Lambert, *Direct likelihood inference on the cause-specific cumulative incidence function: A flexible parametric regression modelling approach.* Stat Med, 2018. **37**(1): p. 82-97.
10. Mozumder, S.I., M.J. Rutherford, and P.C. Lambert, *stpm2cr: A flexible parametric competing risks model using a direct likelihood approach for the cause-specific cumulative incidence function.* Stata J, 2017. **17**(2): p. 462-489.
11. Robins, J.M. and D.M. Finkelstein, *Correcting for noncompliance and dependent censoring in an AIDS Clinical Trial with inverse probability of censoring weighted (IPCW) log-rank tests.* Biometrics, 2000. **56**(3): p. 779-88.
12. Dodd, S., P. Williamson, and I.R. White, *Adjustment for treatment changes in epilepsy trials: A comparison of causal methods for time-to-event outcomes.* Stat Methods Med Res, 2017: p. 962280217735560.
13. Cain, L.E. and S.R. Cole, *Inverse probability-of-censoring weights for the correction of time-varying noncompliance in the effect of randomized highly active antiretroviral therapy on incident AIDS or death.* Stat Med, 2009. **28**(12): p. 1725-38.
14. Hernan, M.A. and J.M. Robins, *Per-Protocol Analyses of Pragmatic Trials.* N Engl J Med, 2017. **377**(14): p. 1391-1398.

Appendix 1. Determining the end date to use when calculating time to event



[a] Mortality will be ascertained primarily via linkage to State death registries. We will use data from ad hoc reporting of death if: (1) we are confident the participant died; (2) their death does not appear in the registry data; and (3) the participant's death was ascertained by D-Health staff prior to the unblinding of participants. Deaths ascertained by D-Health staff after unblinding of participants (that do not appear in registry data) will be censored.

[b] 'date last alive' is the last date at which we know a participant was alive, derived from multiple sources, including Medicare Benefits Scheme, Pharmaceutical Benefits Scheme, survey completion/return dates, diary completed/return dates (cold/flu and falls substudies), communication records, and blood collection dates.

[c] We will assume that when second linkages are performed participants (who are still alive) are living at the last state we have in our records.

Appendix 2. Estimating date participant first became non-compliant with study tablets

If a participant took < 80% of the study tablets overall, we will estimate the date at which they first became non-compliant with study tablets as follows:

1. Let k be the first year (i.e. annual survey) that compliance was missing or the person reported taking < 80% of the year's study tablets.
2. For simplicity, assume that the person's intake of study tablets in that year was "continuous", but then they stopped taking tablets (in that year).
3. Estimate this "stopping" date as being $(k - 1 + C_k/100)*365$ days from randomisation, where k = annual survey identified in Step 1, and C_k = compliance in k^{th} year (set to zero if missing).

If the person is known to have completely stopped tablets prior to date estimated in Step 3, then truncate at date of last tablet.

Appendix 3. Planned main tables

Table 1. Baseline characteristics according to randomisation group

Characteristic	N (%) Vitamin D (N = xx,xxx)	N (%) Placebo (N = xx,xxx)
Age (years)		
60-64	x (x.x)	x (x.x)
65-69	x (x.x)	x (x.x)
70-74	x (x.x)	x (x.x)
≥ 75	x (x.x)	x (x.x)
Sex		
Men	x (x.x)	x (x.x)
Women	x (x.x)	x (x.x)
Body mass index (kg/m²)		
< 18.5	x (x.x)	x (x.x)
18.5 to < 25	x (x.x)	x (x.x)
25 to < 30	x (x.x)	x (x.x)
≥ 30	x (x.x)	x (x.x)
Missing	*x*	*x*
Predicted 25(OH)D concentration (nmol/L)		
< 50	x (x.x)	x (x.x)
≥ 50	x (x.x)	x (x.x)
Ancestry		
British/European	x (x.x)	x (x.x)
Australian/New Zealand	x (x.x)	x (x.x)
Asian	x (x.x)	x (x.x)
Indigenous	x (x.x)	x (x.x)
Mixed/other	x (x.x)	x (x.x)
Missing	*x*	*x*
Highest qualification obtained		
None	x (x.x)	x (x.x)
School or intermediate certificate	x (x.x)	x (x.x)
Higher school or leaving certificate	x (x.x)	x (x.x)
Apprenticeship or certificate	x (x.x)	x (x.x)
University degree or higher	x (x.x)	x (x.x)
Missing	*x*	*X*

Characteristic	N (%) Vitamin D (N = xx,xxx)	N (%) Placebo (N = xx,xxx)
Living alone		
No	x (x.x)	x (x.x)
Yes	x (x.x)	x (x.x)
Missing	*x*	*x*
Alcohol consumption (drinks/week)		
< 1	x (x.x)	x (x.x)
1 to 7	x (x.x)	x (x.x)
> 7 to 14	x (x.x)	x (x.x)
> 14	x (x.x)	x (x.x)
Missing	*x*	*x*
Smoking history		
Never	x (x.x)	x (x.x)
Ex-smoker	x (x.x)	x (x.x)
Current	x (x.x)	x (x.x)
Missing	*x*	*x*
Physical activity (METs/week)		
Low (< 18)	x (x.x)	x (x.x)
Moderate (18 to < 45)	x (x.x)	x (x.x)
High (≥ 45)	x (x.x)	x (x.x)
Missing	*x*	*X*
Time outdoors (hours/week)		
Low (< 8)	x (x.x)	x (x.x)
Moderate (8 to < 18)	x (x.x)	x (x.x)
High (≥ 18)	x (x.x)	x (x.x)
Missing	*x*	*x*
Chronic pain		
No	x (x.x)	x (x.x)
Yes	x (x.x)	x (x.x)
Missing	*x*	*x*
Self-rated overall health		
Excellent or very good	x (x.x)	x (x.x)
Good	x (x.x)	x (x.x)
Fair or poor	x (x.x)	x (x.x)
Missing	*x*	*x*

Characteristic	N (%) Vitamin D (N = xx,xxx)	Placebo (N = xx,xxx)
Self-rated quality of life		
Excellent or very good	x (x.x)	x (x.x)
Good	x (x.x)	x (x.x)
Fair or poor	x (x.x)	x (x.x)
Missing	*x*	*X*
Self-reported history of hypertension		
No	x (x.x)	x (x.x)
Yes	x (x.x)	x (x.x)
Missing	*x*	*x*
Self-reported history of high cholesterol		
No	x (x.x)	x (x.x)
Yes	x (x.x)	x (x.x)
Missing	*x*	*x*
Self-reported history of diabetes		
No	x (x.x)	x (x.x)
Yes	x (x.x)	x (x.x)
Missing	*x*	*x*
Self-reported history of depression		
No	x (x.x)	x (x.x)
Yes	x (x.x)	x (x.x)
Missing	*x*	*x*

Table 2. Effect of supplementation with vitamin D on all-cause mortality. Predicted difference in standardised survival and time-varying hazard ratio at 2, 4, and 6 years post-randomisation, and predicted overall hazard ratio.[a]

Years since randomisation	Survival Difference (95% CI)	Hazard Ratio (95% CI)
2	x.xx (x.xx, x.xx)	x.xx (x.xx, x.xx)
4	x.xx (x.xx, x.xx)	x.xx (x.xx, x.xx)
6	x.xx (x.xx, x.xx)	x.xx (x.xx, x.xx)
Overall HR	..	x.xx (x.xx, x.xx)

[a] Estimates (comparing vitamin D to placebo) from flexible parametric survival models that include randomisation group, age, sex, and state of residence at baseline. Time-varying hazards were predicted using a model that also included an interaction between randomisation group and a restricted cubic spline with one internal knot (placed at the median of uncensored log survival times).

Table 3. Self-reported adverse events[a] according to randomisation group

	No. people[b]		Total no. events (Events per 1000 PYAR[c])		
	Vitamin D (N=)	**Placebo (N=)**	**Vitamin D (PYAR=xx,xxx)**	**Placebo (PYAR=xx,xxx)**	**IRR (95% CI)[d]**
Total AEs	x	x	x (x.x)	x (x.x)	x.xx (x.xx, x.xx)
Non-Serious AEs	x	x	x (x.x)	x (x.x)	x.xx (x.xx, x.xx)
Serious AEs	x	x	x (x.x)	x (x.x)	x.xx (x.xx, x.xx)
Specific AEs					
Hypercalcaemia	x	x	x (x.x)	x (x.x)	x.xx (x.xx, x.xx)
Hyperparathyroidism	x	x	x (x.x)	x (x.x)	x.xx (x.xx, x.xx)
Kidney Stones	x	x	x (x.x)	x (x.x)	x.xx (x.xx, x.xx)
Stomach upset or pain	x	x	x (x.x)	x (x.x)	x.xx (x.xx, x.xx)
Nausea	x	x	x (x.x)	x (x.x)	x.xx (x.xx, x.xx)
Constipation	x	x	x (x.x)	x (x.x)	x.xx (x.xx, x.xx)
Diarrhoea	x	x	x (x.x)	x (x.x)	x.xx (x.xx, x.xx)
Skin rash	x	x	x (x.x)	x (x.x)	x.xx (x.xx, x.xx)

[a] The AEs included in this table are from our ad hoc adverse event reports, with the exception of hypercalcaemia, hyperparathyroidism, and kidney stones, for which we also used data from annual surveys. We report all adverse events irrespective of whether or not they were deemed, at the time of reporting, to be related to the intervention.
[b] Number of people who experienced at least one adverse event
[c] PYAR is time from randomisation until one month after the date last tablet taken
[d] Estimated using unadjusted Poisson regression
Abbreviations: AE, adverse event; CI, confidence interval; IRR, incidence rate ratio; PYAR, person years at risk

Appendix 4. Planned supplementary tables

STable 1. Underlying cause of death according to randomisation group.

Underlying cause of death	ICD-10 Codes	N (%)					
		All deaths		Coded officially[a]		Coded in-house[b]	
		Vitamin D (N = xx)	Placebo (N = xx)	Vitamin D (N = xx)	Placebo (N = xx)	Vitamin D (N = xx)	Placebo (N = xx)
CANCER[c]	**C00-C97**	**x (x)**	**x (x)**	**x (x)**	**x (x)**	**x (x)**	**x (x)**
Of lip, oral cavity and pharynx	C00-C14	x (x)	x (x)	x (x)	x (x)	..	..
Of digestive organs	C15-C26	x (x)	x (x)	x (x)	x (x)	..	..
Of respiratory and intrathoracic organs	C30-C39	x (x)	x (x)	x (x)	x (x)	..	..
Of bone and articular cartilage	C40-C41	x (x)	x (x)	x (x)	x (x)	..	..
Melanoma and other malignant neoplasms of skin	C43-C44	x (x)	x (x)	x (x)	x (x)	..	..
Of mesothelial and soft tissue	C45-C49	x (x)	x (x)	x (x)	x (x)	..	..
Of breast	C50-C50	x (x)	x (x)	x (x)	x (x)	..	..
Of female genital organs	C51-C58	x (x)	x (x)	x (x)	x (x)	..	..
Of male genital organs	C60-C63	x (x)	x (x)	x (x)	x (x)	..	..
Of urinary tract	C64-C68	x (x)	x (x)	x (x)	x (x)	..	..
Of eye, brain and other parts of central nervous system	C69-C72	x (x)	x (x)	x (x)	x (x)	..	..
Of thyroid and other endocrine glands	C73-C75	x (x)	x (x)	x (x)	x (x)	..	..
Of ill-defined, secondary and unspecified sites	C76-C80	x (x)	x (x)	x (x)	x (x)	..	..
Stated or presumed to be primary, of lymphoid, haematopoietic and related tissue	C81-C96	x (x)	x (x)	x (x)	x (x)	..	..
Of independent (primary) multiple sites	C97-C97	x (x)	x (x)	x (x)	x (x)	..	..
DISEASES OF THE CIRCULATORY SYSTEM	**I00-I99**	**x (x)**	**x (x)**	**x (x)**	**x (x)**	**x (x)**	**x (x)**
Acute rheumatic fever	I00-I02	x (x)	x (x)	x (x)	x (x)	..	..
Chronic rheumatic heart diseases	I05-I09	x (x)	x (x)	x (x)	x (x)	..	..

Underlying cause of death	ICD-10 Codes	N (%)					
		All deaths		Coded officially[a]		Coded in-house[b]	
		Vitamin D (N = xx)	Placebo (N = xx)	Vitamin D (N = xx)	Placebo (N = xx)	Vitamin D (N = xx)	Placebo (N = xx)
Hypertensive diseases	I10-I15	x (x)	x (x)	x (x)	x (x)	..	..
Ischaemic heart diseases	I20-I25	x (x)	x (x)	x (x)	x (x)	..	..
Pulmonary heart disease and diseases of pulmonary circulation	I26-I28	x (x)	x (x)	x (x)	x (x)	..	..
Other forms of heart disease	I30-I52	x (x)	x (x)	x (x)	x (x)	..	..
Cerebrovascular diseases	I60-I69	x (x)	x (x)	x (x)	x (x)	..	..
Diseases of arteries, arterioles and capillaries	I70-I79	x (x)	x (x)	x (x)	x (x)	..	..
Diseases of veins, lymphatic vessels and lymph nodes, not elsewhere classified	I80-I89	x (x)	x (x)	x (x)	x (x)	..	..
Other and unspecified disorders of the circulatory system	I95-I99	x (x)	x (x)	x (x)	x (x)	..	..
OTHER CAUSES		**x (x)**	**x (x)**	**x (x)**	**x (x)**	**x (x)**	**x (x)**
EXTERNAL CAUSES		x (x)	x (x)	x (x)	x (x)	x (x)	x (x)
Injury, poisoning and certain other consequences of external causes	S00-T98	x (x)	x (x)	x (x)	x (x)	..	..
External causes of morbidity and mortality	V00-Y98	x (x)	x (x)	x (x)	x (x)	..	..
NON-EXTERNAL CAUSES		x (x)	x (x)	x (x)	x (x)	x (x)	x (x)
Certain infectious and parasitic diseases	A00-B99	x (x)	x (x)	x (x)	x (x)	..	..
In situ neoplasms	D00-D09	x (x)	x (x)	x (x)	x (x)	..	..
Benign neoplasms	D10-D36	x (x)	x (x)	x (x)	x (x)	..	..
Neoplasms of uncertain or unknown behaviour	D37-D48	x (x)	x (x)	x (x)	x (x)	..	..
Diseases of the blood and blood-forming organs and certain disorders involving the immune mechanism	D50-D89	x (x)	x (x)	x (x)	x (x)	..	..
Endocrine, nutritional and metabolic diseases	E00-E90	x (x)	x (x)	x (x)	x (x)	..	..

Underlying cause of death	ICD-10 Codes	N (%)					
		All deaths		Coded officially[a]		Coded in-house[b]	
		Vitamin D (N = xx)	Placebo (N = xx)	Vitamin D (N = xx)	Placebo (N = xx)	Vitamin D (N = xx)	Placebo (N = xx)
Mental and behavioural disorders	F00-F99	x (x)	x (x)	x (x)	x (x)	..	..
Diseases of the nervous system	G00-G99	x (x)	x (x)	x (x)	x (x)	..	..
Diseases of the eye and adnexa	H00-H59	x (x)	x (x)	x (x)	x (x)	..	..
Diseases of the ear and mastoid process	H60-H95	x (x)	x (x)	x (x)	x (x)	..	..
Diseases of the respiratory system	J00-J99	x (x)	x (x)	x (x)	x (x)	..	..
Diseases of the digestive system	K00-K93	x (x)	x (x)	x (x)	x (x)	..	..
Diseases of the skin and subcutaneous tissue	L00-L99	x (x)	x (x)	x (x)	x (x)	..	..
Diseases of the musculoskeletal system and connective tissue	M00-M99	x (x)	x (x)	x (x)	x (x)	..	..
Diseases of the genitourinary system	N00-N99	x (x)	x (x)	x (x)	x (x)	..	..
Congenital malformations, deformations and chromosomal abnormalities	Q00-Q99	x (x)	x (x)	x (x)	x (x)	..	..
Symptoms, signs and abnormal clinical and laboratory findings, not elsewhere classified	R00-R94	x (x)	x (x)	x (x)	x (x)	..	..
Provisional assignment of new diseases of uncertain aetiology or emergency use	U00-U49	x (x)	x (x)	x (x)	x (x)	..	..
Resistance to antimicrobial and antineoplastic drugs	U82-U85	x (x)	x (x)	x (x)	x (x)	..	..
Factors influencing health status and contact with health services	Z00-Z99	x (x)	x (x)	x (x)	x (x)	..	..
ILL-DEFINED OR UNKNOWN	**R95-R99**	**x (x)**	**x (x)**	**x (x)**	**x (x)**	..	..
UNABLE TO BE CODED	**NA**	**x (x)**	**x (x)**	..	..	**x (x)**	**x (x)**
Missing		*x*	*x*				

[a] ICD-10 code for underlying cause of death available

[b] Classified using text description of cause(s) of death into as one of the following: cancer; disease of the circulatory system; other – external cause; other – non-external cause; or unable to be coded.

[c] Malignant neoplasms
Abbreviations: NA, not applicable

STable 2. Retention and compliance with study tablets according to randomisation group

	N (%)	
	Vitamin D (N = xx,xxx)	**Placebo (N = xx,xxx)**
Completed the trial[a]		
No	x (x.x)	x (x.x)
Yes	x (x.x)	x (x.x)
Compliance to end of year 5[b]		
<30%	x (x.x)	x (x.x)
30 to 49%	x (x.x)	x (x.x)
50 to 79%	x (x.x)	x (x.x)
≥80%	x (x.x)	x (x.x)

[a] A participant was classified as having completed the trial if they were still taking tablets at the end of their planned intervention period.

[b] Compliance calculated as number of study tablets taken divided by the number expected to be taken, expressed as a percentage. If a person died without having withdrawn, then the expected number was the number that should have been taken during the period between randomisation and death. Otherwise, the expected number was 60.

STable 3. Intake of vitamin D from off-trial supplements according to randomisation group.

Supplementary Vitamin D (IU/day)	N (%)[a]									
	Annual Survey 1		Annual Survey 2		Annual Survey 3		Annual Survey 4		Annual Survey 5	
	Vitamin D	Placebo	Vitamin D	Placebo	Vitamin D	Placebo	Vitamin D	Placebo	Vitamin D	Placebo
None	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)
>0 to ≤500	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)
>500 to ≤1,000	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)
>1,000	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)
Missing	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)	x (x.x)
Total[b]	x	x	x	x	x	x	x	x	x	x

[a] The denominator for the calculation of % in each of the intake categories excludes the N missing.
[b] The total is the number of participants who should have completed the annual survey (i.e., everyone who was alive at the time the survey was due).

Appendix 5. Selected planned main figures (constructed using dummy data)

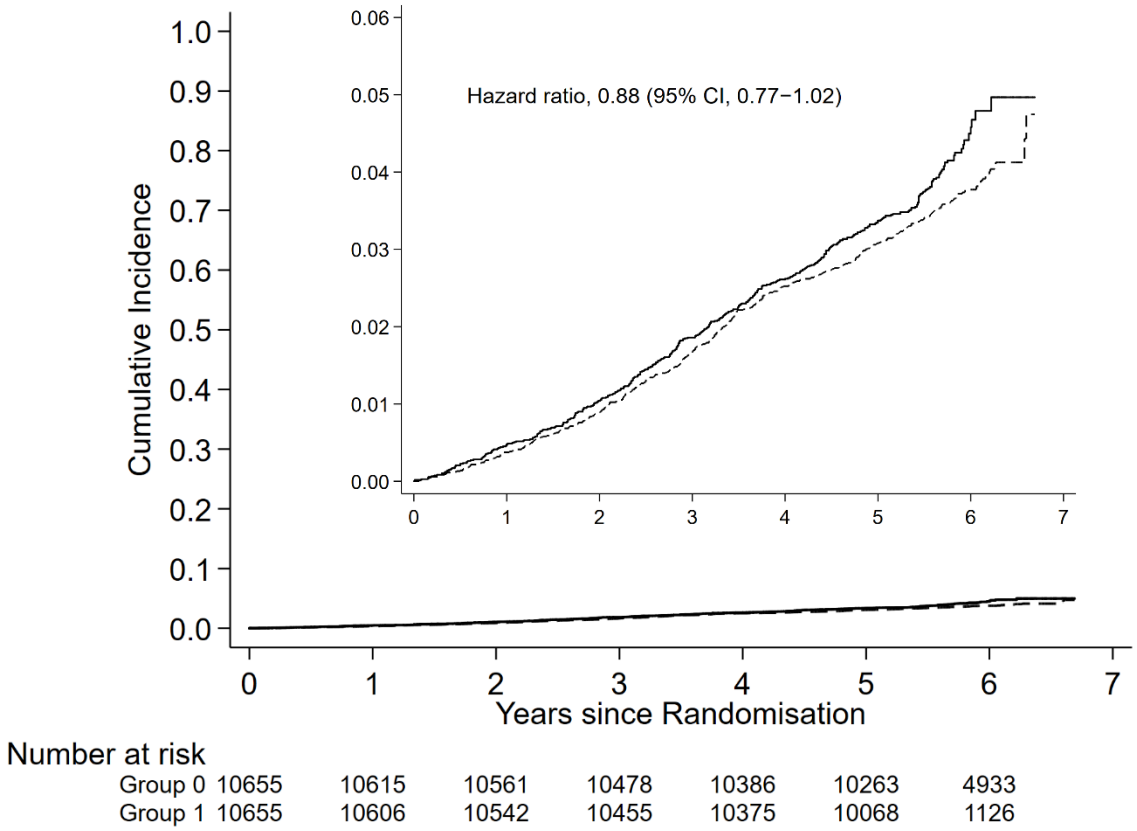


Figure 2. Cumulative incidence of all-cause mortality according to time since randomisation in Group 1 (solid line) and Group 0 (dashed line).

Curves estimated using Kaplan-Meier methods and hazard ratio (Group 1 versus Group 0) estimated using a flexible parametric survival model that included randomisation group, age, sex, and state of residence at baseline. The inset shows the same data on an enlarged y axis.
Abbreviation: CI, confidence interval

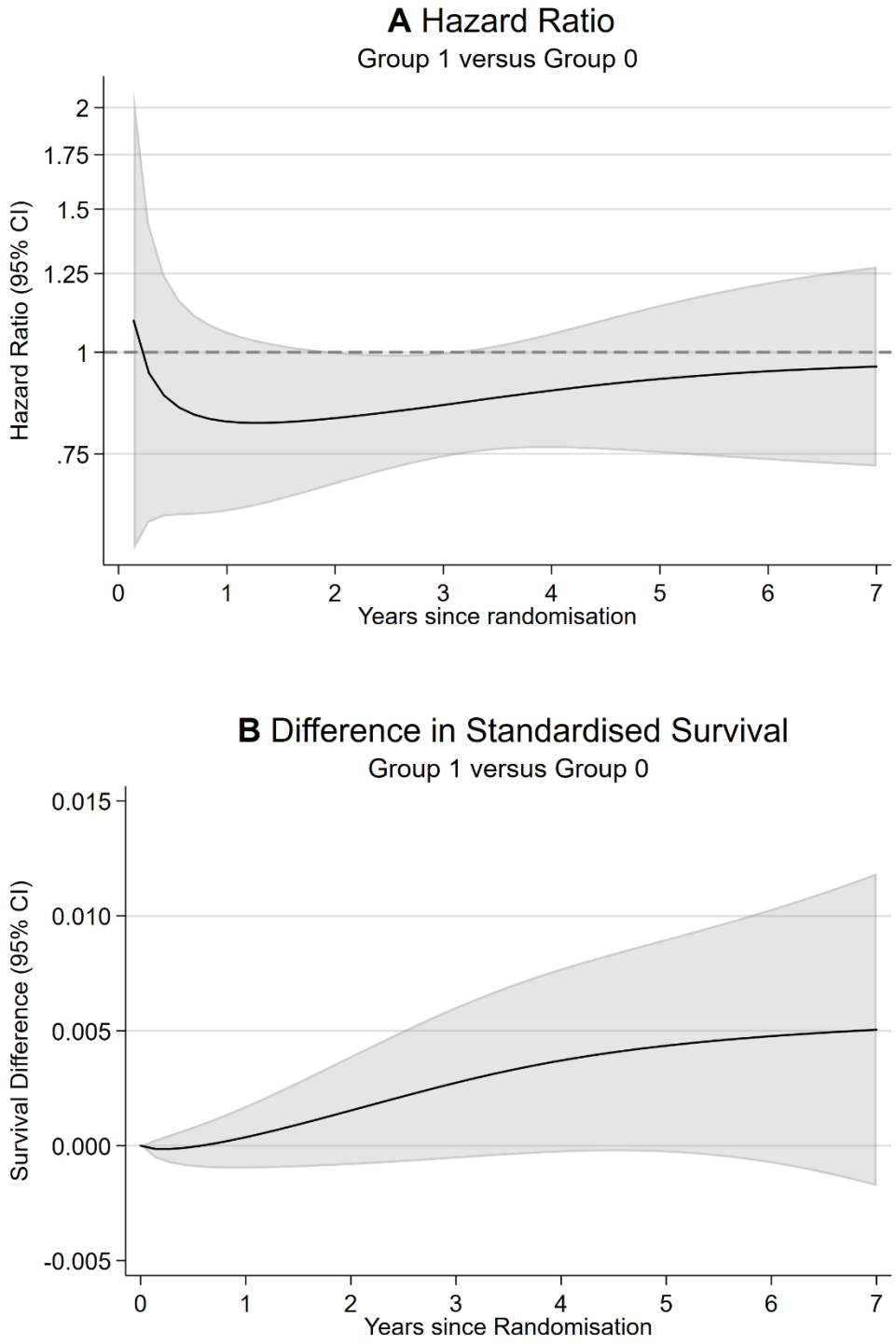


Figure 3. Effect of vitamin D supplementation on all-cause mortality. Panel A shows the time-varying hazard ratio and panel B shows the difference in standardised survival curves as a function of time since randomisation.

Estimates are from a flexible parametric survival model that included randomisation group, age, sex, and state of residence at baseline. The model included an interaction between randomisation group and a restricted cubic spline with one internal knot (placed at the median of uncensored log survival times).
Abbreviation: CI, confidence interval

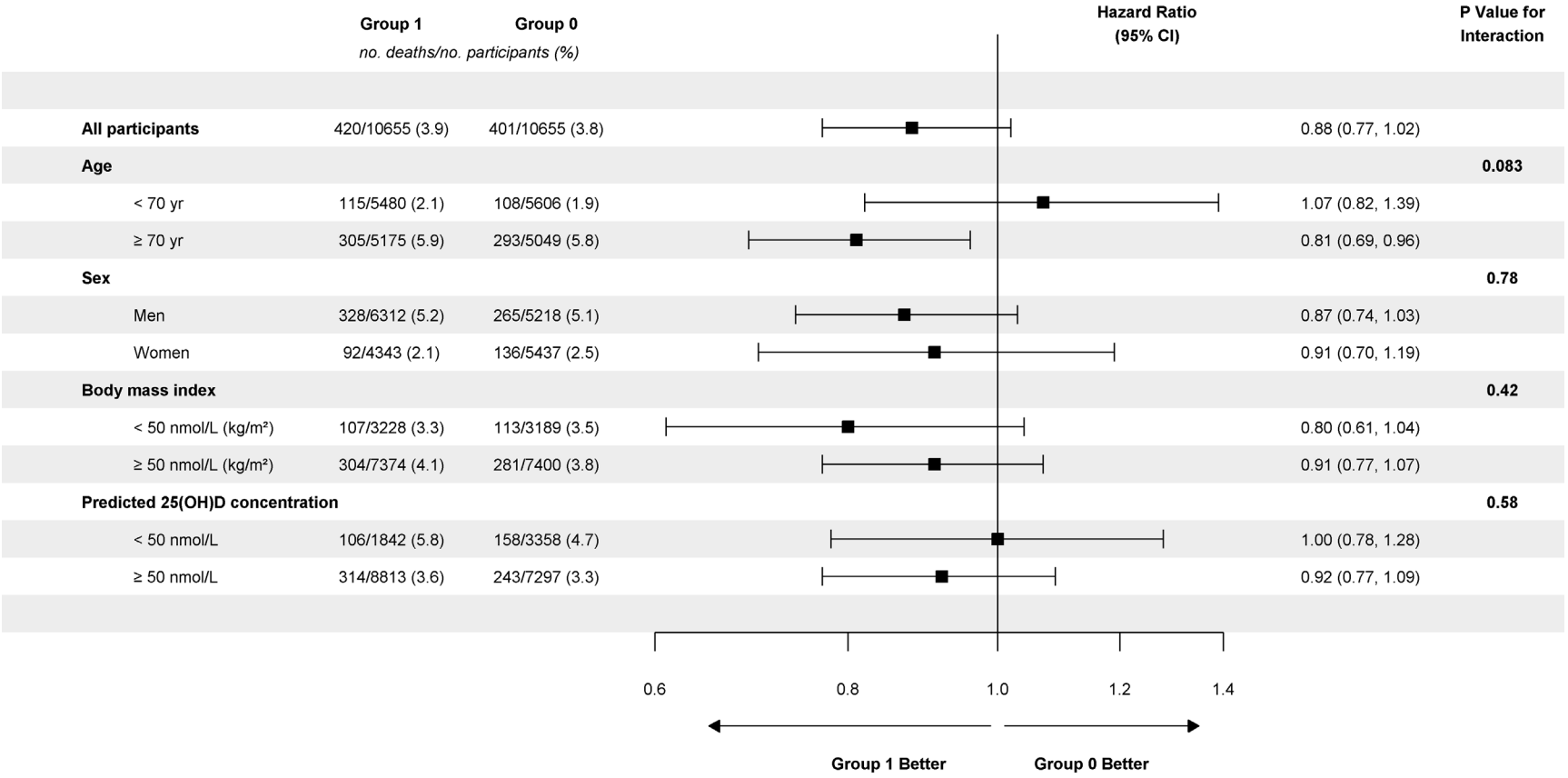
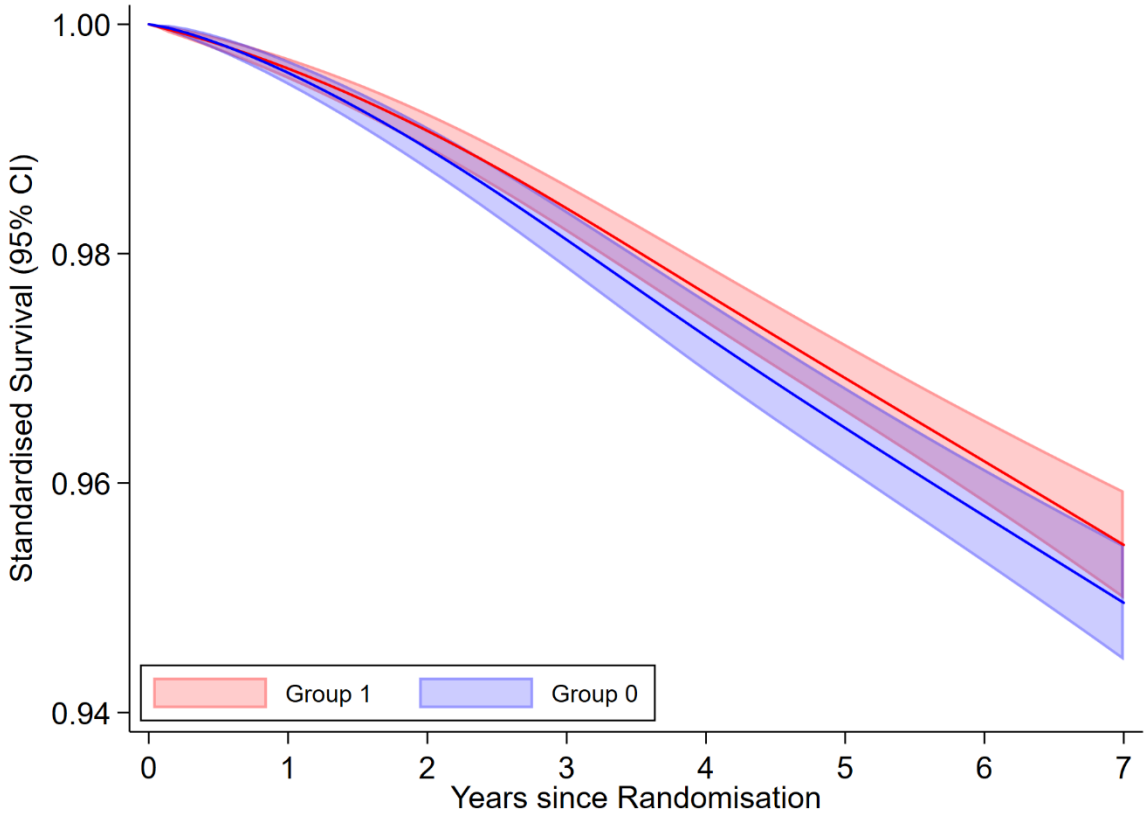


Figure 4. Effect of vitamin D supplementation on all-cause mortality for all participants and by selected baseline characteristics.

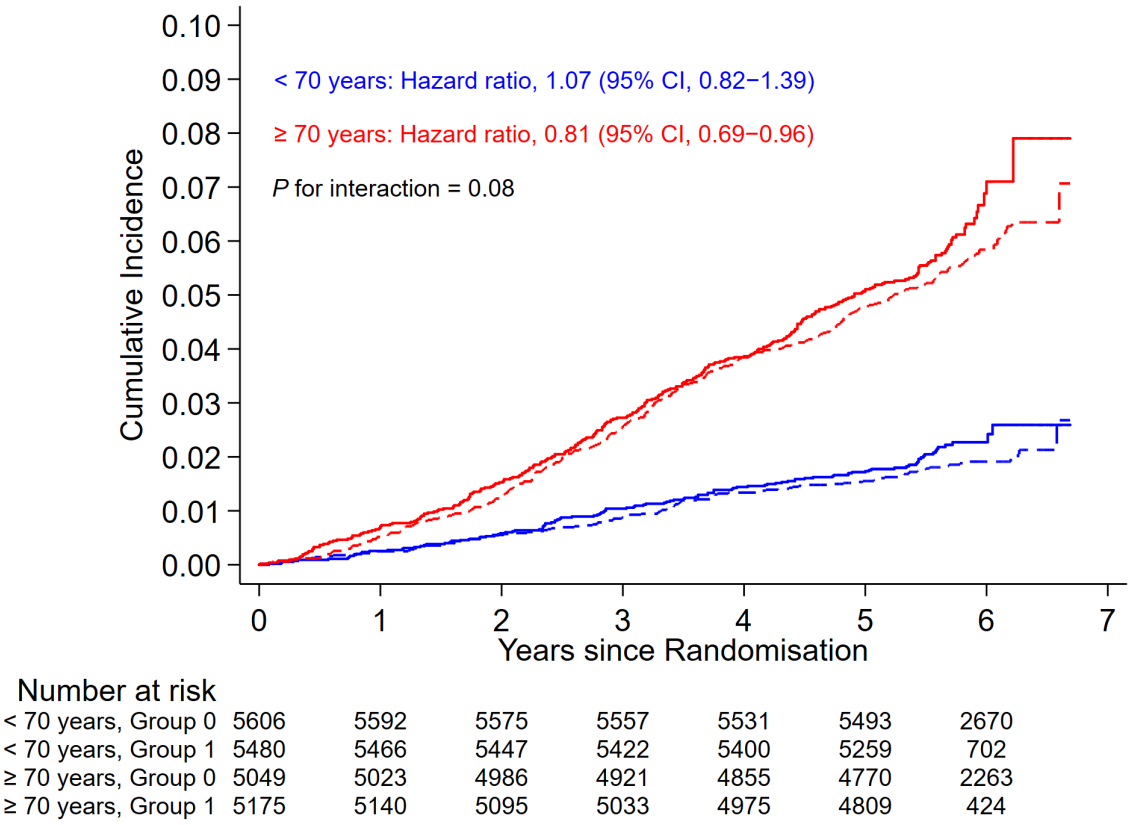
Estimates from flexible parametric survival models. Hazard ratios compare Group 1 to Group 0. All models include randomisation group, age, sex, and state of residence at baseline. Models producing estimates by levels of age, sex, BMI, and predicted 25(OH)D concentration, also include the characteristic of interest (if not already included), and an interaction between randomisation group and the characteristic of interest. P value for interaction is from a likelihood ratio test comparing models with and without the interaction term.
Abbreviation: CI, confidence interval

Appendix 6. Selected planned supplementary figures (constructed using dummy data)



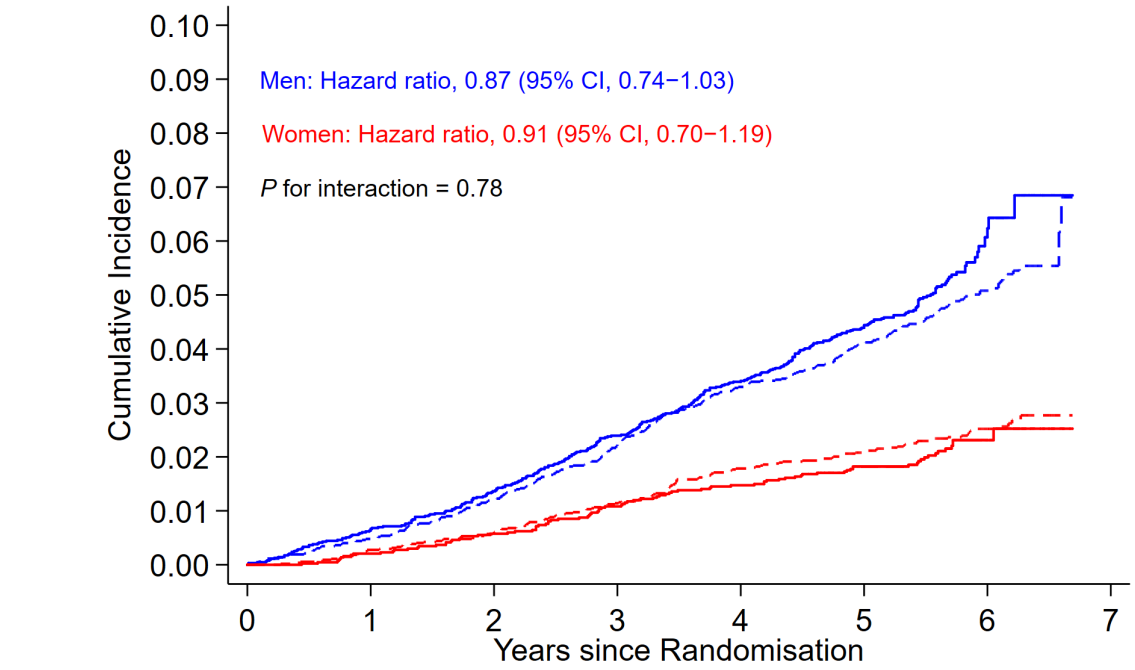
SFigure 1. Standardised survival curves according to randomisation group; estimates from a flexible parametric survival model that included randomisation group, age, sex, and state of residence at baseline, and an interaction between randomisation group and a restricted cubic spline with one internal knot (placed at the median of uncensored log survival times).

Abbreviation: CI, confidence interval



SFigure 2. Cumulative incidence of all-cause mortality according to time since randomisation in Group 1 (solid line) and Group 0 (dashed line) within strata of age (< 70 years; ≥ 70 years).

Curves estimated using Kaplan-Meier methods. Hazard ratios (Group1 versus Group 0) estimated using a flexible parametric survival model that included randomisation group, age, sex, and state of residence at baseline, and an interaction between randomisation group and age. P value for the interaction from a likelihood ratio test comparing the model with and without the interaction term. Abbreviation: CI, confidence interval

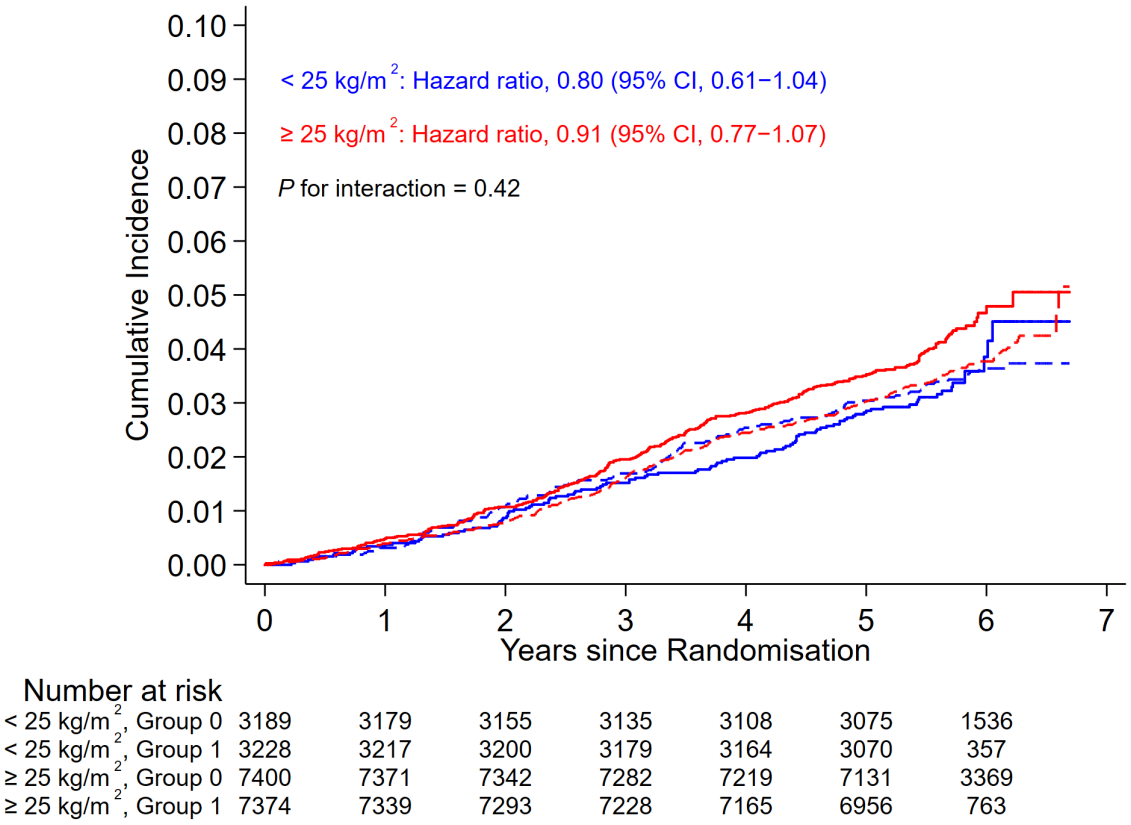


Number at risk

Men, Group 0	5218	5193	5156	5103	5046	4966	2376
Men, Group 1	6312	6272	6224	6159	6096	5903	563
Women, Group 0	5437	5422	5405	5375	5340	5297	2557
Women, Group 1	4343	4334	4318	4296	4279	4165	563

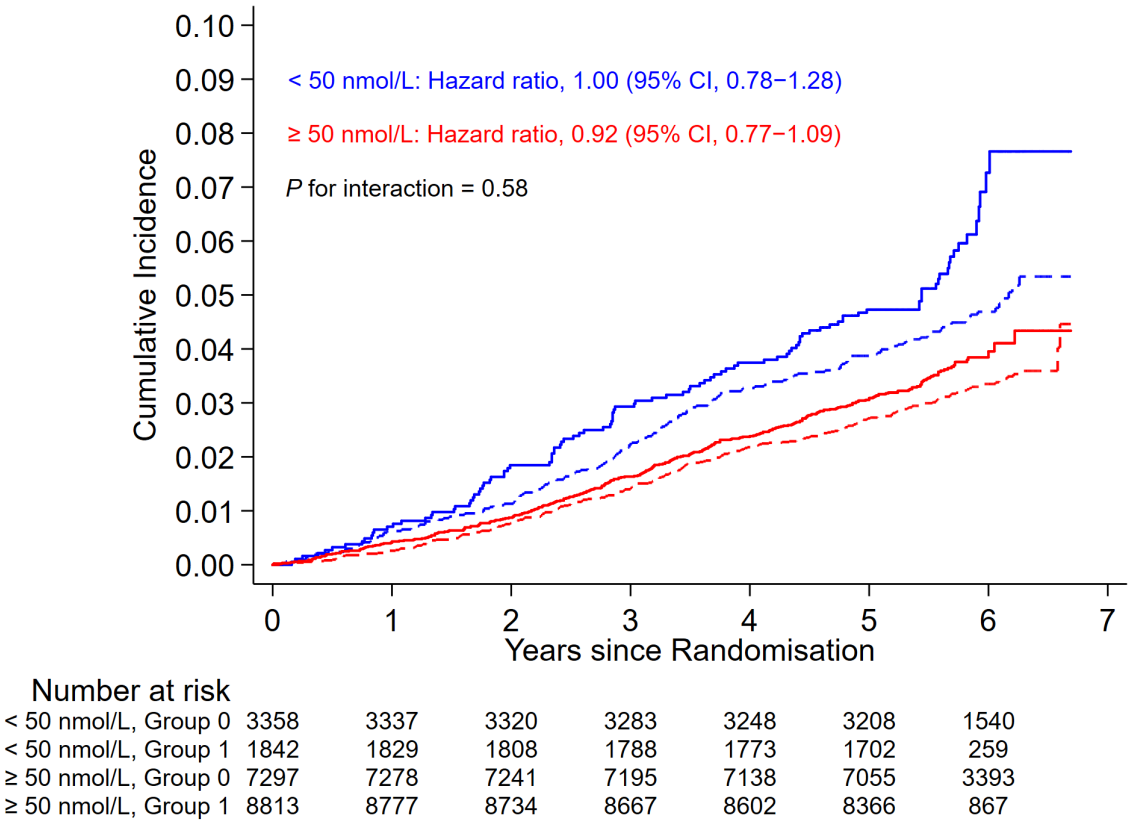
SFigure 3. Cumulative incidence of all-cause mortality according to time since randomisation in Group 1 (solid line) and Group 0 (dashed line) within strata of sex (men; women).

Curves estimated using Kaplan-Meier methods. Hazard ratios (Group1 versus Group 0) estimated using a flexible parametric survival model that included randomisation group, age, sex, and state of residence at baseline, and an interaction between randomisation group and sex. P value for the interaction from a likelihood ratio test comparing the model with and without the interaction term. Abbreviation: CI, confidence interval



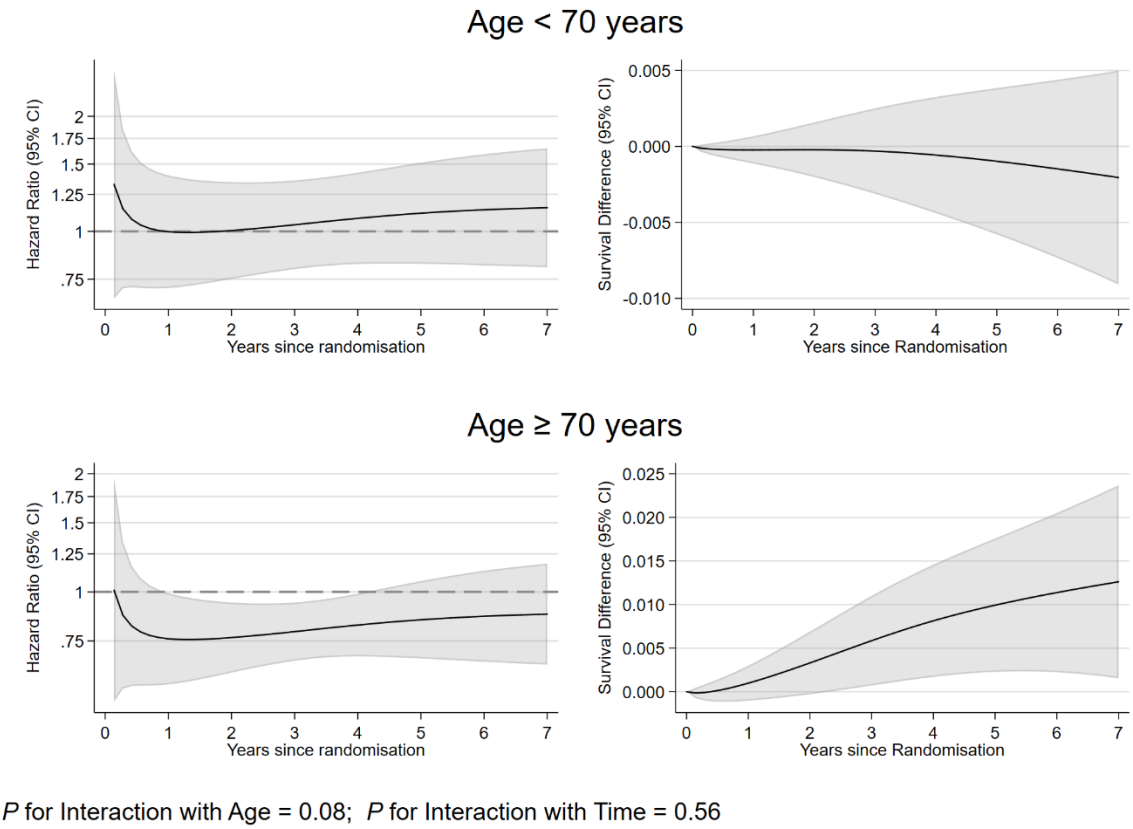
SFigure 4. Cumulative incidence of all-cause mortality according to time since randomisation in Group 1 (solid line) and Group 0 (dashed line) within strata of body mass index (< 25 kg/m^2; ≥ 25 kg/m^2).

Curves estimated using Kaplan-Meier methods. Hazard ratios (Group1 versus Group 0) estimated using a flexible parametric survival model that included randomisation group, BMI, age, sex, and state of residence at baseline, and an interaction between randomisation group and BMI. P value for the interaction from a likelihood ratio test comparing the model with and without the interaction term. Abbreviations: BMI, body mass index; CI, confidence interval



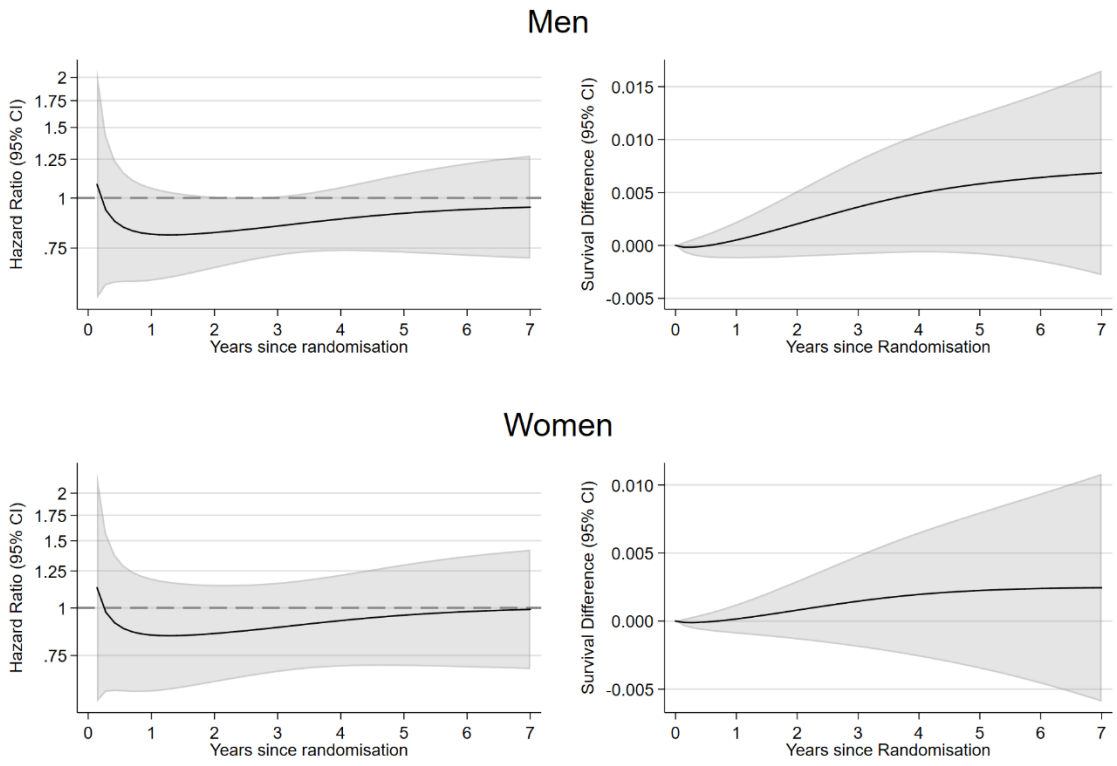
SFigure 5. Cumulative incidence of all-cause mortality according to time since randomisation in Group 1 (solid line) and Group 0 (dashed line) within strata of predicted deseasonalised serum 25(OH)D concentration at baseline (< 50 nmol/L; ≥ 50 nmol/L).

Curves estimated using Kaplan-Meier methods. Hazard ratios (Group1 versus Group 0) estimated using a flexible parametric survival model that included randomisation group, predicted deseasonalised serum 25(OH)D concentration, age, sex, and state of residence at baseline, and an interaction between randomisation group and predicted deseasonalised serum 25(OH)D concentration. P value for the interaction from a likelihood ratio test comparing the model with and without the interaction term.
Abbreviation: CI, confidence interval



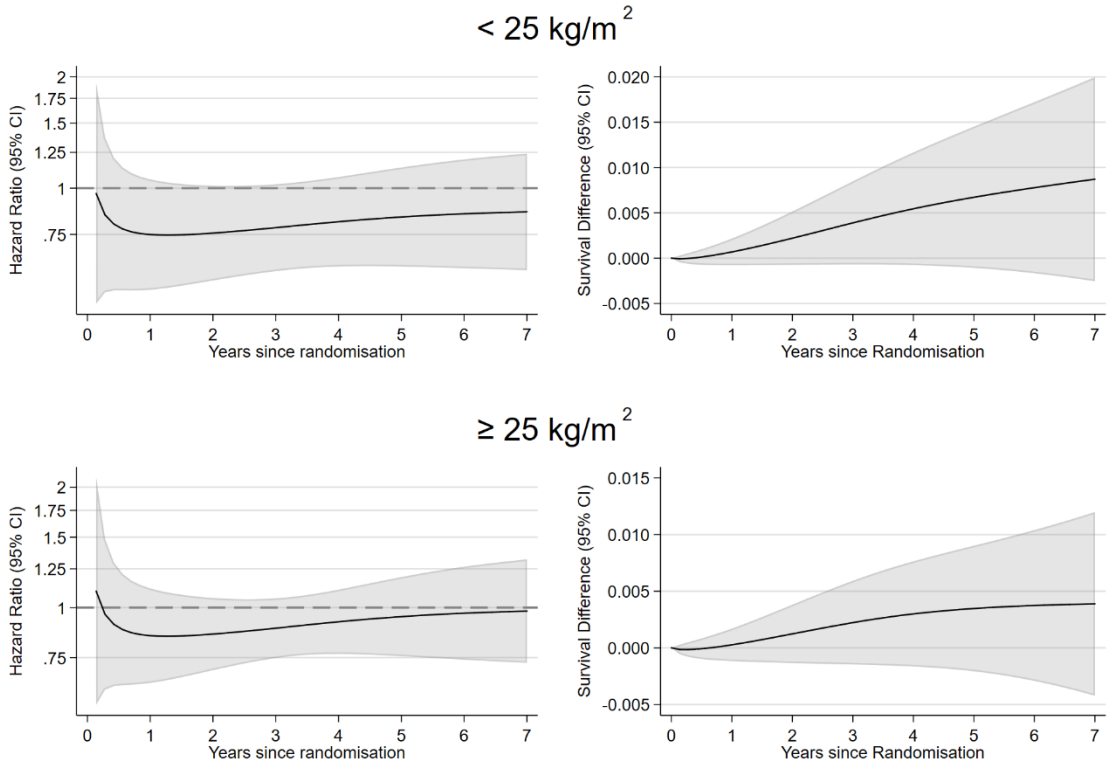
SFigure 6. Effect of vitamin D supplementation on all-cause mortality according to age group (< 70 years (top row); ≥ 70 years (bottom row)). Plots on the left show the time-varying hazard ratio. Plots on the right show the difference in standardised survival curves as a function of time since randomisation.

Estimates (comparing Group 1 to Group 0) are from a flexible parametric survival model that included randomisation group, age, sex, and state of residence at baseline, and an interaction between randomisation group and age. The model included an interaction between randomisation group and a restricted cubic spline with one internal knot (placed at the median of uncensored log survival times). P values for interactions from likelihood ratio tests: P for interaction with age compared models with and without interaction term between randomisation group and age; P for interaction with time compared models that did and did not allow the effect of randomisation group to vary with time. Abbreviation: CI, confidence interval



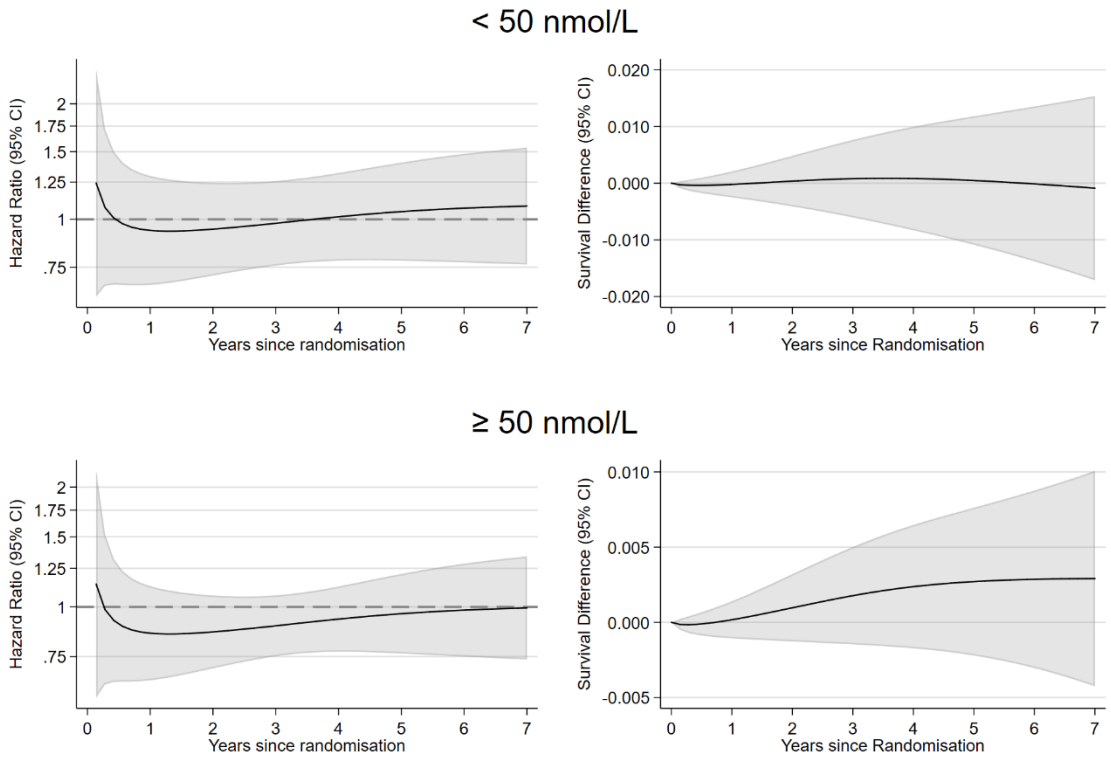
SFigure 7. Effect of vitamin D supplementation on all-cause mortality according to sex (men (top row); women (bottom row)). Plots on the left show the time-varying hazard ratio. Plots on the right show the difference in standardised survival curves as a function of time since randomisation.

Estimates (comparing Group 1 to Group 0) are from a flexible parametric survival model that included randomisation group, age, sex, and state of residence at baseline, and an interaction between randomisation group and sex. The model included an interaction between randomisation group and a restricted cubic spline with one internal knot (placed at the median of uncensored log survival times). P values for interactions from likelihood ratio tests: P for interaction with sex compared models with and without interaction term between randomisation group and sex; P for interaction with time compared models that did and did not allow the effect of randomisation group to vary with time. Abbreviation: CI, confidence interval



SFigure 8. Effect of vitamin D supplementation on all-cause mortality according to body mass index (< 25 kg/m^2 (top row); ≥ 25 kg/m^2 (bottom row)). Plots on the left show the time-varying hazard ratio. Plots on the right show the difference in standardised survival curves as a function of time since randomisation.

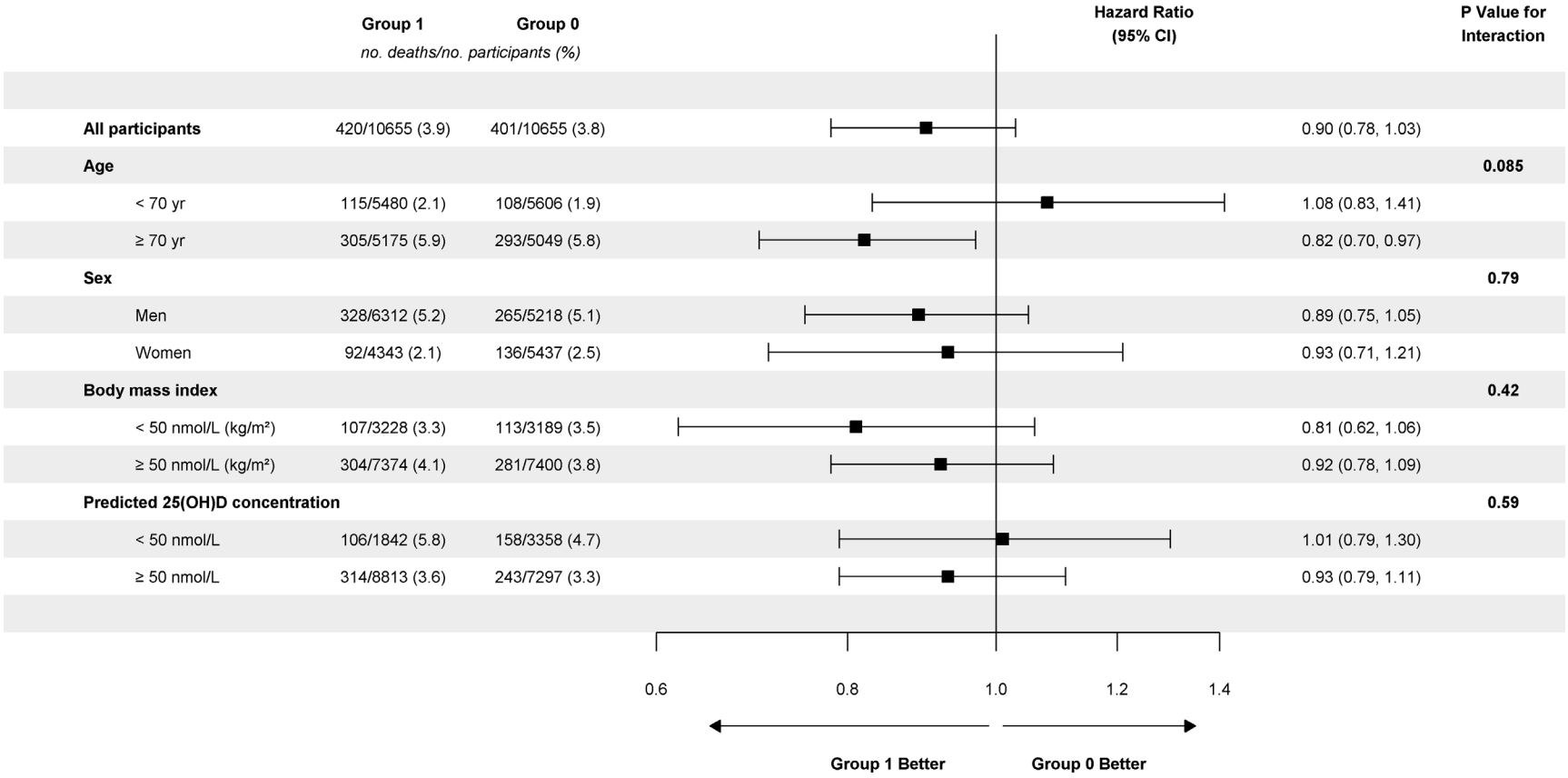
Estimates (comparing Group 1 to Group 0) are from a flexible parametric survival model that included randomisation group, BMI, age, sex, and state of residence at baseline, and an interaction between randomisation group and BMI. The model included an interaction between randomisation group and a restricted cubic spline with one internal knot (placed at the median of uncensored log survival times). P values for interactions from likelihood ratio tests: P for interaction with BMI compared models with and without interaction term between randomisation group and BMI; P for interaction with time compared models that did and did not allow the effect of randomisation group to vary with time.
Abbreviation: BMI, body mass index; CI, confidence interval



P for Interaction with 25(OH)D = 0.57; *P* for Interaction with Time = 0.56

SFigure 9. Effect of vitamin D supplementation on all-cause mortality according to predicted deseasonalised baseline serum 25(OH)D concentration (< 50 nmol/L (top row); ≥ 50 nmol/L (bottom row)). Plots on the left show the time-varying hazard ratio. Plots on the right show the difference in standardised survival curves as a function of time since randomisation.

Estimates (comparing Group 1 to Group 0) are from a flexible parametric survival model that included randomisation group, predicted 25(OH)D concentration, age, sex, and state of residence at baseline, and an interaction between randomisation group and predicted 25(OH)D concentration. The model included an interaction between randomisation group and a restricted cubic spline with one internal knot (placed at the median of uncensored log survival times). P values for interactions from likelihood ratio tests: P for interaction with 25(OH)D concentration compared models with and without interaction term between randomisation group and predicted 25(OH)D concentration; P for interaction with time compared models that did and did allow the effect of randomisation group to vary with time.
Abbreviation: CI, confidence interval



SFigure 10. Effect of vitamin D supplementation on all-cause mortality for all participants and within strata of pre-specified baseline characteristics; hazard ratios estimated using Cox proportional hazards regression models adjusted for age, sex, and state of residence at baseline.

Hazard ratios compare Group 1 to Group 0. All models include randomisation group, age, sex, and state of residence at baseline. Models producing estimates by levels of age, sex, BMI, and predicted 25(OH)D concentration, also include the characteristic of interest (if not already included), and an interaction between randomisation group and the characteristic of interest. P value for interaction is from a likelihood ratio test comparing models with and without the interaction term.
Abbreviation: CI, confidence interval