Version	Date	Changes to planned analyses
2	26 August 2024	**Note:** The changes below were made after we saw the unblinded results based on Version 1 of the analysis plan. **Main analyses** For all non-melanoma outcomes, we decided to include all relevant diagnoses irrespective of the behaviour of the tumour. The only outcome that is affected by this decision is breast cancer; that is we will no longer exclude n=28 diagnoses that were categorised as either 'in situ' or 'uncertain whether benign or malignant'. **Sensitivity analyses** For each non-melanoma outcome, if the number of non-invasive cancer diangoses is >5, then we will repeat the analysis including only diagnoses of invasive cancer. We do not have behaviour for non-melanoma cancers from the Victorian registry; we will assume that all diagnoses were for invasive cancer.

Working Title	The effect of vitamin D supplementation on cancer incidence: results of the randomised, placebo-controlled D-Health Trial		
Date of plan	6 June 2024		
Person conducting analysis	Mary Waterhouse	Email	Mary.Waterhouse@qimrberghofer.edu.au
Potential Authors	Mary Waterhouse, Catherine Baxter, Briony Duarte Romero, Donald McLeod, Bruce Armstrong, Peter Ebeling, Dallas English, Gunter Hartel, Michael Kimlin, Rachel O'Connell, Jolieke Van Der Pols, Alison Venn, Penny Webb, David Whiteman, Rachel Neale		
Background and overview			
Background	In the current study, we will analyse cancer incidence. We will analyse all cancers combined ('any cancer'), all non-melanoma cancers, and some specific cancers (colorectal, lung, invasive breast (women only), prostate (men only), melanoma skin cancer). All cancers combined and colorectal cancer are the secondary endpoints of the D-Health Trial. We have chosen to analyse the other cancer types because they are among the top 4 cancers in males and females in Australia.		
Aims	Our primary aims are to investigate whether randomisation to long-term supplementation with monthly doses of 60,000 IU of vitamin D_3 alters the incidence of: • any cancer (i.e., all cancer excluding cutaneous squamous cell (SCC) carcinoma and cutaneous basal cell carcinoma (BCC)); • any non-melanoma cancer (i.e., all cancer excluding cutaneous melanoma, cutaneous SCC, and cutaneous BCC); • colorectal cancer; • lung cancer; • invasive breast cancer (females only);		

	• prostate cancer (males only); • melanoma skin cancer (i.e., invasive and in situ) (excludes participants from Western Australia [WA]); and • invasive melanoma skin cancer (excludes participants from WA). Our secondary aim is to investigate whether the effect of vitamin D supplementation on (a) incidence of any cancer and (b) incidence of any non-melanoma skin cancer is modified by: • Age (<70 years, ≥70 years); • Sex (males, females); • Body mass index (BMI) (<25 kg/m^2, 25 to <30 kg/m^2; ≥30 kg/m^2); or • Predicted baseline 25(OH)D concentration[1] (<50 nmol/L, ≥50 nmol/L).
Instruments	We will ascertain outcomes using cancer registry data from all 6 states of Australia (New South Wales [NSW], Queensland, Victoria, South Australia [SA], Tasmania, and WA).
Outcomes	
Outcomes	For this analysis, the two main outcomes will be first diagnosis following randomisation of: 1. cancer of any type 2. non-melanoma cancer of any type. Additional outcomes will be first diagnosis following randomisation of: • colorectal cancer; • lung cancer; • invasive breast cancer (females only); • prostate cancer (males only); • melanoma skin cancer (excludes participants from WA); and • invasive melanoma skin cancer (excludes participants from WA).
Follow-up	• Follow-up will begin at randomisation. • For each outcome follow-up will end at the earliest of: (i) the date the outcome first occurs; (ii) 31/12/2019;[1] (iii) 5 years and 1 month after randomisation; or (iv) the date last known to be alive. • In a *sensitivity analysis* of invasive melanoma skin cancer, follow-up will end at the earliest of: (i) date of invasive melanoma skin cancer; (ii) date of in situ melanoma skin cancer; (iii) 31/12/2019;[a] (iiv) 5 years and 1 month after randomisation; or (v) the date last known to be alive. [a] All states provided cancer registry data up to 31/12/2019.
Ascertaining outcomes	We will ascertain outcomes using the ICD-O-3 code for the anatomical site of origin of the cancer at diagnosis. • any cancer (excluding cutaneous SCC and cutaneous BCC): C00 to C80 (excluding C44 with a histology code starting with 805-811) • any non-melanoma cancer (excluding cutaneous melanoma skin cancer, cutaneous SCC and cutaneous BCC: C00 to C80 (excluding C44 with a histology code starting with 805-811 or 872-879) • colorectal cancer: C18 to C20, and C21.8

	• lung cancer: C33 to C34 • invasive breast cancer [females only]: C50 and behaviour was not categorised as either 'in situ' or 'uncertain whether benign or malignant'[a] • prostate cancer [males only]: C61.9 • melanoma skin cancer (invasive and in situ) [excludes participants from WA]: C44 with a histology code beginning with 872-879, and behaviour code of 2 (in situ) or 3 (invasive)[b] • invasive melanoma skin cancer [excludes participants from WA]: C44 with a histology code beginning with 872-879 and a behaviour code of 3 (invasive)[b] Table A1 in Appendix A shows the number of people with at least one diagnosis of any cancer, any non-melanoma cancer, and for each of the specific cancers considered; the data are presented according to sex, and also show the crude person-based incidence rates. We will use the date of diagnosis from the cancer registry data. Since data from the cancer registries of New South Wales, Victoria, and Tasmania included only the year and month of diagnosis, we will approximate the date of diagnosis using the 15th day of the month. [a] Amongst females there were n=306 diagnoses with a primary site code of C50. Of the n=259 diagnoses with behaviour code available, n=28 (11%) were categorised as either 'in situ' or 'uncertain whether benign or malignant'. Behaviour data were not available for n=47 diagnoses (all from the Victorian registry); we have assumed that these were all invasive. [b] Amongst participants who were not living in WA at baseline there were n=837 diagnoses with a primary site code of C44. Of these, n=820 were melanomas of the skin (300 (37%) invasive, 520 (63%) in situ), and n=17 were non-melanoma skin cancers (e.g., Merkel cell carcinoma).
Detectable effect size	
Detectable effect size	The D-Health Trial was designed to enable 80% power to detect a hazard ratio (HR) of 0.91 (α=0.05) for all-cause mortality over a follow-up period of up to 10 years. The analysis of cancer incidence will include 21,308 participants (n=10,648 in the placebo group; n=10,660 in the vitamin D group). Using data from the Australian Institute of Health and Welfare[2] to estimate the probability of having no cancer diagnoses over 5 years amongst placebo participants included in the analysis of cancer incidence, we estimate that we will be able to detect a HR of 0.88 with 80% power (α=0.05) (Appendix B).
Documentation	
Software	SAS Version 9.4, Stata 18, R Version 4.0.3
Datasets	The original dataset will be located here: L:\Lab_RachelN\DHealthDataAnalysis\Projects\cancer incidence\data\original data\ The final dataset including any constructed variables will be located here: L:\Lab_RachelN\DHealthDataAnalysis\Projects\cancer incidence\data\final data\ A copy of the final dataset will also be saved here: R:\Lab_RachelN\Dhealth\Datasets\cancer incidence\
Statistical code	The code that is used to prepare and analyse data will be stored in: L:\Lab_RachelN\DHealthDataAnalysis\Projects\CancerIncidence \code\

	Once finalised, a copy will be placed on R drive: R:\Lab_RachelN\Dhealth\StatisticalCode\CancerIncidence\
Participants and data	
Participants	<u>All outcomes excluding melanoma skin cancer</u> All randomised participants from the D-Health Trial are eligible excluding the five participants who withdrew from the trial and requested deletion of their data (n=21,310). From this cohort, we will further exclude two participants who withdrew consent for linkage to health registers. Hence, the analysis will include 21,308 participants (vitamin D, n=10,660; placebo, n=10,648; males, n=11,530 (54%); females, n=9,778 (46%); mean age 69 (SD 5.5) years). <u>Analyses of melanoma skin cancer</u> The cancer register of WA provided only diagnoses of invasive cancers. We are therefore restricting all analyses of melanoma skin cancer to 17,882 participants (vitamin D, n = 8,922; placebo, n = 8,960) who were not living in WA at baseline. This will enable sensible comparisons across results for melanoma skin cancer, including the sensitivity analysis in which participants are censored if and when they are diagnosed with an in situ melanoma skin cancer (without prior invasive melanoma skin cancer). Note that we have made a simplifying assumption that participants did not more to, or from, WA during follow-up.
Exposure	Randomisation group
Data Cleaning	For 4 participants, the date of death is in the same month and year as diagnosis; this led to potential data anomalies: • For 3 of the 4 participants, their cancer diagnosis was recorded in a registry that provided only the month and year of diagnosis; the default approach for such data is to impute the 15th of the month, but if the person died prior to the 15th, then it is preferable to use the date of death. • For 1 of the participants, their date of death occurs one day prior to the date of diagnosis and the basis of diagnosis is "death certificate"; we are assuming that the date of death (from a state death registry) is accurate. Therefore, we will change the date of diagnosis to be the date of death for 4 participants (1 record each). We are assuming that all other cancer registry data are correct.
Handling missing data	Behaviour code: We do not have behaviour code for non-melanoma cancer diagnoses in the Victorian cancer register. We have assumed that all diagnoses of with a primary site code of C50 were invasive breast cancers. Effect modifiers: BMI is the only covariate with missing data (n missing = 119). We will exclude participants missing BMI from the relevant subgroup analysis.
Maintaining blinding	Investigators will remain blinded to study group allocation until all tables and figures for the analysis have been approved. • Analysis code will be written and debugged using data in which participants will be randomly reassigned to two new groups (of equal size), such that there is no relationship between the new groups and the true treatment allocation.

	• Once all tables and figures (as described in this plan) have been generated and agreed upon by investigators, analyses will be re-run using the correct randomisation codes. Any analyses performed after unblinding will be declared as exploratory.
Proposed analysis	
	We will show participant flow using a CONSORT diagram (**Figure 1**). We will present selected baseline characteristics of included participants according to randomisation group (**STable 1**). Note: This will essentially be identical to the table presented in our manuscript describing the effect of vitamin D supplementation on mortality.[3] **Effect of randomisation to supplementation with vitamin D on cancer incidence** These analyses will follow the intention-to-treat principle. When fitting flexible parametric survival models[4] (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. If the FPSM includes an interaction between randomisation group and time since randomisation, this will be fitted as a restricted cubic spline with one internal knot, placed at the median of uncensored log survival times. The choice of spline functions follows the recommendations made by Royston and Parmar.[4,5] FPSMs will include randomisation group, and the randomisation stratification variables of age (60-64; 65-69; 70-74; ≥75 years), sex (male; female), and state of residence at baseline (Queensland; NSW; Victoria; Tasmania; SA; WA). **<u>Analysis of the main outcomes</u>** For each of the main outcomes (any cancer, any non-melanoma cancer) we will do the following: • We will report the number (%) of people who had at least one diagnosis, and the crude person-based incidence rate (IR) according to randomisation group (**Table 1**). • We will plot the cause-specific cumulative incidence of first cancer diagnosis for each randomisation group, using Aalen-Johansen methods (**Figure 2, panels A and D**). • We will fit two FPSMs. Model 1 will not include any time-varying coefficients; it will be used to estimate an "overall" HR and 95% confidence interval (CI); the overall HR will be reported in **Table 1** and embedded in **Figure 2 (panels A and D)**. Model 2 will include an interaction between randomisation group and time since randomisation, thereby allowing the HR for randomisation group to vary with time. • Using Model 2, we will plot the following as functions of time since randomisation: (a) the estimated HR (**Figure 2, panels B and E**); and (b) the percentage difference in cause-specific standardised cumulative incidence (**Figure 2, panels C and F**). For part (b), we will treat death without prior cancer diagnosis as a competing risk, and we will use the user-written standsurv command in Stata with the competing risks models option. The plots will include 95% CIs. We will embed in the relevant figure 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). We will also report the estimated HR (95% CI), and percentage difference in cause-specific standardised cumulative incidence (95% CI) at 2 and 5 years post-randomisation (**STable 2**). We will use the percentage difference in cause-specific standardised cumulative incidence at 5 years to calculate the number need to treat to avoid one cancer; this will be reported in the manuscript.

Subgroup analyses

For each main outcome we will investigate whether the effect of supplementation on the outcome is modified by the following:

- Age at baseline (<70 years, ≥70 years);
- Sex (males, females);
- BMI at baseline (<25 kg/m^2, 25 to <30 kg/m^2; ≥30 kg/m^2); and
- Predicted deseasonalised baseline 25(OH)D concentration (<50 nmol/L, ≥50 nmol/L).

For each baseline characteristic we will:

- Use Aalen-Johansen methods to plot cause-specific cumulative incidence of cancer for each randomisation group within each stratum of the characteristic (**SFigures 1 and 6**).
- Fit a FPSM that, in addition to the standard variables, includes the baseline characteristic of interest, and an interaction term between randomisation group and the baseline characteristic of interest. Using this model, we will report estimated overall HRs (95% CI) for each level of the baseline characteristic (**Figure 3**); estimates will also be embedded **SFigures 1 and 6**. We will report the p-value from the likelihood ratio test comparing models with and without the interaction term.
- We will investigate whether the effect of randomisation varies with time for a specific level of the baseline characteristic. We will fit two FPSMs to the data from participants within the subgroup of interest. The first FPSM will not include any time-varying coefficients. The second FPSM will include an interaction between randomisation group and time since randomisation. We will use the likelihood ratio test to compare the two models. We will also use the second model to plot the estimated HR as a function of time since randomisation for that level of the baseline characteristic (**SFigures 2-5** [any cancer] and **SFigures 7-10** [any non-melanoma cancer]).

Analyses of additional outcomes

- Analyses of each additional outcome will proceed in the same manner as for the main outcome, with the exception that we will perform subgroup analyses for a baseline covariate only if the main analysis for that cancer suggested an effect.
- The analyses of invasive breast cancer and prostate cancers will be restricted to females and males, respectively.
- Analyses of melanoma skin cancer will be restricted to participants who were not living in WA at baseline.
- Results for colorectal cancer, lung cancer, invasive breast cancer, prostate cancer, any melanoma skin cancer, and invasive melanoma skin cancer will

	be presented in **SFigures 11-16**, respectively. Summary statistics, crude incidence rates, and overall HRs for each additional outcome will be presented in **Table 1**. **Sensitivity analysis** • We will re-analyse invasive melanoma skin cancer such that we censor participants if and when they are diagnosed with an in situ melanoma skin cancer (without prior invasive melanoma skin cancer) (**Table 1**, **STable 2**, **SFigure 17**). **Retention, compliance, intake of off-trial vitamin D supplementation, and adverse events** We have described these outcomes in the mortality paper and therefore we will not include them here.
Significance level	We will use a significance level of 0.05. We will not adjust for multiple testing.
Summary of tables and figures	
Planned main table (Appendix C)	**Table 1:** Effect of vitamin D supplementation on incidence of cancer.
Planned main figures (Appendix D)	**Figure 1.** Participant flow for analyses of cancer incidence (CONSORT flow diagram). **Figure 2.** Effect of vitamin D supplementation on incidence of any cancer (panels A to C) and incidence of non-melanoma cancer (panels D to F) as functions of time since randomisation. **Figure 3.** Effect of vitamin D supplementation on incidence of any cancer (panel A) and incidence of non-melanoma cancer (panel B) for all participants and by selected baseline characteristics.
Planned supplementary tables (Appendix E)	**STable 1.** Baseline characteristics of participants included in the analysis of cancer incidence according to randomisation group. **STable 2.** Effect of supplementation with vitamin D on incidence of cancer. Percentage difference in cause-specific standardised cumulative incidence and time-varying hazard ratio at 2 and 5 years post-randomisation, and predicted overall hazard ratio.
Planned supplementary figures (Appendix F)	<u>Main outcome 1: any cancer</u> **SFigure 1.** Cause-specific cumulative incidence of cancer according to randomisation group (vitamin D, solid line; placebo, dashed line) and time since randomisation, stratified by (A) age, (B) sex, (C) body mass index, and (D) predicted 25(OH)D concentration. **SFigure 2.** Time-dependent effect of vitamin D supplementation on incidence of cancer according to age group (<70 years, ≥70 years). **SFigure 3.** Time-dependent effect of vitamin D supplementation on incidence of cancer according to sex (males, females).

	SFigure 4. Time-dependent effect of vitamin D supplementation on incidence of cancer according to body mass index (<25 kg/m^2, 25 to <30 kg/m^2; ≥30 kg/m^2). **SFigure 5.** Time-dependent effect of vitamin D supplementation on incidence of cancer according to predicted baseline 25(OH)D concentration (<50 nmol/L, ≥50 nmol/L). <u>Main outcome 2: any non-melanoma cancer</u> **SFigure 6.** Cause-specific cumulative incidence of non-melanoma cancer according to randomisation group (vitamin D, solid line; placebo, dashed line) and time since randomisation, stratified by (A) age, (B) sex, (C) body mass index, and (D) predicted 25(OH)D concentration. **SFigure 7.** Time-dependent effect of vitamin D supplementation on incidence of non-melanoma cancer according to age group (<70 years, ≥70 years). **SFigure 8.** Time-dependent effect of vitamin D supplementation on incidence of non-melanoma cancer according to sex (males, females). **SFigure 9.** Time-dependent effect of vitamin D supplementation on incidence of non-melanoma cancer according to body mass index (<25 kg/m^2, 25 to <30 kg/m^2; ≥30 kg/m^2). **SFigure 10.** Time-dependent effect of vitamin D supplementation on incidence of non-melanoma cancer according to predicted baseline 25(OH)D concentration (<50 nmol/L, ≥50 nmol/L). <u>Additional outcomes</u> **SFigure 11.** Effect of vitamin D supplementation on incidence of colorectal cancer as a function of time since randomisation. **SFigure 12.** Effect of vitamin D supplementation on incidence of lung cancer as a function of time since randomisation. **SFigure 13.** Effect of vitamin D supplementation on incidence of invasive breast cancer in females as a function of time since randomisation. **SFigure 14.** Effect of vitamin D supplementation on incidence of prostate cancer in males as a function of time since randomisation. **SFigure 15.** Effect of vitamin D supplementation on incidence of melanoma skin cancer as a function of time since randomisation. **SFigure 16.** Effect of vitamin D supplementation on incidence of invasive melanoma skin cancer as a function of time since randomisation. **SFigure 17.** Sensitivity analysis – Effect of vitamin D supplementation on incidence of invasive melanoma skin cancer as a function of time since randomisation, taking into account in situ melanoma skin cancer.

References

1. Waterhouse M, Baxter C, Duarte Romero B, et al. Predicting deseasonalised serum 25 hydroxy vitamin D concentrations in the D-Health Trial: An analysis using boosted regression trees. *Contemp Clin Trials.* 2021;104:106347.

2. Australian Institute of Health and Welfare. Cancer data in Australia. https://www.aihw.gov.au/reports/cancer/cancer-data-in-australia/contents/cancer-incidence-by-age-visualisation. Updated 31 August 2023. Accessed 22 May 2024.
3. Neale RE, Baxter C, Duarte Romero B, et al. The D-Health Trial: a randomised controlled trial of the effect of vitamin D on mortality. *Lancet Diabetes Endocrinol.* 2022;10(2):120-128.
4. Royston P, Parmar MK. Flexible parametric proportional-hazards and proportional-odds models for censored survival data, with application to prognostic modelling and estimation of treatment effects. *Statistics in medicine.* 2002;21(15):2175-2197.
5. Royston P, Parmar MK. An approach to trial design and analysis in the era of non-proportional hazards of the treatment effect. *Trials.* 2014;15:314.

Appendix A: Descriptive statistics

Table A1. Number of people with at least one cancer diagnosis and crude person-based incidence rates, according to sex.

	Males			Females		
Cancer type	**N**	**PYAR**	**Crude IR[1]**	**N**	**PYAR**	**Crude IR[1]**
Any cancer	1,735	53,187	3,262	905	46,878	1,931
Non-melanoma cancer	1,300	54,297	2,394	717	47,342	1,515
Colorectal cancer	138	56,760	243	98	48,716	201
Lung cancer	95	56,941	167	63	48,845	129
Invasive breast cancer (females only)	..	..	..	270	48,231	560
Prostate cancer (males only)	580	55,683	1,042	..	..	..
Melanoma skin cancer (invasive and in situ)[2]	480	47,775	1005	199	39,473	504
Invasive melanoma skin cancer[2]	204	48,441	421	82	39,734	206

[1] Crude person-based incidence rates are presented per 100,000 PYAR

[2] Restricted to participants who were not living in Western Australia at baseline

Abbreviations: IR, incidence rate; PYAR, person years at risk

Appendix B: Detectable effect size

The analysis of cancer incidence will include 21,308 participants (n=10,648 in the placebo group; n=10,660 in the vitamin D group). We are interested in estimating the difference that we could detect with 80% power, assuming the use of a two-sided test with significance level set at α=0.05.

Table B1 shows incidence rates for all cancers combined in Australia in 2018, according to sex and age group.[2] To estimate how many D-Health placebo participants will experience at least one cancer diagnosis during 5 years of follow-up, we used the data from Table B1, and made the following assumptions:

1. The overwhelming majority of people in Australia who were diagnosed with cancer in 2018 had only one cancer diagnosis in that year (i.e., the cancer-based incidence rates in Table B1 are reasonable approximations to person-based incidence rates); and
2. All D-Health participants had a follow-up of 5 years.

The results of our calculations are shown in Table B2.

Table B1. Cancer incidence in Australia in 2018 according to sex and age.

Age group	Incidence rate (per 100,000 persons)	
	Males	Females
60-64 years	1505.7	1045.5
65-69 years	2157.7	1328.1
70-74 years	2643.8	1617.7
75-84 years	3283.7	1924.3

Table B2. Estimated number of D-Health placebo participants[a] who will experience at least one cancer diagnosis during follow-up, according to sex, age group, and overall.

	Males		Females		Overall	
Age Group	**N**	**n[b]**	**N**	**n[b]**	**N**	**n[b]**
60-64 years	1,227	92	1,397	73	2,624	165
65-69 years	1,434	155	1,480	98	2,914	253
70-74 years	1,665	220	1,228	99	2,893	319
75-84 years	1,437	236	780	75	2,217	311
Total	**5,763**	**703**	**4,885**	**346**	**10,648**	**1,049**

[a] Restricted to N=10,648 D-Health placebo participants who are eligible for inclusion in the analysis of cancer incidence.

[b] n = estimated number of people with at least one cancer diagnosis during follow-up

Hence, we estimate that the probability of having no cancer diagnoses over 5 years of follow-up amongst placebo participants included in the analysis of cancer incidence is (10,648 – 1,049)/10,648 = 0.90. Given our sample sizes and a "survival" probability of 0.90 in the placebo group, a two sided log-rank test with power of 0.8 and significance level of 0.05 could detect a hazard ratio of 0.88.

APPENDICES C-F CONTAIN SAMPLE TABLES AND FIGURES BASED ON "FAKE" DATA*

To generate the "fake" data, we removed the true randomisation and participant identification codes from the original dataset, and then randomly assigned participants to two groups of equal size. There is no relationship between the new groups and the true treatment allocation.

* The one exception is the CONSORT diagram (Appendix D, Figure 1), which is based on real data.

Appendix C – PLANNED MAIN TABLE

Table 1: Effect of vitamin D supplementation on incidence of cancer.

	N (%) [b]		Crude IR per 100,000 PYAR [c]		
Outcome [a]	**Vitamin D (N = 10,655)**	**Placebo (N = 10,653)**	**Vitamin D**	**Placebo**	**HR (95% CI) [d]**
Any cancer	1,280 (12.0)	1,360 (12.8)	2,550	2,727	0.94 (0.87 to 1.01)
Non-melanoma cancer	984 (9.2)	1,033 (9.7)	1,930	2,039	0.95 (0.87 to 1.04)
Specific cancers					
Colorectal cancer	114 (1.1)	122 (1.1)	216	232	0.94 (0.73 to 1.22)
Lung cancer	77 (0.7)	81 (0.8)	145	153	0.96 (0.70 to 1.31)
Invasive breast cancer (females only) [e]	131 (2.7)	139 (2.9)	540	580	0.93 (0.73 to 1.18)
Prostate cancer (males only) [f]	277 (4.8)	303 (5.2)	995	1,088	0.92 (0.78 to 1.08)
Melanoma skin cancer (invasive and in situ) [g]	320 (3.6)	359 (4.0)	734	823	0.89 (0.77 to 1.04)
Invasive melanoma skin cancer [g]	137 (1.5)	149 (1.7)	311	338	0.93 (0.73 to 1.17)
Invasive melanoma skin cancer (sensitivity analysis) [g,h]	125 (1.4)	142 (1.6)	284	322	0.89 (0.70 to 1.13)

[a] Each outcome is first diagnosis following randomisation of the specified cancer type. Outcomes were ascertained using the ICD-O-3 code for the anatomical site of origin of the cancer at diagnosis. Any cancer: C00-C80 (excluding C44 with a histology code starting with 805-811 [cutaneous SCC/BCC]). Non-melanoma cancer: C00-C80 (excluding C44 with a histology code starting with 805-811 [cutaneous SCC/BCC] or 872-879 [cutaneous melanoma]). Colorectal cancer: C18-C20, and C21.8. Lung cancer: C33-C34. Invasive breast cancer: C50 and behaviour not categorised as 'in situ' or 'uncertain whether benign or malignant'. Prostate cancer: C61.9. Melanoma skin cancer: C44 with a histology code starting with 872-879 and behaviour code of 2 (in situ) or 3 (invasive). Invasive melanoma skin cancer: C44 with a histology code starting with 872-879 and behaviour code of 3 (invasive).

[b] The number (%) of people with at least one diagnosis.

[c] Crude person-based incidence rate

[d] Overall HRs estimated using flexible parametric survival models that included randomisation group, age, sex, and state of residence at baseline.

[e] Analysis restricted to females (vitamin D, n = 4,907; placebo, n = 4,871).

[f] Analysis restricted to males (vitamin D, n = 5,748; placebo, n = 5,782).

[g] Analysis restricted to 17,882 participants not living in Western Australia at baseline (vitamin D, n = 8,922; placebo, n = 8,960).

[h] Participants censored if and when they were diagnosed with an in situ melanoma skin cancer (without prior diagnosis of invasive melanoma skin cancer).

Abbreviation: BCC, basal cell carcinoma; CI, confidence interval; HR, hazard ratio; IR, incidence rate; PYAR, person years at risk; SCC, squamous cell carcinoma

Appendix D – PLANNED MAIN FIGURES

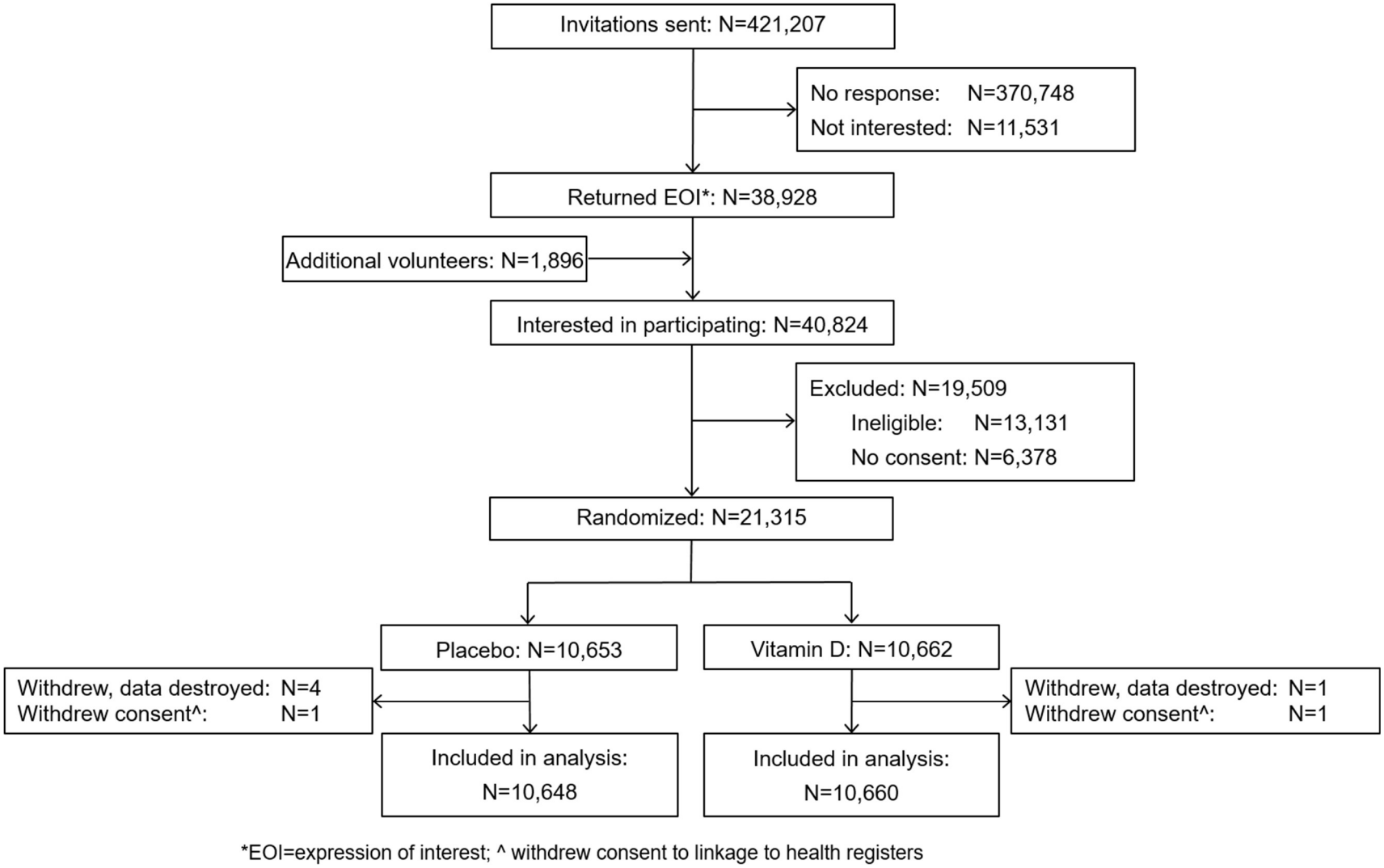


Figure 1. Participant flow for analyses of cancer incidence (CONSORT flow diagram).

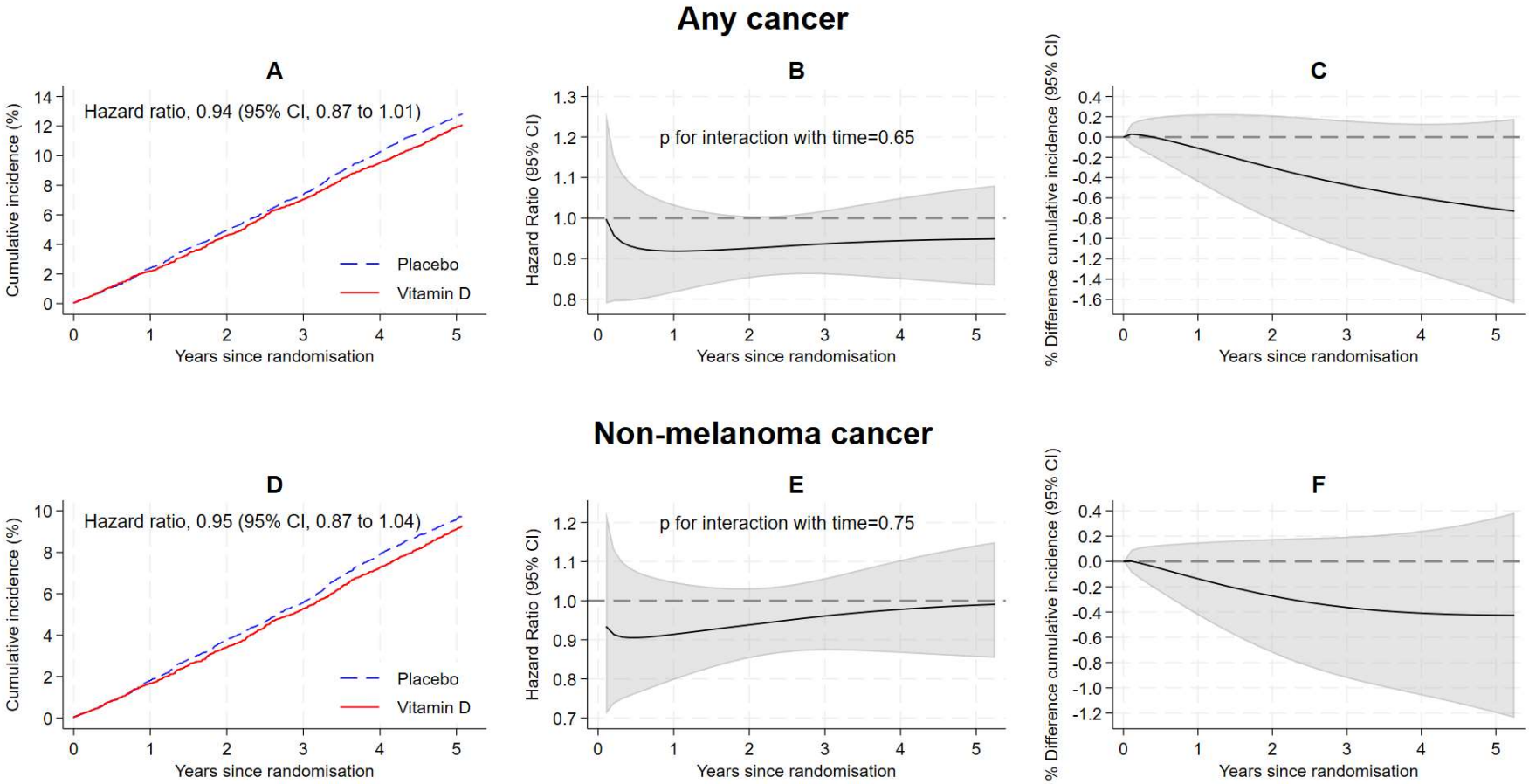


Figure 2. Effect of vitamin D supplementation on incidence of any cancer (panels A to C) and incidence of non-melanoma cancer (panels D to F) as functions of time since randomisation.

Panels A and D show cause-specific cumulative incidence according to randomisation group, estimated using Aalen-Johansen methods, and the overall HR.

Panels B and E show the time-varying HR.

Panels C and F show the percentage difference (vitamin D – placebo) in cause-specific standardised cumulative incidence.

Outcomes were ascertained using the ICD-O-3 code for the anatomical site of origin of the cancer at diagnosis. Any cancer: C00-C80 (excluding C44 with a histology code starting with 805-811 [cutaneous SCC/BCC]). Non-melanoma cancer: C00-C80 (excluding C44 with a histology code starting with 805-811 [cutaneous SCC/BCC] or 872-879 [cutaneous melanoma]).

Analyses included n=21,308 participants. Estimates (vitamin D versus placebo) are from flexible parametric survival models that included randomisation group, age, sex, and state of residence at baseline. Models producing time-varying estimates also included an interaction between randomisation group and time since randomisation. The interaction with time was assessed using a likelihood ratio test that compared models with and without the interaction term.

Death without prior cancer diagnosis was treated as a competing risk. Cause-specific standardised cumulative incidences were standardised to the distribution of age, and state of residence at baseline in the entire cohort.

Abbreviations: BCC, basal cell carcinoma; CI, confidence interval; HR, hazard ratio; ICD-O-3, International Classification of Diseases for Oncology Third Edition; SCC, squamous cell carcinoma

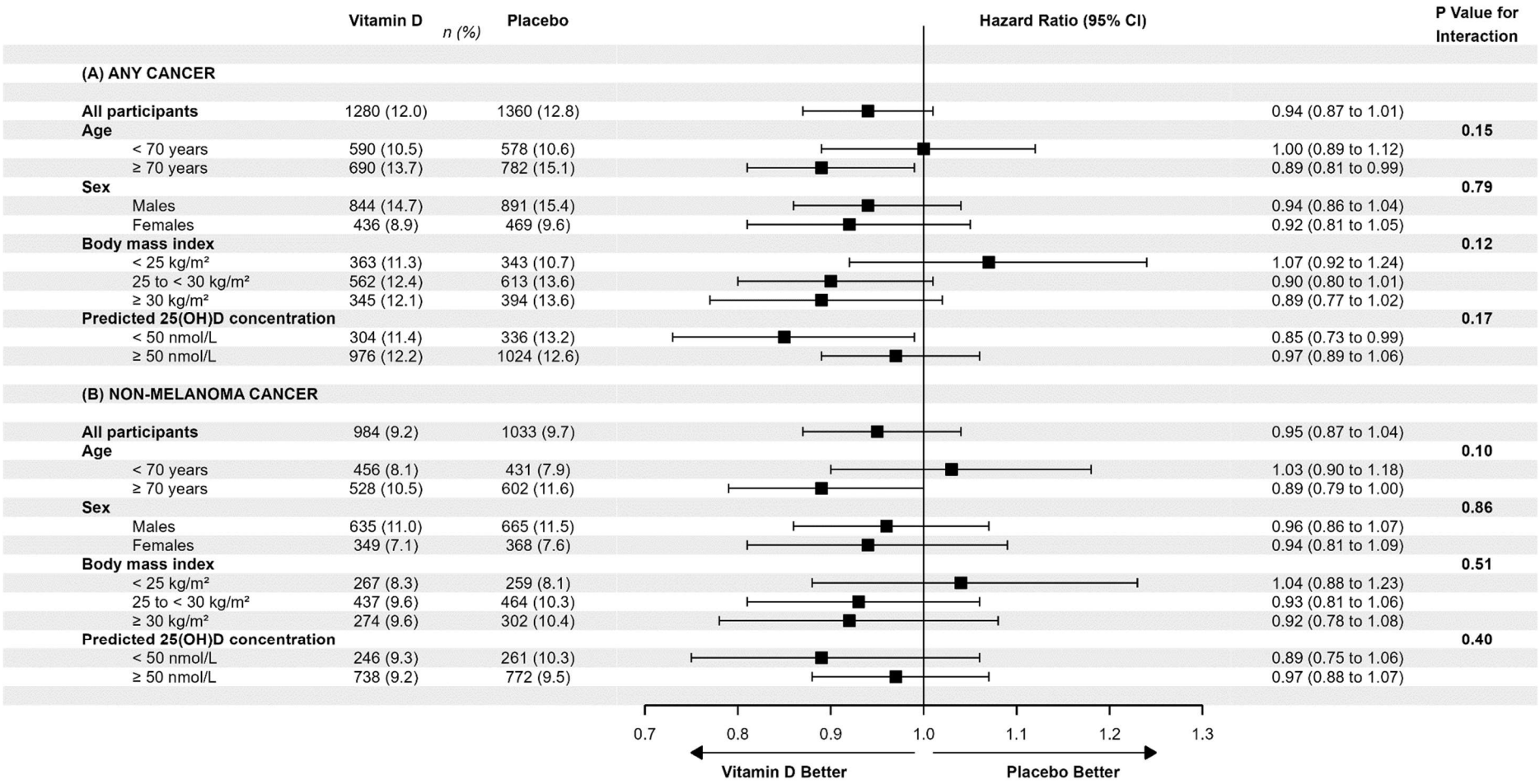


Figure 3. Effect of vitamin D supplementation on incidence of any cancer (panel A) and incidence of non-melanoma cancer (panel B) for all participants and by selected baseline characteristics.

Summary statistics show the number (%) of participants with at least one cancer diagnosis according to randomisation group; denominators can be found in STable 1. Overall hazard ratios (vitamin D versus placebo) are from flexible parametric survival models. All models included randomisation group, age, sex, and state of residence at baseline. Models producing estimates by levels of age, sex, body mass index, and predicted 25(OH)D concentration, included the characteristic of interest 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 E – PLANNED SUPPLEMENTARY TABLES

STable 1. Baseline characteristics of participants included in the analysis of cancer incidence according to randomisation group.

Characteristic	N (%) Vitamin D (N = 10,655)	Placebo (N = 10,653)
Age (years)		
60-64	2,668 (25.0)	2,583 (24.2)
65-69	2,944 (27.6)	2,890 (27.1)
70-74	2,859 (26.8)	2,936 (27.6)
≥ 75	2,184 (20.5)	2,244 (21.1)
Sex		
Males	5,748 (53.9)	5,782 (54.3)
Females	4,907 (46.1)	4,871 (45.7)
Body mass index (kg/m²)		
< 25	3,209 (30.3)	3,207 (30.3)
25 to < 30	4,535 (42.8)	4,494 (42.4)
≥ 30	2,851 (26.9)	2,893 (27.3)
Missing	*60*	*59*
Predicted 25(OH)D concentration (nmol/L)		
< 50	2,656 (24.9)	2,543 (23.9)
≥ 50	7,999 (75.1)	8,110 (76.1)
Ancestry		
British/European	9,726 (92.9)	9,722 (92.9)
Australian/New Zealander	358 (3.4)	369 (3.5)
Asian	125 (1.2)	116 (1.1)
Indigenous	78 (0.7)	73 (0.7)
Mixed/other	184 (1.8)	181 (1.7)
Missing	*184*	*192*
Highest qualification obtained		
None	1,050 (10.0)	1,094 (10.4)
School or intermediate certificate	1,779 (16.9)	1,776 (16.9)
Higher school or leaving certificate	1,469 (13.9)	1,496 (14.2)
Apprenticeship or certificate	3,541 (33.6)	3,490 (33.2)
University degree or higher	2,692 (25.6)	2,658 (25.3)
Missing	*124*	*139*
Alcohol consumption (drinks/week)		
< 1	2,513 (24.5)	2,533 (24.7)
1 to 7	4,565 (44.6)	4,539 (44.3)
> 7 to 14	1,881 (18.4)	1,873 (18.3)
> 14	1,281 (12.5)	1,298 (12.7)
Missing	*415*	*410*
Smoking history		
Never	5,820 (55.1)	5,771 (54.6)
Ex-smoker	4,305 (40.8)	4,332 (41.0)
Current	432 (4.1)	463 (4.4)
Missing	*98*	*87*
Self-rated overall health		
Excellent or very good	5,868 (55.9)	5,778 (55.2)
Good	3,741 (35.7)	3,776 (36.1)
Fair or poor	884 (8.4)	916 (8.7)
Missing	*162*	*183*

STable 2. Effect of supplementation with vitamin D on incidence of cancer. Percentage difference in cause-specific standardised cumulative incidence and time-varying hazard ratio at 2 and 5 years post-randomisation, and predicted overall hazard ratio.[a,b,c]

Years since randomisation	% Difference cumulative incidence (95% CI)	Hazard Ratio (95% CI)
Any cancer		
2	-0.30 (-0.82 to 0.21)	0.93 (0.85 to 1.01)
5	-0.71 (-1.58 to 0.16)	0.95 (0.84 to 1.08)
Overall HR	..	0.94 (0.87 to 1.01)
Any non-melanoma cancer		
2	-0.27 (-0.72 to 0.18)	0.94 (0.85 to 1.03)
5	-0.42 (-1.20 to 0.35)	0.99 (0.86 to 1.14)
Overall HR	..	0.95 (0.87 to 1.04)
Colorectal cancer		
2	-0.05 (-0.20 to 0.11)	0.84 (0.62 to 1.15)
5	-0.06 (-0.34 to 0.21)	1.09 (0.68 to 1.76)
Overall HR	..	0.94 (0.73 to 1.22)
Lung cancer		
2	0.01 (-0.11 to 0.12)	1.05 (0.68 to 1.61)
5	-0.03 (-0.26 to 0.20)	0.86 (0.51 to 1.45)
Overall HR	..	0.96 (0.70 to 1.31)
Invasive breast cancer (females only)		
2	-0.35 (-0.76 to 0.06)	1.09 (0.81 to 1.46)
5	-0.19 (-0.84 to 0.45)	1.14 (0.76 to 1.71)
Overall HR	..	0.93 (0.73 to 1.18)
Prostate cancer (males only)		
2	-0.04 (-0.47 to 0.39)	0.91 (0.76 to 1.09)
5	-0.41 (-1.20 to 0.38)	0.86 (0.66 to 1.13)
Overall HR	..	0.92 (0.78 to 1.08)
Any melanoma skin cancer (invasive and in situ)		
2	-0.05 (-0.36 to 0.26)	0.88 (0.75 to 1.04)
5	-0.41 (-0.96 to 0.15)	0.84 (0.66 to 1.07)
Overall HR	..	0.89 (0.77 to 1.04)
Invasive melanoma skin cancer		
2	0.03 (-0.17 to 0.24)	0.92 (0.71 to 1.18)
5	-0.11 (-0.48 to 0.25)	0.81 (0.55 to 1.18)
Overall HR	..	0.93 (0.73 to 1.17)
Invasive melanoma skin cancer (sensitivity analysis)[d]		
2	0.02 (-0.19 to 0.23)	0.87 (0.67 to 1.13)
5	-0.17 (-0.52 to 0.18)	0.76 (0.51 to 1.12)
Overall HR	..	0.89 (0.70 to 1.13)

[a] Each outcome is the first diagnosis following randomisation of the specified cancer type. Outcomes were ascertained using the ICD-O-3 code for the anatomical site of origin of the cancer at diagnosis. Any cancer: C00-C80 (excluding C44 with histology code starting with 805-811 [cutaneous SCC/BCC]). Any non-melanoma cancer: C00-C80 (excluding C44 with a histology code starting with 805-811 [cutaneous SCC/BCC] or 872-879 [cutaneous melanoma]). Colorectal cancer: C18-C20, and C21.8. Lung cancer: C33-C34. Invasive breast cancer: C50 and behaviour not categorised as 'in situ' or 'uncertain whether benign or malignant'. Prostate cancer: C61.9. Any melanoma skin cancer: C44 with histology code starting with 872-879 and behaviour code of 2 (in situ) or 3 (invasive). Invasive melanoma skin cancer: C44 with histology code starting with 872-879 and behaviour code of 3 (invasive).

[b] 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 estimates (i.e., estimates at 2

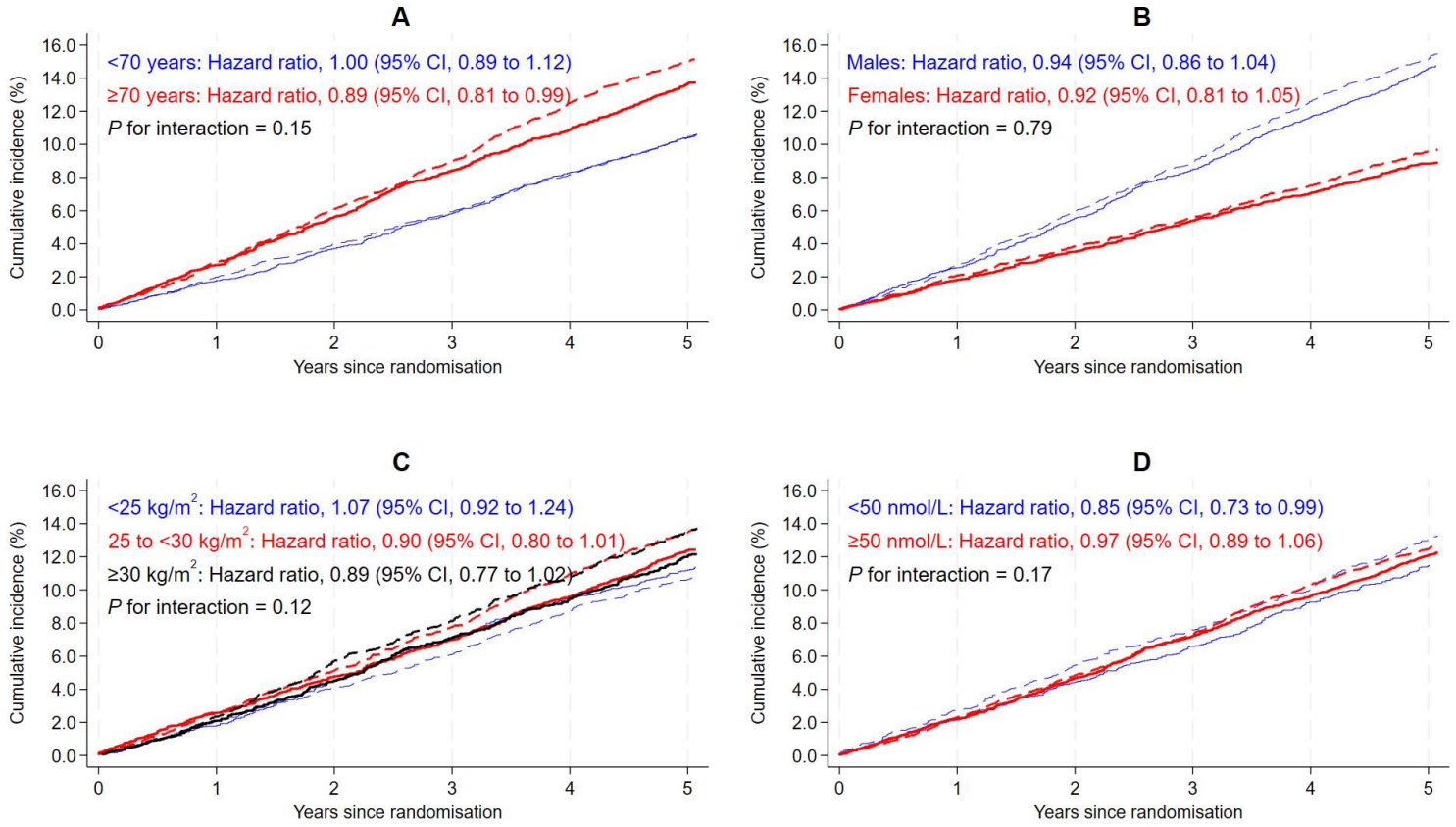
and 5 years post randomisation) were predicted using a model that also included an interaction between randomisation group and time since randomisation. Cause-specific standardised cumulative incidences were estimated treating death (without prior cancer of interest) as a competing risk, and probabilities were standardised to the distribution of age, sex, and state of residence at baseline in the entire cohort.

[c] Analyses of any cancer, any non-melanoma cancer, colorectal cancer, and lung cancer included n=21,308 participants (vitamin D, n=10,655; placebo, n=10,653). Analyses of invasive breast cancer and prostate cancer included n=9,778 females (vitamin D, n=4,907; placebo, n=4,871) and n=11,530 males (vitamin D, n=5,748; placebo, n=5,782), respectively. Analyses of melanoma skin cancer included 17,822 participants who were not living in Western Australia at baseline (vitamin D, n=8,922; placebo, n = 8,960).

[d] Participants were censored if and when they were diagnosed with an in situ melanoma (without prior diagnosis of invasive melanoma).

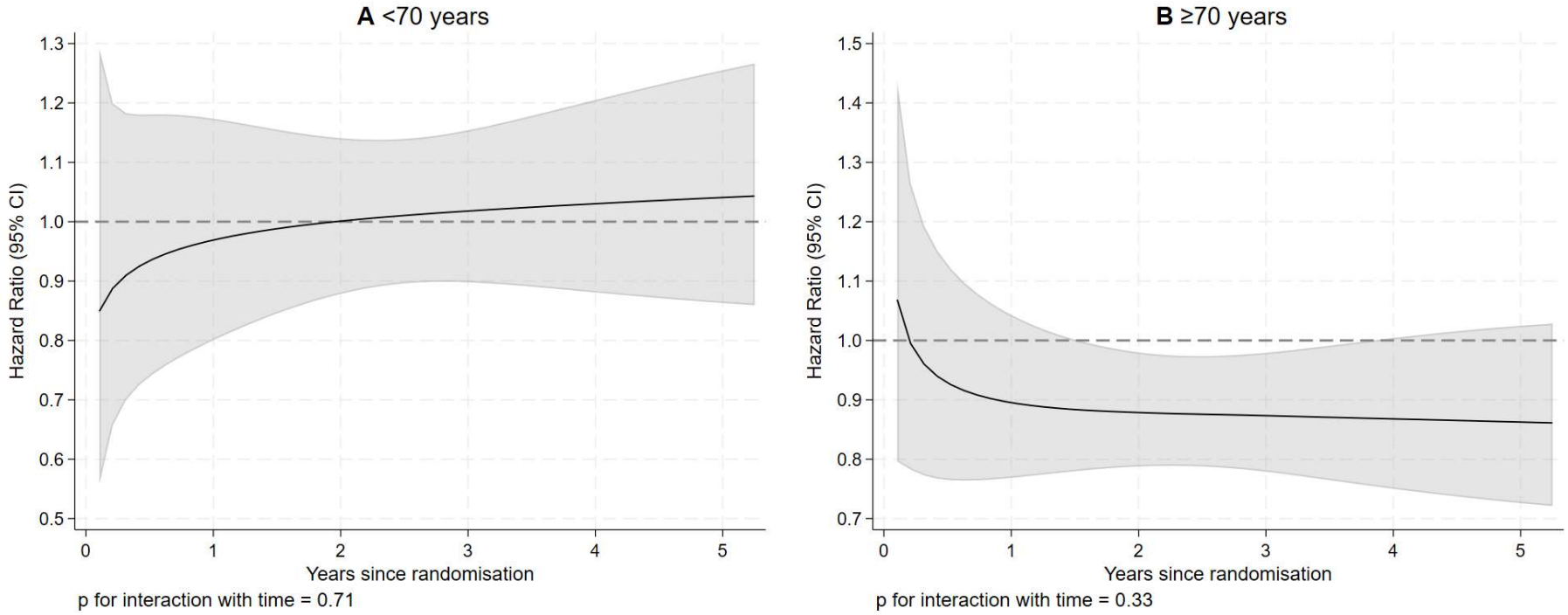
Abbreviations: BCC, basal cell carcinoma; CI, confidence interval; HR, hazard ratio; ICD-O-3, International Classification of Diseases for Oncology Third Edition; SCC, squamous cell carcinoma

Appendix F – PLANNED SUPPLEMENTARY FIGURES



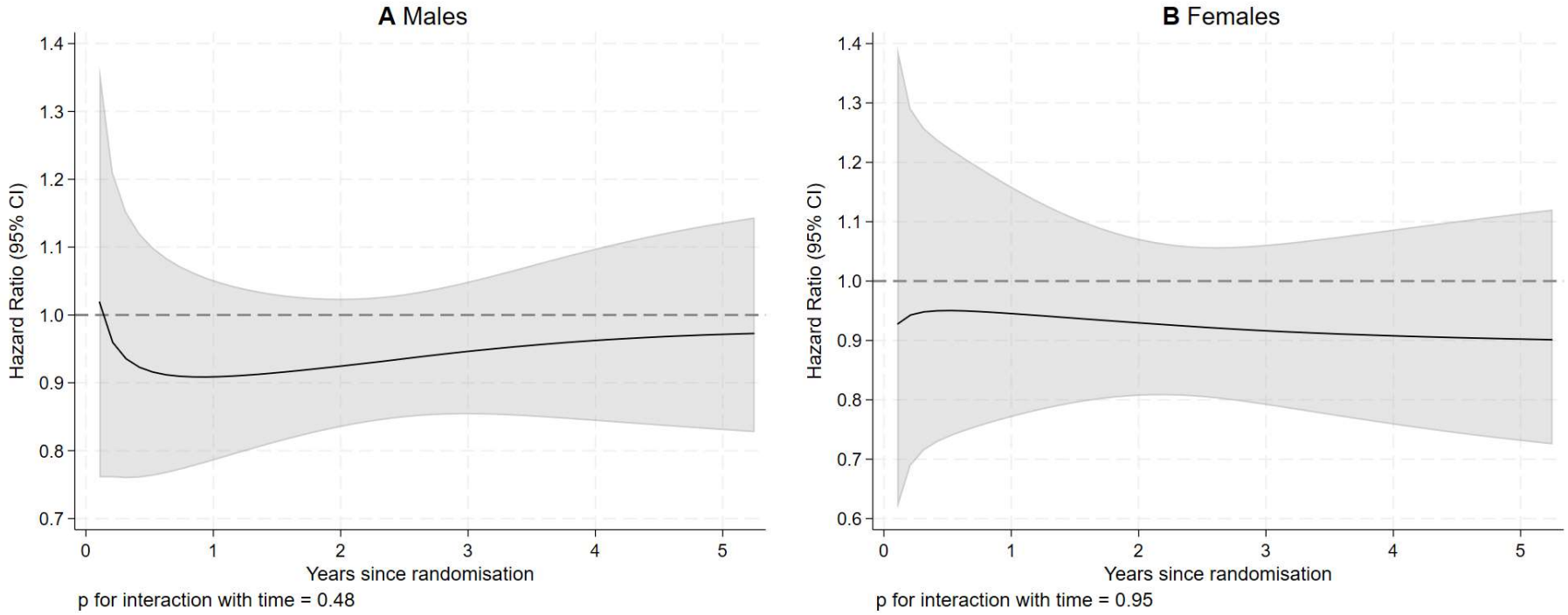
SFigure 1. Cause-specific cumulative incidence of cancer according to randomisation group (vitamin D, solid line; placebo, dashed line) and time since randomisation, stratified by (A) age, (B) sex, (C) body mass index, and (D) predicted 25(OH)D concentration.

The outcome is first diagnosis following randomisation of cancer of any type, excluding cutaneous squamous cell carcinoma and cutaneous basal cell carcinoma (ICD-O-3 codes C00-C80, excluding C44 with a histology code starting with 805-811). Curves estimated using Aalen-Johansen methods, treating death without prior cancer diagnosis as a competing risk. The overall hazard ratios (vitamin D versus placebo) were 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 the baseline characteristic of interest. P value for the interaction is from a likelihood ratio test comparing models with and without the interaction term. Abbreviations: CI, confidence interval; ICD-O-3, International Classification of Diseases for Oncology Third Edition



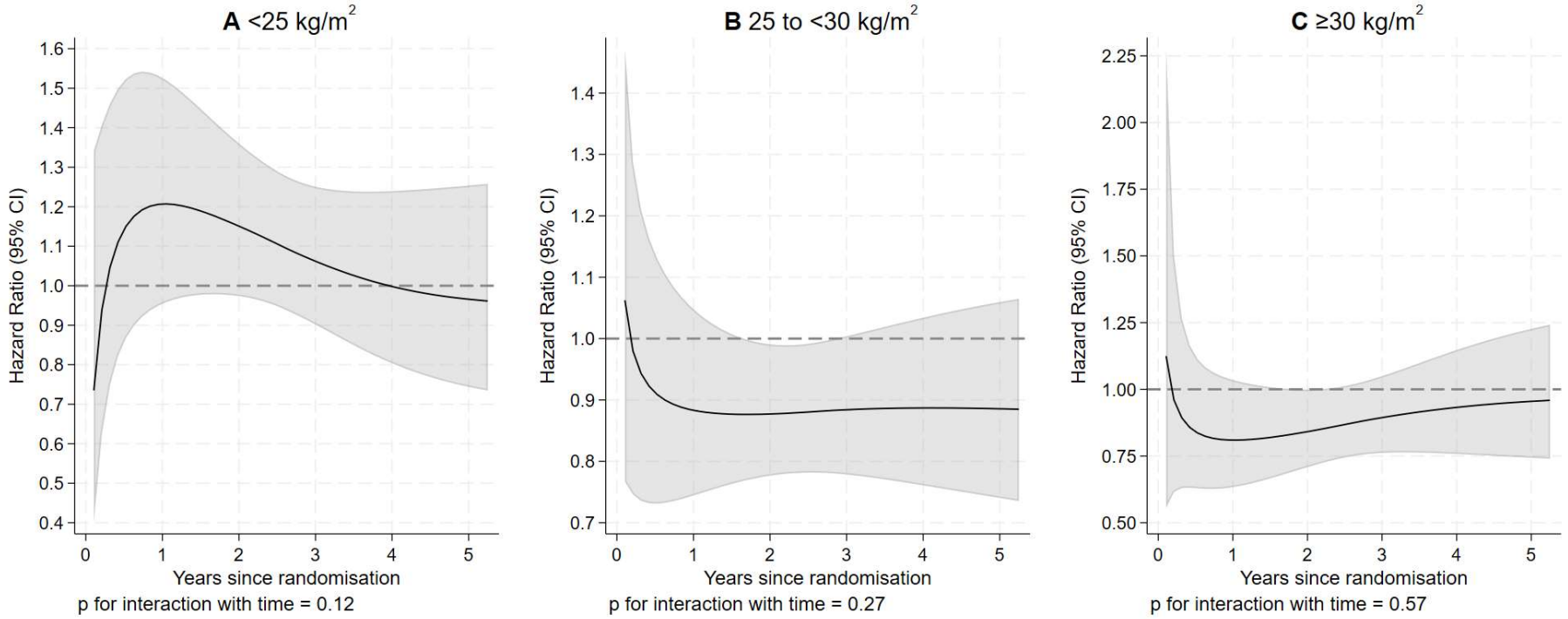
SFigure 2. Time-dependent effect of vitamin D supplementation on incidence of cancer according to age group (<70 years, ≥70 years).

The outcome is first diagnosis following randomisation of cancer of any type, excluding cutaneous squamous cell carcinoma and cutaneous basal cell carcinoma (ICD-O-3 codes C00-C80, excluding C44 with a histology code starting with 805-811). Time-varying hazard ratios (vitamin D vs placebo) are from flexible parametric survival models that included randomisation group, sex, state of residence at baseline, and an interaction between randomisation group and time since randomisation. Models were fitted separately within each subgroup of age. The p value for the interaction with time is from a likelihood ratio test comparing models with and without the interaction term. Abbreviations: CI, confidence interval; ICD-O-3, International Classification of Diseases for Oncology Third Edition



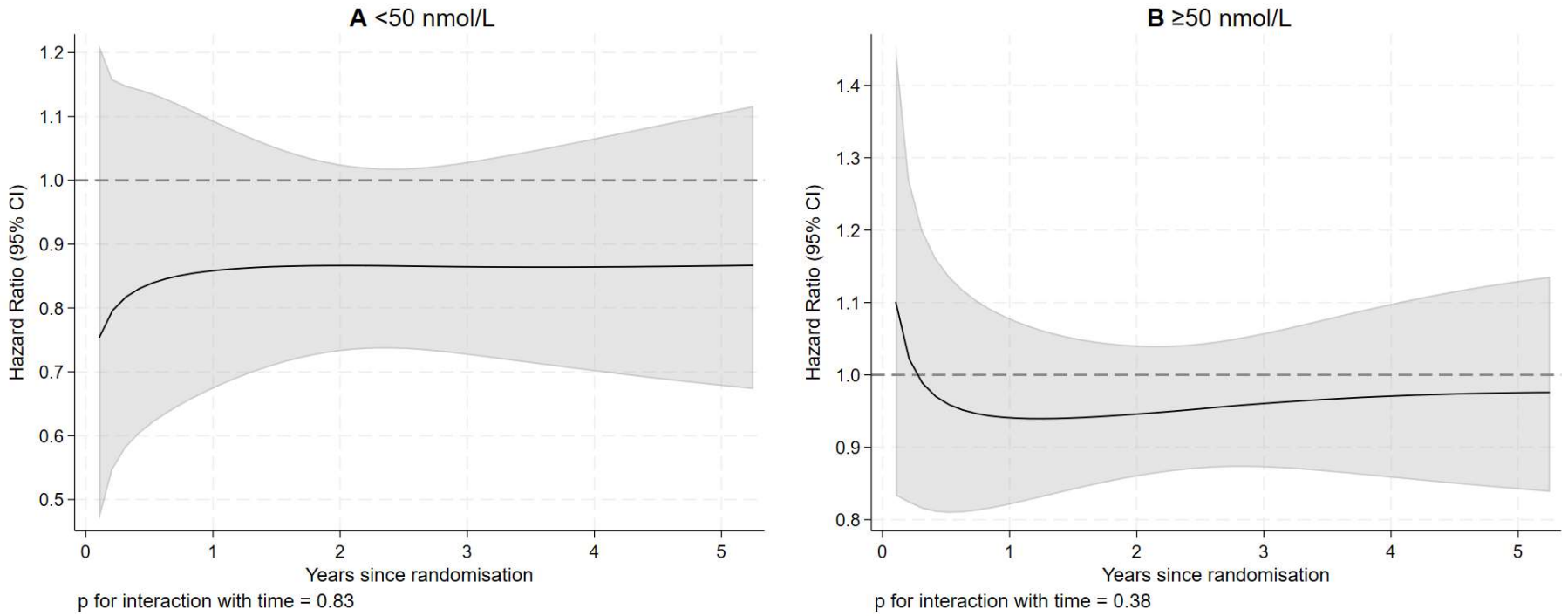
SFigure 3. Time-dependent effect of vitamin D supplementation on incidence of cancer according to sex (males, females).

The outcome is first diagnosis following randomisation of cancer of any type, excluding cutaneous squamous cell carcinoma and cutaneous basal cell carcinoma (ICD-O-3 codes C00-C80, excluding C44 with a histology code starting with 805-811). Time-varying hazard ratios (vitamin D vs placebo) are from flexible parametric survival models that included randomisation group, age, state of residence at baseline, and an interaction between randomisation group and time since randomisation. Models were fitted separately within males and females. The p value for the interaction with time is from a likelihood ratio test comparing models with and without the interaction term. Abbreviations: CI, confidence interval; ICD-O-3, International Classification of Diseases for Oncology Third Edition



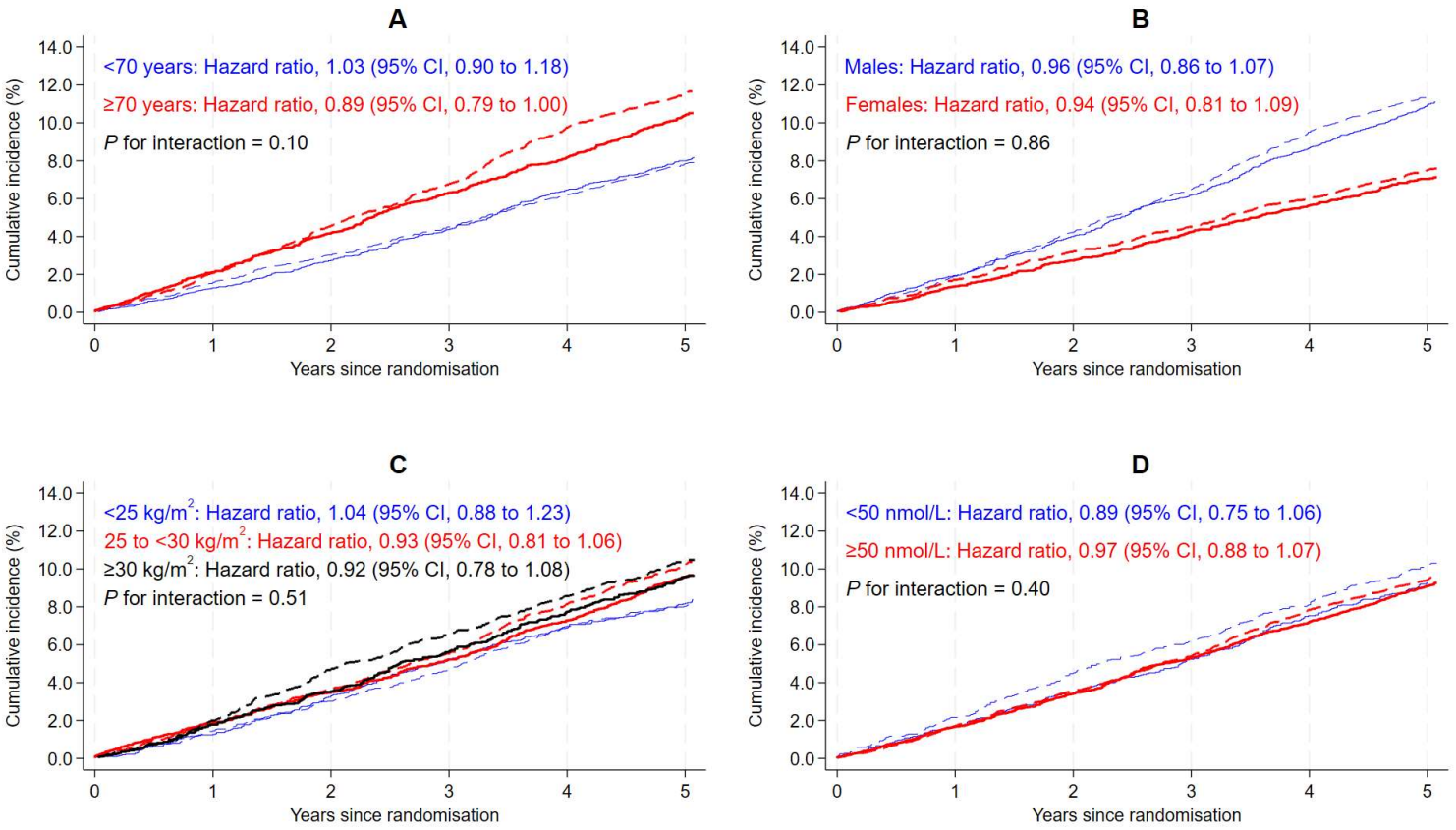
SFigure 4. Time-dependent effect of vitamin D supplementation on incidence of cancer according to body mass index (<25 kg/m^2, 25 to <30 kg/m^2; ≥30 kg/m^2).

The outcome is first diagnosis following randomisation of cancer of any type, excluding cutaneous squamous cell carcinoma and cutaneous basal cell carcinoma (ICD-O-3 codes C00-C80, excluding C44 with a histology code starting with 805-811). Time-varying hazard ratios (vitamin D vs placebo) are from flexible parametric survival models that included randomisation group, age, sex, state of residence at baseline, and an interaction between randomisation group and time since randomisation. Models were fitted separately within each subgroup of body mass index. The p value for the interaction with time is from a likelihood ratio test comparing models with and without the interaction term. Abbreviations: CI, confidence interval; ICD-O-3, International Classification of Diseases for Oncology Third Edition



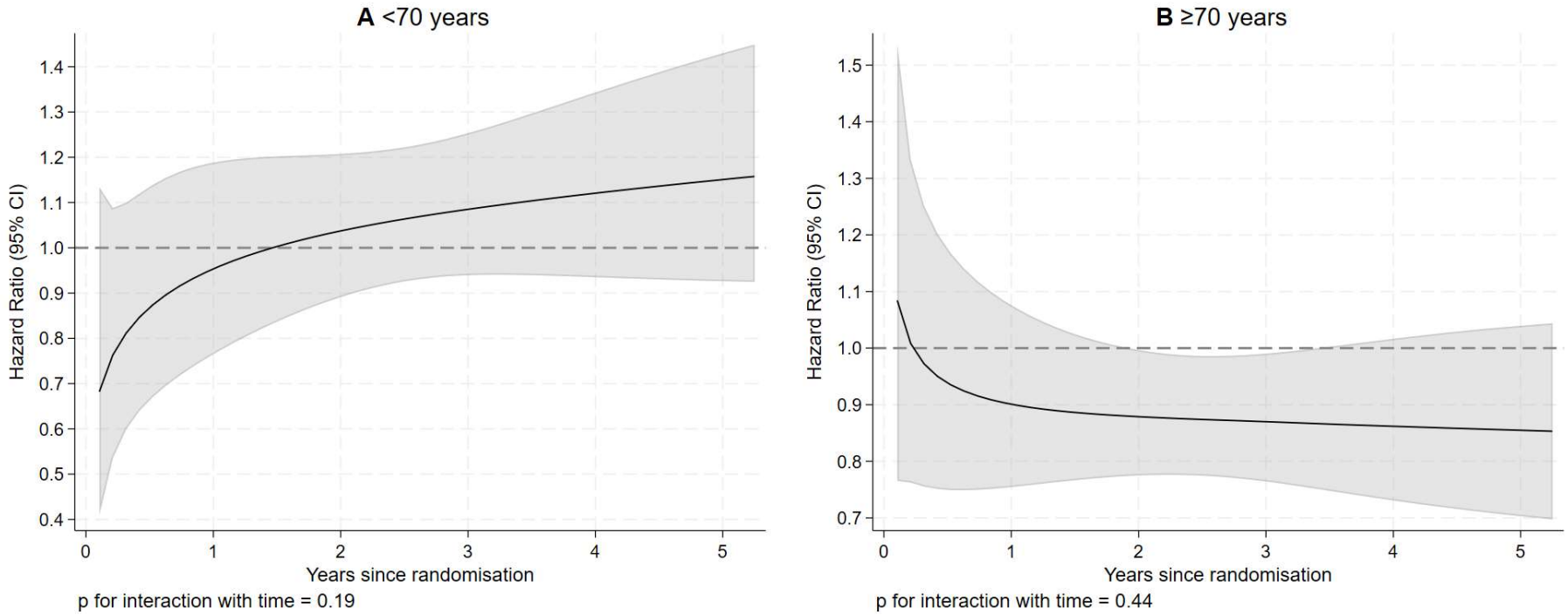
SFigure 5. Time-dependent effect of vitamin D supplementation on incidence of cancer according to predicted baseline 25(OH)D concentration (<50 nmol/L, ≥50 nmol/L).

The outcome is first diagnosis following randomisation of cancer of any type, excluding cutaneous squamous cell carcinoma and cutaneous basal cell carcinoma (ICD-O-3 codes C00-C80, excluding C44 with a histology code starting with 805-811). Time-varying hazard ratios (vitamin D vs placebo) are from flexible parametric survival models that included randomisation group, age, sex, state of residence at baseline, and an interaction between randomisation group and time since randomisation. Models were fitted separately within each subgroup of predicted baseline 25(OH)D concentration. The p value for the interaction with time is from a likelihood ratio test comparing models with and without the interaction term. Abbreviations: CI, confidence interval; ICD-O-3, International Classification of Diseases for Oncology Third Edition



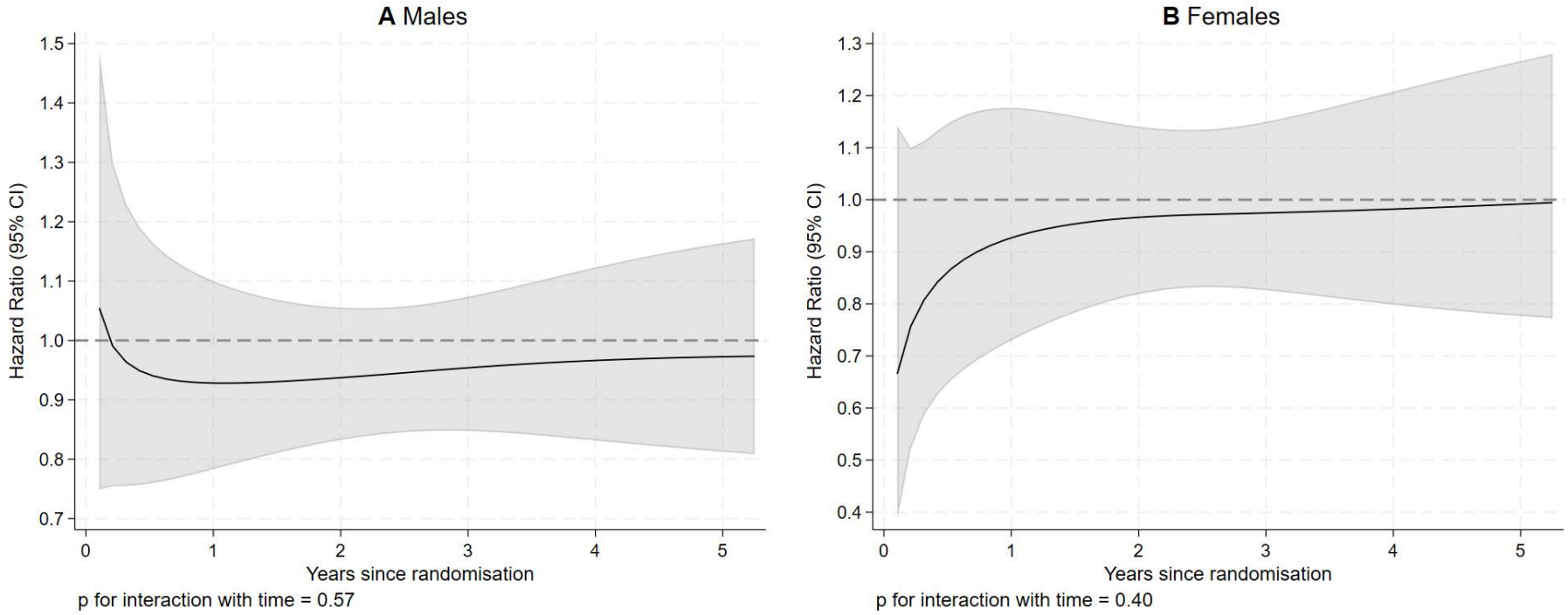
SFigure 6. Cause-specific cumulative incidence of **non-melanoma cancer** according to randomisation group (vitamin D, solid line; placebo, dashed line) and time since randomisation, stratified by (A) age, (B) sex, (C) body mass index, and (D) predicted 25(OH)D concentration.

The outcome is first diagnosis following randomisation of cancer of any type, excluding cutaneous melanoma, cutaneous squamous cell carcinoma, and cutaneous basal cell carcinoma (ICD-O-3 codes C00-C80, excluding C44 with a histology code starting with 805-811 or 872-879). Curves estimated using Aalen-Johansen methods, treating death without prior diagnosis non-melanoma cancer as a competing risk. The overall hazard ratios (vitamin D versus placebo) were 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 the baseline characteristic of interest. P value for the interaction is from a likelihood ratio test comparing models with and without the interaction term. Abbreviations: CI, confidence interval; ICD-O-3, International Classification of Diseases for Oncology Third Edition



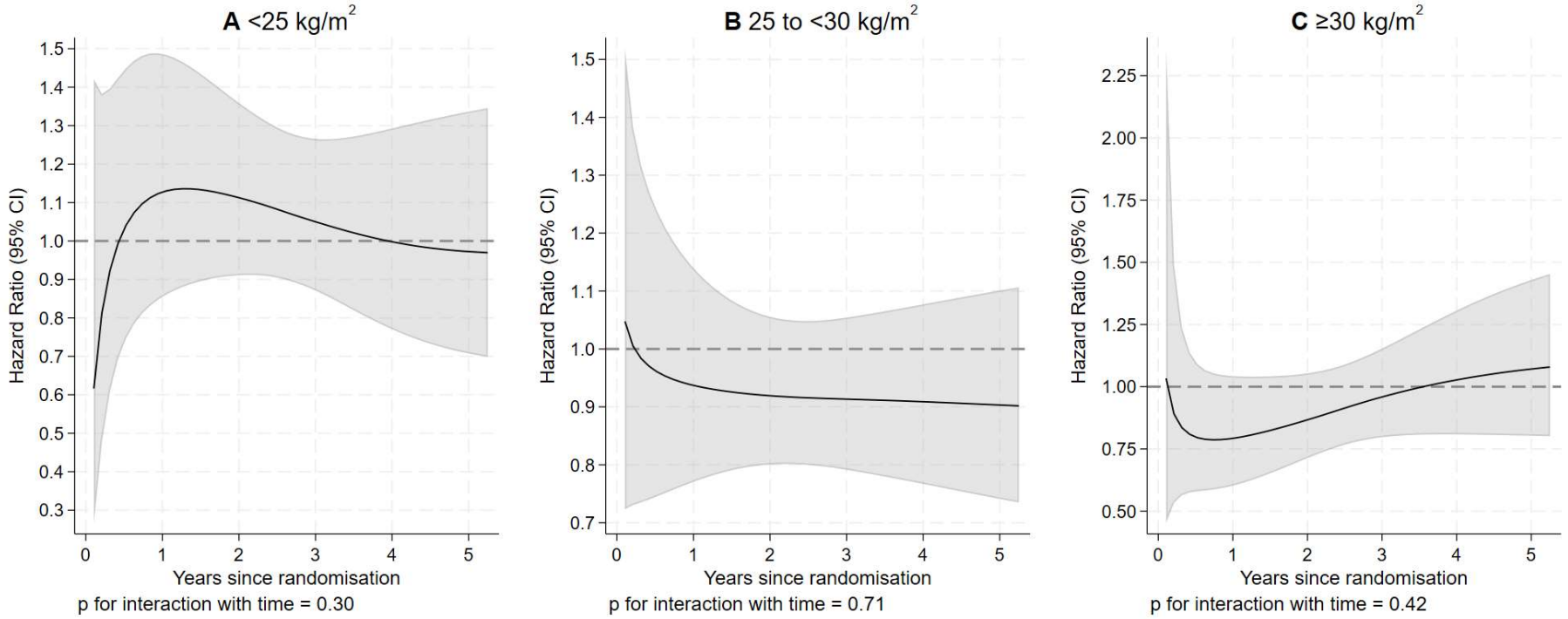
SFigure 7. Time-dependent effect of vitamin D supplementation on incidence of **non-melanoma cancer** according to age group (<70 years, ≥70 years).

The outcome is first diagnosis following randomisation of cancer of any type, excluding cutaneous melanoma, cutaneous squamous cell carcinoma, and cutaneous basal cell carcinoma (ICD-O-3 codes C00-C80, excluding C44 with a histology code starting with 805-811 or 872-879). Time-varying hazard ratios (vitamin D vs placebo) are from flexible parametric survival models that included randomisation group, sex, state of residence at baseline, and an interaction between randomisation group and time since randomisation. Models were fitted separately within each subgroup of age. The p value for the interaction with time is from a likelihood ratio test comparing models with and without the interaction term. Abbreviations: CI, confidence interval; ICD-O-3, International Classification of Diseases for Oncology Third Edition



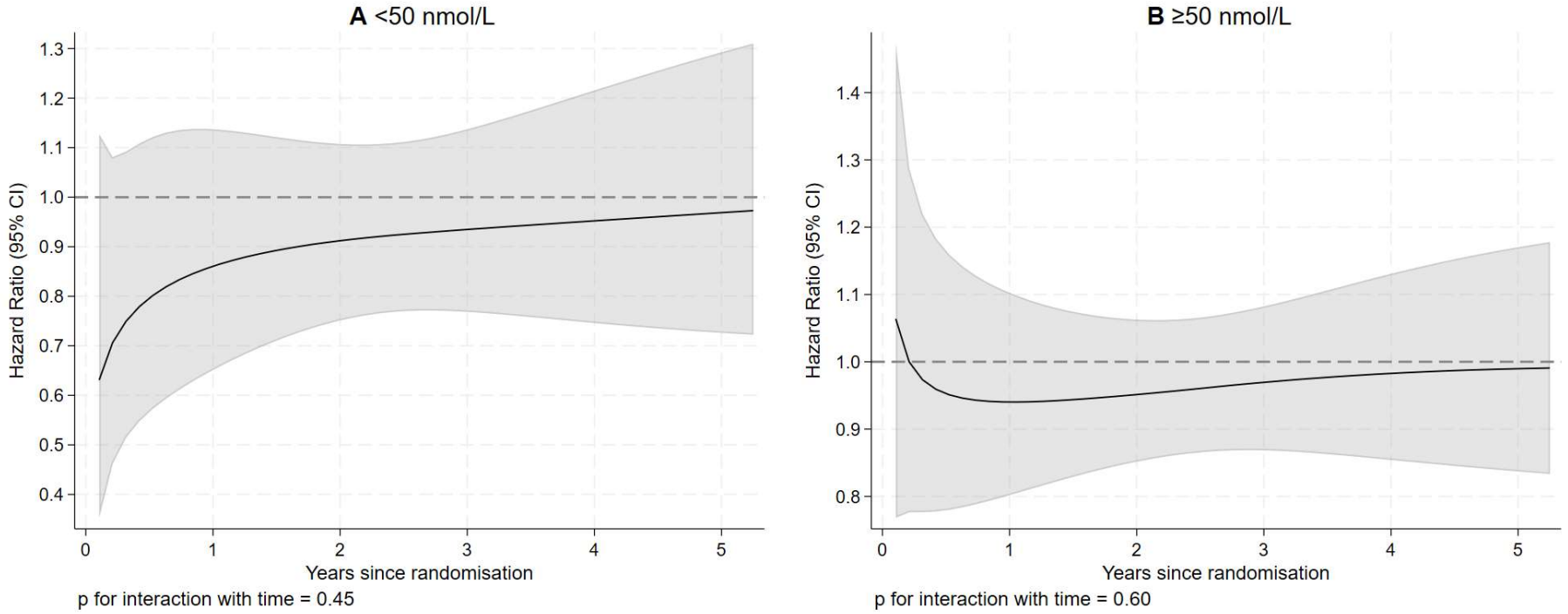
SFigure 8. Time-dependent effect of vitamin D supplementation on incidence of **non-melanoma cancer** according to sex (males, females).

The outcome is first diagnosis following randomisation of cancer of any type, excluding cutaneous melanoma, cutaneous squamous cell carcinoma, and cutaneous basal cell carcinoma (ICD-O-3 codes C00-C80, excluding C44 with a histology code starting with 805-811 or 872-879). Time-varying hazard ratios (vitamin D vs placebo) are from flexible parametric survival models that included randomisation group, age, state of residence at baseline, and an interaction between randomisation group and time since randomisation. Models were fitted separately within males and females. The p value for the interaction with time is from a likelihood ratio test comparing models with and without the interaction term. Abbreviations: CI, confidence interval; ICD-O-3, International Classification of Diseases for Oncology Third Edition



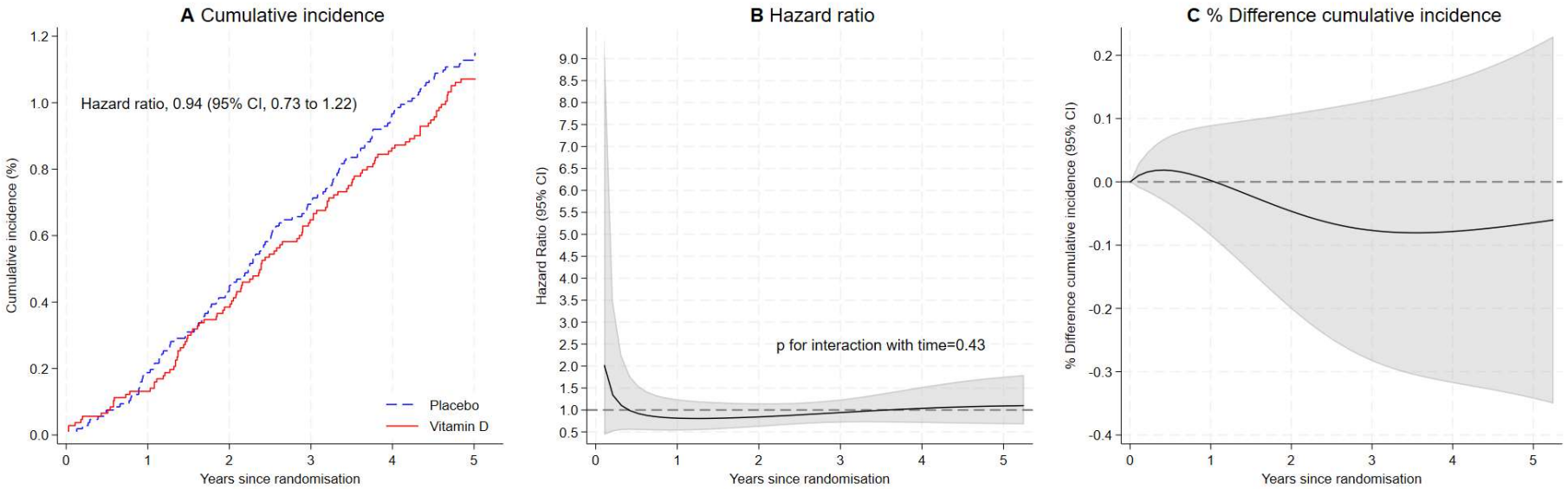
SFigure 9. Time-dependent effect of vitamin D supplementation on incidence of **non-melanoma cancer** according to body mass index (<25 kg/m^2, 25 to <30 kg/m^2; ≥30 kg/m^2).

The outcome is first diagnosis following randomisation of cancer of any type, excluding cutaneous melanoma, cutaneous squamous cell carcinoma, and cutaneous basal cell carcinoma (ICD-O-3 codes C00-C80, excluding C44 with a histology code starting with 805-811 or 872-879). Time-varying hazard ratios (vitamin D vs placebo) are from flexible parametric survival models that included randomisation group, age, sex, state of residence at baseline, and an interaction between randomisation group and time since randomisation. Models were fitted separately within each subgroup of body mass index. The p value for the interaction with time is from a likelihood ratio test comparing models with and without the interaction term. Abbreviations: CI, confidence interval; ICD-O-3, International Classification of Diseases for Oncology Third Edition



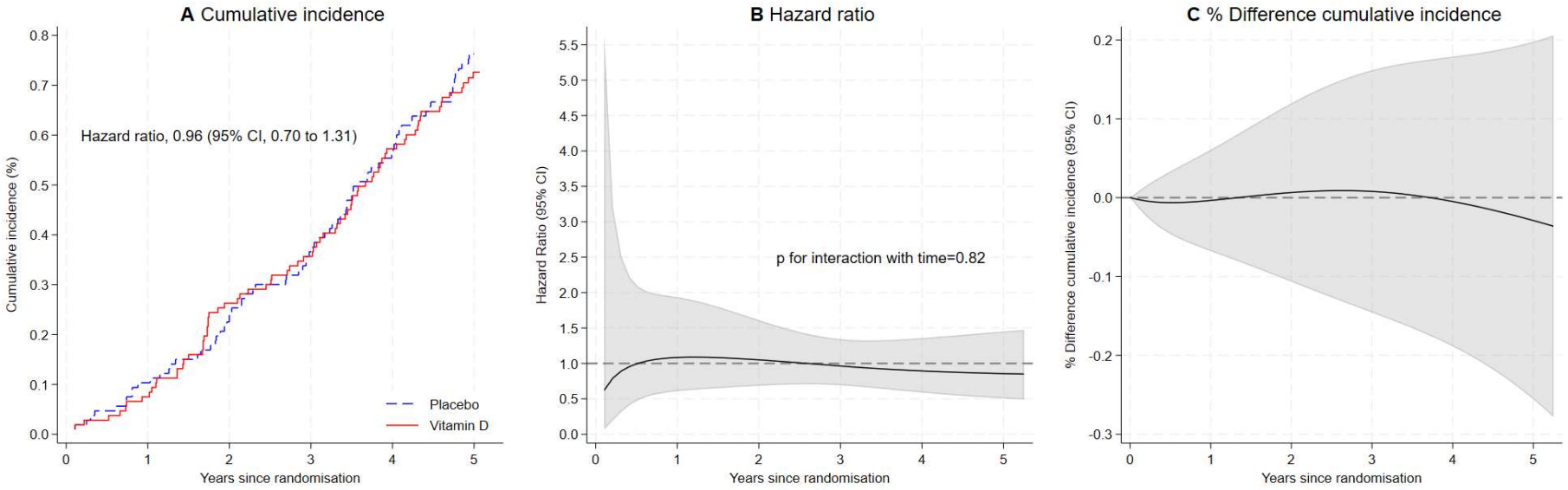
SFigure 10. Time-dependent effect of vitamin D supplementation on incidence of **non-melanoma cancer** according to predicted baseline 25(OH)D concentration (<50 nmol/L, ≥50 nmol/L).

The outcome is first diagnosis following randomisation of cancer of any type, excluding cutaneous melanoma, cutaneous squamous cell carcinoma, and cutaneous basal cell carcinoma (ICD-O-3 codes C00-C80, excluding C44 with a histology code starting with 805-811 or 872-879). Time-varying hazard ratios (vitamin D vs placebo) are from flexible parametric survival models that included randomisation group, age, sex, state of residence at baseline, and an interaction between randomisation group and time since randomisation. Models were fitted separately within each subgroup of predicted baseline 25(OH)D concentration. The p value for the interaction with time is from a likelihood ratio test comparing models with and without the interaction term. Abbreviations: CI, confidence interval; ICD-O-3, International Classification of Diseases for Oncology Third Edition



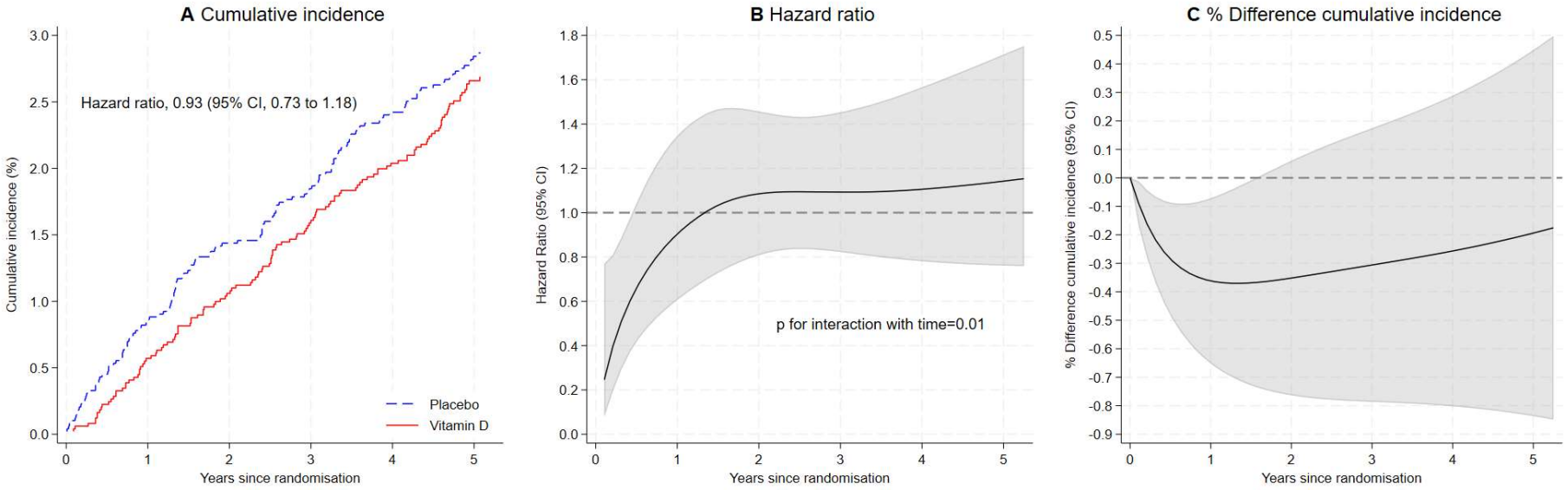
SFigure 11. Effect of vitamin D supplementation on incidence of **colorectal cancer** as a function of time since randomisation. Panel A shows the cause-specific cumulative incidence, panel B shows the hazard ratio, and panel C shows the percentage difference in the cause-specific standardised cumulative incidence.

The outcome was first diagnosis following randomisation of colorectal cancer (ICD-O-3 codes C18-C20 and C21.8). Cumulative incidence curves (panel A) were estimated using Aalen-Johansen methods, treating death without prior diagnosis of colorectal cancer as a competing risk. The overall hazard ratio (vitamin D versus placebo) shown in panel A was estimated using a flexible parametric survival model that included randomisation group, age, sex, and state of residence at baseline. The time-varying estimates (panels B and C) are from a flexible parametric survival model that also included an interaction between randomisation group and time since randomisation. Cause-specific standardised cumulative incidences were estimated treating death without prior diagnosis of colorectal cancer as a competing risk, and probabilities were standardised to the distribution of age, sex, and state of residence at baseline in the entire cohort. The interaction with time was assessed using a likelihood ratio test that compared models with and without the interaction term. Abbreviations: CI, confidence interval; ICD-O-3, International Classification of Diseases for Oncology Third Edition



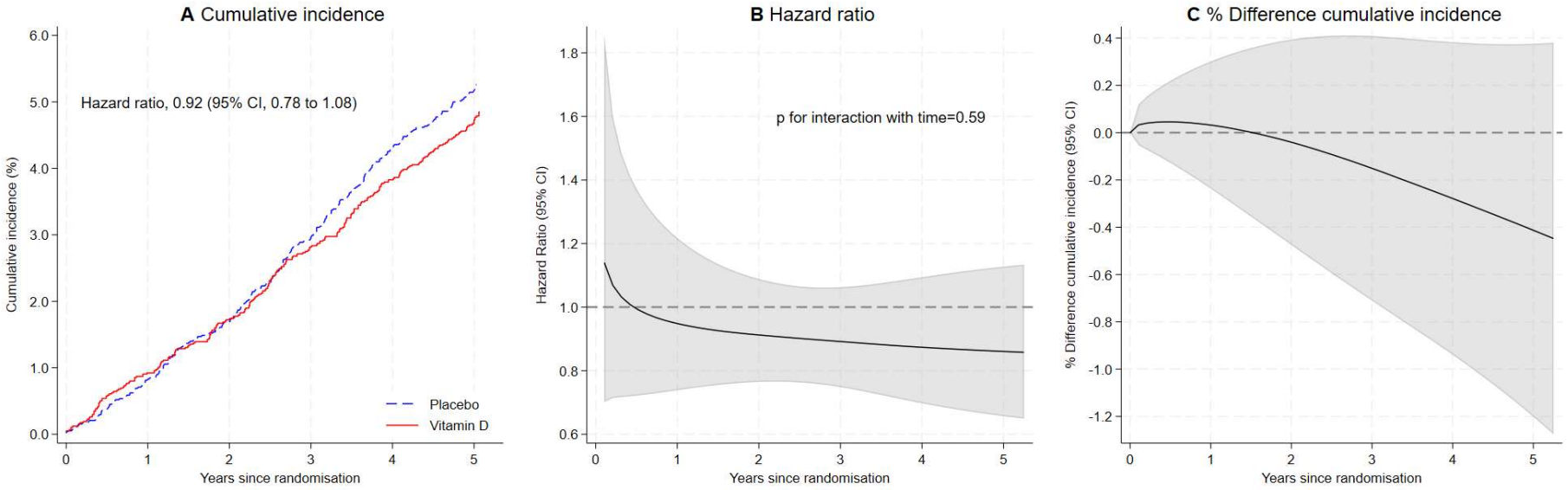
SFigure 12. Effect of vitamin D supplementation on incidence of **lung cancer** as a function of time since randomisation. Panel A shows the cause-specific cumulative incidence, panel B shows the hazard ratio, and panel C shows the percentage difference in the cause-specific standardised cumulative incidence.

The outcome was first diagnosis following randomisation of lung cancer (ICD-O-3 codes C33-C34). Cumulative incidence curves (panel A) were estimated using Aalen-Johansen methods, treating death without prior diagnosis of lung cancer as a competing risk. The overall hazard ratio (vitamin D versus placebo) shown in panel A was estimated using a flexible parametric survival model that included randomisation group, age, sex, and state of residence at baseline. The time-varying estimates (panels B and C) are from a flexible parametric survival model that also included an interaction between randomisation group and time since randomisation. Cause-specific standardised cumulative incidences were estimated treating death without prior diagnosis of lung cancer as a competing risk, and probabilities were standardised to the distribution of age, sex, and state of residence at baseline in the entire cohort. The interaction with time was assessed using a likelihood ratio test that compared models with and without the interaction term. Abbreviations: CI, confidence interval; ICD-O-3, International Classification of Diseases for Oncology Third Edition



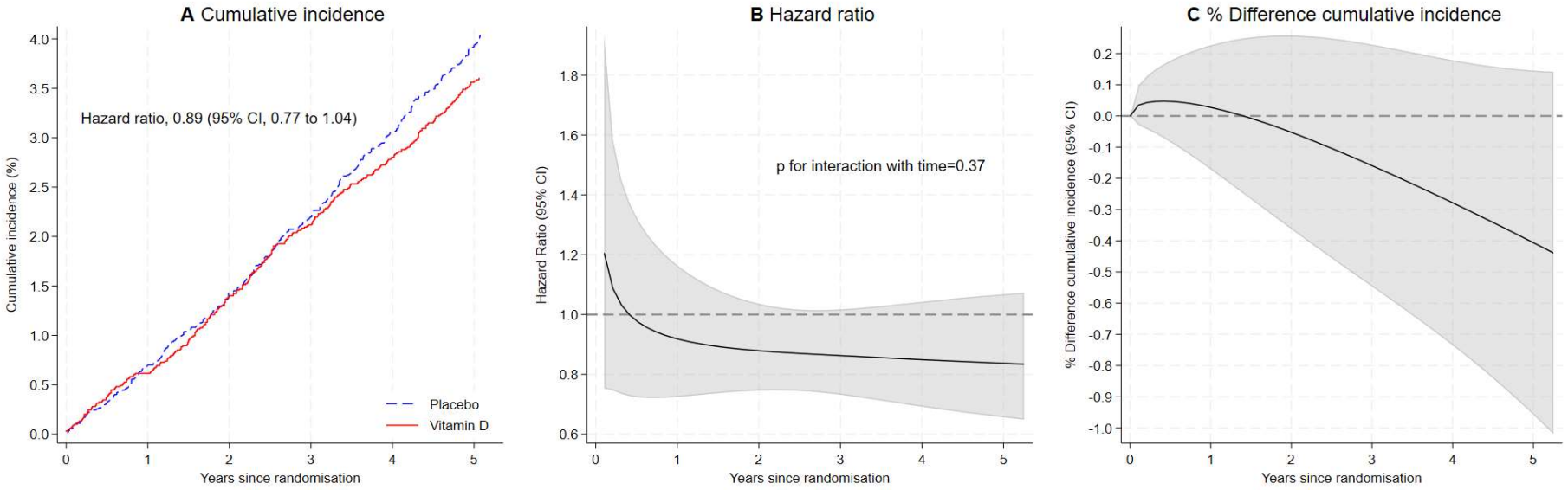
SFigure 13. Effect of vitamin D supplementation on incidence of **invasive breast cancer in females** as a function of time since randomisation. Panel A shows the cause-specific cumulative incidence, panel B shows the hazard ratio, and panel C shows the percentage difference in the cause-specific standardised cumulative incidence.

This analysis was restricted to 9,778 females. The outcome was first diagnosis following randomisation of invasive breast cancer (ICD-O-3 code of C50 and behaviour not categorised as 'in situ' or 'uncertain whether benign or malignant'). Cumulative incidence curves (panel A) were estimated using Aalen-Johansen methods, treating death without prior diagnosis of invasive breast cancer as a competing risk. The overall hazard ratio (vitamin D versus placebo) shown in panel A was estimated using a flexible parametric survival model that included randomisation group, age, and state of residence at baseline. The time-varying estimates (panels B and C) are from a flexible parametric survival model that also included an interaction between randomisation group and time since randomisation. Cause-specific standardised cumulative incidences were estimated treating death without prior diagnosis of invasive breast cancer as a competing risk, and probabilities were standardised to the distribution of age, and state of residence at baseline in the entire cohort of females. The interaction with time was assessed using a likelihood ratio test that compared models with and without the interaction term. Abbreviations: CI, confidence interval; ICD-O-3, International Classification of Diseases for Oncology Third Edition



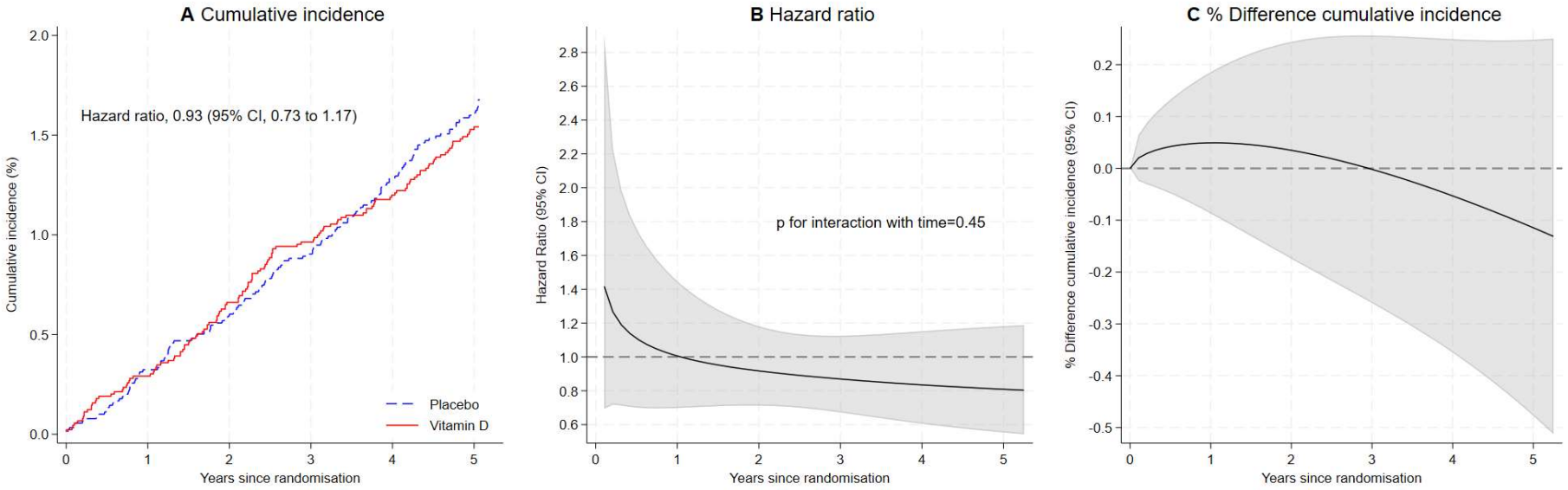
SFigure 14. Effect of vitamin D supplementation on incidence of **prostate cancer in males** as a function of time since randomisation. Panel A shows the cause-specific cumulative incidence, panel B shows the hazard ratio, and panel C shows the percentage difference in the cause-specific standardised cumulative incidence.

This analysis was restricted to 11,530 males. The outcome was first diagnosis following randomisation of prostate cancer (ICD-O-2 code C61.9). Cumulative incidence curves (panel A) were estimated using Aalen-Johansen methods, treating death without prior diagnosis of prostate cancer as a competing risk. The overall hazard ratio (vitamin D versus placebo) shown in panel A was estimated using a flexible parametric survival model that included randomisation group, age, and state of residence at baseline. The time-varying estimates (panels B and C) are from a flexible parametric survival model that also included an interaction between randomisation group and time since randomisation. Cause-specific standardised cumulative incidences were estimated treating death without prior diagnosis of prostate cancer as a competing risk, and probabilities were standardised to the distribution of age, and state of residence at baseline in the entire cohort of males. The interaction with time was assessed using a likelihood ratio test that compared models with and without the interaction term. Abbreviations: CI, confidence interval; ICD-O-3, International Classification of Diseases for Oncology Third Edition



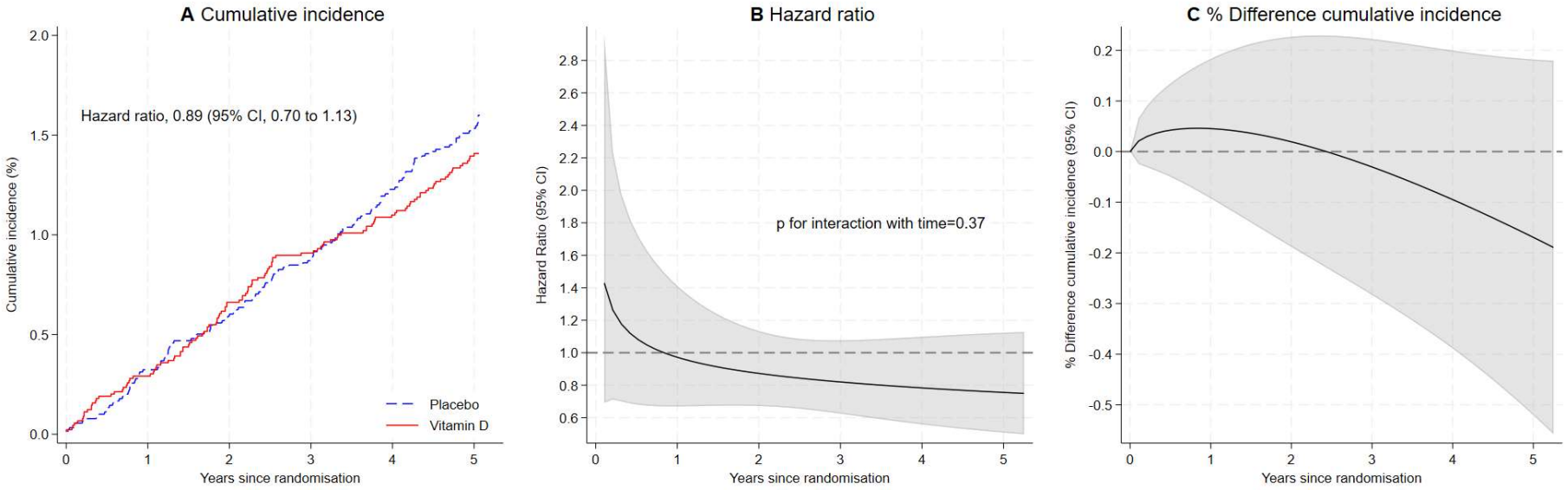
SFigure 15. Effect of vitamin D supplementation on incidence of **melanoma skin cancer** (invasive and in situ) as a function of time since randomisation. Panel A shows the cause-specific cumulative incidence, panel B shows the hazard ratio, and panel C shows the percentage difference in the cause-specific standardised cumulative incidence.

The outcome was first diagnosis following randomisation of melanoma skin cancer (ICD-O-3 code of C44 with histology code starting with 872-879 and behaviour code of 2 [in situ] or 3 [invasive]). The analysis was restricted to 17,882 participants who were not living in Western Australia at baseline. Cumulative incidence curves (panel A) were estimated using Aalen-Johansen methods, treating death without prior diagnosis of melanoma skin cancer as a competing risk. The overall hazard ratio (vitamin D versus placebo) shown in panel A was estimated using a flexible parametric survival model that included randomisation group, age, sex, and state of residence at baseline. The time-varying estimates (panels B and C) are from a flexible parametric survival model that also included an interaction between randomisation group and time since randomisation. Cause-specific standardised cumulative incidences were estimated treating death without prior diagnosis of melanoma skin cancer as a competing risk, and probabilities were standardised to the distribution of age, sex, and state of residence at baseline in the entire cohort. The interaction with time was assessed using a likelihood ratio test that compared models with and without the interaction term. Abbreviations: CI, confidence interval; ICD-O-3, International Classification of Diseases for Oncology Third Edition



SFigure 16. Effect of vitamin D supplementation on incidence of **invasive melanoma skin cancer** as a function of time since randomisation. Panel A shows the cause-specific cumulative incidence, panel B shows the hazard ratio, and panel C shows the percentage difference in the cause-specific standardised cumulative incidence.

The outcome was first diagnosis following randomisation of invasive melanoma skin cancer (ICD-O-3 code of C44 with histology code starting with 872-879 and behaviour code of 3 [invasive]). This analysis was restricted to 17,882 participants who were not living in Western Australia at baseline. Cumulative incidence curves (panel A) were estimated using Aalen-Johansen methods, treating death without prior diagnosis of invasive melanoma skin cancer as a competing risk. The overall hazard ratio (vitamin D versus placebo) shown in panel A was estimated using a flexible parametric survival model that included randomisation group, age, sex, and state of residence at baseline. The time-varying estimates (panels B and C) are from a flexible parametric survival model that also included an interaction between randomisation group and time since randomisation. Cause-specific standardised cumulative incidences were estimated treating death without prior diagnosis of invasive melanoma skin cancer as a competing risk, and probabilities were standardised to the distribution of age, sex, and state of residence at baseline in the entire cohort. The interaction with time was assessed using a likelihood ratio test that compared models with and without the interaction term. Abbreviations: CI, confidence interval; ICD-O-3, International Classification of Diseases for Oncology Third Edition



SFigure 17. ***Sensitivity analysis*** – Effect of vitamin D supplementation on incidence of **invasive melanoma skin cancer** as a function of time since randomisation, taking into account in situ melanoma skin cancer. Panel A shows the cause-specific cumulative incidence, panel B shows the hazard ratio, and panel C shows the percentage difference in the cause-specific standardised cumulative incidence.

The outcome was first diagnosis following randomisation of invasive melanoma skin cancer (ICD-O-3 code of C44 with histology code starting with 872-879 and behaviour code of 3 [invasive]). This analysis was restricted to 17,882 participants who were not living in Western Australia at baseline. Follow-up ended at the earliest of: (i) date of invasive melanoma skin cancer; (ii) date of in situ melanoma skin cancer; (iii) 31/12/2019;1 (iiv) 5 years and 1 month after randomisation; or (v) the date last known to be alive. Cumulative incidence curves (panel A) were estimated using Aalen-Johansen methods, treating death without prior diagnosis of invasive melanoma skin cancer as a competing risk. The overall hazard ratio (vitamin D versus placebo) shown in panel A was estimated using a flexible parametric survival model that included randomisation group, age, sex, and state of residence at baseline. The time-varying estimates (panels B and C) are from a flexible parametric survival model that also included an interaction between randomisation group and time since randomisation. Cause-specific standardised cumulative incidences were estimated treating death without prior diagnosis of invasive melanoma skin cancer as a competing risk, and probabilities were standardised to the distribution of age, sex, and state of residence at baseline in the entire cohort. The interaction with time was assessed using a likelihood ratio test that compared models with and without the interaction term. Abbreviations: CI, confidence interval; ICD-O-3, International Classification of Diseases for Oncology Third Edition