

SUPPLEMENTARY DATA

Appendix 1. Search method

(((((("vitamin d"[MeSH Terms] OR "vitamin d"[All Fields] OR "ergocalciferols"[MeSH Terms] OR "ergocalciferols"[All Fields]) AND incident[All Fields]) AND ("breast neoplasms"[MeSH Terms] OR ("breast"[All Fields] AND "neoplasms"[All Fields]) OR "breast neoplasms"[All Fields] OR ("breast"[All Fields] AND "cancer"[All Fields]) OR "BC"[All Fields])) AND ("25-hydroxyvitamin D"[Supplementary Concept] OR "25-hydroxyvitamin D"[All Fields] OR "25 hydroxyvitamin d"[All Fields] OR "calcifediol"[MeSH Terms] OR "calcifediol"[All Fields])) OR 25[All Fields] AND ("hydroxide ion"[Supplementary Concept] OR "hydroxide ion"[All Fields] OR "oh"[All Fields]) AND D[All Fields]) AND ("epidemiology"[Subheading] OR "epidemiology"[All Fields] OR "incidence"[All Fields] OR "incidence"[MeSH Terms]) AND ("free full text"[sb] AND ("2007/01/01"[PDAT] : "2017/01/01"[PDAT]) AND "humans"[MeSH Terms] AND English[lang] AND "female"[MeSH Terms] AND "adult"[MeSH Terms])).

Appendix 2. Quality score

The quality score is composed of 5 items, and each item was allocated 0, 1 or 2 points. This allowed a total score between 0 and 8 points, 8 representing the highest quality. The following items are included in the score:

1. Study design

0 for case-control studies

1 for cross-sectional studies

2 for longitudinal studies (retrospective or prospective) or non-randomized intervention studies

2. Population

Observational studies

0 if n <500

1 if n 500 to 2000

2 if n >2000

3. Outcome (see table below)

	0	1	2
BC	If not reported or if diagnosis was based on a proxy for disease (i.e. use of medication that is not specific to the disease or symptoms)	If diagnosis was self-reported or retrieved from reports (e.g. medical records) not collected specifically for study or was based on the use of disease-specific medication (e.g. antidiabetic medication)	If diagnosis was made during study using adequate test or official registries were checked by a researcher

SUPPLEMENTARY DATA

4. Adjustments

0 if findings are not controlled** for at least age and gender

1 if findings are controlled for age and gender

2 if an intervention study is adequately randomized or if findings are additionally controlled for other key covariates (e.g. age at menarche, parity, family history of BC etc.)

** *'Controlled for' here refers to: adjusted for in the statistical analyses (e.g. with multiple regression); stratified for in the analyses (e.g. pre- and post-menopausal separately); or narrow selection criteria of study participants on this covariate (e.g. Pre-menopausal women only).*

SUPPLEMENTARY DATA

Begg's Test

adj. Kendall's Score (P-Q) = 68

Std. Dev. of Score = 50.62

Number of Studies = 28

z = 1.34

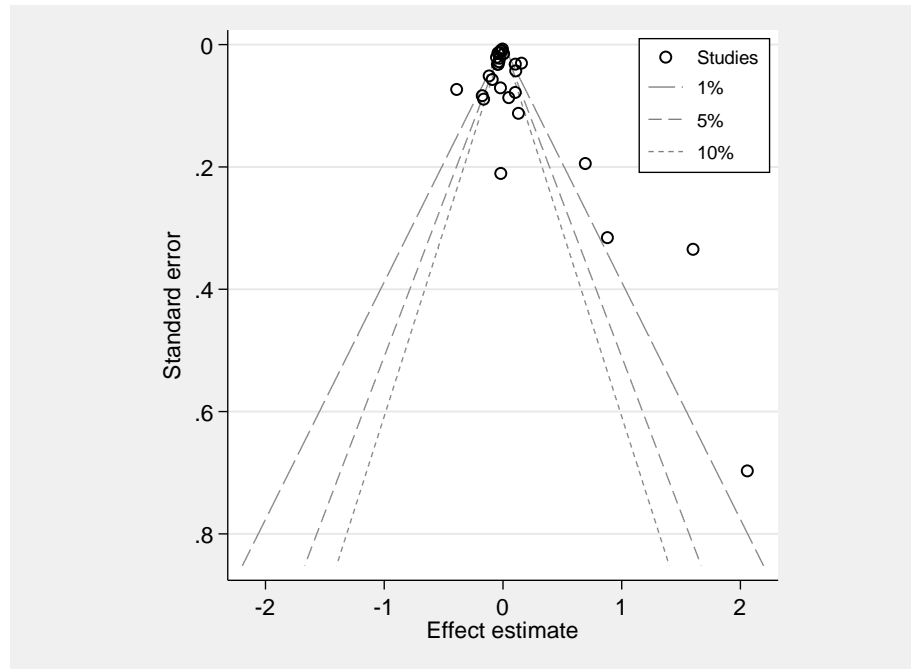
Pr > |z| = 0.179

z = 1.32 (continuity corrected)

Pr > |z| = 0.186 (continuity corrected)

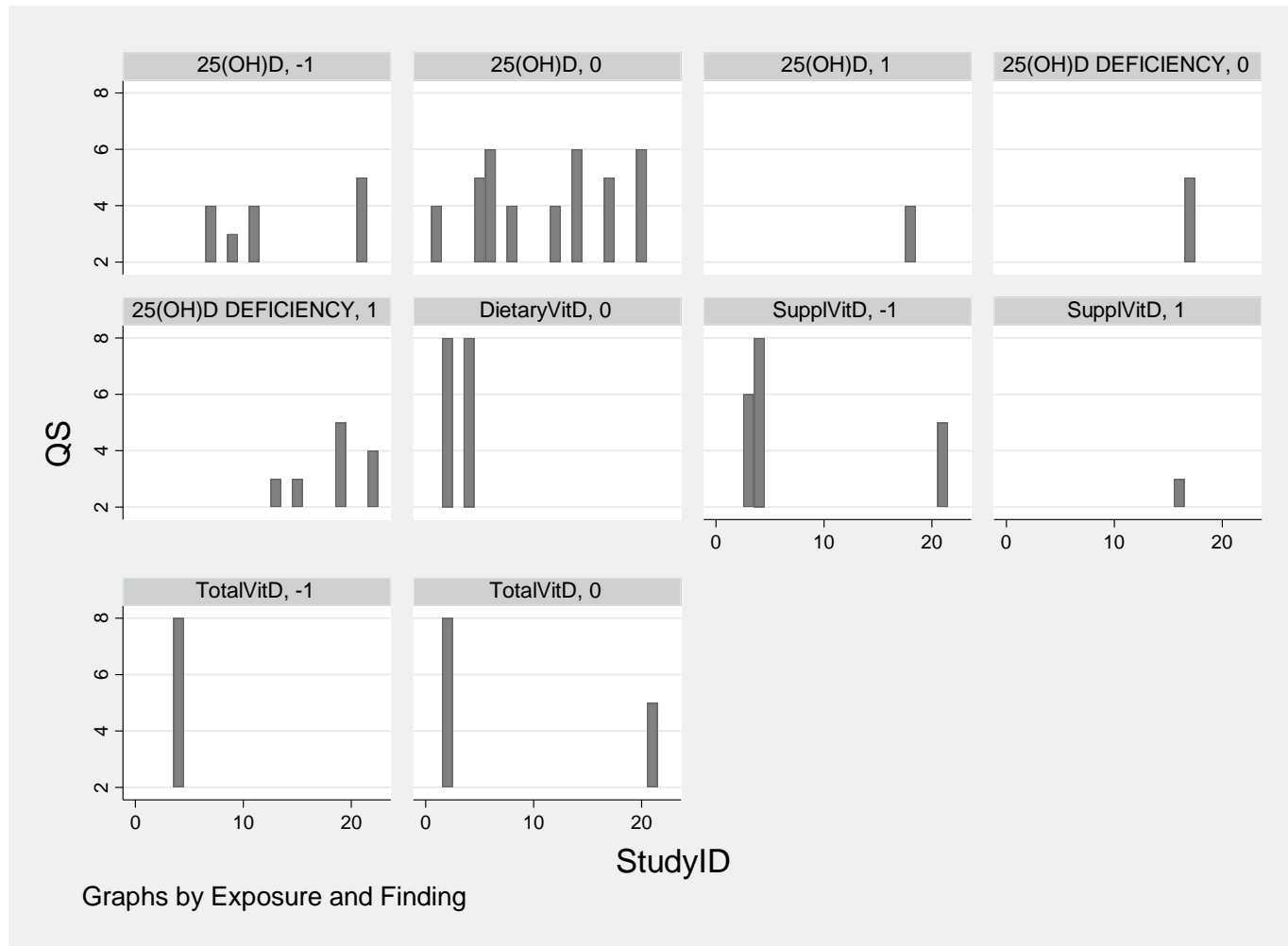
Egger's test

Std_Eff	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
slope	-.0235691	.0135	-1.75	0.093	-.0513188	.0041805
bias	.7586765	.6614072	1.15	0.262	-.6008655	2.118218



Supplemental Figure 1. Begg's funnel plot of all data-points included in the meta-analysis (n=28 data-points), 2000-2018

SUPPLEMENTARY DATA



Supplemental Figure 2 Harvest plot of quality scores (QS) by type of exposures and qualitative finding

Note: QS are presented by exposure/qualitative finding pair and annotated by Study ID. Qualitative finding can be -1="Inverse association", 0="null finding" or 1="Positive association".