**Supplementary Method: Analysis plans**

For each research question a set of planned analyses were pre-specified. A series of sex-adjusted and sex-stratified regression models were adjusted for baseline age, and fully adjusted models also accounted for age, residence area, only child status, family structure, self-perceived family economic status, BMI at T1, physical activity at T1 and self-perceived pubertal timing. The following four research questions were tested:

**1. Is vitamin D levels or status cross-sectionally associated with depressive symptoms in early adolescence?**

This was investigated cross-sectionally at both T1 and T3 in the entire sample using (a) a linear regression to test association between vitamin D levels (continuous) and depression symptoms (b) a logistic regression to test the relationship vitamin D levels (continuous) and presence of depression symptoms (dichotomous) and (c) a logistic regression to test the link between vitamin D insufficiency (dichotomous) and depression symptoms (dichotomous).

**2. Is baseline vitamin D levels or status longitudinally associated with depressive symptoms in early adolescence?**

Firstly, logistics regression models were performed to test if vitamin D levels and VDD at baseline predicted incident depression (dichotomous) at T2 and T3 during 2-year follow-up after excluding those who had depression at the baseline. Also, risk ratios (RRs) derived from multivariable logistic regression were used to assess the associations between baseline vitamin D status (continuous and VDD) and the number of depressive symptoms and trajectory of depression symptoms across three time points in the entire sample.

**3. Are longitudinal changes in vitamin D levels associated with depressive symptoms in early adolescence?**

Conditional changes (independent of baseline) in each biomarker were calculated by residuals from sex-specific linear regression models for the biomarker at follow-up on the baseline vitamin D levels. Firstly, risk ratios (RRs) derived from multivariable logistic regression were used to assess the associations between changes in 25(OH)D concentrations and the cumulative incident depression after excluding those who had depression at the baseline. Then, odds ratios (ORs) derived from multivariable logistic regression were used to assess the associations between trajectories of vitamin D status and the number of depressive symptoms, as well as trajectory of depression symptoms across three time points in the entire sample.

**4. Are longitudinal trajectories of vitamin D status associated with depressive symptoms in early adolescence?**

Based on the vitamin D status at baseline and follow-up, we identified four possible trajectories: (1) persistent non-deficiency (reference); (2) baseline non-deficiency to follow-up deficiency; (3) baseline deficiency to follow-up non-deficiency; (4) persistent deficiency. Firstly, the relationships between trajectories of vitamin D status and cumulative incident depression were determined by odds ratios (ORs) derived from logistic regressions after excluding those who had depression at the baseline. Odds ratios (ORs) derived from multivariable logistic regression were used to assess the associations between trajectories of vitamin D status and the number of depressive symptoms and trajectory of depression symptoms across three waves in the entire sample.