

Supplementary materials

Supplementary Material 1. Characteristics of included studies (random sample of 10).

<i>Study label (author year)</i>	<i>First author</i>	<i>Year</i>	<i>Country</i>	<i>Total N</i>	<i>N (men)</i>	<i>N (women)</i>	<i>Measurement tool</i>
<i>Altaş 2022</i>	Altaş	2022	Turkey	330	157	173	Self-Care Management Protection Scale (SCMP-G)
<i>Bai 2009</i>	Bai	2009	Taiwan	165	87	78	Diabetes Self-Care Scale (DSC)
<i>Cortez 2017</i>	Cortez	2017	Brazil	238	80	158	ESM, DES
<i>Gao 2013</i>	Gao	2013	China	222	85	137	DMSES (Chinese version)
<i>Gucciardi 2008</i>	Gucciardi	2008	Canada	275	132	143	Diabetes Education Self-Efficacy Scale
<i>McEwen 2010</i>	McEwen	2010	USA	21	4	17	SDSCA
<i>Misra 2009</i>	Misra	2009	USA	180	86	94	Perceived Difficulty (DQIP-derived)
<i>Mogre 2017</i>	Mogre	2017	Ghana	187	52	135	SDSCA
<i>Russo 2015</i>	Russo	2015	Italy	415,294	227,169	188,125	LDL-C control
<i>Saleh 2017</i>	Saleh	2016	Bangladesh	458	204	254	Modified DAS and DKT

Note: Studies were selected at random from the full corpus for descriptive illustration only; all pooled estimates and diagnostics in the manuscript are based on the complete dataset. NOS = Newcastle–Ottawa Scale. Instruments are reported as named in the source studies (e.g., SDSCA = Summary of Diabetes Self-Care Activities; DMSES = Diabetes Management Self-Efficacy Scale; DES = Diabetes Empowerment Scale; DKT = Diabetes Knowledge Test; DAS = Diabetes Attitude Scale; DQIP = Diabetes Quality Improvement Project).

Supplementary Material 2. Primary multilevel model (REML)

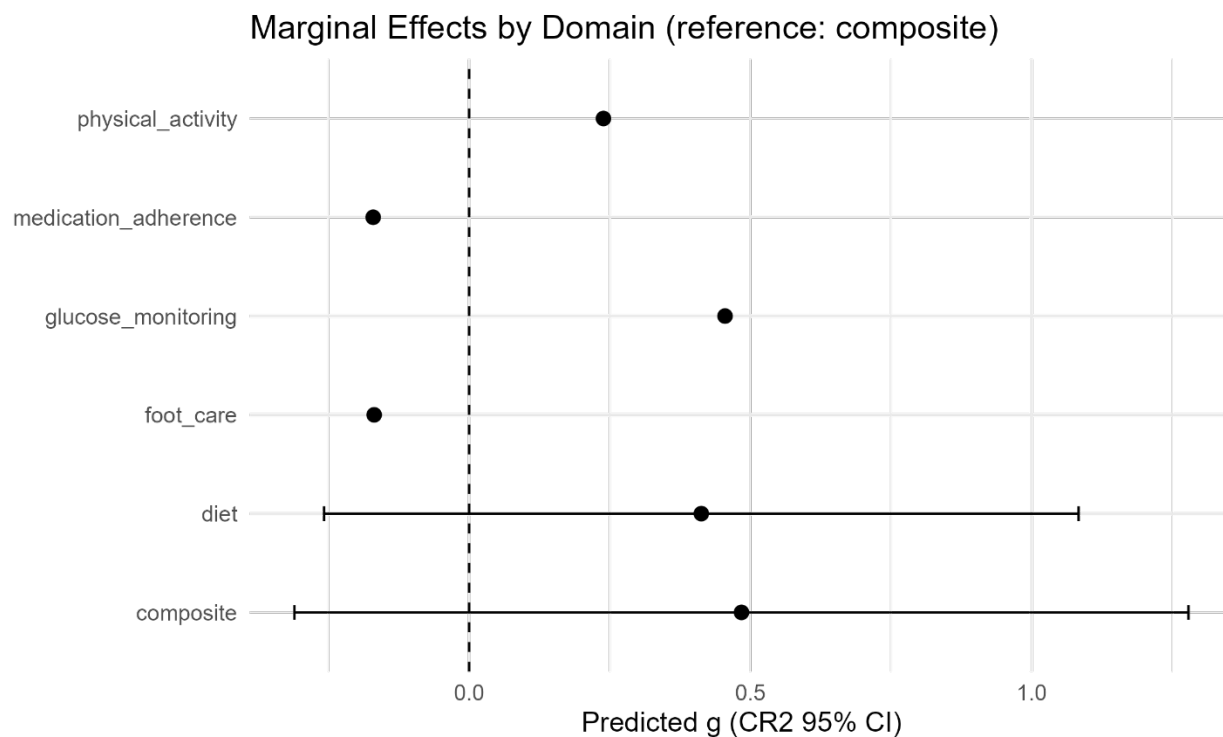
<i>Statistic</i>	<i>Estimate</i>
<i>Pooled Hedges' g (men – women)</i>	0.369
<i>95% CI</i>	–0.108 to 0.847
<i>95% prediction interval</i>	–1.669 to 2.407
τ^2 <i>between-study</i>	0.510933
τ^2 <i>within-study</i>	0.510933
I^2 <i>between-study (%)</i>	47.6
I^2 <i>within-study (%)</i>	47.6
I^2 <i>total (%)</i>	95.3
k (<i>studies</i>)	18
k (<i>effects</i>)	18
<i>Estimation method</i>	REML

Note: Hedges' g is adjusted for small-sample bias and defined as men – women (positive = higher scores in men).

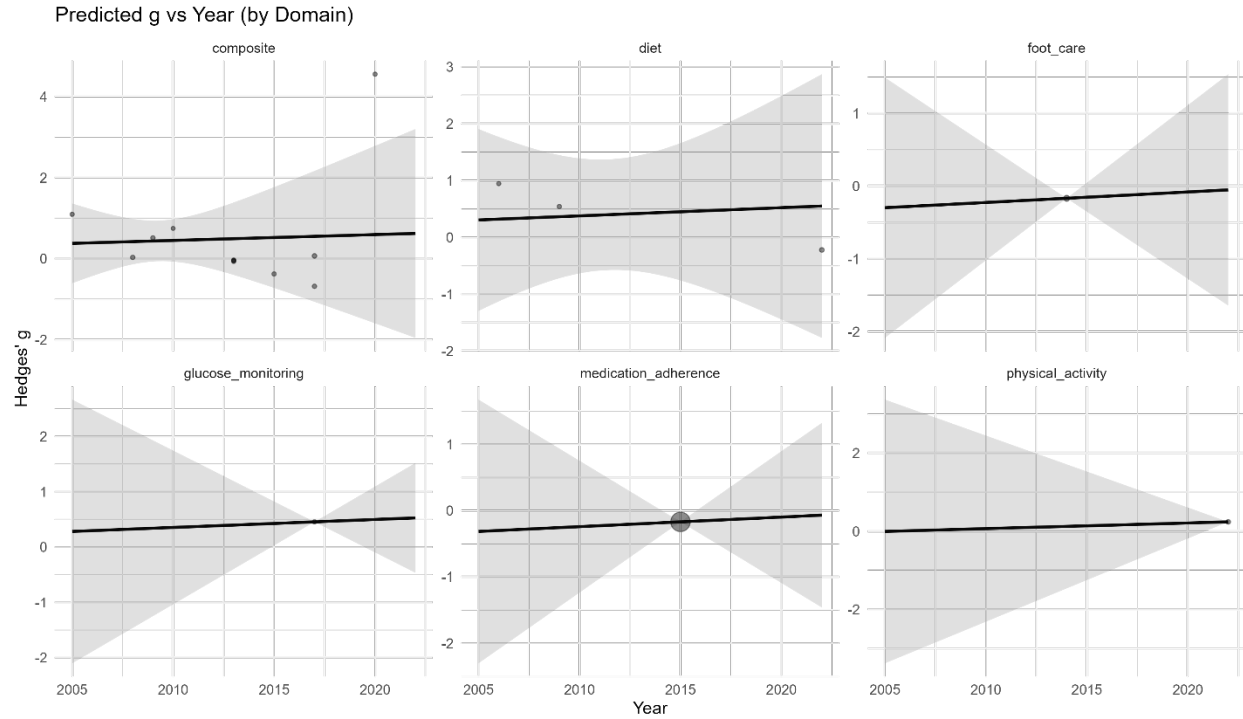
Supplementary Material 3: Moderator models (CR2 SEs). coefficients, p-values, $\Delta AICc$, pseudo- R^2 , and k

<i>Model</i>	<i>Term</i>	<i>Est.</i>	<i>CR2 SE</i>	<i>p</i>	<i>$\Delta AICc$</i>	<i>Pseudo-R^2 (%)</i>	<i>k(studies)</i>	<i>k(effects)</i>
<i>Domain</i>	outcome_domaindiet	-0.071	0.531	0.901	13.396	-47.61	18	18
<i>Domain</i>	outcome_domainfoot_care	-0.653	0.406	0.138	13.396	-47.61	18	18
<i>Domain</i>	outcome_domainglucose_monitoring	-0.029	0.406	0.944	13.396	-47.61	18	18
<i>Domain</i>	outcome_domainmedication_adherence	-0.655	0.406	0.137	13.396	-47.61	18	18
<i>Domain</i>	outcome_domainphysical_activity	-0.245	0.406	0.559	13.396	-47.61	18	18
<i>MeasurementType</i>	measurement_typegeneral	0.289	0.542	0.602	1.098	-6.05	18	18
<i>Region</i>	regionEurope	-0.996	0.707	0.310	17.799	-36.794	18	18
<i>Region</i>	regionLatin America & Caribbean	-1.406	0.699	0.101	17.799	-36.794	18	18
<i>Region</i>	regionMiddle East & North Africa	-0.714	0.736	0.446	17.799	-36.794	18	18
<i>Region</i>	regionNorth America	-0.061	0.726	0.935	17.799	-36.794	18	18
<i>Region</i>	regionSouth Asia	-0.656	0.699	0.392	17.799	-36.794	18	18
<i>Region</i>	regionSub-Saharan Africa	-0.268	0.699	0.718	17.799	-36.794	18	18
<i>Year</i>	year_c	0.002	0.064	0.973	1.412	-7.933	18	18

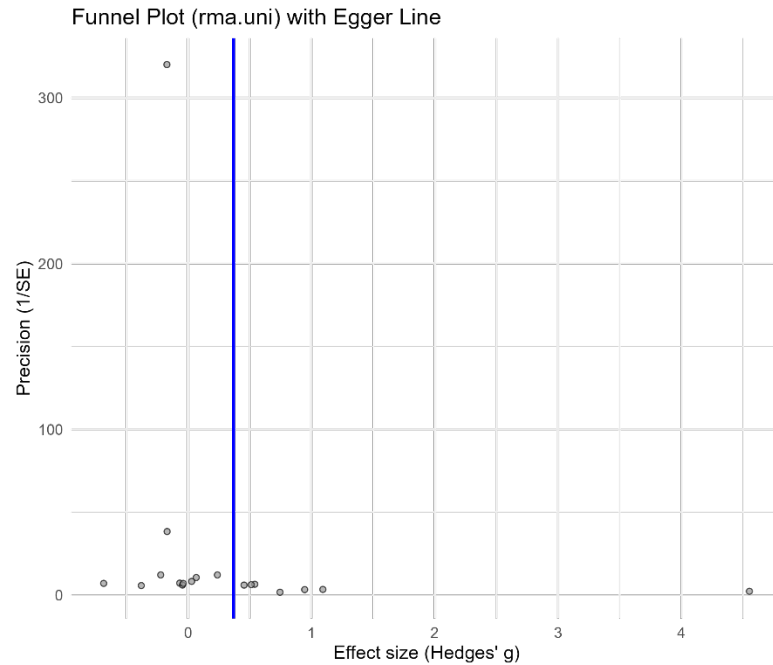
Note: Hedges' g is adjusted for small-sample bias and defined as men – women. Pseudo- R^2 reflects proportional reduction in the between-cluster variance relative to the Step-2 intercept-only model. $\Delta AICc$ compares each moderator model with the primary multilevel model; negative pseudo- R^2 values reflect greater residual variance under the moderator specification given the small sample.



Supplementary Material 4. Marginal effects by behavioral domain; Predicted Hedges' g with 95% simultaneous confidence intervals; ribbons reflect study weights. Intervals widely overlapping zero indicate limited evidence of domain differences at this stage.



Supplementary Material 5. Predicted Hedges' g versus year, faceted by domain; Bubble size is proportional to precision ($1/SE$). Slopes are near flat with wide bands in sparsely populated domains, consistent with the near-zero-year effect in Supplementary Material 4.



Supplementary Material 6. Contour-enhanced funnel plot with meta-analytic mean (blue); Scatter of Hedges' g versus precision ($1/SE$). The vertical blue line marks the pooled mean from the primary model. Asymmetry is visible with a relative excess of low-precision positive effects. Error note: one extreme low-precision positive point is flagged as a potential influence; its presence aligns with the significant Egger result reported in Table 4.

Supplementary Material 7. Small-study effects: Egger regression (precision on effect, Knapp–Hartung)

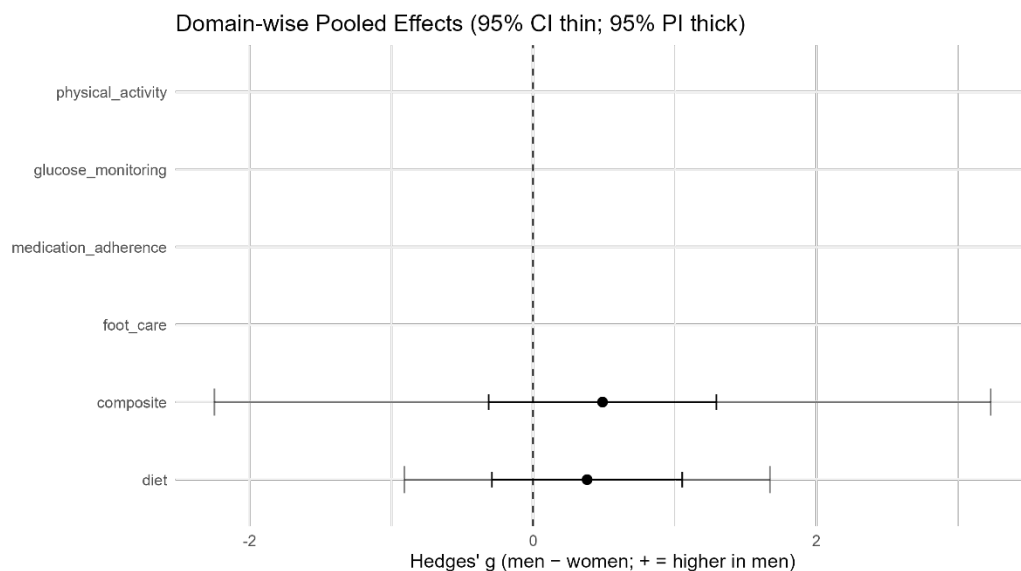
<i>Measure</i>	<i>Value</i>
<i>Test statistic (z)</i>	3.489
<i>p-value</i>	0.003
<i>Outcome metric</i>	Hedges' g (men – women)
<i>Meta-analytic framework</i>	REML
<i>Predictor</i>	Precision (1/SE)
<i>Inference</i>	Knapp–Hartung

Note: A significant Egger test indicates funnel asymmetry compatible with small-study effects; it does not, by itself, establish publication or reporting bias.

Supplementary Material 8. Sensitivity matrix (pooled Hedges' g and Δ vs. primary estimate)

<i>pooled_g</i>	<i>delta_vs_primary</i>
0.369	0.000
0.405	0.035
0.393	0.023
0.109	-0.260
0.398	0.029
0.397	0.027
0.362	-0.008
0.339	-0.031
0.391	0.021
0.405	0.035
0.364	-0.006
0.397	0.027
0.354	-0.016
0.330	-0.040
0.416	0.046
0.421	0.052
0.151	-0.218

Note: Values are pooled Hedges' g and the change relative to the Step 2 primary model. Where only pooled g and Δ were available in the audit export, CI/PI fields are not shown. Positive Δ indicates an increase relative to the primary



Supplementary Material 9. Domain-wise pooled effects with 95% prediction intervals; Pooled Hedges' g by domain with 95% prediction intervals. Composite and diet show positive point estimates with wide intervals crossing zero. Empty rows denote domains without sufficient effects; no estimates were generated for those domains.

Supplementary Material 10. Summary of Findings (GRADE)

<i>construct</i>	<i>pooled_g</i>	<i>certainty</i>
<i>Overall (self-efficacy & self-management combined)</i>	0.369	Very low
<i>Domain: diet</i>	0.382	Very low
<i>Domain: composite</i>	0.490	Very low
<i>Domain: foot_care</i>	—	Low
<i>Domain: medication_adherence</i>	—	Low
<i>Domain: glucose_monitoring</i>	—	Low
<i>Domain: physical_activity</i>	—	Low

Note: Hedges' g is defined as men – women. Dashes indicate domains without estimable pooled effects at this step. Certainty ratings reflect imprecision (prediction intervals crossing zero), inconsistency (heterogeneity from Step 2), and small-study effects (Step 4).