

Table S1 Quality evaluation of included studies

| Study | S1 | S2 | S3 | S4 | C | O1 | O2 | O3 | Sum |
|----------------|----|----|---------------------|----|----|----|---------|----------|-----|
| Song, 2007 | ★ | ★ | ★(Laboratory tests) | ★ | ★★ | ★ | ★(8.8) | ★(90.4%) | 9 |
| Hinnouho, 2015 | ★ | ★ | ★(Laboratory tests) | ★ | ★★ | ★ | ★(17.4) | —(NA) | 8 |
| Andersen, 2015 | ★ | ★ | ★(Interview) | ★ | ★★ | ★ | ★(5.6) | —(NA) | 8 |
| Guo, 2016 | ★ | ★ | ★(Laboratory tests) | ★ | ★★ | ★ | ★(18.7) | —(NA) | 8 |
| Rishi, 2017 | ★ | ★ | ★(Interview) | ★ | ★★ | ★ | ★(5.4) | —(NA) | 8 |
| Lee, 2018 | ★ | ★ | ★(Laboratory tests) | ★ | ★★ | ★ | ★(7.4) | —(NA) | 8 |
| Nathalie, 2018 | — | ★ | ★(Questionnaires) | ★ | ★★ | ★ | ★(24) | ★(74.2%) | 8 |
| Li, 2019 | ★ | ★ | ★(Laboratory tests) | ★ | ★★ | ★ | —(3.6) | —(NA) | 7 |

We herein selected "age, gender" as the most important adjusting factors. A mean follow-up duration of at least 5 years was predefined as long enough for outcome to occur in our study. It was regarded as adequate when the follow-up rate was at least 70%. NA: not available; S1: Representativeness of the exposed cohort; S2: Selection of the non-exposed cohort; S3: Ascertainment of exposure; S4: Demonstration that outcome of interest was not present at start of study; C: Comparability of cohorts on the basis of the design or analysis; O1: Assessment of outcome; O2: Was follow-up long enough for outcomes to occur?; O3: Adequacy of follow up of cohorts.

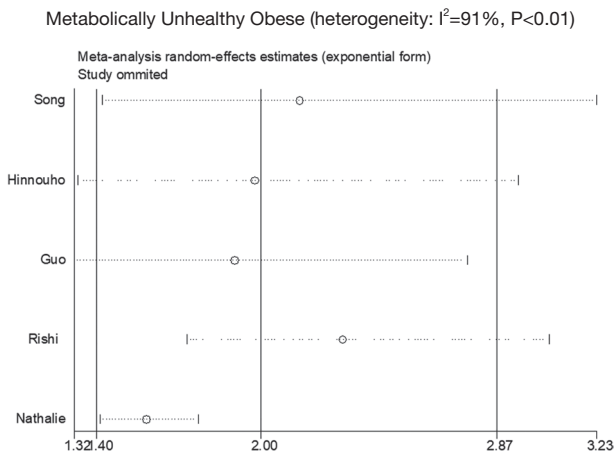
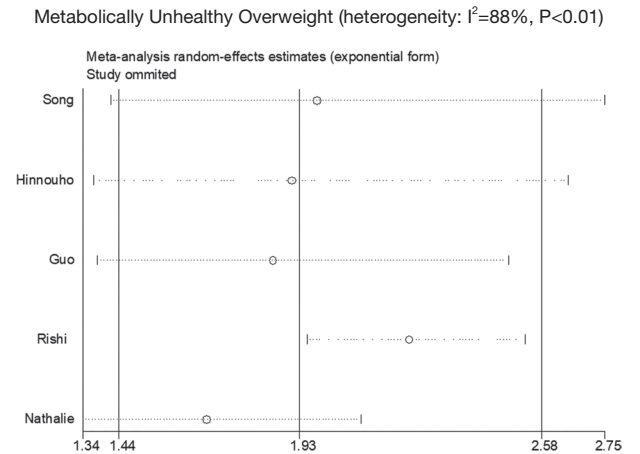
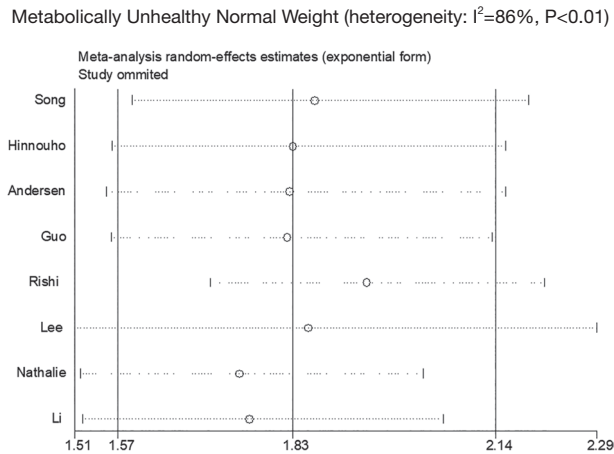
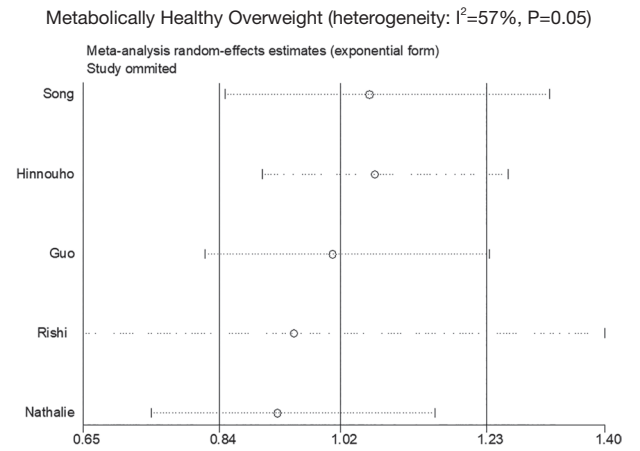
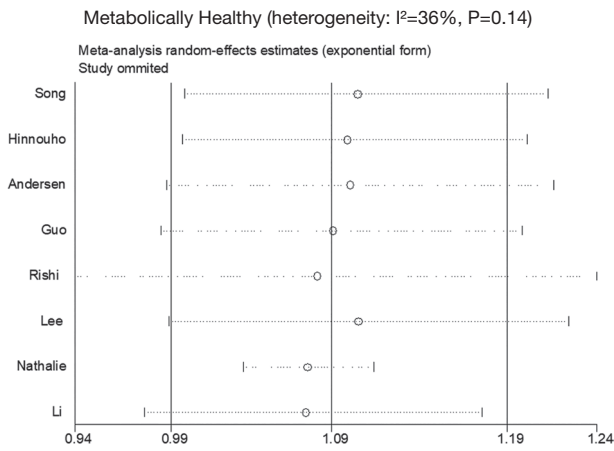


Figure S1 Results of sensitivity analysis after excluding one single study.

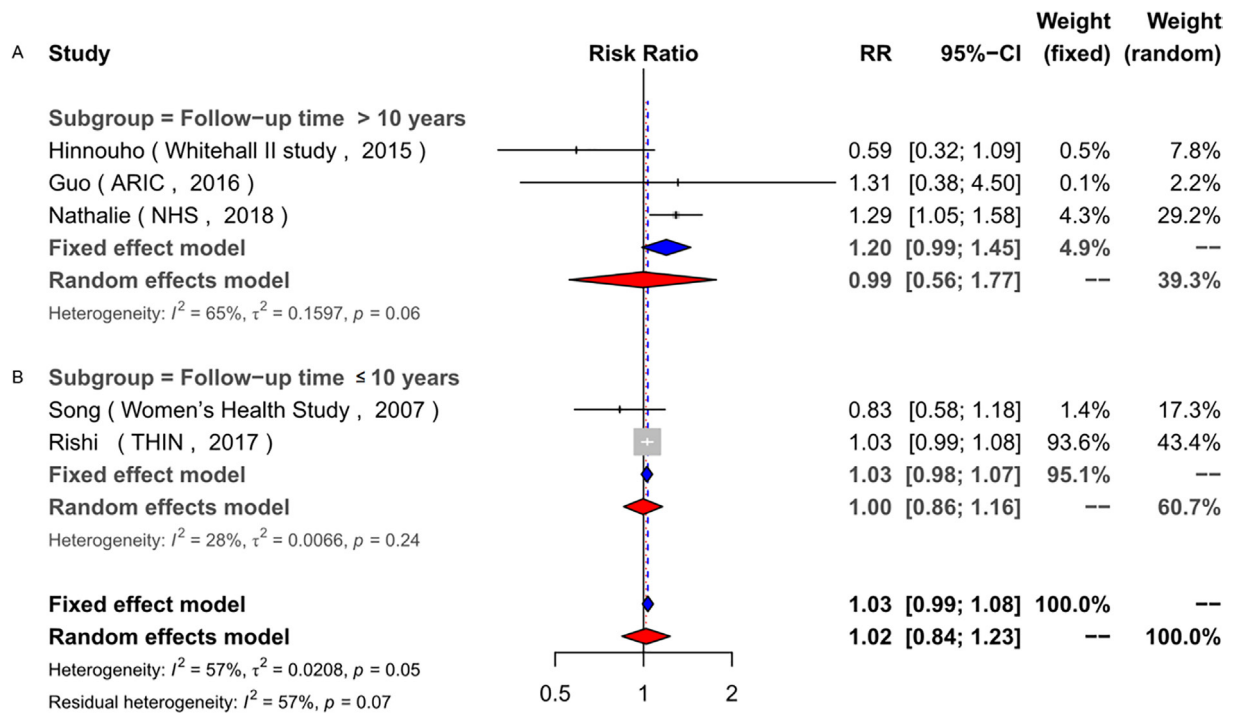


Figure S2 Subgroup analyses' results of the Metabolically Healthy Overweight phenotype.

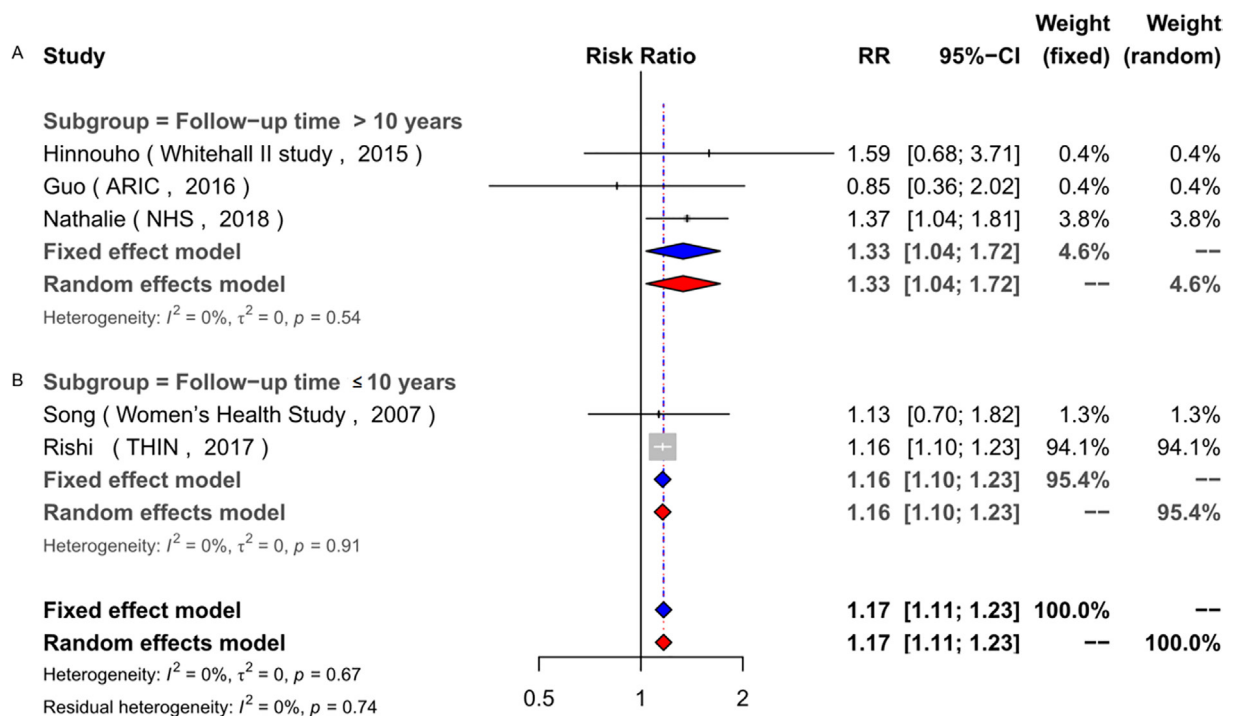


Figure S3 Subgroup analyses' results of the Metabolically Healthy Obesity phenotype.

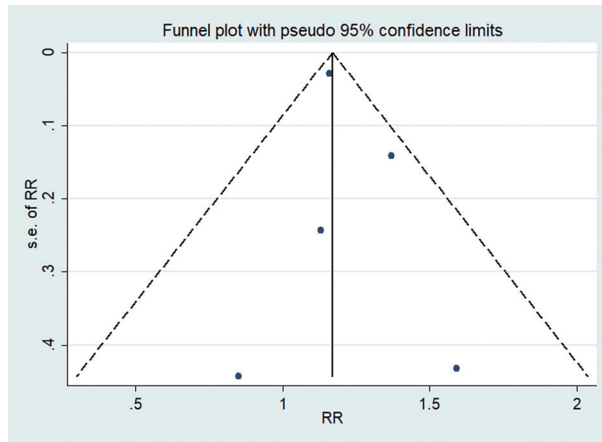


Figure S4 The funnel plot of studies assessing the association between MHO group and the risk of stroke.

| MHO group and risk of stroke compared to MHNW for Stroke | | | |
|---|--|-----------------------------------|----------------------------------|
| Patient or population: Stroke Setting: Intervention: MHO group and risk of stroke Comparison: MHNW | | | |
| Outcomes | Nº of participants (studies) Follow up | Certainty of the evidence (GRADE) | Relative effect (95% CI) |
| RR | (5 observational studies) | ⊕⊕⊕⊙ MODERATE | RR 1.17 (1.11 to 1.23) |
| *The risk in the intervention group (and its 95% confidence interval) is based on the assumed risk in the comparison group and the relative effect of the intervention (and its 95% CI). CI: Confidence interval; RR: Risk ratio | | | |
| GRADE Working Group grades of evidence High certainty: We are very confident that the true effect lies close to that of the estimate of the effect Moderate certainty: We are moderately confident in the effect estimate: The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different Low certainty: Our confidence in the effect estimate is limited: The true effect may be substantially different from the estimate of the effect Very low certainty: We have very little confidence in the effect estimate: The true effect is likely to be substantially different from the estimate of effect | | | |

Figure S5 The credibility of result categorized with GRADEpro GDT.

Table S2 Factors with a change of statistical significance after conducting sensitivity analysis

| Body mass index-metabolic status phenotypes | Result | RR | LCI | UCI | I ² | P | Possible source of heterogeneity | A or B is more credible |
|---|--|------|------|------|----------------|-------|---|-------------------------|
| Metabolically Healthy and BMI ≥ 25 kg/m ² | Primary result | 1.09 | 0.99 | 1.19 | 36% | 0.14 | | |
| | Result after excluding Nathalie Eckel (2018) | 1.07 | 1.04 | 1.11 | 0% | 0.55 | The diagnosis of metabolic abnormalities is not based on laboratory tests | B |
| Metabolically Healthy Overweight | Primary result | 1.02 | 0.84 | 1.23 | 57% | 0.05 | | |
| | Result after excluding Nathalie Eckel (2018) | 0.93 | 0.74 | 1.16 | 36% | 0.20 | The diagnosis of metabolic abnormalities is not based on laboratory tests | B |
| Metabolically Unhealthy Normal Weight | Primary result | 1.83 | 1.57 | 2.14 | 86% | 0.01 | | |
| | Result after excluding Rishi Caleyachetty (2017) | 1.95 | 1.71 | 2.22 | 52% | 0.05 | The diagnosis of metabolic abnormalities is not based on laboratory tests | B |
| Metabolically Unhealthy Overweight | Primary result | 1.93 | 1.44 | 2.58 | 88% | <0.01 | | |
| | Result after excluding Rishi Caleyachetty (2017) | 2.23 | 1.95 | 2.54 | 0% | 0.58 | The diagnosis of metabolic abnormalities is not based on laboratory tests | B |
| Metabolically Unhealthy Obese | Primary result | 2.00 | 1.40 | 2.87 | 91% | <0.01 | | |
| | Result after excluding Rishi Caleyachetty (2017) | 2.30 | 1.73 | 3.06 | 30% | 0.23 | The diagnosis of metabolic abnormalities is not based on laboratory tests | B |

BMI, body mass index; A, the primary result; B, the result after excluding one literature; RR, risk factor; LCI, low confidence interval; UCI, upper confidence interval.