**Supplementary Table 4.** Results of the quality assessment of eligible studies according to the cohort subscale of the Newcastle-Ottawa scale

| Study | Selection | Comparability | Outcome | Total score |
|---|---|---|---|---|
| Tanizaki et al. (2000)[79] | ★★★★ | ★★ | ★★★ | 9 |
| van der Veen et al. (2015)[80] | ☆★★★ | ★★ | ★★☆ | 7 |
| Haring et al. (2017)[61] | ★★★☆ | ★★ | ★★☆ | 7 |
| Cermakova et al. (2017)[54] | ★★★☆ | ★☆ | ★★★ | 7 |
| Davis et al. (1998)[57] | ☆★★★ | ☆☆ | ★★★ | 6 |
| Fox et al. (2005)[59] | ★★★☆ | ★★ | ★☆☆ | 6 |
| Ferreira et al. (2017)[58] | ☆★★☆ | ★★ | ★★☆ | 6 |
| Johansen et al. (2018)[21] | ★★★☆ | ★★ | ★☆☆ | 6 |
| Hirose et al. (2011)[64] | ★★★☆ | ★★ | ★☆☆ | 6 |
| Moore et al. (2018)[20] | ★★★☆ | ★★ | ★☆☆ | 6 |
| Nakanishi et al. (2017)[22] | ★★★☆ | ★★ | ★☆☆ | 6* |
| Pirinen et al. (2017)[74†] | ☆★★☆ | ★★ | ★☆☆ | 5 |
| Lee et al. (2004)[69] | ☆★★☆ | ★★ | ★☆☆ | 5 |
| Butenaerts et al. (2016)[53] | ☆★★☆ | ★★ | ★☆☆ | 5 |
| Martinez-Vea et al. (2006)[72] | ☆★★☆ | ★★ | ★☆☆ | 5 |
| Jeerakathil et al. (2004)[66] | ★★★☆ | ★☆ | ★☆☆ | 5 |
| Vedala et al. (2019)[55] | ☆★★☆ | ★★ | ★☆☆ | 5 |
| Kohara et al. (1999)[68] | ☆★★☆ | ★☆ | ★☆☆ | 4 |
| Das et al. (2008)[56] | ★★★☆ | ★☆ | ☆☆☆ | 4 |
| Henskens et al. (2009)[63] | ☆★★☆ | ★☆ | ★☆☆ | 4 |
| Ryu et al. (2014)[75] | ☆★☆☆ | ★★ | ★☆☆ | 4 |
| Selvetella et al. (2003)[76] | ☆★★☆ | ☆☆ | ★☆☆ | 3 |
| Kawamoto et al. (1991)[67] | ☆★★☆ | ☆☆ | ★☆☆ | 3 |
| Ikeda et al. (1994)[65] | ☆★★☆ | ☆☆ | ★☆☆ | 3 |
| Görner et al. (2007)[60] | ☆★★☆ | ☆☆ | ★☆☆ | 3 |
| Sierra et al. (2002)[78] | ☆★★☆ | ☆☆ | ★☆☆ | 3 |
| Shimada et al. (1990)[77] | ☆★★☆ | ☆☆ | ★☆☆ | 3 |
| Hénon et al. (1996)[62] | ☆★☆☆ | ★☆ | ★☆☆ | 3 |
| Longstreth et al. (1996)[71] | ☆★☆☆ | ★☆ | ★☆☆ | 3 |
| Lee et al. (2018)[70] | ☆★★☆ | ☆☆ | ★☆☆ | 3 |
| Mounier-Vehier et al. (1993)[73] | ☆★☆☆ | ☆☆ | ☆☆☆ | 1 |

Selection items include: representativeness of the exposed cohort, selection of the non-exposed cohort, ascertainment of exposure, and exclusion of outcome presence at start of study. Comparability items include: comparability for age and for hypertension. Outcome items include: assessment of outcome, length of follow-up, and adequacy of follow-up cohorts. Articles are sorted in order of decreasing total quality score.

*Only 5 points are scored regarding the outcome "lacunes" due to inadequate assessment of outcome; †Case-control study. No point is awarded for the representativeness of the exposed cohort criterion.