Version of Record: https://www.sciencedirect.com/science/article/pii/S0091743522001797
Manuscript_54d87f9bf548756d0d7f8daee9533d62

Word count
Abstract: 206
Text: 3005

Should we abandon annual physical examination? -- A meta-analysis of annual physical examination and all-cause mortality in adults based on observational studies

Rashmi Pathak [1,2], Di Kang [1,3], Yuanyuan Lu [1], Fahad Mansuri [1],

Stephanie Kasen [4,5], Yunlong Deng [6], Henian Chen [1]

1. Study Design and Data Analysis Center, College of Public Health, University of South Florida, USA
2. University of Oklahoma Health Science Center, USA
3. Department of Bioinformatics and Biostatistics, The Moffitt Cancer Center & Research Institute, USA
4. Department of Psychiatry, Columbia University, New York, NY, USA
5. Division of Epidemiology, New York State Psychiatric Institute, New York, NY, USA
6. The 3rd Xiangya Hospital, Central South University, Changsha, Hunan Province, China

Corresponding Author: Henian Chen, PhD.
Study Design and Data Analysis Center
College of Public Health, University of South Florida
13201 Bruce B. Downs Blvd., MDC 56, Tampa, FL 33612, USA
Email: hchen1@usf.edu
Telephone: 813-974-4285
Fax: 813-974-4719

Abstract

Several meta-analyses based on randomized clinical trials data have failed to find an association between the annual physical examination (APE) and reduced mortality; however, no comparable meta-analysis based on observational data exists. We conducted a meta-analysis of observational studies comparing APE versus non-APE in adults for all-cause mortality. English-language searches of four databases (PubMed, CINAHL, EMBASE, and Google Scholar) between the years 2000 to 2019 yielded seven observational studies that investigated APE versus non-APE in healthy adults in relation to all-cause mortality. Random effects models were used to calculate pooled hazard ratios and 95% confidence intervals (CI), and to incorporate variation between studies. During follow-up periods that ranged from two to 25 years, there were 35,055 deaths among 633,957 participants. APE was significantly associated with a 45% lower hazard of all-cause mortality, with pooled hazard ratio of 0.55 (95% CI 0.48 to 0.64, P<0.01) for all participants. This meta-analysis of seven observational studies in the past 20 years provides evidence of an association between APE and a lower hazard of all-cause mortality, a finding that contrasts with findings based on meta-analyses of randomized clinical trials data. Nonetheless, at present the evidence available about the effectiveness or ineffectiveness of APE on all-cause mortality still needs further study.

Key Words: Annual physical examination; All-cause mortality; Observational study; Hazard ratio; Meta-analysis

Introduction

The annual physical examination (APE) of asymptomatic individuals, typically conducted in conjunction with a review of health history, laboratory testing, and a dialogue on health behaviors and risks (Birtwhistle et al., 2017), is a widely-employed and traditional strategy for facilitating preventive health care (Bjerregaard et al., 2017; Goodyear-Smith, 2013; Goto et al., 2019; Grosios et al., 2010; Oboler et al., 2002; Seiler et al., 2014). APE is used to promote secondary and tertiary prevention efforts to control the rate of disease, and for the early detection and treatment of disease. Since 1948, when Great Britain put the APE into practice, the British National Health Service Act has ensured the availability of preventive health services for the public (Grosios et al., 2010). Moreover, the British government regularly provides free health check-ups specifically for adults aged 40-75 at low risk with the goal of lowering the rate of preventable deaths by detecting and treating cardiovascular risk factors (e.g., hypertension, high cholesterol, obesity, etc.) (Goodyear-Smith, 2013). In Germany, all adults 35 years and older are provided with a preventive health check-up funded by health insurance since 1989 (Kahl et al., 1999). Other countries (e.g., in the Netherlands and Australia) also have instituted such health check-up policies (Gotzsche et al., 2014; Si et al., 2014). In Japan, all adults aged ≥40 are eligible for free APEs; however, working individuals are required by law to have an APE (Goto et al., 2019). In the United States (US), the Patient Protection and Affordable Care Act (ACA), implemented in 2010, mandated that public and private insurances cover clinical preventive services, including an APE (Seiler et al., 2014).

APE has been used as a preventive measure to reduce chronic disease burden in the adult population, mainly targeting risk factors such as body mass index (BMI), blood pressure, serum glucose, and lifestyle behaviours to reduce the risk of developing a chronic disease. APE may also utilize surrogate outcomes to assess the risk of developing a chronic

disease, for instance, hypertension as an indicator of heart disease, or increased creatinine levels as an indicator of kidney disease. APE is often used as a tool to detect such potential harbingers of more serious maladies in order to lessen the growing burden of chronic disease in the adult population. The usefulness of APE in contrast with its high cost in terms of time, financial and professional resources has been debated. Many clinicians and researchers question the effectiveness of APE as a means of improving health outcomes and reducing mortality (Birtwhistle et al., 2017; Chacko and Anderson, 2007; Goroll, 2015; Howard-Tripp, 2011; Mavriplis, 2011; Mehrotra and Prochazka, 2015). Others argue that there is insufficient information to even determine the ineffectiveness of APE and predict that abandoning APE will disproportionately affect low-income and elderly adults. (Himmelstein and Phillips, 2016).

In 2007, Boulware et al. conducted a systematic review of 50 articles (both clinical trials and observational studies) from 1973 to 2004 in order to examine the effectiveness of physical health examination. The study showed that a physical health examination improved delivery of some preventive services; however, outcomes were heterogeneous and study type included in the review were not well defined (Boulware et al., 2007). In 2012, Krogsbøll et al. published a review of 14 clinical trials and concluded that participants receiving general health checks did not show lower total mortality, as indicated by a risk ratio of 0.99 (95% CI 0.95 to 1.03) based on 9 trials with total mortality outcome data (155899 people and 11940 deaths). That review has been cited in several medical research articles, including the 2016 "European Guidelines on cardiovascular disease prevention in clinical practice," (Piepoli et al., 2016), and may have had a disproportionate effect on informing such guidelines as well as on clinical and research strategies such as APE (Birtwhistle et al., 2017). On the other hand, a 2014 review of six clinical trials comparing end-point differences between the intervention and control arm reported that general practice-based health checks were

associated with statistically significant improvement in total cholesterol, blood pressure and BMI; however, there was no difference in risk for mortality (Si et al., 2014). Clinical trials may not be the most appropriate study design for establishing how effective a tool an APE is in contributing to lower rates of mortality due to their limited sample sizes, short follow-up periods, contamination, spill-over effect, and inability to blind participants. Most of clinical trials conducted in Europe and the USA and most of the observational studies from Asian population. There are no systematic reviews that solely use observational studies to examine the effectiveness of the APE as a predictor of risk of disease or of all-cause mortality. To address this gap, the primary purpose of this study is to examine whether APE is associated with all-cause mortality rates using results based on observational studies.

Methods

Search strategy

In order to carry out a meta-analysis of observational studies that examined the associations of APE with risk of all-cause mortality, we followed guidelines in the Meta-analysis of Observational Studies in Epidemiology (MOOSE) (Stroup et al., 2000) for performing and reporting the results of the present meta-analysis. A systematic literature search was conducted using a combination of PubMed, CINAHL, EMBASE and Google Scholar. Appropriate keywords or groupings of keywords such as general health check-ups, primary intervention, annual physical examination, well-care visits, primary prevention or screening, and mortality, were used to perform the literature search. We identified additional articles by manually searching the reference lists from recent reviews and the extracted papers, limiting our search to publications in the English language in the past 20 years, that is between the years 2000 through 2019.

Study selection

The pre-specified inclusion criteria were: 1) Types of study: observational studies (case-control, cohort, and cross-sectional); 2) Participants: healthy general adult populations with age ≥ 18 years; 3) Intervention: APE; 4) Comparison: no APE or usual care; 5) Outcomes: all-cause mortality; 6) Effect size: hazard ratio (HR). Three independent investigators (RP, DK, YL) conducted an initial screening of all the abstracts and then evaluated all potentially relevant articles based on full text reviews. Studies were excluded if they did not meet all criteria.

Data extraction and quality assessment

Three authors (RP, DK, YL) independently performed the extraction of data from all the seven eligible studies using a standardized procedure developed by the Cochrane Eyes and Vision Group, and results were compared using a common measure of association. We used hazard ratio (HR) as a measure of the association. The outcome of interest in this study was all-cause mortality. We recorded the following characteristics in all identified studies: first author, publication year, age, sex, sample size, number (%) of deaths, duration of follow-up, numbers of APE, and study place. Regarding inclusion of studies and interpretation of data, a fourth investigator (FM) was consulted to resolve any discrepancies. Disagreements were settled through consensus with all four authors and the research team. We used the Risk Of Bias In Non-randomized Studies of Interventions (ROBINS-I) assessment tool to formally assess the risk of bias in the studies included in our analysis (Sterne et al., 2016). The pre-intervention domains include confounding, the bias in selecting study participants, and bias in the classification of interventions were assessed for the studies included in the analysis. Likewise, the post-intervention domains of the tool, examining the biases due to deviations in intervention, missing data, outcome measurement, and reporting of results were assessed.

Data analysis

Review Manager 5.3 (RevMan 2014) analysis software (Collaboration, 2014) and R metafor package (Viechtbauer, 2010) were used to analyse the data. This meta-analysis used HR and the 95% confidence interval (CI) to measure the effect size of each study. Estimates of HR were pooled using a random-effect model, and were weighted using the generic inverse variance. Cochran's Q test and Higgins I^2 index statistics were used to measure the heterogeneity among studies. All statistical tests were two-sided and statistical significance was defined as $p< 0.05$. Gender is a potential effect modifier of APE and mortality association and some of the studies included in the analysis have reported results for gender sub-groups. Accordingly, 95% CIs for estimates of overall HR and of HR stratified by gender are reported here. Begg's test and Egger's test were used to evaluate publication bias (Viechtbauer, 2010). This meta-analysis was based on published papers and met our institution's guidelines for protection of human subjects concerning their safety and privacy.

Results

Literature search

Figure 1 shows the study selection process and results of the literature search. We identified a total of 2816 publications in the initial search, of which 1625 were published between 2000 and 2019. Of those 1625 publications, an additional 1298 were removed. Further review of the remaining 327 publications resulted in the elimination of an additional 252 papers. Among the 75 publications that were not yet excluded, 68 papers either did not have a control group or the effect size measure was not hazard ratio (or both). The remaining seven studies (Chiou and Chang, 2002; Henny et al., 2012; Hozawa et al., 2010; Igarashi et al., 2019; Iwasa et al., 2007; Lee et al., 2015; Suh et al., 2017) met the full list of study

inclusion criteria and were included in our systematic review and meta-analysis. Table 1 shows the results of the risk of bias analysis from the ROBINS-I tool (Sterne et al., 2016).

Study characteristics

Table 2 shows the characteristics of the seven studies in this systematic review, all of which examined the impact of APE on all-cause mortality in a healthy population over the age of 40. A total of 633,957 participants participated. The length of follow-up ranged from 2 to 25 years with a median of six years. The total numbers of participants (from 854 to 443,337) and deaths (from 78 to 21,880, respectively) varied widely across the seven studies. Five studies reported HRs for the entire sample only (Chiou and Chang, 2002; Igarashi et al., 2019; Iwasa et al., 2007; Lee et al., 2015; Suh et al., 2017), one study reported HRs for the entire sample and separately for males and females (Hozawa et al., 2010), and one study reported HR by gender only, and not for the entire sample (Henny et al., 2012). We carried out the meta-analysis for all participants 'based on the six studies (Chiou and Chang, 2002; Hozawa et al., 2010; Igarashi et al., 2019; Iwasa et al., 2007; Lee et al., 2015; Suh et al., 2017) and by gender using two studies (Henny et al., 2012; Hozawa et al., 2010).

APE and hazard of all-cause mortality

Table 3 shows the results of the pooled analyses for the seven included studies in this meta-analysis. The association between APE and hazard of all-cause mortality in the total samples was evaluated in six studies (Chiou and Chang, 2002; Hozawa et al., 2010; Igarashi et al., 2019; Iwasa et al., 2007; Lee et al., 2015; Suh et al., 2017), comprising 607,858 participants and 31,498 deaths. The pooled hazard ratio of all-cause mortality was 0.55 (95% confidence interval 0.48 to 0.64; P<0.01), with significant heterogeneity (P<0.01, I^2=81%) (see Figure 2). No significant publication bias was detected (Begg's test P=0.27, Egger's test P=0.25). The association between APE and hazard of all-cause mortality by gender was evaluated in two studies (Henny et al., 2012; Hozawa et al., 2010), comprising 74,874

participants and 10,842 deaths. The pooled hazard ratios of all-cause mortality for males and females were 0.71 (95% CI 0.67 to 0.75; P<0.01) and 0.72 (95% CI 0.60 to 0.86; P<0.01) (see Figure 3).

Discussion

Findings based on this meta-analysis of seven observational studies conducted over the past 20 years indicate that the APE is significantly associated with a 45% lower hazard of all-cause mortality across all adult participants (age ≥ 18 years). Utilization of an APE among asymptomatic individuals (also referred to as an annual health check-up or medical examination), was proposed initially by Horace Dobell, a British physician, in 1861; since then, it has been used in several countries as a tool to enhance preventive health care (Bjerregaard et al., 2017; Hoebel et al., 2014; Rotarou and Sakellariou, 2018; Schülein et al., 2017). To our knowledge, this study is the first meta-analysis in the literature to analyse the association between APE and all-cause mortality based only on observational studies. In general, randomized clinical trials have been the preferred study design to assess evidence of treatment effectiveness. Two meta-analyses based on clinical trials data concluded that general health checks are not beneficial in reducing the risk of all-cause mortality (Krogsbøll et al., 2012; Si et al., 2014). Krogsbøll et al. (2012) included participants from the age group 18-65 years old, whereas our meta-analysis was comprised of studies of adults who were at least 40 years old, a population more likely to harbor undetected risk factors for cardiovascular disease, diabetes-2, cancer, prostate disease, and other maladies that are found once middle-age sets in. Most clinical trials included in Krogsbøll et al. (2012) meta-analysis were conducted before 1975. Consequently, treatment availability and effectiveness at the time of those trials most likely differed substantially from treatment availability and effectiveness in more recent years. In the meta-analysis by Si, et al., (2014) the primary

endpoints of the six studies analysed are disease risk factors (e.g., high cholesterol, high blood pressure) as opposed to mortality. In their limitations section, Si, et al. (2014) note that the four studies that included mortality were likely under-powered. Randomized clinical trials might not be ideal for assessing the association between APE and mortality. Certain weaknesses often inherent in clinical trials data such as limited sample sizes, short follow-up periods, contamination, spill-over effects, and inability to mask participants, make clinical trials data less appropriate than data obtained from observational studies for assessing the association between APE and mortality (Herbert et al., 2018). Consequently, although often used to advocate for the elimination of APE as a preventive treatment strategy, meta-analyses only based on clinical trials are not the ideal study design for demonstrating the potential effectiveness of APE in reducing mortality

The findings here should be interpreted in light of the study limitations: The first is the inclusion of only seven studies in the meta-analysis, a limitation due predominantly to the lack of a control group (an inclusion criterion) in many published observational studies. Second, only two studies in our meta-analysis used advanced methods such as propensity score matching to adjust for confounders, raising some concern that the results of other studies included in this analysis may be prone to selection bias and residual confounding. Third, observational studies with non-significant results ($p<0.05$) are less likely to be published as compared to clinical trials with non-significant results, and may lead to the misconception that observational studies are more likely to detect a statistically significant association between APE and mortality than similar studies based on clinical trials (Herbert et al., 2018). Fourth, six of the seven studies in this meta-analysis were conducted in Asia (see Table 2), where most observational studies are conducted; that is in contrast to randomized clinical trials, which are conducted predominantly in Europe and the USA. Consequently, our findings may not be generalizable to participants of other nationalities. On

the other hand, this study also has notable strengths: It is the first meta-analysis to use observational study data to examine the association between APE and all-cause mortality. We reviewed all the publications in the English language in the past 20 years. The median follow-up period in our studies was 6 years (range 2-25 years) ensuring a follow-up period sufficiently long enough to assess mortality in study participants.

In summary, contrasting results, whether across observational studies and clinical trial randomized studies or even within each of those two study types, may due in part to variations in length of follow-up periods; study population demographics (e.g., age, ethnicity or nationality); unadjusted health-related baseline differences between cases and controls; and potential bias in publication criteria applied to studies based on clinical trial data versus studies based on observational data. To address that going forward future studies should consider examining potential patient factors such as age or socioeconomic status that could moderate the association between APE and mortality. Another aspect of APE to consider would be the frequency with which it is used. Although frequency is explicit in the term Annual Physical Examination, it may be a more efficient use of APE to vary frequency according to age or gender or other relevant patient quality. That would contribute to increased specificity as to whether the effectiveness of APE as a treatment strategy differs across populations. In conclusion, past meta-analyses conducted using randomized clinical trial data have concluded that APE in healthy adults does not reduce risk of mortality (Krogsbøll et al., 2012; Si et al., 2014), whereas our meta-analysis using observational data shows that APE significantly reduces the hazard of mortality in adult populations. Nonetheless, at present the evidence available about the effectiveness or ineffectiveness of APE on mortality still needs further study. We strongly recommend that APE not be abandoned. Rather, we suggest that in planning a study of APE and all-cause mortality based on either clinical trials or observational studies, more careful consideration be given to

implementing strategies that can counteract the potential weaknesses characteristic of each of those design types: For example, to strengthen clinical trials, lengthening follow-up periods and applying less restrictive eligibility criteria that exclude certain populations, especially those more likely to benefit from APE (e.g., the elderly), avoiding a spill-over effect (e.g., by boosting interview compliance in controls), whereas to bolster observational studies, adjusting for confounders by using advanced statistical methods (e.g., propensity score matching). Moreover, given the unique strengths of each of those designs, we recommend that future meta-analyses combine data from both clinical trials and observational studies. Accordingly, findings based on those two types of design paradigms would then allow for a more accurate accounting of just how useful a tool APE is for reducing all-cause mortality and, given the alternative yet complementary strengths of each of those designs, generate more conclusive evidence.

The authors declare no competing interests.
This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

Birtwhistle, R., Bell, N.R., Thombs, B.D., Grad, R., Dickinson, J.A., 2017. Periodic preventive health visits: a more appropriate approach to delivering preventive services: From the Canadian Task Force on Preventive Health Care. Can Fam Physician 63:824-26.

Bjerregaard, A.L., Maindal, H.T., Bruun, N.H., Sandbaek, A., 2017. Patterns of attendance to health checks in a municipality setting: the Danish 'Check Your Health Preventive Program'. Prev Med Rep 5:175-82.

Boulware, L.E., Marinopoulos, S., Phillips, K.A., Hwang, C.W., Maynor, K., Merenstein, D., Wilson, R.F., Barnes, G.J., Bass, E.B., et al., 2007. Systematic review: the value of the periodic health evaluation. Annals of internal medicine 146:289-300.

Chacko, K.M., Anderson, R.J., 2007. The annual physical examination: important or time to abandon? Am J Med 120:581-3.

Chiou, C.-J., Chang, H.-Y., 2002. Do the elderly benefit from annual physical examination? An example from Kaohsiung City, Taiwan. Preventive medicine 35:264-70.

Collaboration, C., 2014. RevMan 5.3 user guide. The Cochrane Collaboration 5.

Dobell, H., 1861. Lectures on the germs and vestiges of disease: and on the prevention of the invasion and fatality of disease by periodical examinations; delivered at the royal infirmary for diseases of the chest. John Churchill.

Goodyear-Smith, F., 2013. Government's plans for universal health checks for people aged 40-75. BMJ 347:f4788.

Goroll, A.H., 2015. Toward trusting therapeutic relationships—in favor of the annual physical. N Engl J Med 373:1487-9.

Goto, E., Ishikawa, H., Okuhara, T., Kiuchi, T., 2019. Relationship of health literacy with utilization of health-care services in a general Japanese population. Prev Med Rep 14:100811.

Gotzsche, P.C., Jorgensen, K.J., Krogsboll, L.T., 2014. General health checks don't work. BMJ 348:g3680.

Grosios, K., Gahan, P.B., Burbidge, J., 2010. Overview of healthcare in the UK. EPMA J 1:529-34.

Henny, J., Paulus, A., Helfenstein, M., Godefroy, T., Guéguen, R., 2012. Relationship between the achievement of successive periodic health examinations and the risk of dying. Appraisal of a prevention scheme. J Epidemiol Community Health 66:1092-96.

Herbert, R.D., Kasza, J., Bø, K., 2018. Analysis of randomised trials with long-term follow-up. BMC Medical Research Methodology 18:48.

Himmelstein, D.U., Phillips, R.S., 2016. Should We Abandon Routine Visits? There Is Little Evidence for or Against. Ann Intern Med 164:498-9.

Hoebel, J., Starker, A., Jordan, S., Richter, M., Lampert, T., 2014. Determinants of health check attendance in adults: findings from the cross-sectional German Health Update (GEDA) study. BMC Public Health 14:913.

Howard-Tripp, M., 2011. Should we abandon the periodic health examination?: YES. Canadian Family Physician 57:158-60.

Hozawa, A., Kuriyama, S., Watanabe, I., Kakizaki, M., Ohmori-Matsuda, K., Sone, T., Nagai, M., Sugawara, Y., Nitta, A., et al., 2010. Participation in health check-ups and mortality using propensity score matched cohort analyses. Preventive medicine 51:397-402.

Igarashi, Y., Okuno, T., Kodera, K., Iritani, O., Hamazaki, Y., Himeno, T., Yano, H., Okuro, M., Morita, T., et al., 2019. Non‐participation in health checkup and Kihon Checklist predicts loss of certification‐free survival in community‐dwelling older adults. Geriatrics & Gerontology International 19:1206-14.

Iwasa, H., Yoshida, H., Kim, H., Yoshida, Y., Kwon, J., Sugiura, M., Furuna, T., Suzuki, T., 2007. A mortality comparison of participants and non-participants in a comprehensive health examination among elderly people living in an urban Japanese community. Aging clinical and experimental research 19:240-45.

Kahl, H., Holling, H., Kamtsiuris, P., 1999. [Utilization of health screening studies and measures for health promotion]. Gesundheitswesen 61 Spec No:S163-8.
Krogsbøll, L.T., Jørgensen, K.J., Larsen, C.G., Gøtzsche, P.C., 2012. General health checks in adults for reducing morbidity and mortality from disease: Cochrane systematic review and meta-analysis. Bmj 345, e7191.
Lee, H., Cho, J., Shin, D.W., Lee, S.-P., Hwang, S.-S., Oh, J., Yang, H.-K., Hwang, S.-H., Son, K.Y., et al., 2015. Association of cardiovascular health screening with mortality, clinical outcomes, and health care cost: a nationwide cohort study. Preventive medicine 70:19-25.
Mavriplis, C.A., 2011. Should we abandon the periodic health examination?: NO. Canadian Family Physician 57:159-61.
Mehrotra, A., Prochazka, A., 2015. Improving Value in Health Care--Against the Annual Physical. N Engl J Med 373:1485-7.
Oboler, S.K., Prochazka, A.V., Gonzales, R., Xu, S., Anderson, R.J., 2002. Public expectations and attitudes for annual physical examinations and testing. Ann Intern Med 136:652-9.
Piepoli, M.F., Hoes, A.W., Agewall, S., Albus, C., Brotons, C., Catapano, A.L., Cooney, M.T., Corrà, U., Cosyns, B., et al., 2016. 2016 European Guidelines on cardiovascular disease prevention in clinical practice: The Sixth Joint Task Force of the European Society of Cardiology and Other Societies on Cardiovascular Disease Prevention in Clinical Practice (constituted by representatives of 10 societies and by invited experts)Developed with the special contribution of the European Association for Cardiovascular Prevention & Rehabilitation (EACPR). Eur Heart J 37:2315-81.
Rotarou, E.S., Sakellariou, D., 2018. Determinants of utilisation rates of preventive health services: evidence from Chile. BMC public health 18:839.
Schülein, S., Taylor, K.J., Schriefer, D., Blettner, M., Klug, S.J., 2017. Participation in preventive health check-ups among 19,351 women in Germany. Preventive medicine reports 6:23-26.
Seiler, N., Malcarney, M.B., Horton, K., Dafflitto, S., 2014. Coverage of clinical preventive services under the Affordable Care Act: from law to access. Public Health Rep 129:526-32.
Si, S., Moss, J.R., Sullivan, T.R., Newton, S.S., Stocks, N.P., 2014. Effectiveness of general practice-based health checks: a systematic review and meta-analysis. Br J Gen Pract 64:e47-53.
Sterne, J.A., Hernán, M.A., Reeves, B.C., Savović, J., Berkman, N.D., Viswanathan, M., Henry, D., Altman, D.G., Ansari, M.T., et al., 2016. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. bmj 355.
Stroup, D.F., Berlin, J.A., Morton, S.C., Olkin, I., Williamson, G.D., Rennie, D., Moher, D., Becker, B.J., Sipe, T.A., et al., 2000. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Jama 283:2008-12.
Suh, Y., Lee, C.J., Cho, D.-K., Cho, Y.-H., Shin, D.-H., Ahn, C.-M., Kim, J.-S., Kim, B.-K., Ko, Y.-G., et al., 2017. Impact of National Health Checkup Service on hard atherosclerotic cardiovascular disease events and all-cause mortality in the general population. The American Journal of Cardiology 120:1804-12.
Viechtbauer, W., 2010. Conducting meta-analyses in R with the metafor package. Journal of statistical software 36:1-48.

Table 1. Risk of bias analysis of the studies

	Pre/at- intervention domains			Post-intervention domains			
Study	**Bias due to confounding**	**Bias due to selection of participants**	**Bias in classification of intervention**	**Bias due to deviations from intervention**	**Bias due to missing data**	**Bias in measurement of outcomes**	**Bias in selection of the reported result**
Chiou and Chang, 2002	Moderate	Low	Moderate	Low	Low	Moderate	Moderate
Iwasa et al., 2007	Serious	Serious	Low	Serious	Moderate	Low	Low
Hozawa et al., 2010	Serious	Moderate	Moderate	Low	Low	Moderate	Moderate
Henny et al., 2012	Serious	Moderate	Serious	Low	Moderate	Moderate	Moderate
Lee et al., 2015	Moderate	Serious	Low	Low	Low	Moderate	Moderate
Suh et al., 2017	Serious	Serious	Low	Low	Serious	Moderate	Low
Igarashi et al., 2019	Moderate	Low	Low	Low	Low	Low	Moderate

Table 2: Characteristics of seven studies included in the meta-analysis

Author (year)	Sample size	No of deaths	Follow-up (years)	Age (years)	No. of checks	Area
Chiou and Chang, 2002	1193	130	6	65-89	Study continued for 5 year	Taiwan
Iwasa et al., 2007	854	78	3	70 - 84	2 times	Japan
Hozawa et al., 2010	48775	7285	11	> 40	1 time	Japan
Henny et al., 2012	26099	3557	20-25	40-59	Study continued for 5 year	France
Lee et al., 2015	443337	21880	5	> 40	4 times	South Korea
Suh et al., 2017	110550	1933	2	40-69	Study continued for 5 year	South Korea
Igarashi et al., 2019	3149	192	6	≥ 65	1 time	Japan

Table 3: Meta-analysis of annual physical exam and risk of all-cause mortality for overall, male and female

	Studies	No.	Death/participants	Pooled HR (95% CI) P value	Heterogeneity (I^2) P value	Begg's test, Egger's test
Overall	Chiou and Chang, 2002 Hozawa et al., 2010 Igarashi et al., 2019 Iwasa et al., 2007 Lee et al., 2015 Suh et al., 2017	6	31498/607858	0.55 (0.48 to 0.64) <0.01	81 <0.01	0.27, 0.25
Male	Henny et al., 2012 Hozawa et al., 2010	2	7118/36254	0.71 (0.67 to 0.75) <0.01	0 0.43	1.00, NA
Female	Henny et al., 2012 Hozawa et al., 2010	2	3724/38620	0.72 (0.60 to 0.86) <0.01	78 0.03	1.00, NA

Notes: HR: Hazard Ratio; CI: Confidence Interview

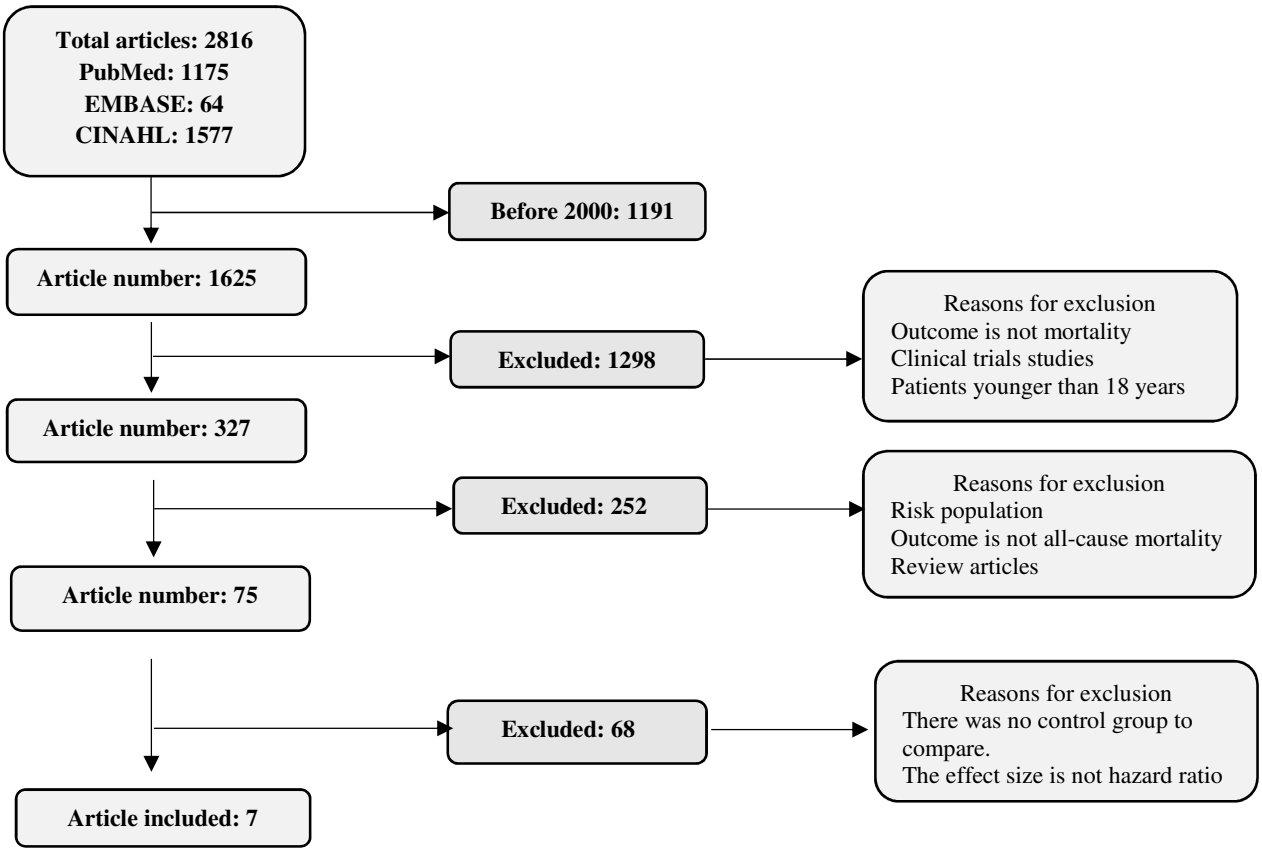
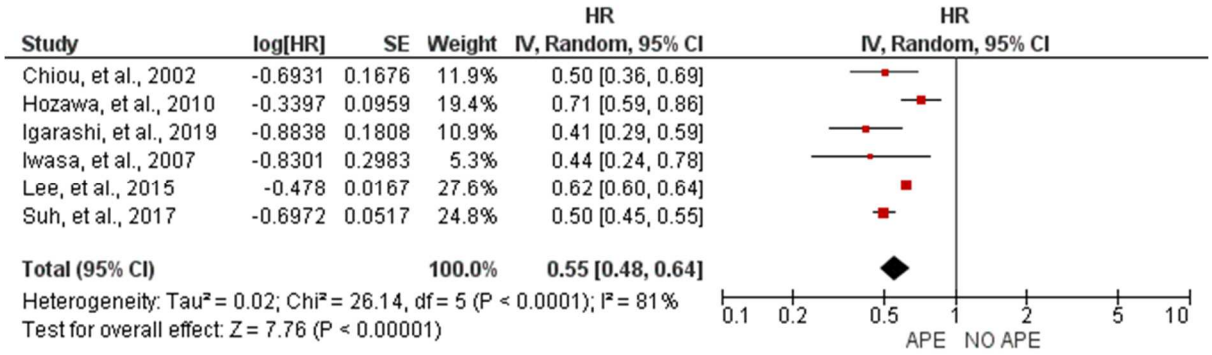
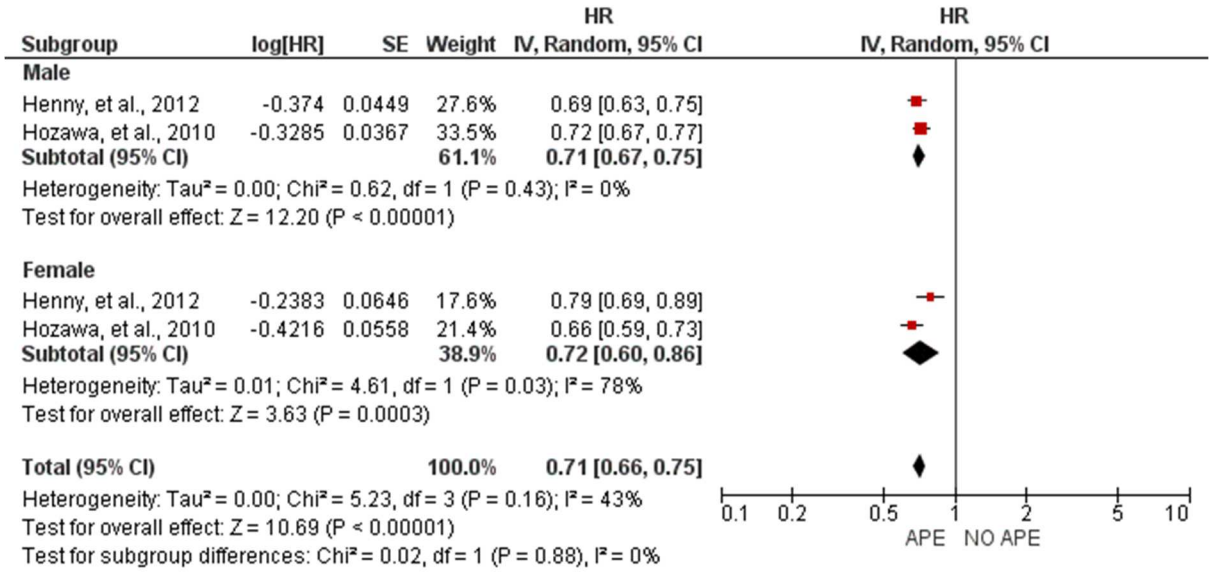


Figure 1. Summary of literature search and review process



Notes: **HR**: Hazard Ratio; **CI**: Confidence Interview; **APE**: Annual Physical Examination

Figure 2: Risk of all-cause mortality associated with annual physical examination (APE)



Notes: **HR**: Hazard Ratio; **CI**: Confidence Interview; **APE**: Annual Physical Examination

Figure 3: Risk of all-cause mortality associated with annual physical examination (APE) by gender