Supplementary Material

Materials and Methods

Cohort Selection and Medical Comorbidities

The Mayo Clinic Cohort Study of Oophorectomy and Aging-2 (MOA-2) is a population-based cohort study of Olmsted County, MN women. The data were collected using the medical records-linkage system of the Rochester Epidemiology Project as reported in detail elsewhere.(1-3) Bilateral oophorectomy was identified by screening the electronic indexes of the Rochester Epidemiology Project for procedure codes for oophorectomy from the International Classification of Diseases, 9th Edition (ICD-9) from January 1, 1988 through December 31, 2007, as previously described.(4, 5) Women who underwent bilateral oophorectomy due to ovarian or other estrogen-sensitive malignancies or due to a high risk of ovarian cancer were excluded.

Assessment of Kidney Function to Apply the Diagnosis of Chronic Kidney Disease (CKD)

Serum creatinine (Cr) and protein measurements were reviewed prior to the index date or >30 days after the index date, to exclude the immediate effects of surgery. Estimated glomerular filtration rate (eGFR) was calculated from IDMS calibrated serum creatinine measurements using the CKD-EPI equation.(6) Both inpatient and outpatient creatinine tests were included. Dipstick measurements of proteinuria were not included. Spot urine protein-to-Cr ratios, protein-to-osmolarity ratios, albumin-to-Cr ratios, and 24-hour urine collections for albumin and total protein excretion were all acceptable measurements (protein-to-osmolarity ratios are routinely reported by the Mayo Clinic laboratory on all spot urine measurements).(7) However, spot urine measurements that were contaminated by non-glomerular hematuria or by an active

urinary tract infection were excluded. Abnormal urinary sediment was defined as presence of casts or dysmorphic hematuria. Women with reduced eGFR levels or markers of kidney damage due to separate episodes of acute kidney injury were not considered to have CKD. Women with structural abnormalities of the kidney identified by imaging, but with normal eGFR and no markers of kidney damage, were not considered to have CKD. This exclusion criterion was used to reduce the possible surveillance bias caused by the higher frequency of abdominal imaging in the oophorectomy group. The onset of CKD was the date of the second qualifying eGFR, abnormal protein measurement, or abnormal urinary sediment (meeting the time gap specified), and CKD was considered incident if onset was on or after the index date.

Statistical Methods – Stratified Analyses and Sensitivity Analyses

Analyses were performed overall and in strata by age at index date (≤45 vs. 46-49 years), and by use of estrogen therapy following the oophorectomy (within age strata; use to the 46th birthday vs. not; or use to the 50th birthday vs. not). Robust sandwich covariance estimates were used to account for women who were included in both cohorts (e.g., referent women who subsequently underwent bilateral oophorectomy). The proportional hazards assumption was assessed using graphical methods and the addition of time-dependent coefficients to the Cox models. None of the models violated this assumption.

We performed three sets of sensitivity analyses to 1) exclude women who did not have any serum Cr measurements in their records and women with an index date before 1994 (only for the eGFR-based definition of CKD), 2) to exclude women with any of the 17 chronic conditions or with onset of CKD by either definition prior to index date, and 3) to censor referent women at the date of subsequent bilateral oophorectomy if performed before age 50 years. The first sensitivity analysis was performed to account for missing Cr data either due to lack of routine

testing for a given woman or due to the date ranges of electronically available data. The second and third sets of sensitivity analyses were determined *a priori*.

Results

Menopausal Status and Estrogen Therapy

Among the 1,653 women who underwent oophorectomy, 157 (10%) had hysterectomy prior to bilateral oophorectomy. Among the 1,653 referent women, 270 (16%) were menopausal at index date (166 surgical due to hysterectomy, 98 natural, and 6 due to radiation or chemotherapy). An additional 944 (57%) referent women became menopausal during follow-up (764 natural, 104 surgical due to bilateral oophorectomy, 67 surgical due to hysterectomy or thermal ablation, and 9 due to radiation or chemotherapy).

The majority of women received some estrogen therapy after bilateral oophorectomy (n=1,494, 90%), whereas only 479 referent women (29%) had estrogen therapy after the index date. Oral estrogen was most commonly used, either alone or in combination with patch usage. A total of 1,265/1,494 (85%) women who underwent oophorectomy and 409/479 (85.4%) referent women took oral estrogen after index. The median (IQR) cumulative years of estrogen therapy after index was 7.1 (4.0-11.0) in the bilateral oophorectomy cohort and 4.9 (IQR 2.5-8.4) in the referent cohort. Among women who were ≤45 years of age at oophorectomy (n=1,031), 63% were on estrogen therapy on their 46th birthday. Among women who were 46-49 years at the time of oophorectomy (n=622), 72% were on estrogen therapy on their 50th birthday.

Exploration of Potential Surveillance Bias

The number of Cr tests and contacts with the healthcare system (visits) were obtained for each woman, excluding any visits or tests performed after the onset of CKD, and excluding a ±1

year time window around oophorectomy (or index date). The density of Cr tests and visits were calculated for each woman as the number of events divided by the corresponding length of medical record. Although both densities were similar in the oophorectomy and referent cohorts, the differences were statistically significant (because of the large sample size). Before index, the median density of Cr tests was 0 per year for both groups (oophorectomy IQR 0-0.11; referent IQR 0-0.07; $P<0.001$), whereas the median visit density was 1.9 per year (IQR 1.1-4.0) for the oophorectomy group and 1.6 (IQR 1.0-3.2; $P<0.001$) for the referent group. After index, the median density of Cr tests was 0.58 per year (IQR 0.23-1.13) for the oophorectomy group and 0.38 (IQR 0.11-0.83; $P<0.001$) for the referent group, whereas the median visit density was 8.4 per year (IQR 5.2-14.2) for the oophorectomy group and 6.4 (IQR 3.8-10.4; $P<0.001$) for the referent group.

Causes of CKD from Medical Records

A total of 77 women who underwent bilateral oophorectomy and 54 referent women had a diagnosis of prevalent or incident CKD in their medical records. The most common cause listed in the records was diabetes mellitus (30 women who underwent bilateral oophorectomy, 21 referent women). CKD was attributed to medications, such as non-steroidal anti-inflammatory drugs or antibiotics, in 16 women who underwent bilateral oophorectomy and in 6 referent women. Kidney biopsies were performed in 6 women who underwent bilateral oophorectomy and in 2 referent women; 7 of 8 biopsies showed glomerular diseases. Six women who underwent bilateral oophorectomy and 2 referent women also developed end-stage renal disease.

References

1. St Sauver JL, Grossardt BR, Yawn BP, Melton LJ, 3rd, Rocca WA. Use of a medical records linkage system to enumerate a dynamic population over time: the Rochester epidemiology project. Am J Epidemiol. 2011;173(9):1059-68.

2. Rocca WA, Yawn BP, St Sauver JL, Grossardt BR, Melton LJ, 3rd. History of the Rochester Epidemiology Project: half a century of medical records linkage in a US population. Mayo Clin Proc. 2012;87(12):1202-13.

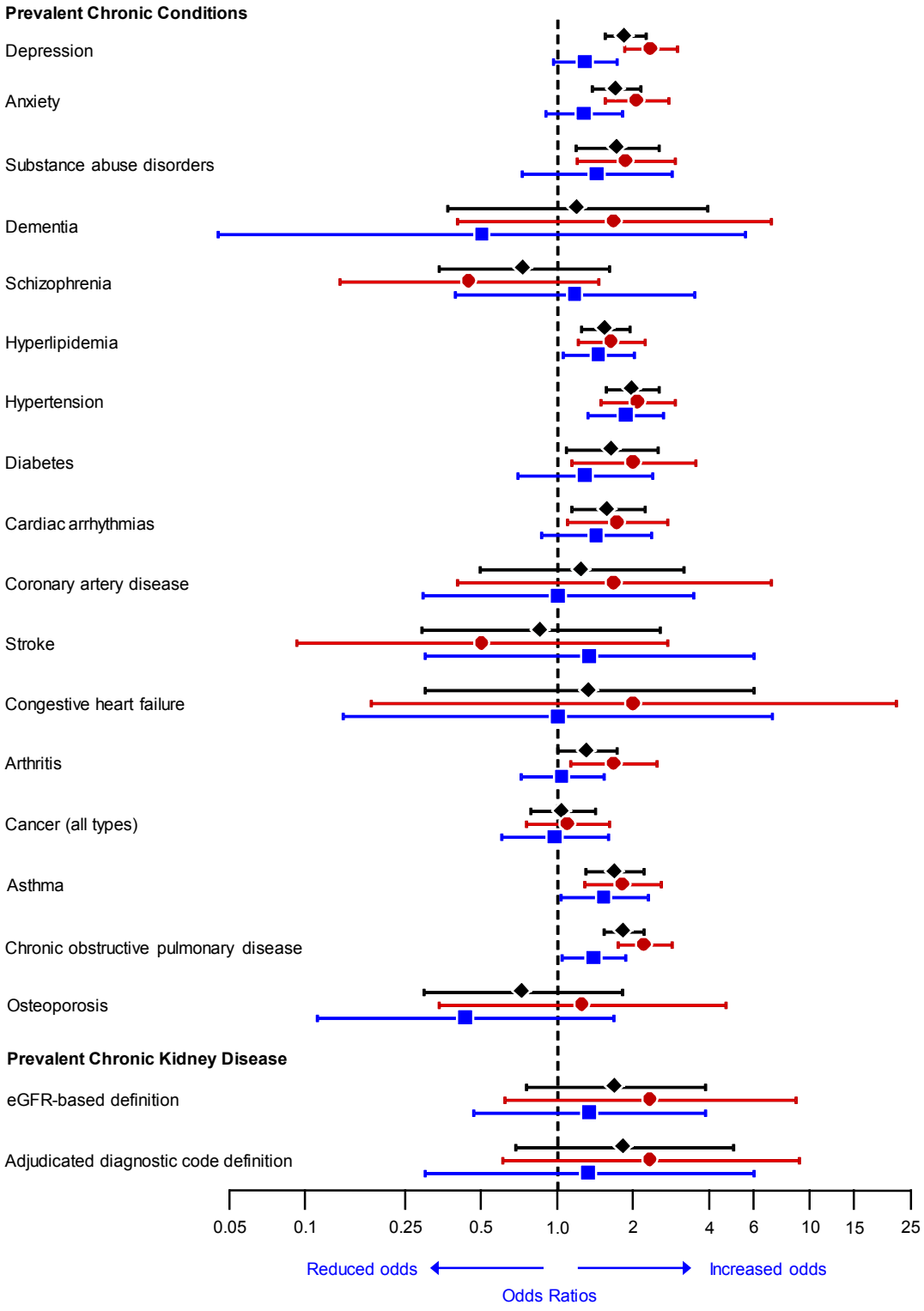
3. St Sauver JL, Grossardt BR, Yawn BP, Melton LJ, 3rd, Pankratz JJ, Brue SM, et al. Data resource profile: the Rochester Epidemiology Project (REP) medical records-linkage system. Int J Epidemiol. 2012;41(6):1614-24.

4. Rocca WA, Gazzuola-Rocca L, Smith CY, Grossardt BR, Faubion SS, Shuster LT, et al. Accelerated Accumulation of Multimorbidity After Bilateral Oophorectomy: A Population-Based Cohort Study. Mayo Clin Proc. 2016.

5. Rocca WA, Gazzuola Rocca L, Smith CY, Grossardt BR, Faubion SS, Shuster LT, et al. Cohort profile: the Mayo Clinic Cohort Study of Oophorectomy and Aging-2 (MOA-2) in Olmsted County, Minnesota (USA). BMJ Open. 2017;7(11):e018861.

6. Levey AS, Stevens LA, Schmid CH, Zhang YL, Castro AF, 3rd, Feldman HI, et al. A new equation to estimate glomerular filtration rate. Ann Intern Med. 2009;150(9):604-12.

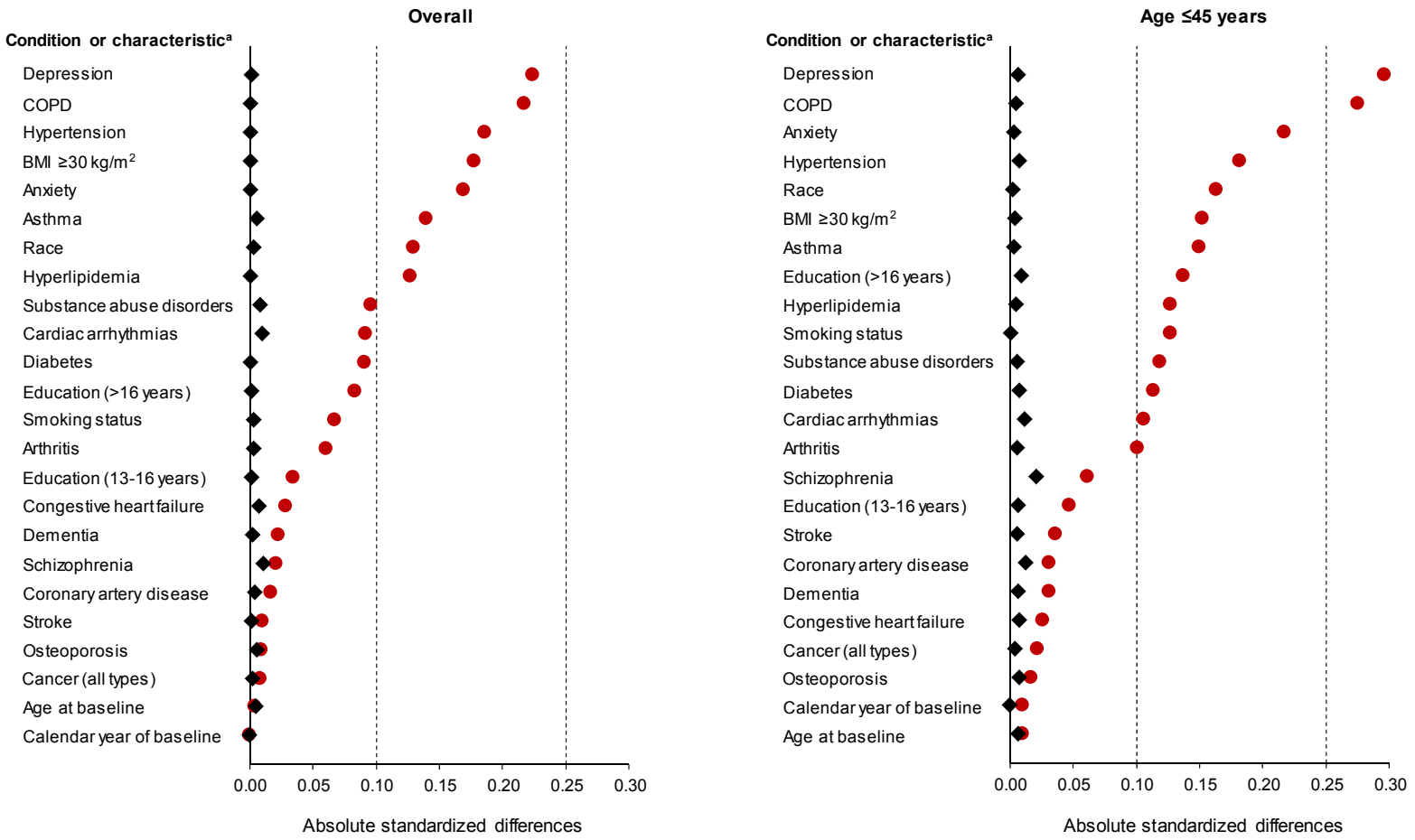
7. Wilson DM, Anderson RL. Protein-osmolality ratio for the quantitative assessment of proteinuria from a random urinalysis sample. Am J Clin Pathol. 1993;100(4):419-24.



Supplementary Figure 1. Case-control analysis of the chronic conditions present at index date.

The odds ratios, confidence intervals, and p-values were calculated using conditional logistic regression models (matched pairs). Analyses are presented overall (black diamond) and in strata by age at oophorectomy (red circles for age ≤45 years and blue squares age 46-49 years). Each one of the 17 chronic conditions was defined as having two diagnostic codes separated by more than 30 days prior to the oophorectomy or index date. Estimated glomerular filtration rate (eGFR)-based chronic kidney disease (CKD) was defined as having 2 or more eGFR values <60 ml/min/1.73 m^2

(calculated from serum creatinine using the CKD-EPI equation) separated by more than 90 days and prior to the oophorectomy or index date. CKD by adjudication of diagnostic codes was defined as having 2 or more eGFR measurements <45 ml/min/1.73 m^2 separated by more than 90 days and prior to the oophorectomy or index date, or as having evidence of kidney damage (proteinuria or active urinary sediment) on at least 2 occasions >90 days apart prior to the oophorectomy or index date.



Supplementary Figure 2. Balance of characteristics at baseline obtained using inverse probability weights for chronic kidney disease defined using eGFR levels, overall (left panel) and in women who underwent bilateral oophorectomy at age ≤45 years (right panel).

The red circles indicate the absolute standardized differences before balancing and the black diamonds indicate these differences after balancing using IPW. After the IPW adjustment, all of the standardized differences were below the recommended threshold of 0.10, denoting negligible imbalance of these characteristics. For each condition or characteristic, the absolute standardized difference is defined as the absolute value of the difference in means for that characteristic between women with and without bilateral oophorectomy, divided by the pooled standard deviation for that characteristic. The weights were derived from propensity scores estimated from logistic regression models including 17 chronic conditions present at baseline, years of education (unknown or ≤12, 13-16, >16), race (white vs nonwhite), body mass index (unknown or <30 vs ≥30 kg/m^2), cigarette smoking (current or former vs never), and age and calendar year at baseline (continuous). These models were fit overall, and separately in each stratum to maximize the balance at the index date. Less than 1% of women with oophorectomy and no referent women fell outside the overlapping ranges (i.e., a wide region of common support). The propensity scores ranged between 0.15 and 0.87 for the overall women with oophorectomy and between 0.21-0.83 for referent women and between 0.06 and 0.91 for the women who underwent oophorectomy at age ≤45 years and between 0.13-0.88 for referent women. Weights greater than 10 were trimmed by setting these weights to the value of the 99th percentile for their respective group (bilateral oophorectomy or referent women). The weights were then stabilized to reduce variability by dividing each weight by the mean

weight for their respective group. The stabilized IPW ranged between 0.6 and 3.3 for the overall group (0.6-3.3 for women with oophorectomy, 0.6-3.0 for referent women) and between 0.5 and 4.3 for the women age ≤45 years at baseline (0.5-2.2 for women with oophorectomy 0.6-4.3 for referent women).

Abbreviations: BMI, body mass index; COPD, chronic obstructive pulmonary disease; eGFR, estimated glomerular filtration rate; IPW, inverse probability weights.

[a] Characteristics are presented in descending order of magnitude for each panel.