SUPPLEMENTAL MATERIAL

Data S1.

SUPPLEMENTAL METHODS

Inverse probability of weighting

We used inverse probability of weighting to balance the distribution of covariates between two patient groups. If *e* denotes the estimated propensity score (i.e. $e=\hat{P}$) $(Z=1 | x)$, where the patient x is included in patient group 1; then, $1-e = \hat{P}(Z=0 | x)$, then the original sample is weighted by the following weights: Z/e+(1−Z)/ 1−e where Z represents the patient group. For instance, women $(Z=1)$ are assigned a weight equal to the reciprocal of the propensity score (1/e), while men $(Z=0)$ are assigned a weight equal to the reciprocal of one minus the propensity score (1/1-e). The weighting procedure for each sample balances the covariate distributions between two patient groups. 18

Nearest neighbor imputation algorithms

Nearest neighbor (NN) imputation algorithms are efficient methods to fill in missing data where each missing value on some records is replaced by a value obtained from related cases in the whole set of records. Thus, imputation for clinical features was conducted using the average of measured values from k records (kNN) .¹⁹

NN algorithms are similarity-based methods that rely on distance metrics and results may change in relation to the similarity measure used to evaluate the distance between recipients and donors. In our work, we used the following norm as metric to evaluate distance:

$(\Sigma$ ni=1|xi−yi|p)1/p

Before imputation of the recipient Xi, the full set with no missing data C(X) was filtered to select a subset of features relevant to the missing variable to be imputed (Xi miss). To this end, $C(X)$ was considered as a dataset in the context of a regression problem, where the variable with the missing

data (Xmiss) was set as the class variable and the other q variables $(X1, X2, ..., Xq)$ as predictors. We also applied the RReliefF algorithm. The set was, therefore, filtered to select a subset $Cs(X) \subset C(X)$ where $(X1, X2, ..., Xs) \subset (X1, X2, ..., Xq)$ and $s < q$. In the present context, we set the number of neighbors for RReliefF equal to 10 and set s as 10 %, 20 % or 30 % of q. As C(X) is invariant to Xi, the filtering step was performed only once before the NN imputation step that, on the contrary was performed separately for each Xi.

More specifically, to impute the missing value in i-th column, we find k-nearest neighbor columns from i-th column (in terms of Euclidean distance) and replace the missing value with weighted mean of the k-nearest neighbor columns. Weights are inversely proportional to the Euclidean distance from i-th column.

Interaction test

The comparison of two estimated quantities, each with its standard error, is a general method that can be applied widely. ²⁰ These measures were always analyzed on the log scale because the distributions of the log ratios tend to be those closer to normal than of the ratios themselves. If the estimates are *E*1 and *E*2 with standard errors SE(*E*1) and SE(*E*2), then the difference *d*=*E*1 - *E*2 has standard error $SE(d) = \ddot{O}[SE(E1)2 + SE(E2)2]$ i.e., the square root of the sum of the squares of the separate standard errors. The ratio $z = d/SE(d)$ gives a test of the null hypothesis that in the population the difference *d* is zero, by comparing the value of *z* to the standard normal distribution. The 95% confidence interval for the difference is *d*-1.96SE(*d*) to *d*+1.96SE(*d*).

SUPPLEMENTAL RESULTS

Interaction tests

In our study, the estimated women-to-men RR ratio for obstructive CAD among nondiabetics was 0.43 (95%CI 0.36– 0.51) and diabetics was 0.89 (0.43–1.83), but are the relative risks from the

subgroups significantly different from each other? We show how to answer this question by using the interaction test based on the summary data quoted. **(Table S4)**. We obtained the logs of the odds ratios (relative risks) and their confidence intervals (rows 2 and 4). As 95% confidence intervals were obtained as 1.96 standard errors either side of the estimate, the SE of each log relative risk was obtained by dividing the width of its confidence interval by 2×1.96 (row 6). The estimated difference in log relative risks was d=E1- E2= 0.5696 (row 7) and its standard error 0.1958 (row 8). From these two values, we tested the interaction and estimated the ratio of the relative risks (with confidence interval). The test of interaction was the ratio of d to its standard error: $z = 2.9091$, which gives p value=0.0018 when we referred it to a table of the normal distribution (row 10). The estimated interaction effect was \exp =1.7676 (row 11). The confidence interval for this effect was 1.2042 to 2.5945 on the log scale (row 9). Transforming back to the relative risk scale, we got 1.2042 to 2.5945 (row 12). There was thus good evidence to support different outcome effects of diabetes on obstructive CAD between sexes. A similar approach was used for comparing any other sex difference. (**Tables S5, S6, and S9**).

Table S1. Baseline characteristics of the overall population sorted by sex and CAD status in patients with acute coronary syndrome at index event.

Table S2. Use of medications and PCI within 24 hours from hospitalization sorted by sex (women versus men) and CAD status in the overall population of patients with acute coronary syndromes.

ACE indicates angiotensin-converting enzyme; ARBs, angiotensin II receptor blockers; CAD, coronary artery disease; GP, glycoprotein; LMWH, low molecular weight heparin; PCI, percutaneous coronary intervention.

Table S3. Use of medications and reperfusion therapies within 24 hours from hospitalization sorted by sex (women versus men) and CAD status in patients with STEMI.

ACE indicates angiotensin-converting enzyme; ARBs, angiotensin II receptor blockers; CAD, coronary artery disease; GP, glycoprotein; LMWH, low molecular weight heparin; PCI, percutaneous coronary intervention; STEMI, ST-segment elevation myocardial infarction

Table S5. Interaction test calculations for comparing two estimated risk ratios (relative risks of women versus men) by inverse probability of weighting: diabetes, current smoking, hypercholesterolemia, hypertension for obstructive CAD.

Table S6. Interaction test: calculations for comparing two estimated RR ratios (women versus men) by inverse probability of weighting: STEMI in obstructive versus nonobstructive CAD in patients with acute coronary syndrome at index event.

Table S7. Interaction test: calculations for comparing two estimated RR ratios (women versus men) by inverse probability of weighting: 30-day mortality in obstructive versus nonobstructive CAD in patients with acute coronary syndrome at index event.

Table S8. Inverse probability of weighting: outcomes sorted by sex (women versus men) in patients with obstructive CAD who underwent primary PCI.

Table S9. Inverse probability of weighting: outcomes sorted by sex (women versus men) and CAD status in patients with acute coronary syndrome at index event. Analysis restricted the cohort of obstructive CAD patients having 70% or greater stenosis

BP indicates blood pressure; CAD, coronary artery disease.

Obstructive CAD was defined as a 70% or more narrowing of the luminal diameter.

Table S10. Interaction test: calculations for comparing two estimated RR ratios (women versus men) by inverse probability of weighting: 30-day mortality in obstructive (stenosis ≥70%) versus nonobstructive CAD in patients with acute coronary syndrome at index event.

Figure S1. **Study Flow Chart.**

CABG indicates coronary artery bypass graft; CAD, coronary artery disease; NSTE-ACS, non-ST elevation acute coronary syndromes; PCI, percutaneous coronary intervention; STEMI, STelevation myocardial infarction.