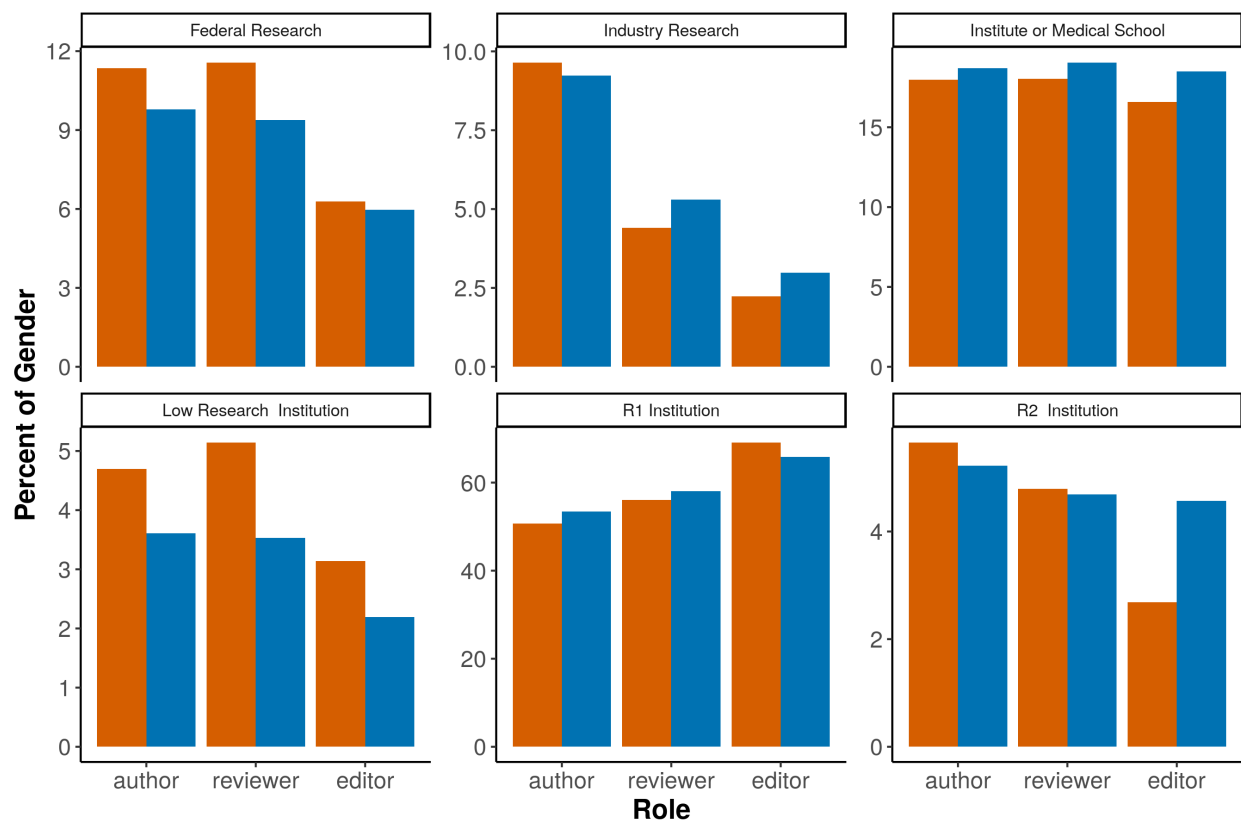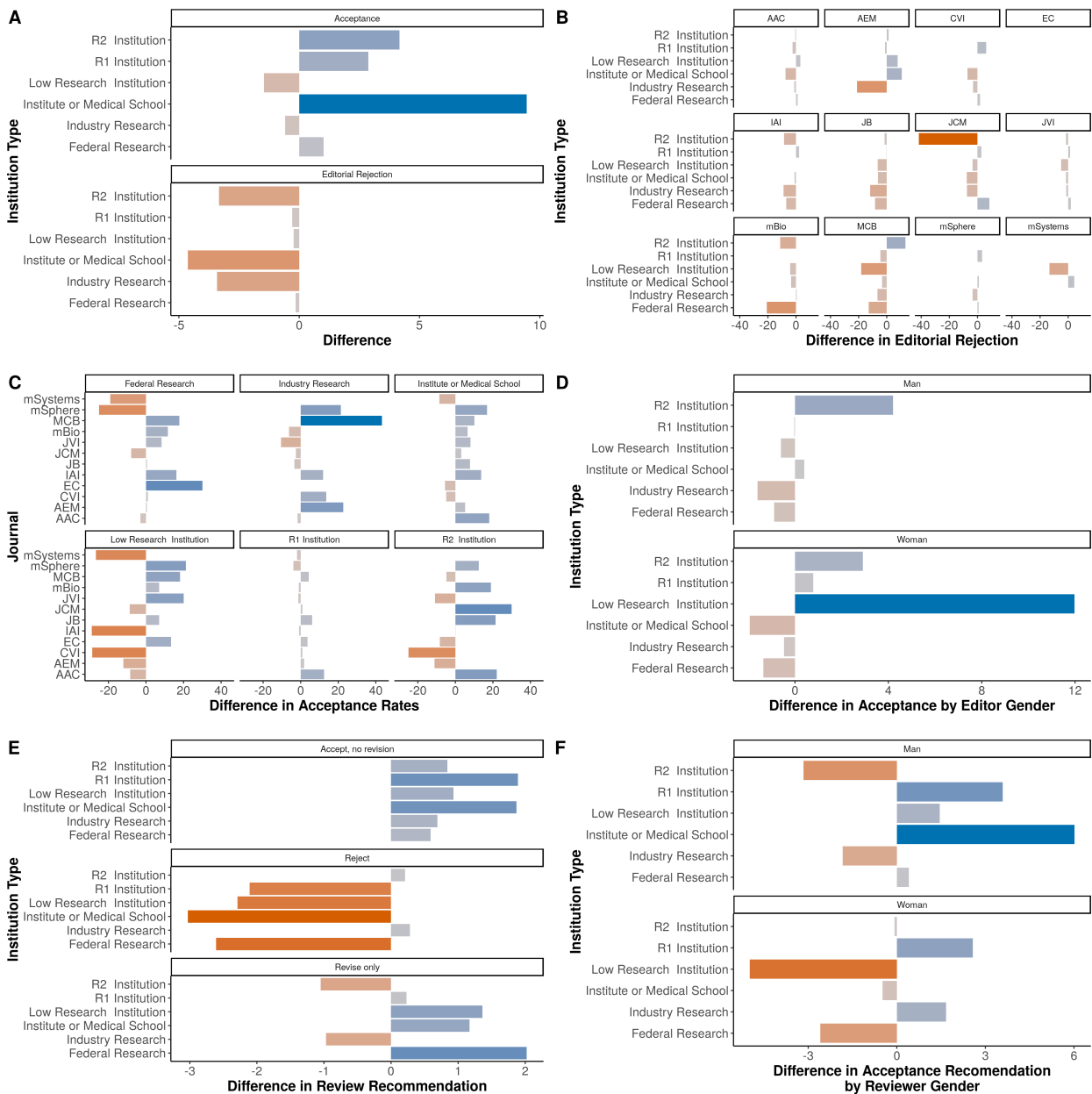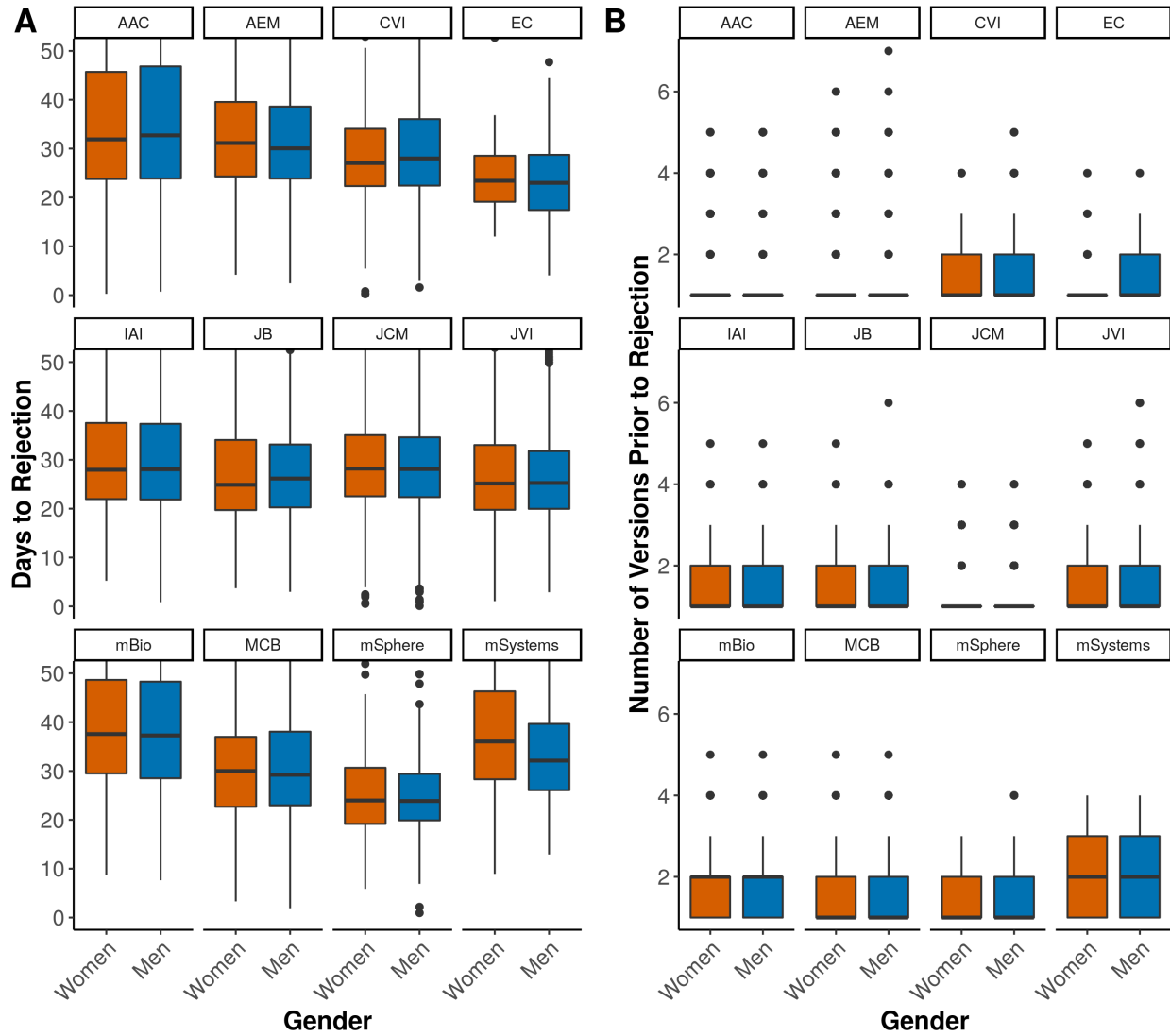# Supplementary

Gender representation

Supplemental Figure . Break down of individuals in each role by their gender and US institute type. Authors are restricted to senior authors (e.g., corresponding, last).
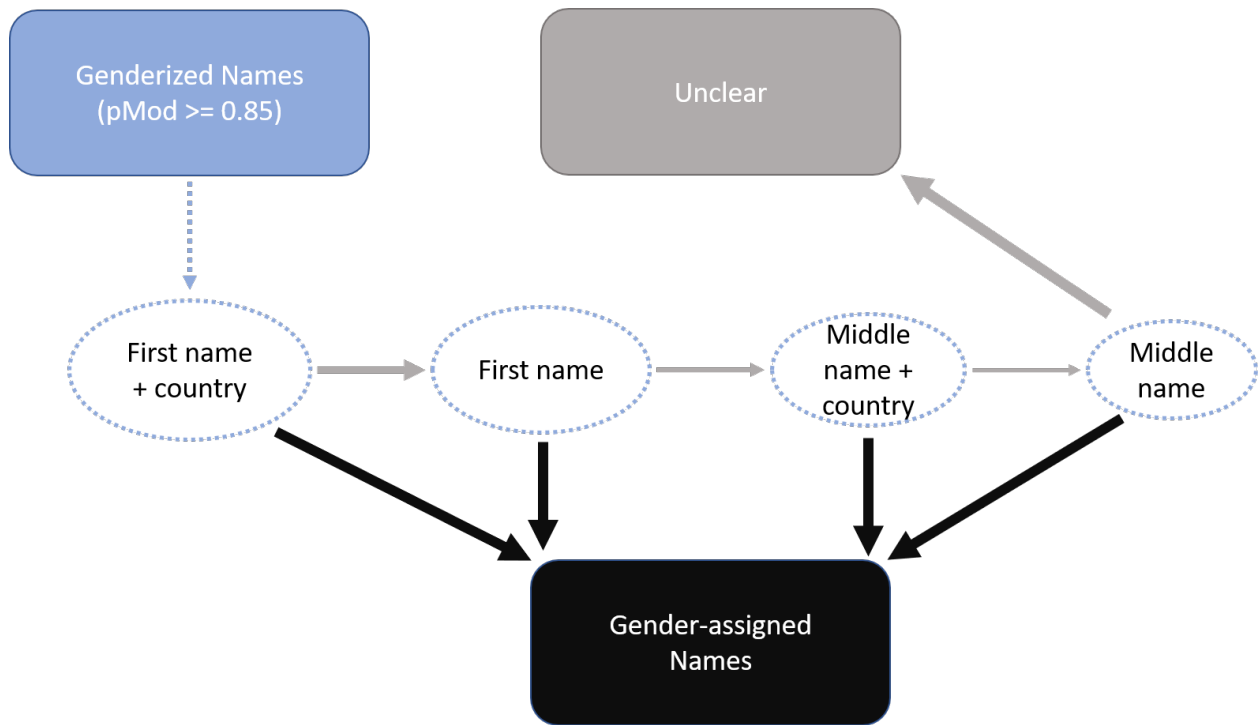
Gender bias

Supplemental Figure . Difference in acceptance and rejections by institution type.

Supplemental Figure . Total days and versions between submission and final rejection at each journal.

Gender prediction and assignment

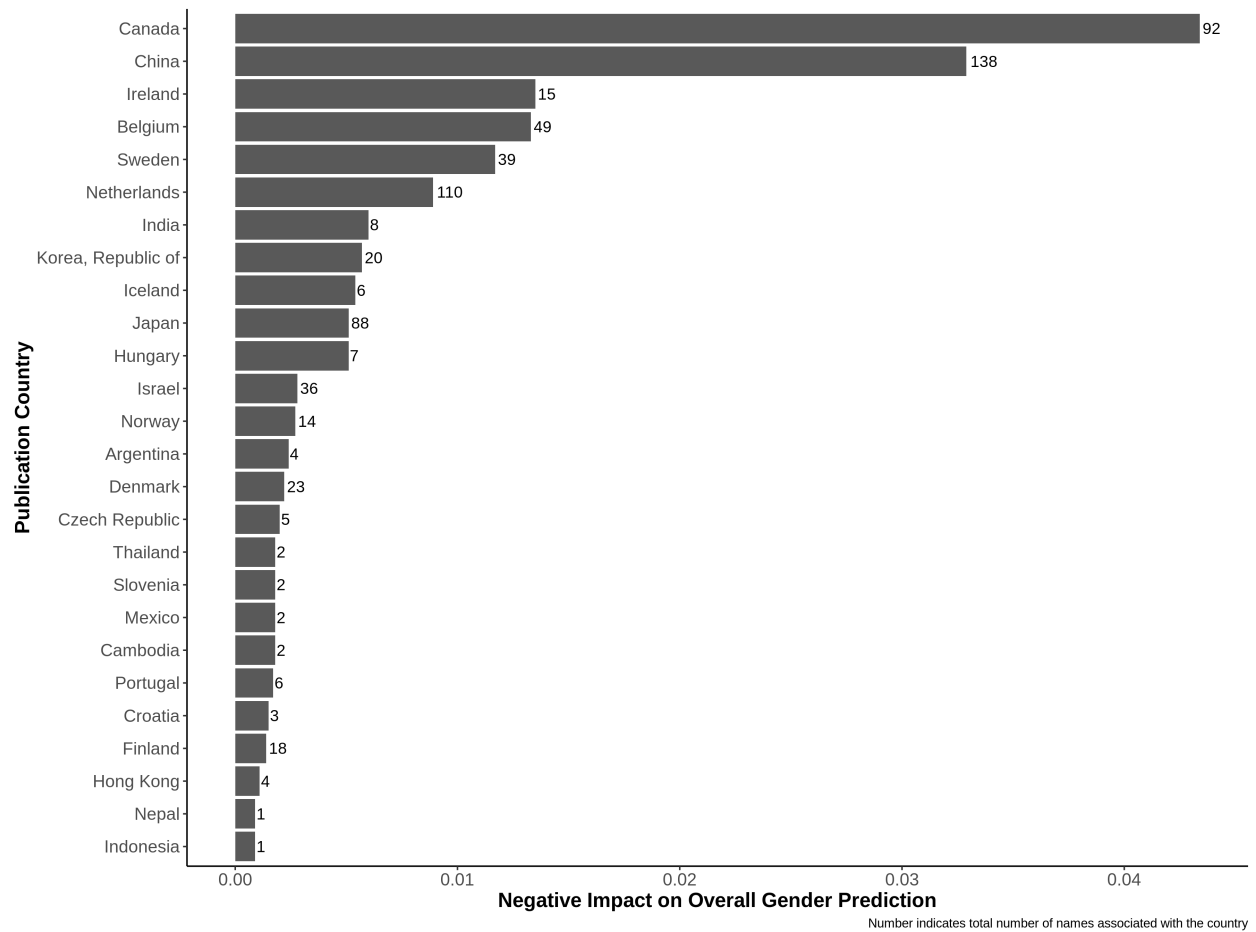Supplemental Figure . Schematic of gender prediction and assignment.

Validating gender analysis

Supplemental Table 1. sensitivity/specificity/accuracy of genderize thresholds. Bolded text denotes the accuracy of the threshold used in all further analyses.
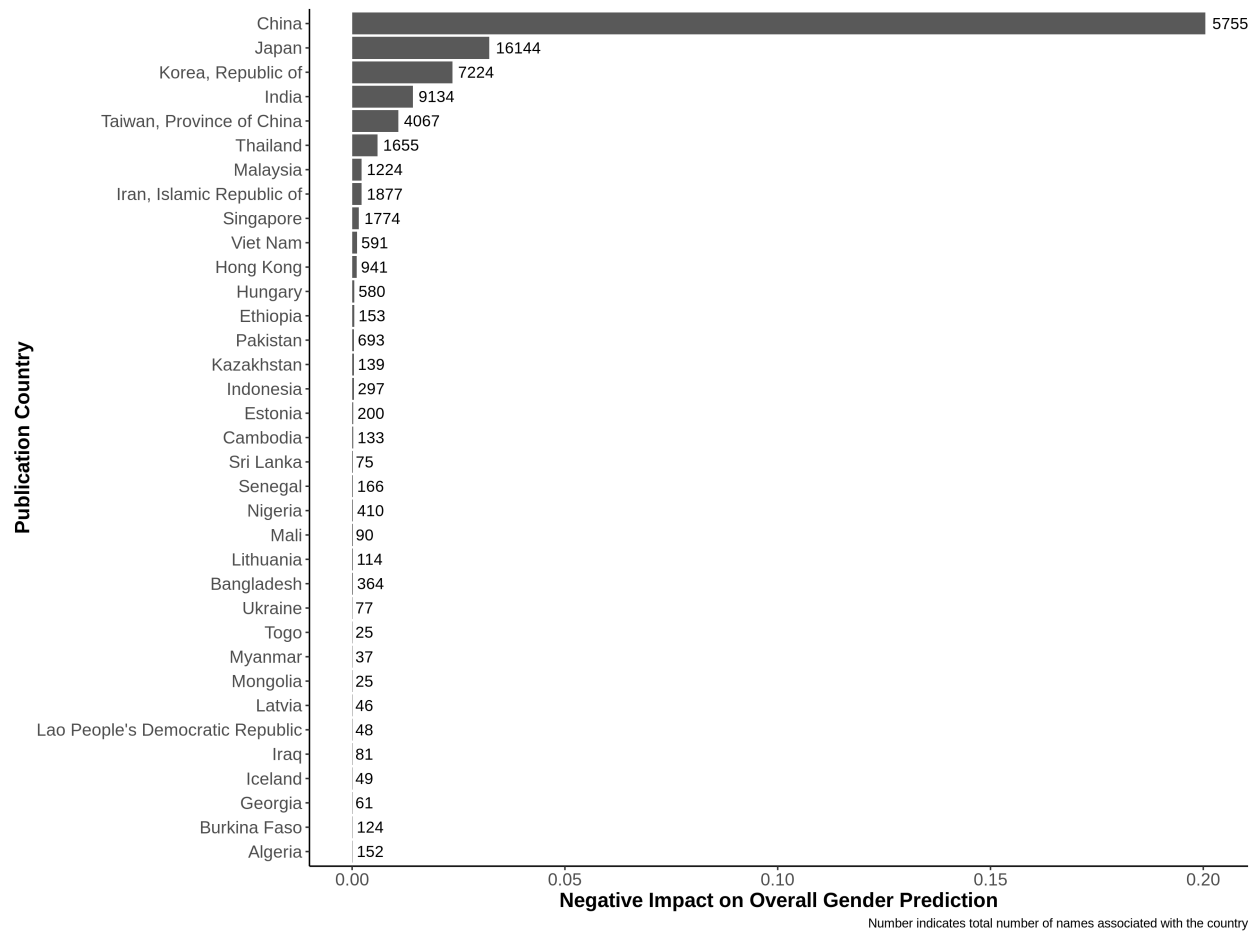
Measure	First Names			Plus Country Data		
	p0.5	p0.85	pmod0.85	p0.5	p0.85	pmod0.85
Sensitivity	0.8943	0.9516	0.971	0.9055	0.9471	0.9669
Specificity	0.9339	0.9593	0.972	0.9265	0.9553	0.9727
Accuracy	0.9110	0.9549	**0.9714**	0.9146	0.9507	**0.9695**

$$Impact_C = \left| \frac{(\% \, Unpredicted_C - \% \, Unpredicted_{Total}) \; x \; \left(\frac{Observations_C}{Observations_{Total}} \right)}{\% \, Unpredicted_{Total}} \right|$$

Supplemental Figure . Equation for calculating negative bias by genderize. C indicates an individual country.

Supplemental Figure . The negative impact of each country on the overall gender prediction of the validation dataset. Number indicates the total number of names associated with each country.

Supplemental Figure . The negative impact of each country on the overall gender prediction of the full dataset. Number indicates the total number of names associated with each country.