NISTIR 8280

Ongoing Face Recognition Vendor Test (FRVT)
Part 3: Demographic Effects

Annex 6 : Cross-race and sex false match rates in United States mugshot images

2019/12/19

NIST
National Institute of
Standards and Technology
U.S. Department of Commerce

1 Overview

This annex includes figures that show cross-race false match rates. Each panel on the pages is a heatmap corresponding to one algorithm. It shows a matrix of values. The value in a cell is the FMR obtained when images of persons identified by the row label are compared with images of persons identified by the column label. The labels come straight from the textual metadata that accompanied the images at the time they were supplied to NIST.

2 Data

The images are all high-quality frontal portraits collected in immigration offices. All images have a white background and are in close approximation to ISO/IEC 39794-5 / ICAO specifications. As such, potential quality related drivers of high false match rates (such as blur) can be expected to be absent.

The total number of images is 8 201 768. The total number of persons is 2 824 627. The total number of impostor comparisons is just over 108 million (108,000 690). Note that we could have executed many more impostor comparisons, via full cross-comparison. We elected not to do so first to save resources and secondly to include only limited re-use of images. Thus each image is involved in no more than 300 impostor comparisons. The numbers within the demographic groups are tabulated below. The set yielded genuine comparisons also - their use in computing FNMR is covered in Annexes Annex 12 and Annex 13 .

Race Label	Sex Label	Mated Comparison Count	Impostor Comparison Count	
1	A	F	10 995	3 000 000
2	A	M	139 342	3 000 001
3	B	F	263 910	3 000 007
4	B	M	1 954 864	3 000 009
5	I	F	26 699	3 000 000
6	I	M	268 364	3 000 006
7	W	F	362 816	3 000 012
8	W	M	1 033 237	3 000 017
9	Total		4061227	108000690

3 Fixed Threshold

A false match is declared if the comparison score is equal to, or exceeds, a threshold. For any given algorithm, this same value applies to all comparisons in all cells. The threshold value could be any value germane to that comparison algorithm. The threshold value was taken as the smallest value that gives FMR ≤ 0.001 on white male mugshots.

4 Plot

The plots encode false match rate on a logarithmic scale. A FMR of 0.01 corresponds to a value of -2, for example. Large negative values indicate low (good) false match rates, colored blue. Small negative values indicate high (poor) false match rates, colored red.

False positive: Incorrect association of two subjects
False negative: Failed association of one subject

1:1 FMR   1:N FPIR
1:1 FNMR   1:N FNIR

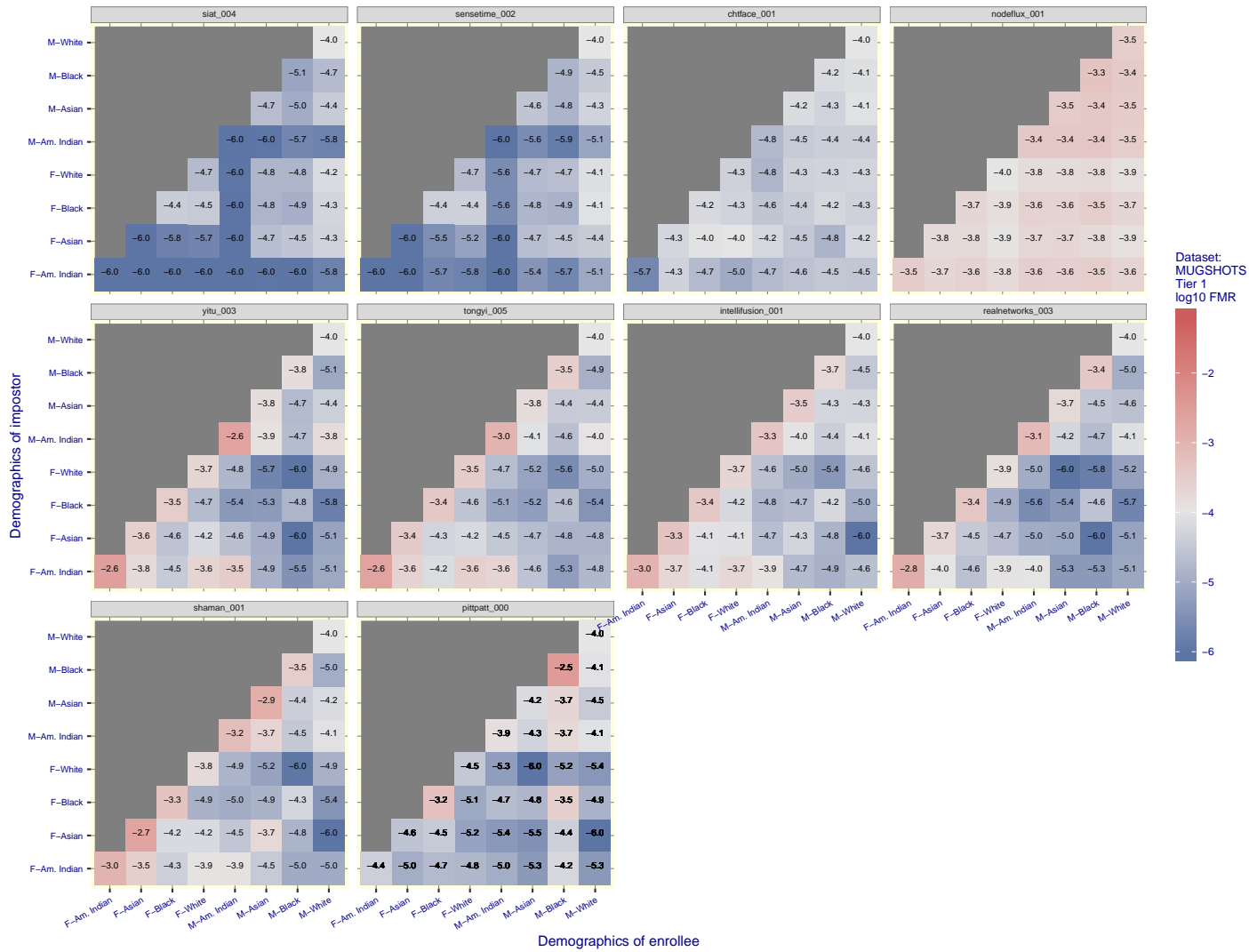T ≫ 0   → FMR, FPIR → 0
                → FNMR, FNIR → 1

Figure 1: For domestic mugshot photos for which one of four race labels and a sex label is available, the figure shows false match rates when the verification algorithm shown in the panel's header strip is used to compare randomly selected photos from the groups identified in the respective rows and columns. The threshold is set to a fixed value everywhere. The value is the smallest threshold that gives FMR ≤ 0.0001 on white male impostors. Each cell depicts FMR on a logarithmic scale. The text value is $\log_{10}(FMR)$ with large negative values encoding superior false match rates. The panels appear in order of the highest FMR on African American females, lowest first.
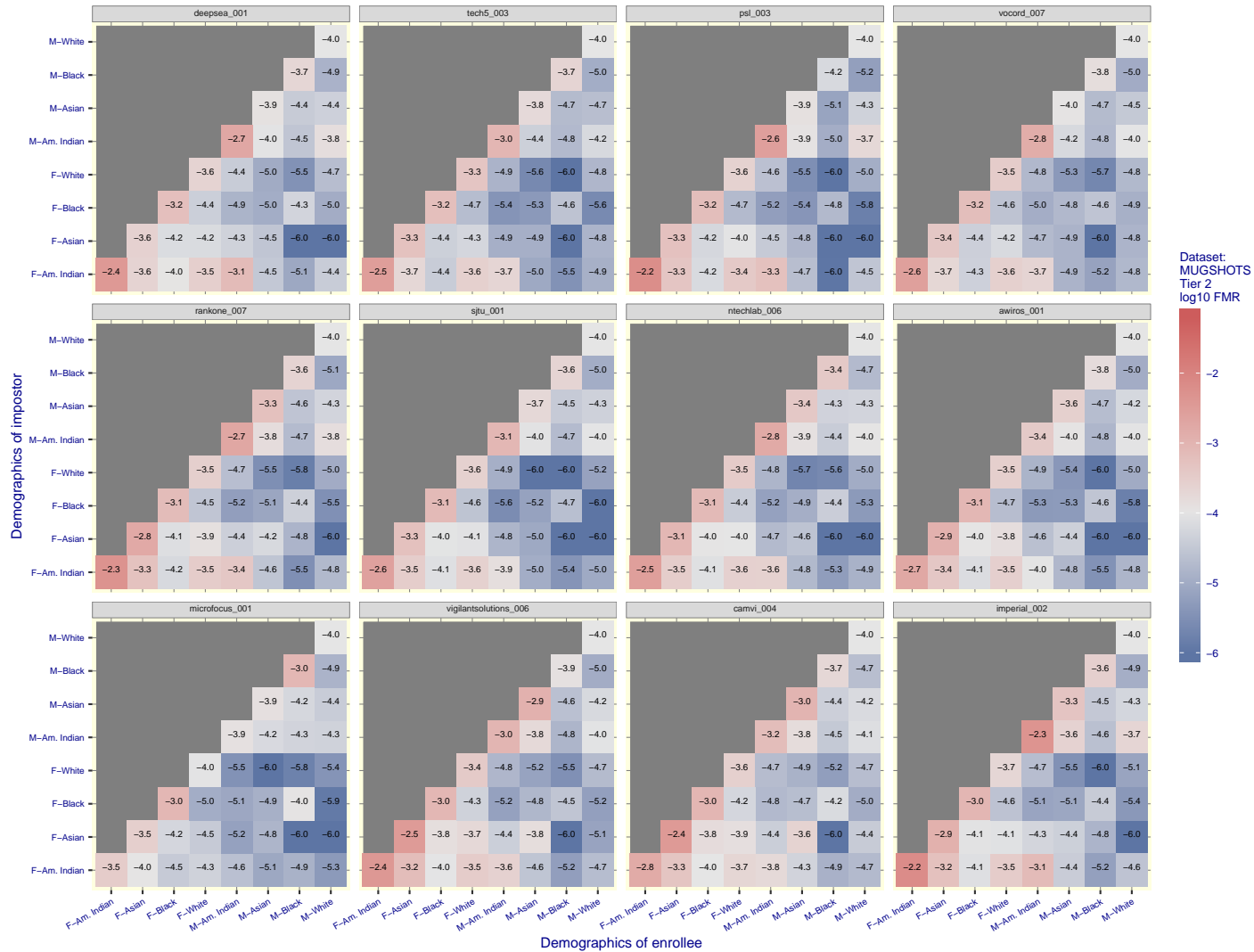
Figure 2: For domestic mugshot photos for which one of four race labels and a sex label is available, the figure shows false match rates when the verification algorithm shown in the panel's header strip is used to compare randomly selected photos from the groups identified in the respective rows and columns. The threshold is set to a fixed value everywhere. The value is the smallest threshold that gives FMR ≤ 0.0001 on white male impostors. Each cell depicts FMR on a logarithmic scale. The text value is $\log_{10}(FMR)$ with large negative values encoding superior false match rates. The panels appear in order of the highest FMR on African American females, lowest first.
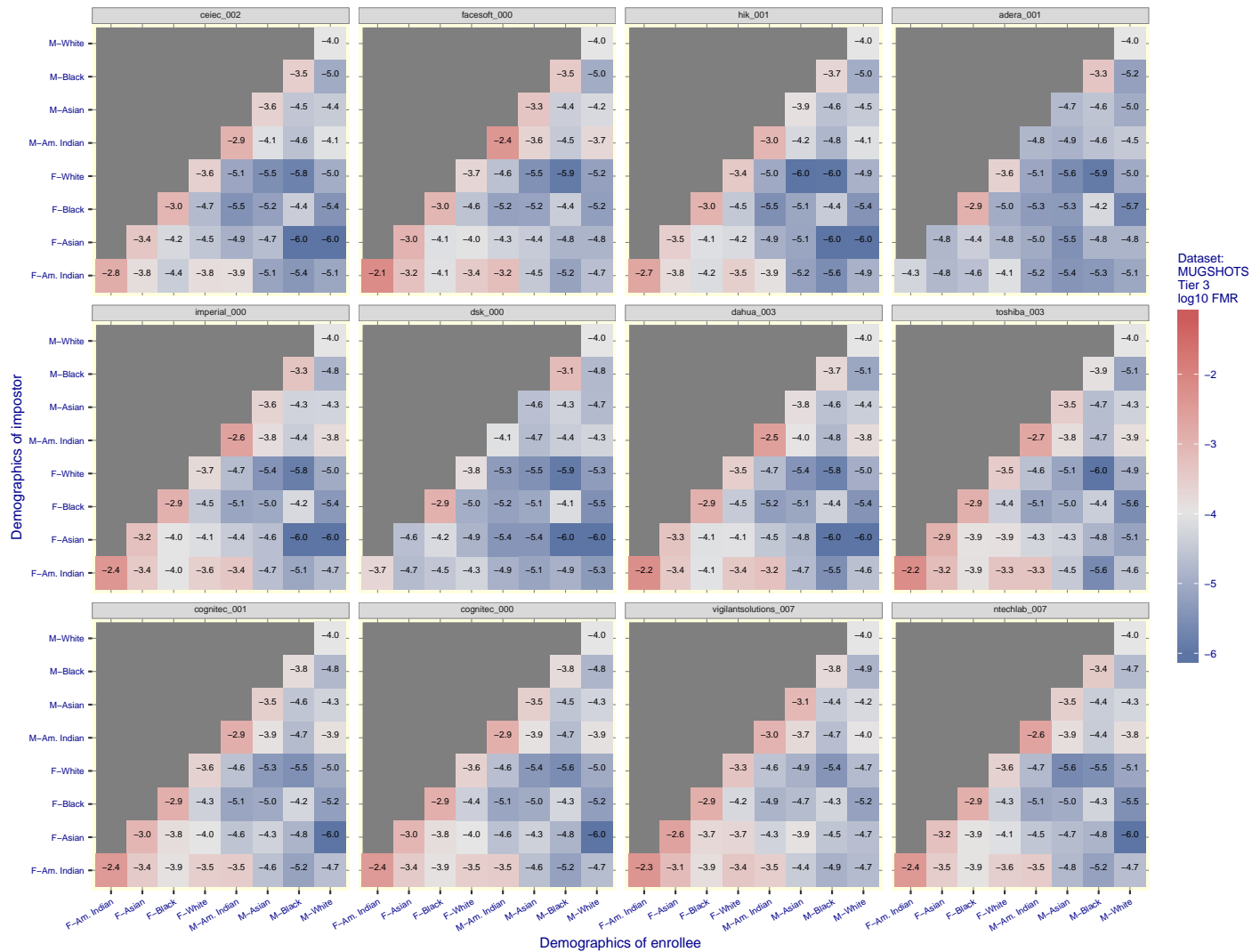
Figure 3: For domestic mugshot photos for which one of four race labels and a sex label is available, the figure shows false match rates when the verification algorithm shown in the panel's header strip is used to compare randomly selected photos from the groups identified in the respective rows and columns. The threshold is set to a fixed value everywhere. The value is the smallest threshold that gives FMR ≤ 0.0001 on white male impostors. Each cell depicts FMR on a logarithmic scale. The text value is $\log_{10}(FMR)$ with large negative values encoding superior false match rates. The panels appear in order of the highest FMR on African American females, lowest first.
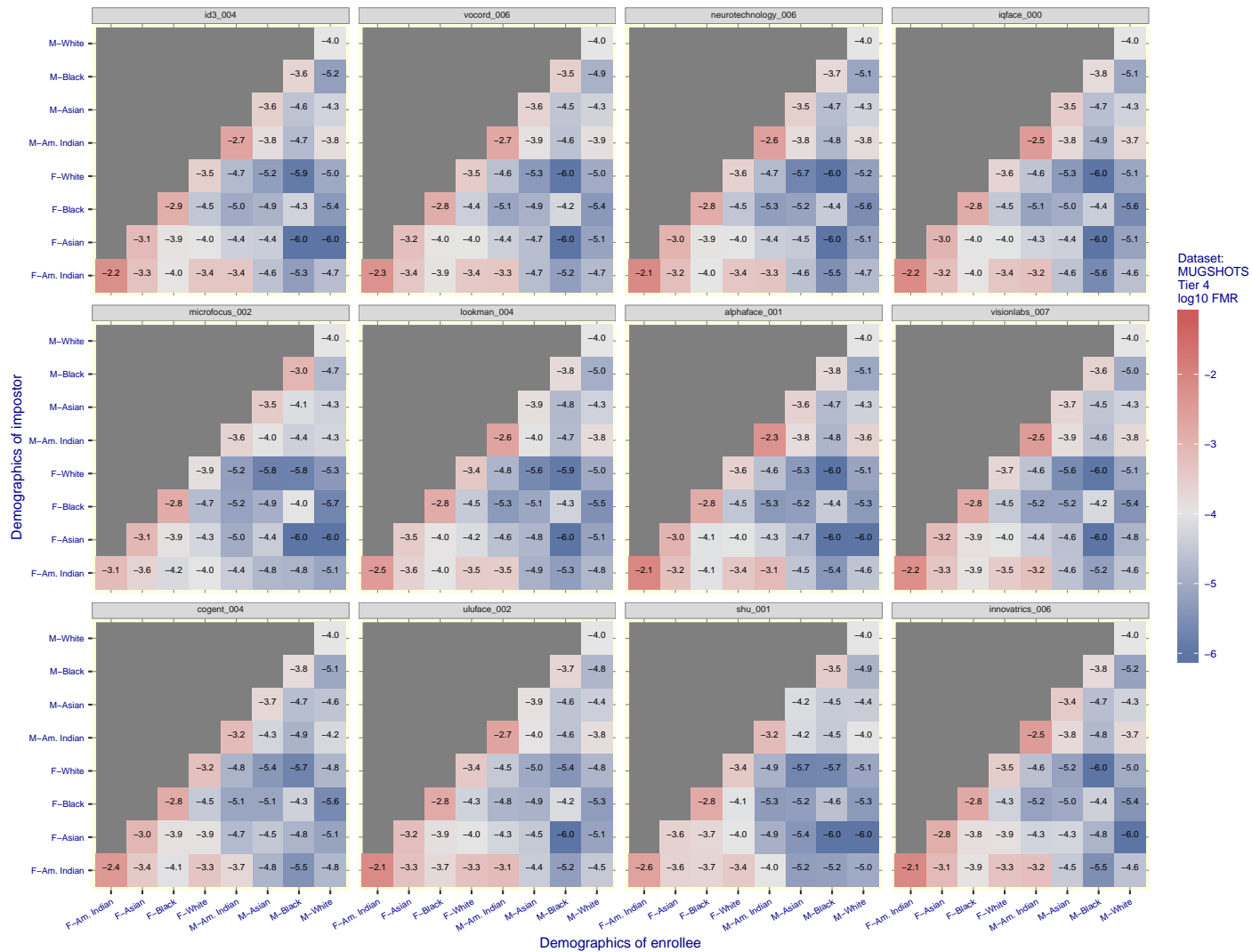
Figure 4: For domestic mugshot photos for which one of four race labels and a sex label is available, the figure shows false match rates when the verification algorithm shown in the panel's header strip is used to compare randomly selected photos from the groups identified in the respective rows and columns. The threshold is set to a fixed value everywhere. The value is the smallest threshold that gives FMR ≤ 0.0001 on white male impostors. Each cell depicts FMR on a logarithmic scale. The text value is $\log_{10}(FMR)$ with large negative values encoding superior false match rates. The panels appear in order of the highest FMR on African American females, lowest first.
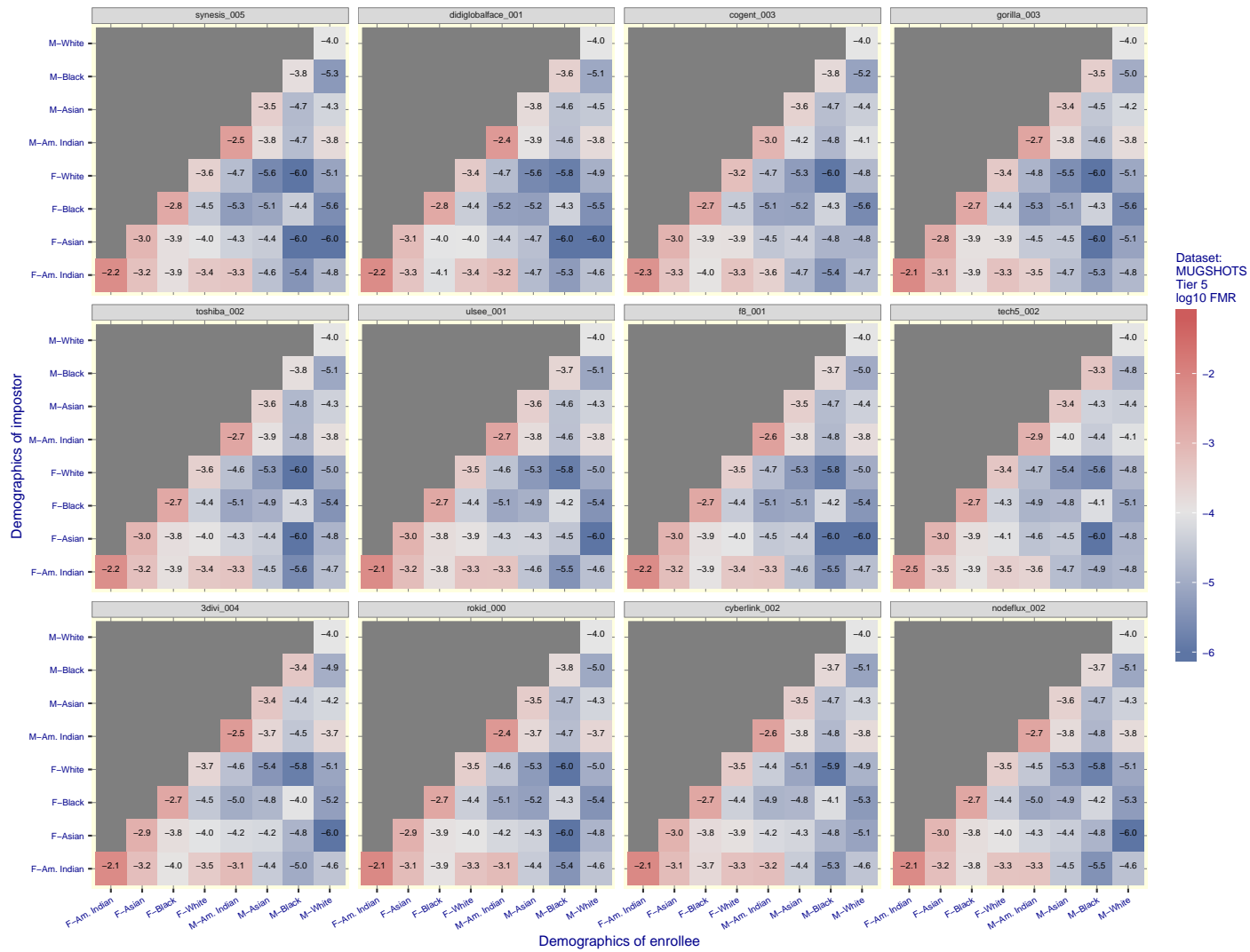
Figure 5: *For domestic mugshot photos for which one of four race labels and a sex label is available, the figure shows false match rates when the verification algorithm shown in the panel's header strip is used to compare randomly selected photos from the groups identified in the respective rows and columns. The threshold is set to a fixed value everywhere. The value is the smallest threshold that gives FMR ≤ 0.0001 on white male impostors. Each cell depicts FMR on a logarithmic scale. The text value is $\log_{10}(FMR)$ with large negative values encoding superior false match rates. The panels appear in order of the highest FMR on African American females, lowest first.*
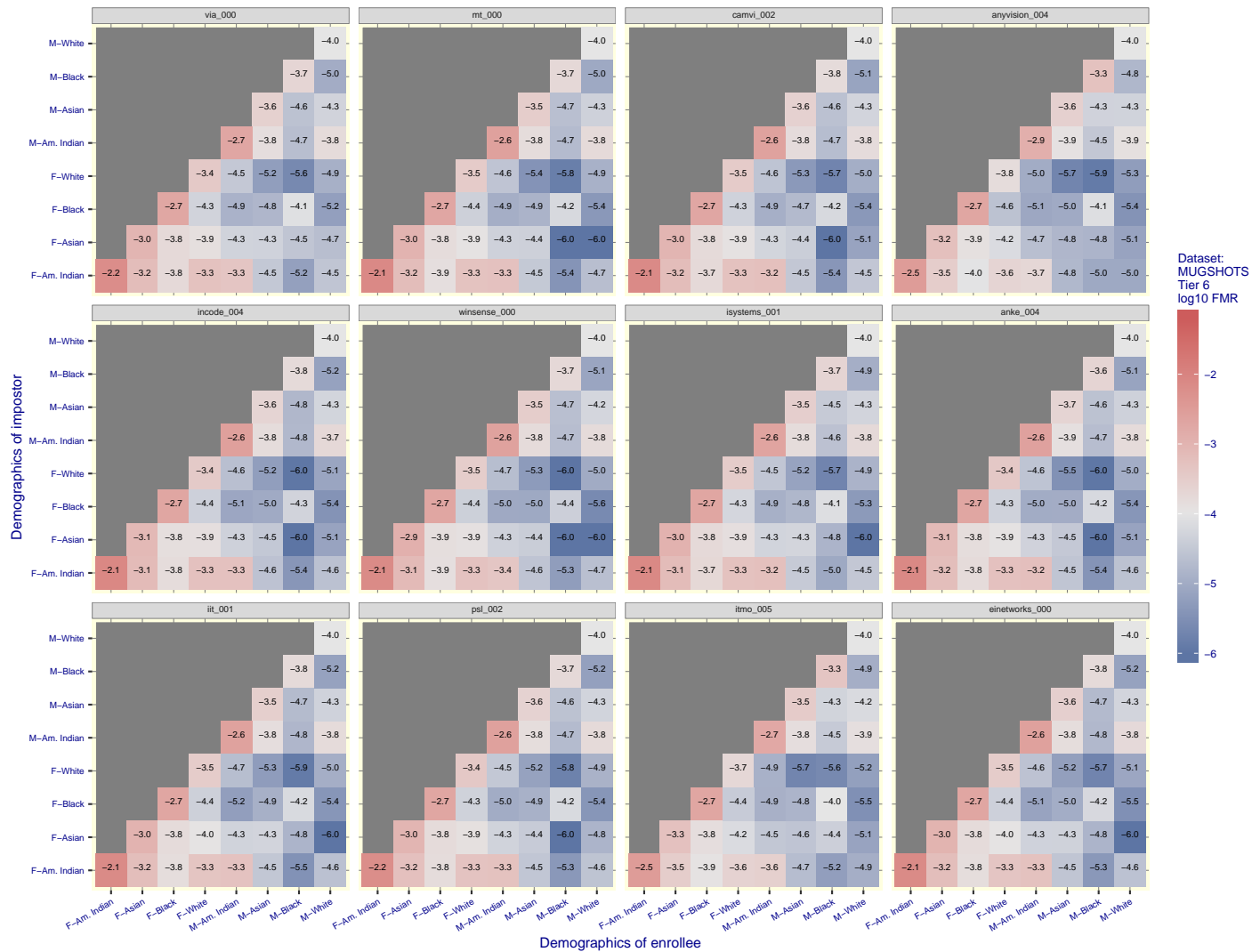
Figure 6: *For domestic mugshot photos for which one of four race labels and a sex label is available, the figure shows false match rates when the verification algorithm shown in the panel's header strip is used to compare randomly selected photos from the groups identified in the respective rows and columns. The threshold is set to a fixed value everywhere. The value is the smallest threshold that gives FMR ≤ 0.0001 on white male impostors. Each cell depicts FMR on a logarithmic scale. The text value is $\log_{10}(FMR)$ with large negative values encoding superior false match rates. The panels appear in order of the highest FMR on African American females, lowest first.*
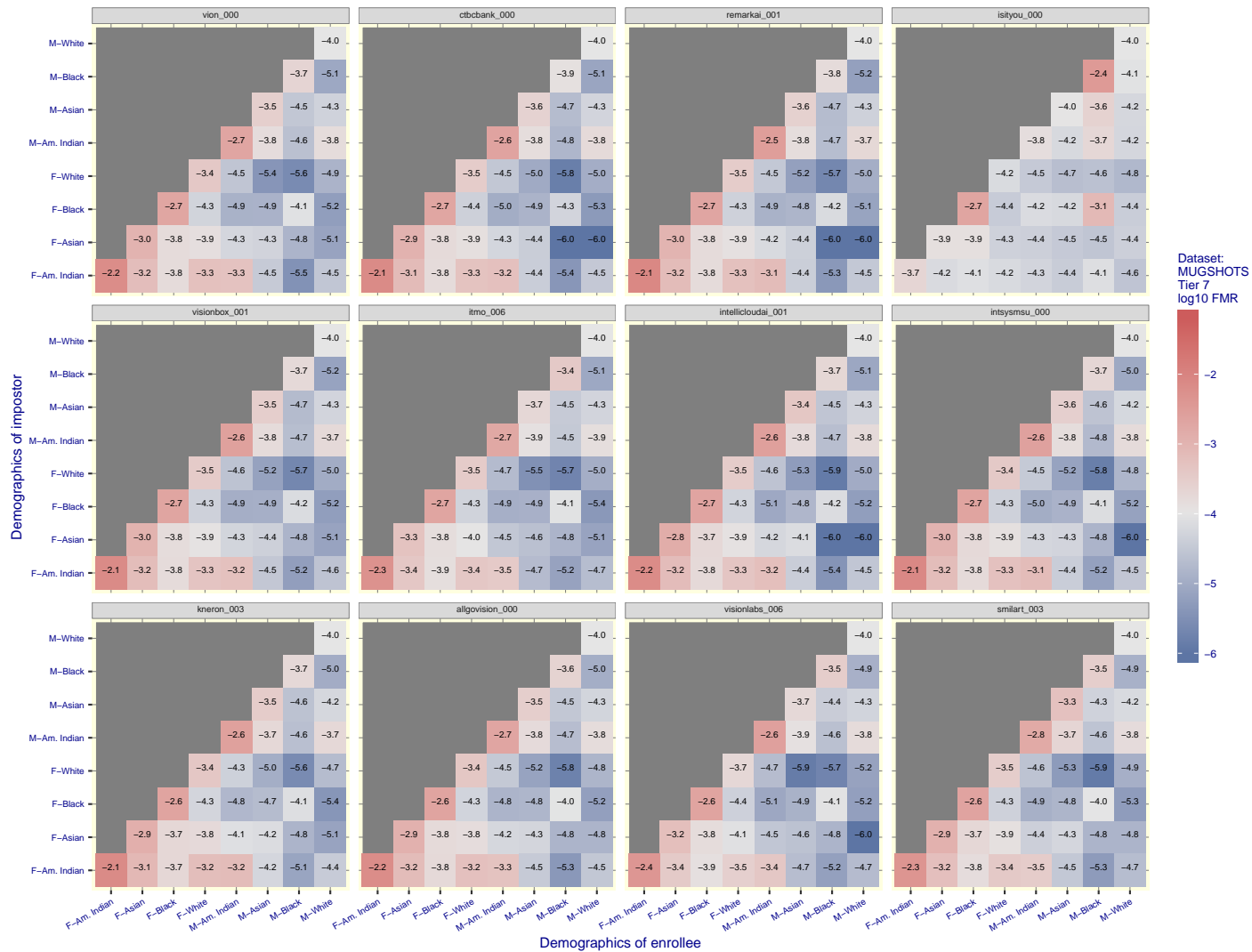
Figure 7: For domestic mugshot photos for which one of four race labels and a sex label is available, the figure shows false match rates when the verification algorithm shown in the panel's header strip is used to compare randomly selected photos from the groups identified in the respective rows and columns. The threshold is set to a fixed value everywhere. The value is the smallest threshold that gives FMR ≤ 0.0001 on white male impostors. Each cell depicts FMR on a logarithmic scale. The text value is $\log_{10}(FMR)$ with large negative values encoding superior false match rates. The panels appear in order of the highest FMR on African American females, lowest first.
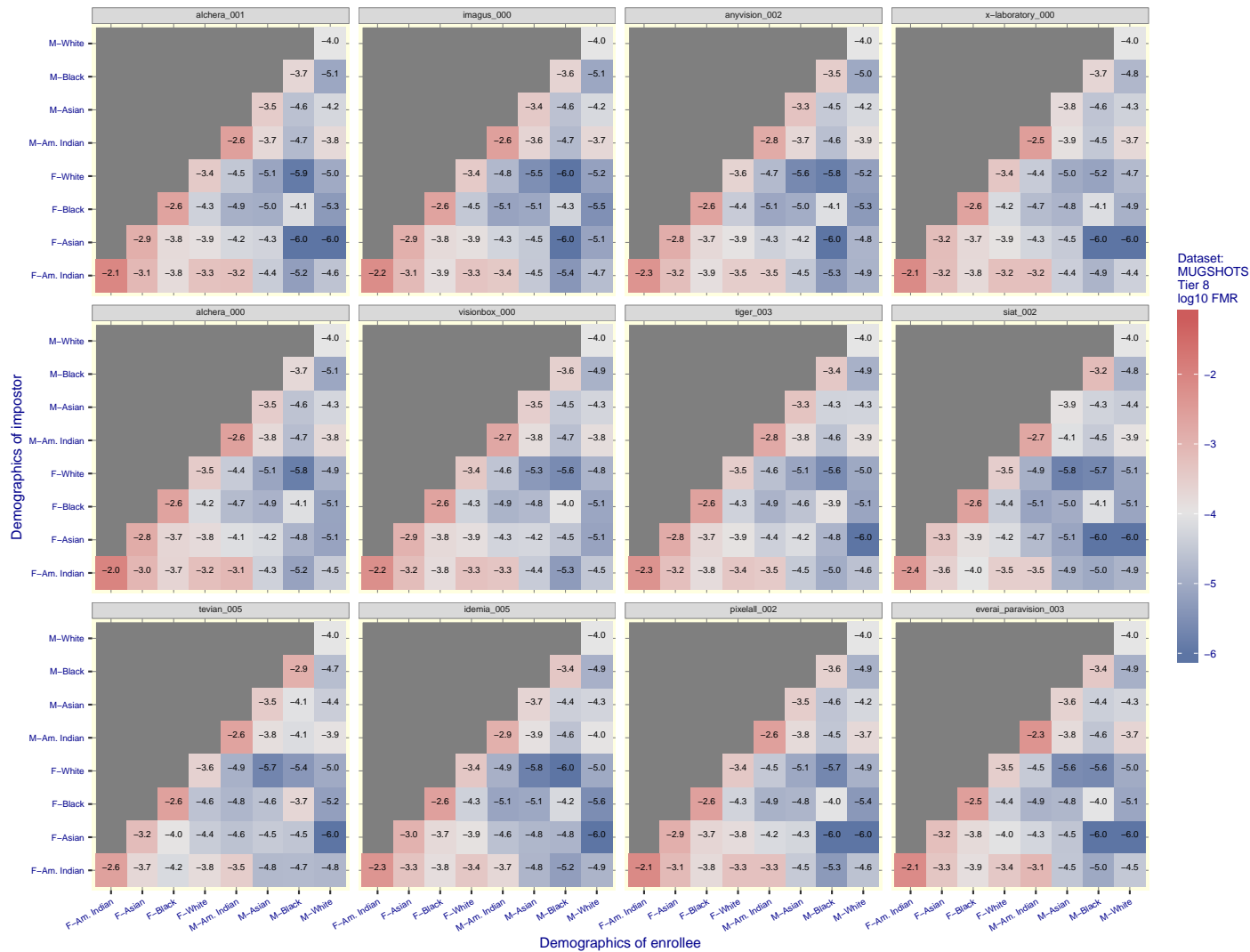
Figure 8: For domestic mugshot photos for which one of four race labels and a sex label is available, the figure shows false match rates when the verification algorithm shown in the panel's header strip is used to compare randomly selected photos from the groups identified in the respective rows and columns. The threshold is set to a fixed value everywhere. The value is the smallest threshold that gives FMR ≤ 0.0001 on white male impostors. Each cell depicts FMR on a logarithmic scale. The text value is $\log_{10}(FMR)$ with large negative values encoding superior false match rates. The panels appear in order of the highest FMR on African American females, lowest first.
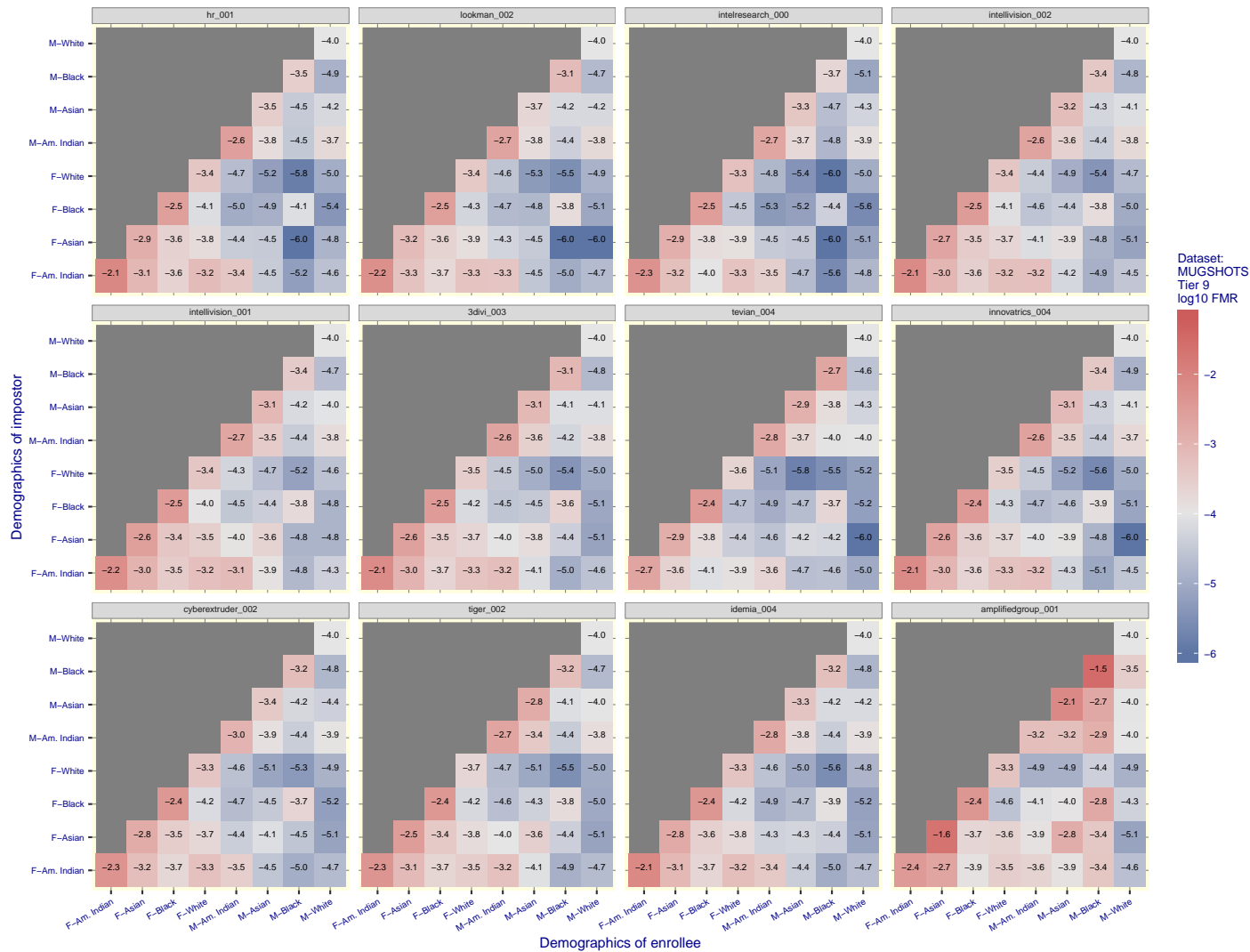
Figure 9: For domestic mugshot photos for which one of four race labels and a sex label is available, the figure shows false match rates when the verification algorithm shown in the panel's header strip is used to compare randomly selected photos from the groups identified in the respective rows and columns. The threshold is set to a fixed value everywhere. The value is the smallest threshold that gives FMR ≤ 0.0001 on white male impostors. Each cell depicts FMR on a logarithmic scale. The text value is $\log_{10}(FMR)$ with large negative values encoding superior false match rates. The panels appear in order of the highest FMR on African American females, lowest first.
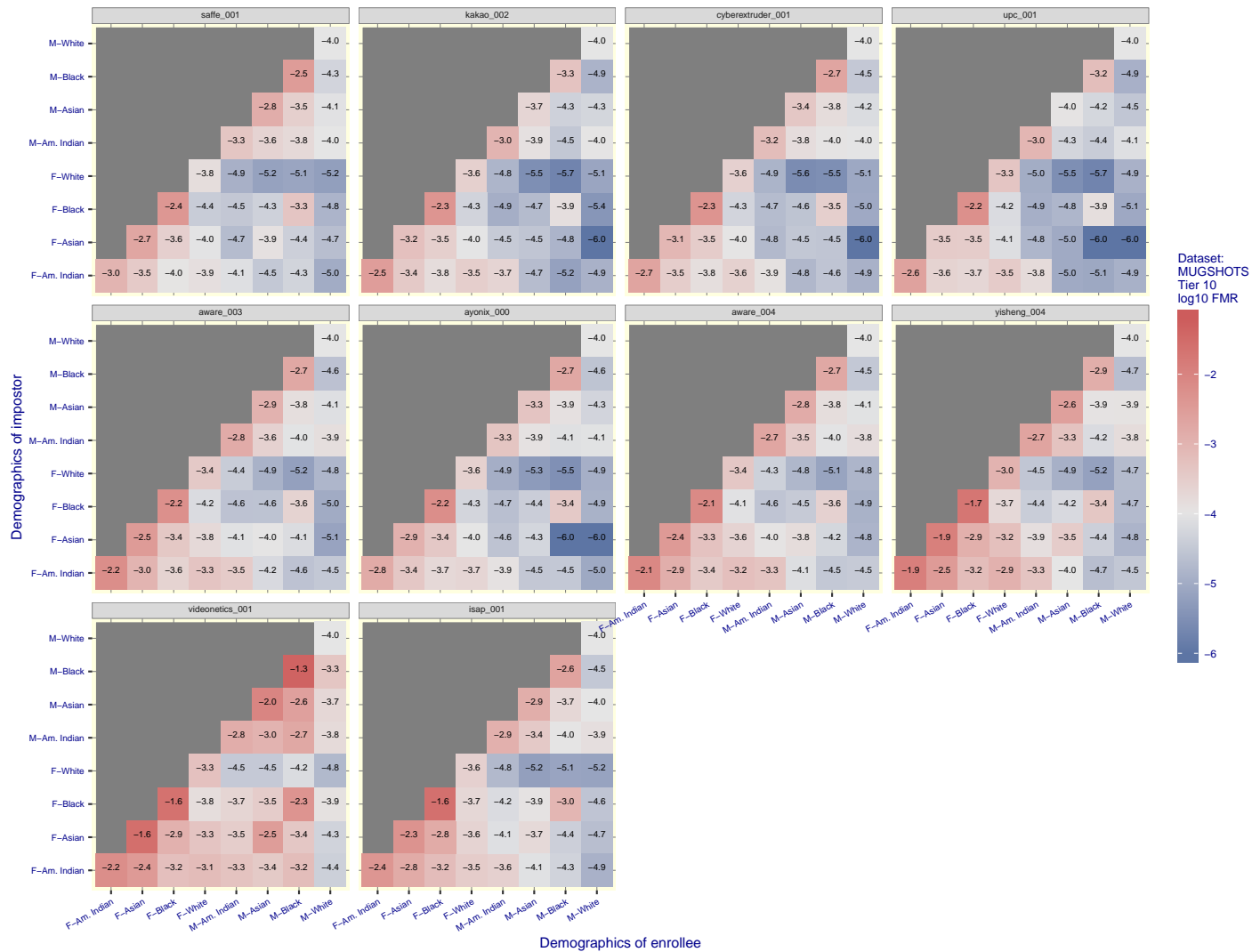
Figure 10: For domestic mugshot photos for which one of four race labels and a sex label is available, the figure shows false match rates when the verification algorithm shown in the panel's header strip is used to compare randomly selected photos from the groups identified in the respective rows and columns. The threshold is set to a fixed value everywhere. The value is the smallest threshold that gives FMR ≤ 0.0001 on white male impostors. Each cell depicts FMR on a logarithmic scale. The text value is $\log_{10}(FMR)$ with large negative values encoding superior false match rates. The panels appear in order of the highest FMR on African American females, lowest first.