NISTIR XXXX Draft

Ongoing Face Recognition Vendor Test (FRVT)

Part 1: Verification

Patrick Grother
Mei Ngan
Kayee Hanaoka
Information Access Division
Information Technology Laboratory

This publication is available free of charge from:
https://www.nist.gov/programs-projects/face-recognition-vendor-test-frvt-ongoing

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

DISCLAIMER

ABOUT THIS REPORT

This report is a draft NIST Interagency Report, and is open for comment. It documents the verification-track of the ongoing Face Recognition Vendor Test. The report will be updated continuously as new algorithms are evaluated, as new datasets are added, and as new analyses are included. Comments and suggestions should be directed to frvt@nist.gov.

Contents

List of Tables

List of Figures

1 Changelog

2017-08-22

- ▷ Added results for three additional algorithms, rankone-002, neurotechnology-001, and itmo-002.
- ▷ The algorithms dermalog-001 and rank-001 have been retired - FRVT lists only two algorithms per organization
- ▷ The algorithm tupel-001 has been retired as it is not operable on all datasets
- ▷ Clarified the tradeoff Figures 1 and 2 plot only genuine comparison durations.
- ▷ Corrected image type label in section 3.5

2017-08-07

- ▷ Added results for 5 new algorithms
- ▷ Added Figure 3 giving simulated example images.
- ▷ Added Figure 1 showing an alternative view of the same tradeoff data in Figure 2
- ▷ Added Figure 5 showing accuracy on visa images just for low FMR.
- ▷ Added Figure 20 showing impostor distribution shifts from certain country pairs. Section 4.6.1 in this and prior reports documents high false match rates for individuals from certain countries. That effect, however, is often not confined to anomalously high impostor scores in the tails of the distribution, but arises from systematic shifts of the whole distribution. These shifts sometimes reach 2σ.

2017-07-29

- ▷ Added results for 8 new algorithms
- ▷ Added results for a child-exploitation dataset
- ▷ Added Table 2 a standalone tabulation of false non-match rates
- ▷ We have received additional CPU algorithms - Results should appear August 4, 2017
- ▷ We have received additional GPU algorithms - Results to appear as computational resources are released from the Face Recognition Prize Challenge

2017-06-19

- ▷ Added five new algorithms, three of which remain in-process
- ▷ Added results for a "wild" dataset of images similar to non-cooperative photojournalism images
- ▷ Added Table 3 a standalone tabulation of failure to enrol rates
- ▷ Added Fig. 2 showing tradeoff between FNMR, template size, template generation time, and match duration.
- ▷ Added Fig. 16 showing how FMR is concentrated in certain images.
- ▷ Restated cross-region false match rates at nominal FMR = 0.0001 instead of 0.001
- ▷ Improved DET legends.

	Developer	Short	Seq.	Validation	Config[1]	Template		GPU	Comparison Time (ns)[3]	
	Name	Name	Num.	Date	Data (KB)	Size (B)	Time (ms)[2]		Genuine	Impostor
1	3DiVi	3divi	000	2017-03-16	169360	[8]512 ± 0	[13]285 ± 52	Yes	[1]378 ± 20	[2]375 ± 19
2	Ayonix	ayonix	000	2017-06-22	58505	[11]1036 ± 0	[1]18 ± 2	No	[2]621 ± 23	[3]620 ± 26
3	Cyberextruder	cyberex	001	2017-08-02	121211	[6]256 ± 0	[28]893 ± 25	No	[6]1083 ± 16	[8]1079 ± 19
4	Dermalog	dermalog	002	2017-02-22	0	[13]1043 ± 0	[3]81 ± 0	No	[18]22169 ± 114	[21]22105 ± 146
5	Dermalog	dermalog	003	2017-07-10	0	[12]1043 ± 0	[9]121 ± 22	No	[19]22957 ± 93	[22]22808 ± 131
6	Digital Barriers	barriers	000	2017-05-31	157794	[17]2056 ± 0	[5]104 ± 0	No	[14]13232 ± 166	[17]13226 ± 146
7	Digital Barriers	barriers	001	2017-07-20	236934	[18]2056 ± 0	[14]294 ± 1	No	[13]12311 ± 164	[15]12347 ± 197
8	ID3 Technology	id3	001	2017-08-04	225574	[10]520 ± 0	[10]238 ± 19	No	[5]1058 ± 32	[7]1049 ± 28
9	ID3 Technology	id3	002	2017-08-04	225574	[9]520 ± 0	[18]482 ± 34	No	[7]1100 ± 59	[6]1048 ± 32
10	Innovatrics	innova	000	2017-07-25	0	[3]146 ± 0	[19]578 ± 5	No	[10]4964 ± 63	[11]4665 ± 262
11	Innovatrics	innova	001	2017-07-25	0	[7]284 ± 0	[23]645 ± 5	No	[11]5506 ± 131	[12]4975 ± 308
12	Is It You	isityou	000	2017-06-26	48010	[29]19200 ± 0	[8]113 ± 5	No	[30]237517 ± 1318	[30]237374 ± 1279
13	ITMO University	itmo	001	2017-06-12	1923215	[31]37997 ± 0	[32]959 ± 18	No	[20]28901 ± 1492	[23]27517 ± 186
14	ITMO University	itmo	002	2017-08-07	1923215	[25]4162 ± 0	[21]611 ± 17	No	[12]7423 ± 96	[13]7451 ± 94
15	Morpho	morpho	000	2017-07-11	100806	[1]116 ± 0	[7]109 ± 1	No	[4]993 ± 31	[5]1000 ± 34
16	Neurotechnology	neurotech	000	2017-03-22	62129	[28]7148 ± 0	[22]611 ± 48	No	[28]74288 ± 2194	[28]72879 ± 2640
17	Neurotechnology	neurotech	001	2017-08-07	280771	[21]2718 ± 0	[26]881 ± 46	No	[27]69356 ± 684	[27]69140 ± 579
18	N-Tech Lab	ntech	000	2017-03-13	191530	[22]2906 ± 1	[12]278 ± 13	No	[22]30787 ± 142	[24]30846 ± 77
19	N-Tech Lab	ntech	001	2017-05-10	691296	[27]6744 ± 1	[20]587 ± 11	No	[25]67692 ± 833	[26]67486 ± 244
20	Rank One Computing	rankone	000	2017-03-21	0	[2]144 ± 0	[4]82 ± 9	No	[24]39932 ± 468	[14]8722 ± 171
21	Rank One Computing	rankone	002	2017-08-18	0	[5]224 ± 0	[2]75 ± 1	No	[26]69113 ± 3802	[1]369 ± 26
22	Samtech InfoNet Limited	samtech	000	2017-05-02	109774	[16]2056 ± 0	[11]262 ± 2	No	[9]4550 ± 26	[10]4541 ± 28
23	TongYi Transportation Technology	tongyi	001	2017-04-01	625339	[20]2058 ± 0	[15]310 ± 20	No	[16]17769 ± 74	[19]17750 ± 63
24	TongYi Transportation Technology	tongyi	002	2017-07-15	625336	[19]2058 ± 0	[16]356 ± 35	No	[21]29816 ± 281	[20]17799 ± 127
25	VCognition	vcog	001	2017-03-28	86103	[23]4126 ± 0	[6]108 ± 17	Yes	[15]16320 ± 197	[18]16426 ± 425
26	VCognition	vcog	002	2017-06-12	3229434	[32]61504 ± 5	[17]357 ± 25	No	[31]296154 ± 3077	[31]296436 ± 4183
27	Vigilant Solutions	vigilant	000	2017-03-30	352218	[30]31540 ± 0	[27]884 ± 23	No	[17]18201 ± 94	[16]13030 ± 83
28	Vigilant Solutions	vigilant	001	2017-06-13	344685	[15]1544 ± 0	[30]921 ± 2	No	[3]644 ± 13	[4]649 ± 16
29	VisionLabs	visionlabs	001	2017-06-12	343661	[4]204 ± 0	[31]943 ± 8	No	[8]1395 ± 45	[9]1148 ± 53
30	Vocord	vocord	001	2017-04-21	616989	[26]6194 ± 0	[29]908 ± 16	No	[32]1094730 ± 64282	[32]1107193 ± 66523
31	Vocord	vocord	002	2017-06-07	918292	[14]1330 ± 0	[25]782 ± 36	Yes	[29]83063 ± 517	[29]83072 ± 714
32	Shanghai Yitu Technology	yitu	000	2017-05-23	2211068	[24]4130 ± 0	[24]672 ± 2	No	[23]35352 ± 114	[25]37848 ± 1773

Notes	
1	The size of configuration data does not capture static data included in the libraries. We do not include the size of the libraries because some algorithms include common ancilliary libraries for image processing (e.g. openCV) or numerical computation (e.g. blas).
2	The median template creation times are measured on Intel®Xeon®CPU E5-2630 v4 @ 2.20GHz processors or, in the case of GPU-enabled implementations, NVidia Tesla K40.
3	The median comparison durations, in nanoseconds, are estimated using std::chrono::high_resolution_clock which on the machine in (2) counts clock ticks of duration 1ns. Precision is somewhat worse than that however. The ± value is the median absolute deviation times 1.48 for Normal consistency.

Table 1: Summary of algorithms and properties included in this report. The red superscripts give ranking for the quantity in that column.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

	Algorithm Name	FALSE NON-MATCH RATE (FNMR)													
		CONSTRAINED, COOPERATIVE								LESS CONSTRAINED, NON-COOPERATIVE					
		VISA		VISA		MUGSHOT		WEBCAM		SELFIE		WILD		CHILD EXP	
	FMR	1E-06		0.0001		0.0001		0.0001		0.0001		0.0001		0.01	
1	3divi-000	0.133	7	0.029	6	0.037	9	0.002	7	0.055	14	0.547	4	0.553	3
2	ayonix-000	0.487	27	0.230	28	0.309	28	0.172	27	0.360	28	0.807	20	0.843	16
3	cyberextruder-001	0.255	18	0.076	20	0.165	22	0.029	19	0.144	24	0.853	23	-	31
4	dermalog-002	0.315	22	0.122	23	0.241	26	0.109	26	0.179	26	0.709	12	0.985	21
5	dermalog-003	0.280	20	0.112	22	0.202	25	0.041	23	0.115	23	0.693	9	0.845	18
6	digitalbarriers-000	0.463	26	0.161	25	0.184	24	0.045	24	0.170	25	0.741	17	0.771	14
7	digitalbarriers-001	0.502	28	0.155	24	0.041	12	0.029	20	0.115	22	0.678	8	-	25
8	id3-001	0.250	17	0.063	18	0.036	8	0.002	8	0.040	10	0.765	18	-	32
9	id3-002	0.239	16	0.057	17	0.032	7	0.003	10	0.037	8	0.810	21	-	23
10	innovatrics-000	0.191	12	0.034	10	0.046	17	0.001	4	0.040	9	0.720	13	-	24
11	innovatrics-001	0.183	11	0.034	9	0.043	13	0.001	5	0.043	11	0.643	6	-	28
12	isityou-000	0.703	30	0.414	30	0.680	30	0.690	31	-	32	1.000	31	-	22
13	itmo-001	0.441	24	0.174	27	0.256	27	0.069	25	0.199	27	0.932	28	-	27
14	itmo-002	0.287	21	0.050	15	0.047	18	0.036	21	0.069	20	0.892	25	-	26
15	morpho-000	0.134	8	0.026	4	0.028	4	0.007	13	0.012	2	0.893	26	0.846	19
16	neurotechnology-000	0.237	15	0.052	16	0.062	19	0.005	12	0.052	12	0.943	29	0.845	17
17	neurotechnology-001	0.222	14	0.044	13	0.046	16	0.001	6	0.017	7	0.817	22	-	29
18	ntechlab-000	0.086	6	0.027	5	0.044	14	0.003	9	0.014	5	0.367	2	0.533	2
19	ntechlab-001	0.083	5	0.025	3	0.030	5	0.003	11	0.014	4	0.319	1	0.472	1
20	rankone-000	0.276	19	0.071	19	0.177	23	0.021	17	0.092	21	0.723	14	0.787	15
21	rankone-002	0.217	13	0.049	14	0.032	6	0.000	2	0.052	13	0.705	11	-	30
22	samtech-000	0.443	25	0.161	26	0.044	15	0.021	18	0.063	18	0.878	24	0.765	13
23	tongyitrans-001	0.072	4	0.038	12	0.041	11	0.009	14	0.063	16	0.704	10	0.743	9
24	tongyitrans-002	0.066	3	0.030	7	0.039	10	0.010	15	0.063	17	0.725	15	0.746	10
25	vcog-001	0.892	31	0.409	29	-	32	0.302	28	0.427	29	-	32	0.686	6
26	vcog-002	0.903	32	0.504	32	0.692	31	0.559	30	0.666	31	0.778	19	0.752	11
27	vigilantsolutions-000	0.688	29	0.415	31	0.595	29	0.401	29	0.643	30	0.915	27	0.894	20
28	vigilantsolutions-001	0.348	23	0.105	21	0.101	21	0.016	16	0.061	15	0.729	16	0.730	8
29	visionlabs-001	0.180	10	0.030	8	0.024	3	0.001	3	0.014	6	0.591	5	0.561	4
30	vocord-001	0.141	9	0.035	11	0.063	20	0.036	22	0.069	19	0.654	7	0.695	7
31	vocord-002	0.034	2	0.013	1	0.019	2	-	32	0.012	3	0.948	30	0.762	12
32	yitu-000	0.033	1	0.021	2	0.017	1	0.000	1	0.012	1	0.431	3	0.586	5

Table 2: FNMR is the proportion of mated comparisons below a threshold set to achieve the FMR given in the header on the fourth row. FMR is the proportion of impostor comparisons at or above that threshold. Note that the webcam and selfie values apply to images collected on the same day, and that will often yield optimisticly low FNMR values.

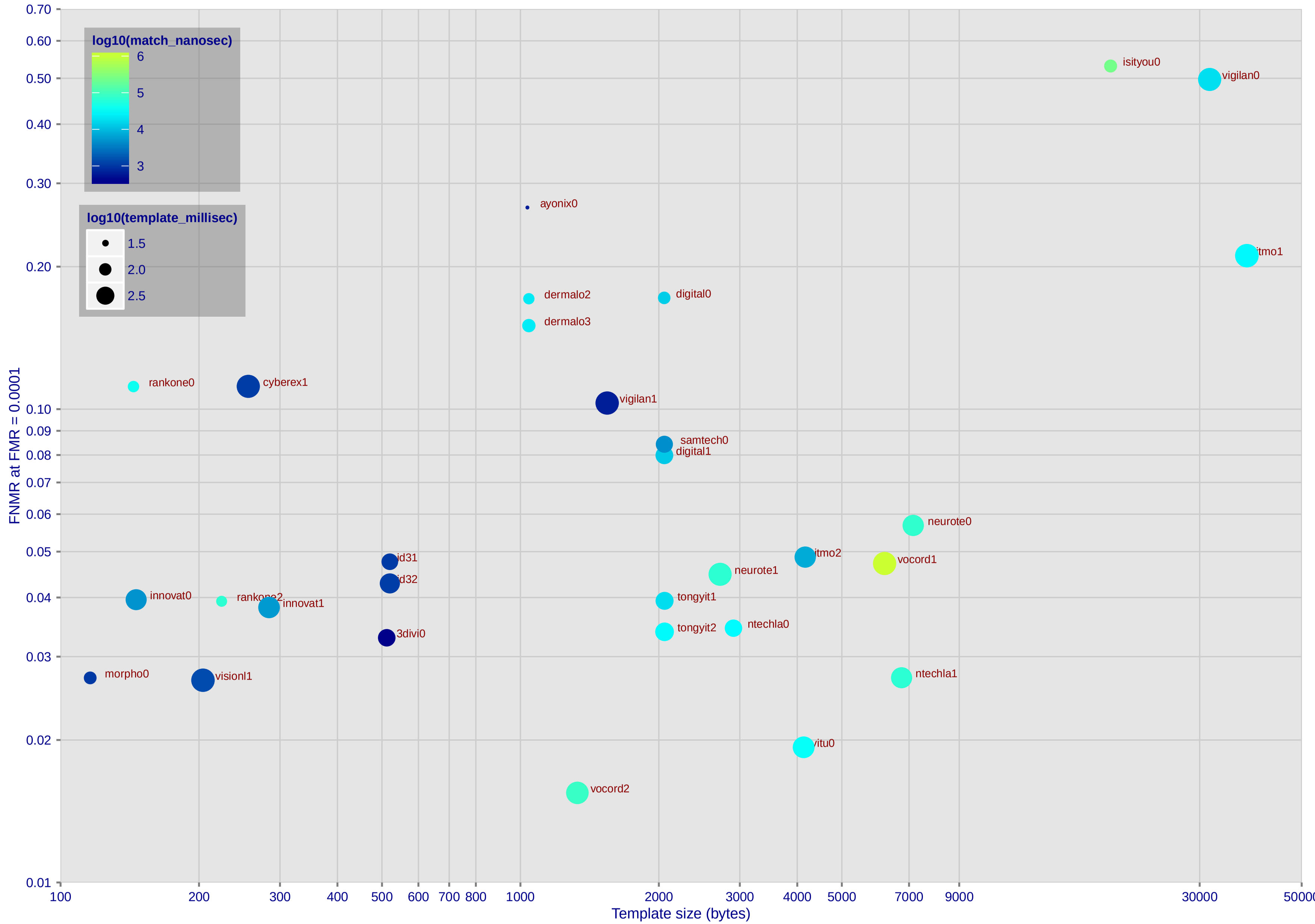


Figure 1: The points show false non-match rates (FNMR) versus the duration of the template generation operation. FNMR is the geometric mean of FNMR values for visa and mugshot images (from Figs. 4 and 6) at a false match rate (FMR) of 0.0001. Template generation time is a median estimated over 640 x 480 pixel portraits. The size of the points encodes template generation time - which spans one order of magnitude. The color of the points encodes one-to-one genuine template comparison duration - which span three orders of magnitude.

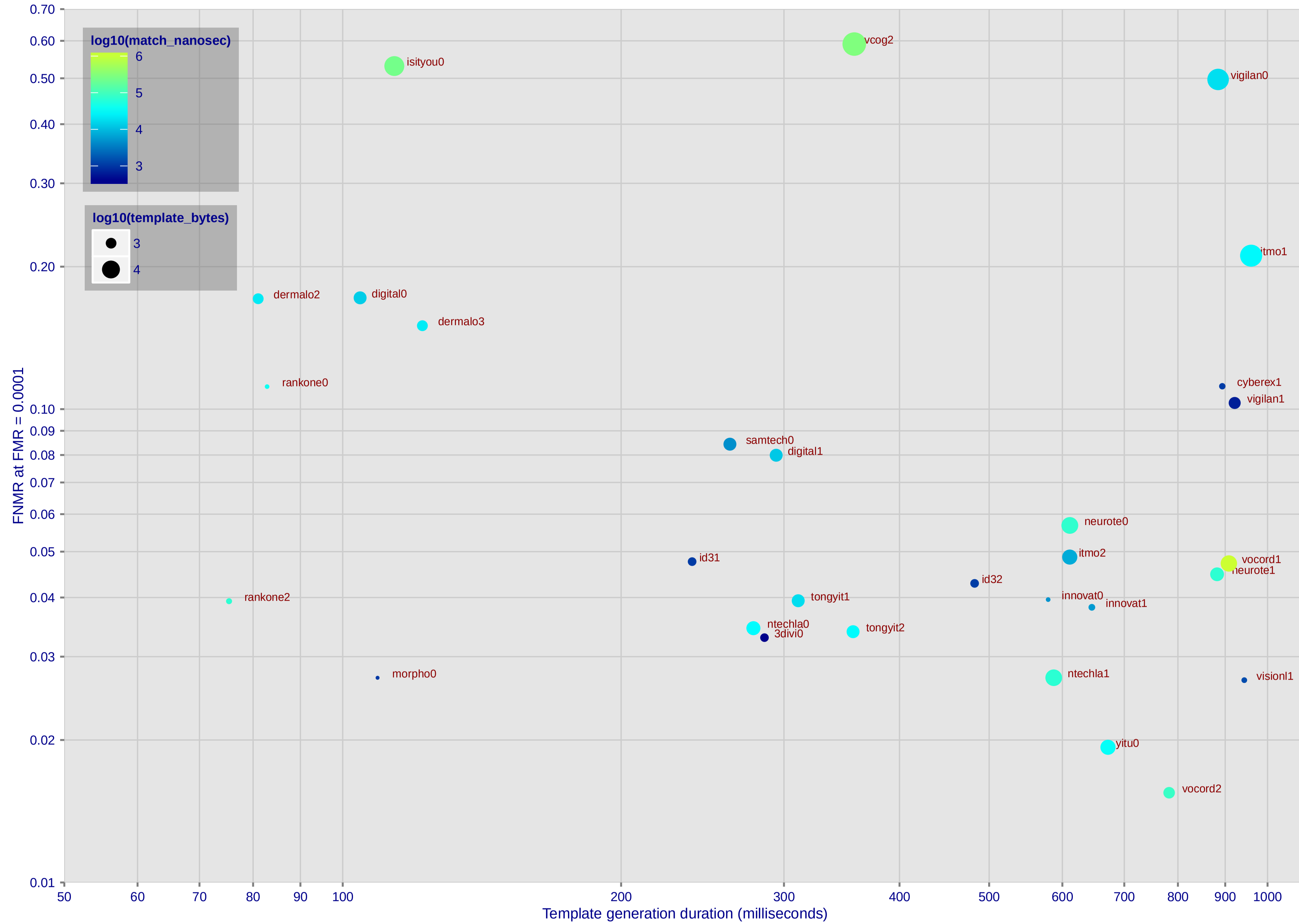


Figure 2: The points show false non-match rates (FNMR) versus the duration of the template generation operation. FNMR is the geometric mean of FNMR values for visa and mugshot images (from Figs. 4 and 6) at a false match rate (FMR) of 0.0001. Template generation time is a median estimated over 640 x 480 pixel portraits. The size of the points encodes template size - which span two orders of magnitude. The color of the points encodes one-to-one genuine template comparison duration - which span three orders of magnitude.

2 Metrics

2.1 Core accuracy

Given a vector of N genuine scores, u, the false non-match rate (FNMR) is computed as the proportion below some threshold, T:

$$\text{FNMR}(T) = 1 - \frac{1}{N}\sum_{i}^{N} H(u_i - T) \tag{1}$$

where $H(x)$ is the unit step function, and $H(0)$ taken to be 1.

Similarly, given a vector of N impostor scores, v, the false match rate (FMR) is computed as the proportion above T:

$$\text{FMR}(T) = \frac{1}{N}\sum_{i}^{N} H(v_i - T) \tag{2}$$

The threshold, T, can take on any value. We typically generate a set of thresholds from quantiles of the observed impostor scores, v, as follows. Given some interesting false match rate range, $[\text{FMR}_L, \text{FMR}_U]$, we form a vector of K thresholds corresponding to FMR measurements evenly spaced on a logarithmic scale

$$T_k = Q_v(1 - \text{FMR}_k) \tag{3}$$

where Q is the quantile function, and FMR_k comes from

$$\log_{10}\text{FMR}_k = \log_{10}\text{FMR}_L + \frac{k}{K}\ [\log_{10}\text{FMR}_U - \log_{10}\text{FMR}_L] \tag{4}$$

Error tradeoff characteristics are plots of FNMR(T) vs. FMR(T). These are plotted with $\text{FMR}_U \to 1$ and FMR_L as low as is sustained by the number of impostor comparisons, N. This is somewhat higher than the "rule of three" limit $3/N$ because samples are not independent, due to re-use of images.

3 Datasets

3.1 Child exploitation images

- ▷ The number of images is $O(10^4)$.
- ▷ The number of subjects is $O(10^3)$.
- ▷ The number of subjects with two images $O(10^3)$.
- ▷ The images are operational. They are taken from ongoing investigations of child exploitation crimes. The images are arbitrarily unconstrained. Pose varies considerably around all three axes, including subject lying down. Resolution varies very widely. Faces can be occluded by other objects, including hair and hands. Lighting varies, although the images are intended for human viewing. Mis-focus is rare. Images are given to the algorithm without any cropping; faces may occupy widely varying areas.
- ▷ The images are usually large from contemporary cameras. The mean interocular distance (IOD) is 70 pixels.
- ▷ The images are of subjects from several countries, due to the global production of this imagery.
- ▷ The images are of children, from infancy to late adolescence.
- ▷ All of the images are live capture, none are scanned. Many have been cropped.
- ▷ When these images are input to the algorithm, they are labelled as being of type ”EXPLOITATION” - see Table 4 of the FRVT API.

3.2 Visa images

- ▷ The number of images is $O(10^5)$.
- ▷ The number of subjects is $O(10^5)$.
- ▷ The number of subjects with two images $O(10^4)$.
- ▷ The images have geometry in reasonable conformance with the ISO/IEC 19794-5 Full Frontal image type. Pose is generally excellent.
- ▷ The images are of size 252x300 pixels. The mean interocular distance (IOD) is 69 pixels.
- ▷ The images are of subjects from greater than 100 countries, with significant imbalance due to visa issuance patterns.
- ▷ The images are of subjects of all ages, including children, again with imbalance due to visa issuance demand.
- ▷ Many of the images are live capture. A substantial number of the images are photographs of paper photographs.
- ▷ When these images are input to the algorithm, they are labelled as being of type ”ISO” - see Table 4 of the FRVT API.

3.3 Mugshot images

- ▷ The number of images is $O(10^6)$.
- ▷ The number of subjects is $O(10^5)$.
- ▷ The number of subjects with two images $O(10^5)$.

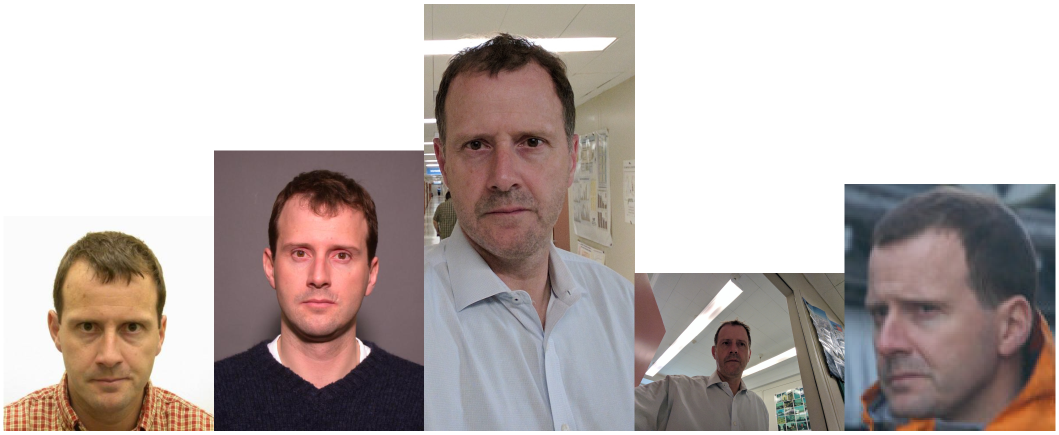
- ▷ The images have geometry in reasonable conformance with the ISO/IEC 19794-5 Full Frontal image type.
- ▷ The images are of variable sizes. The median IOD is 104 pixels. The mean IOD is 123 pixels.
- ▷ The images are of subjects from the United States.
- ▷ The images are of adults.
- ▷ The images are all live capture.
- ▷ When these images are input to the algorithm, they are labelled as being of type ”mugshot” - see Table 4 of the FRVT API.

3.4 Selfie images

- ▷ The number of images is below 500.
- ▷ The number of subjects is below 500.
- ▷ All subjects have a selfie image, and a portrait image.
- ▷ The portrait images are in reasonable conformance with the ISO/IEC 19794-5 Full Frontal image type.
- ▷ The selfie images vary: taken with camera above and below eye level, with one hand or two hands. Pitch angles vary more than yaw angles, which are frontal. Some perspective distortion is evident.
- ▷ The images have mean IOD of 140 pixels.
- ▷ The images are of subjects from the United States.
- ▷ The images are of adults.
- ▷ The images are all live capture.
- ▷ When these images are input to the algorithm, they are labelled as being of type ”WILD” - see Table 4 of the FRVT API.

3.5 Webcam images

- ▷ The number of images is below 1500.
- ▷ The number of subjects is below 1500.
- ▷ All subjects have a webcam image, and a portrait image.
- ▷ The portrait images are in reasonable conformance with the ISO/IEC 19794-5 Full Frontal image type.
- ▷ The webcam images are taken with camera at a typical head height, with mild pitch angles, low yaw angles, but some variation in range, such that low perspective distortion is sometimes evident.
- ▷ The images have mean IOD of 68 pixels (sd=12).
- ▷ The images are of subjects from the United States.
- ▷ The images are of adults.
- ▷ The images are all live capture.
- ▷ When these images are input to the algorithm, they are labelled as being of type ”MUGSHOT” - see Table 4 of the FRVT API.



(a) Visa (b) Mugshot (c) Webcam (d) Selfie (e) Wild

Figure 3: The figure gives simulated samples of image types used in this report.

3.6 Wild images

- ▷ The number of images is O(10^5).
- ▷ The number of subjects is O(10^3).
- ▷ The number of subjects with two images O(10^3).
- ▷ The images include many photojournalism-style images. Images are given to the algorithm using a variable but generally tight crop of the head. Resolution varies very widely. The images are very unconstrained, with wide yaw and pitch pose variation. Faces can be occluded, including hair and hands.
- ▷ The images are of adults.
- ▷ All of the images are live capture, none are scanned.
- ▷ When these images are input to the algorithm, they are labelled as being of type ”WILD” - see Table 4 of the FRVT API.

4 Results

4.1 Test goals

- ▷ To state overall accuracy.
- ▷ To compare algorithms.

4.2 Test design

Method: For visa images:

- ▷ The comparisons are of visa photos against visa photos.
- ▷ The number of genuine comparisons is O(10^4).
- ▷ The number of impostor comparisons is O(10^{10}).

- ▷ The comparisons are fully zero-effort, meaning impostors are paired without attention to sex, age or other covariates. However, later analysis is conducted on subsets.
- ▷ The number of persons is O(10^5).
- ▷ The number of images used to make 1 template is 1.
- ▷ The number of templates used to make each comparison score is two corresponding to simple one-to-one verification.

For mugshot images:

- ▷ The comparisons are of mugshot photos against mugshot photos.
- ▷ The number of genuine comparisons is O(10^5).
- ▷ The number of impostor comparisons is O(10^7).
- ▷ The comparisons are fully zero-effort, meaning impostors are paired without attention to sex, age or other covariates.
- ▷ The number of persons is O(10^6).
- ▷ The number of images used to make 1 template is 1.
- ▷ The number of templates used to make each comparison score is two corresponding to simple one-to-one verification.

For selfie images:

- ▷ The comparisons are of selfie photos against portrait photos.
- ▷ The number of genuine comparisons is O(10^2).
- ▷ The number of impostor comparisons is O(10^8) selfies are compared with portraits of O(10^6) other subjects.
- ▷ The comparisons are fully zero-effort, meaning impostors are paired without attention to sex, age or other covariates.
- ▷ The number of persons is O(10^6).
- ▷ The number of images used to make 1 template is 1.
- ▷ The number of templates used to make each comparison score is two corresponding to simple one-to-one verification.

For webcam images:

- ▷ The comparisons are of webcam photos against portrait photos.
- ▷ The number of genuine comparisons is O(10^3).
- ▷ The number of impostor comparisons is O(10^9) webcams are compared with portraits of O(10^6) other subjects.
- ▷ The comparisons are fully zero-effort, meaning impostors are paired without attention to sex, age or other covariates.
- ▷ The number of persons is O(10^6).
- ▷ The number of images used to make 1 template is 1.

▷ The number of templates used to make each comparison score is two corresponding to simple one-to-one verification.

For child exploitation images:

▷ The comparisons are of unconstrained child exploitation photos against others of the same type.

▷ The number of genuine comparisons is O(10^4).

▷ The number of impostor comparisons is O(10^7).

▷ The comparisons are fully zero-effort, meaning impostors are paired without attention to sex, age or other covariates.

▷ The number of persons is O(10^3).

▷ The number of images used to make 1 template is 1.

▷ The number of templates used to make each comparison score is two corresponding to simple one-to-one verification.

▷ We produce two performance statements. First, is a DET as used for visa and mugshot images. The second is a cumulative match characteristic (CMC) summarizing a simulated one-to-many search process. This is done as follows.

- We regard M enrollment templates as items in a gallery.
- These M templates come from $M > N$ individuals, because multiple images of a subject are present in the gallery under separate identifiers.
- We regard the verification templates as search templates.
- For each search we compute the rank of the highest scoring mate.
- This process should properly be conducted with a 1:N algorithm, such as those tested in NIST IR 8009. We use the 1:1 algorithms in a simulated 1:N mode here to a) better reflect what a child exploitation analyst does, and b) to do show algorithm efficacy is better than that revealed in the verification DETs.

4.3 Failure to enrol

Algorithm	Failure to Enrol Rate[1]											
Name	CHILD-EXPLOIT		MUGSHOT		SELFIES		VISA		WEBCAM		WILD	
3divi-000	0.2019	12	0.0019	18	0.0202	18	0.0008	12	0.0020	15	0.2070	19
ayonix-000	0.0000	1	0.0109	31	0.0751	29	0.0137	35	0.0109	21	0.0000	2
cyberextruder-001	-	32	0.0036	22	0.0376	20	0.0029	24	0.0205	26	0.5730	32
dermalog-002	0.9109	26	0.0045	26	0.0954	32	0.0013	15	0.0471	30	0.3979	26
dermalog-003	0.0434	6	0.0007	7	0.0000	1	0.0025	23	0.0007	6	0.0701	11
digitalbarriers-000	0.5469	23	0.0043	24	0.0925	30	0.0019	21	0.0184	25	0.5170	31
digitalbarriers-001	-	32	0.0044	25	0.0925	31	0.0018	20	0.0232	27	0.4593	30
id3-001	-	32	0.0043	23	0.0260	19	0.0043	32	0.0014	12	0.2090	20
id3-002	-	32	0.0030	21	0.0202	17	0.0032	26	0.0020	14	0.1496	14
innovatrics-000	-	32	0.0013	11	0.0087	13	0.0004	8	0.0027	16	0.3337	24
innovatrics-001	-	32	0.0013	12	0.0087	15	0.0004	9	0.0027	17	0.3337	25
isityou-000	0.4714	21	0.0022	20	0.0665	27	0.0010	13	0.0116	23	0.4586	29
itmo-001	-	32	0.0114	32	0.0694	28	0.0047	33	0.0546	32	0.7410	36
itmo-002	-	32	0.0068	29	0.0636	26	0.0029	25	0.0498	31	0.6985	34
morpho-000	0.0000	2	0.0000	1	0.0000	4	0.0000	2	0.0000	4	0.0000	5
neurotechnology-000	0.0000	4	0.0000	3	0.0000	5	0.0000	3	0.0000	5	0.0163	8
neurotechnology-001	-	32	0.0000	2	0.0000	2	0.0000	1	0.0000	1	0.0163	7
ntechlab-000	0.2496	15	0.0015	16	0.0058	11	0.0016	18	0.0007	7	0.1099	13
ntechlab-001	0.0926	10	0.0009	8	0.0029	7	0.0005	10	0.0007	8	0.0584	10
rankone-000	0.0187	5	0.0005	5	0.0000	6	0.0003	6	0.0007	10	0.2349	21
rankone-002	-	32	0.0001	4	0.0000	3	0.0000	4	0.0000	3	0.0855	12
samtech-000	0.5474	24	0.0052	27	0.0491	24	0.0042	31	0.0252	28	0.7023	35
tongyitrans-001	0.0000	3	0.0068	28	0.0462	22	0.0040	29	0.0055	19	0.0000	3
tongyitrans-002	0.3609	19	0.0078	30	0.0462	23	0.0040	30	0.0055	20	0.0000	4
vcog-001	0.1579	11	-	32	0.0058	10	0.0018	19	0.0000	2	-	32
vcog-002	0.2209	13	0.0021	19	0.0087	14	0.0019	22	0.0007	9	0.1672	16
vigilantsolutions-000	0.5580	25	0.0018	17	0.0462	21	0.0007	11	0.0109	22	0.5927	33
vigilantsolutions-001	0.3585	18	0.0010	10	0.0116	16	0.0004	7	0.0048	18	0.3262	23
visionlabs-001	0.2699	16	0.0014	13	0.0058	12	0.0014	17	0.0020	13	0.1803	17
vocord-001	0.4732	22	0.0158	33	0.0520	25	0.0038	28	0.0348	29	0.4494	28
vocord-002	0.3782	20	0.0015	14	0.0029	9	0.0037	27	0.0171	24	0.1992	18
yitu-000	0.3475	17	0.0015	15	0.0029	8	0.0013	16	0.0014	11	0.1591	15

Table 3: FTE is the proportion of failed template generation attempts. Failures can occur because the software throws an exception, or because the software electively refuses to process the input image. This would typically occur if a face is not detected. FTE is measured as the number of function calls that give a non-zero error code, OR that give a "small" template. This is defined as one whose size is less than 0.3 times the median template size. This second rule is needed because some algorithms incorrectly fail to return a non-zero error code when template generation fails.
[1] *The effects of FTE are included in the accuracy results of this report by regarding any template comparison involving a failed template enrollment to produce a low similarity score. Thus higher FTE results in higher FNMR.*

4.4 Recognition accuracy

Core algorithm accuracy is stated via:

- **Cooperative subjects**
  - The summary table of Figure 2;
  - The visa image DETs of Figures 4 and 5;
  - The mugshot DETs of Figure 6 ;
  - The selfie-portrait DETs of Figure 7;
  - The webcam-portrait DETs of Figure 8;
- **Cooperative subjects**
  - The photojournalism DET of Figure 9
  - The child-exploitation DET of Figure 10;
  - The child-exploitation CMC of Figure 11.

Figure 15 shows dependence of false match rate on algorithm score threshold. This allows a deployer to set a threshold to target a particular false match rate appropriate to the security objectives of the application.

Figure 17 likewise shows FMR(T) but for mugshots, and specially four subsets of the population.

Note that in both the mugshot and visa sets false match rates vary with the ethnicity, age, and sex, of the enrollee and impostor - see section 4.6.

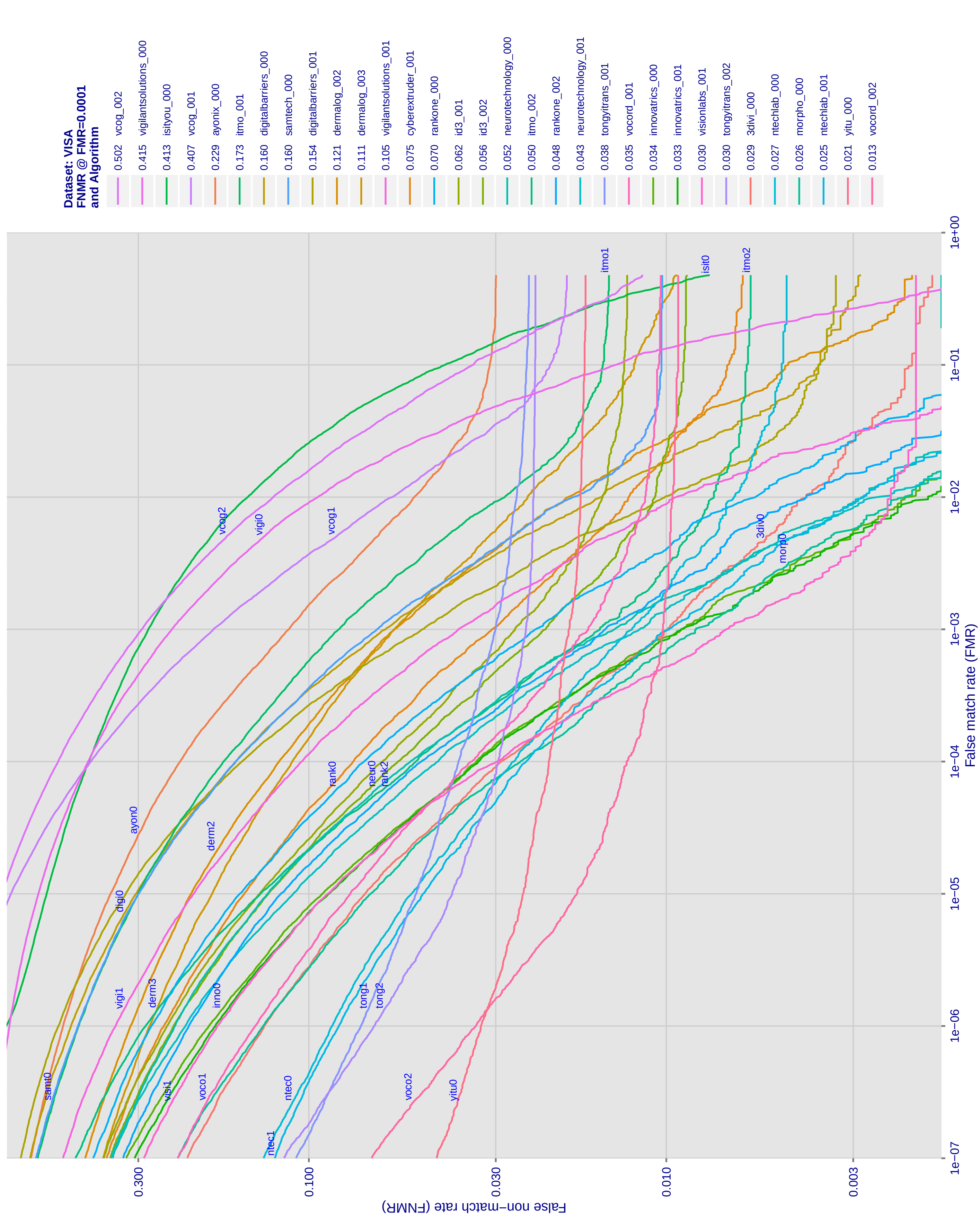


Figure 4: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

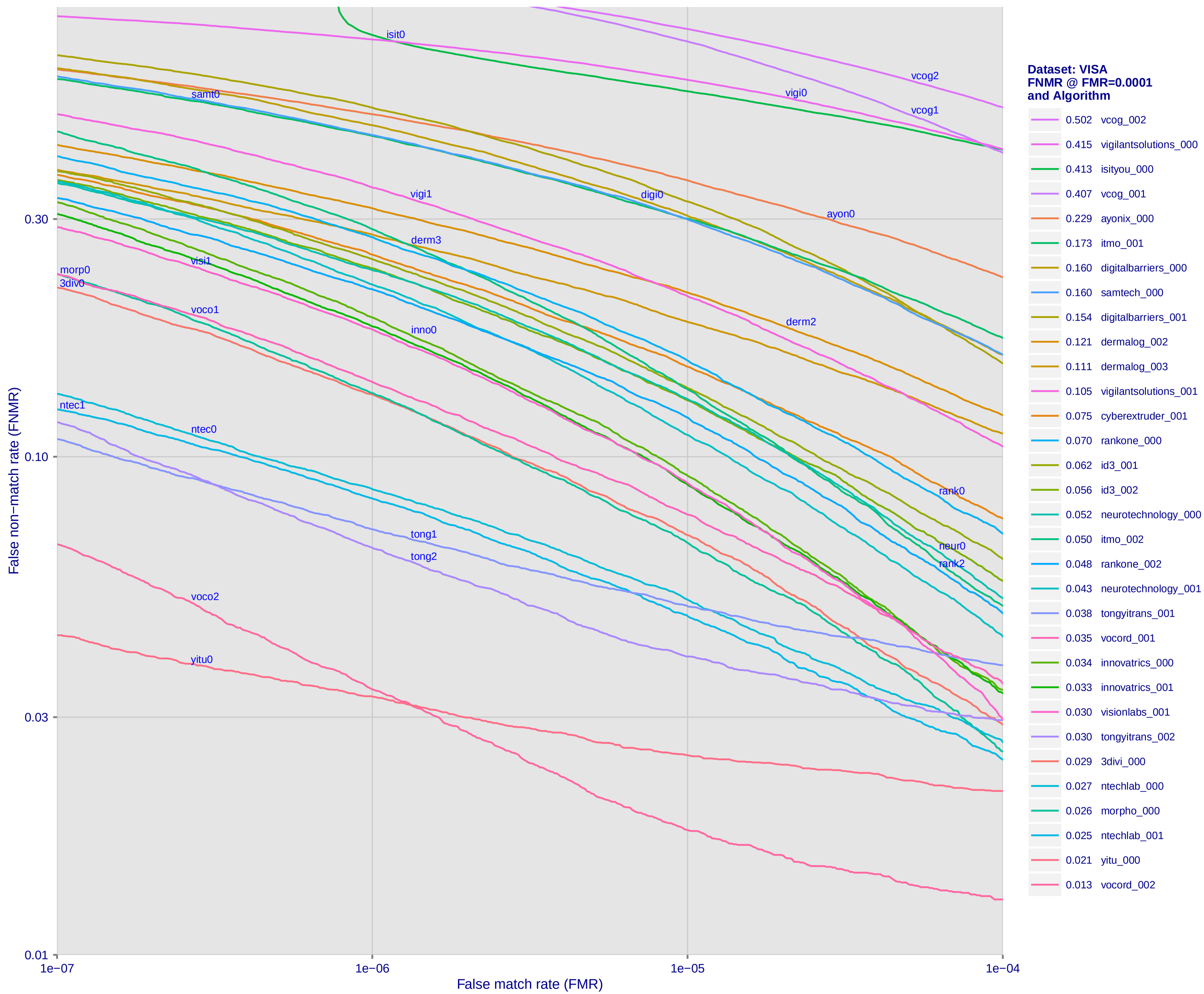


Figure 5: For the visa images, but now just for low FMR, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

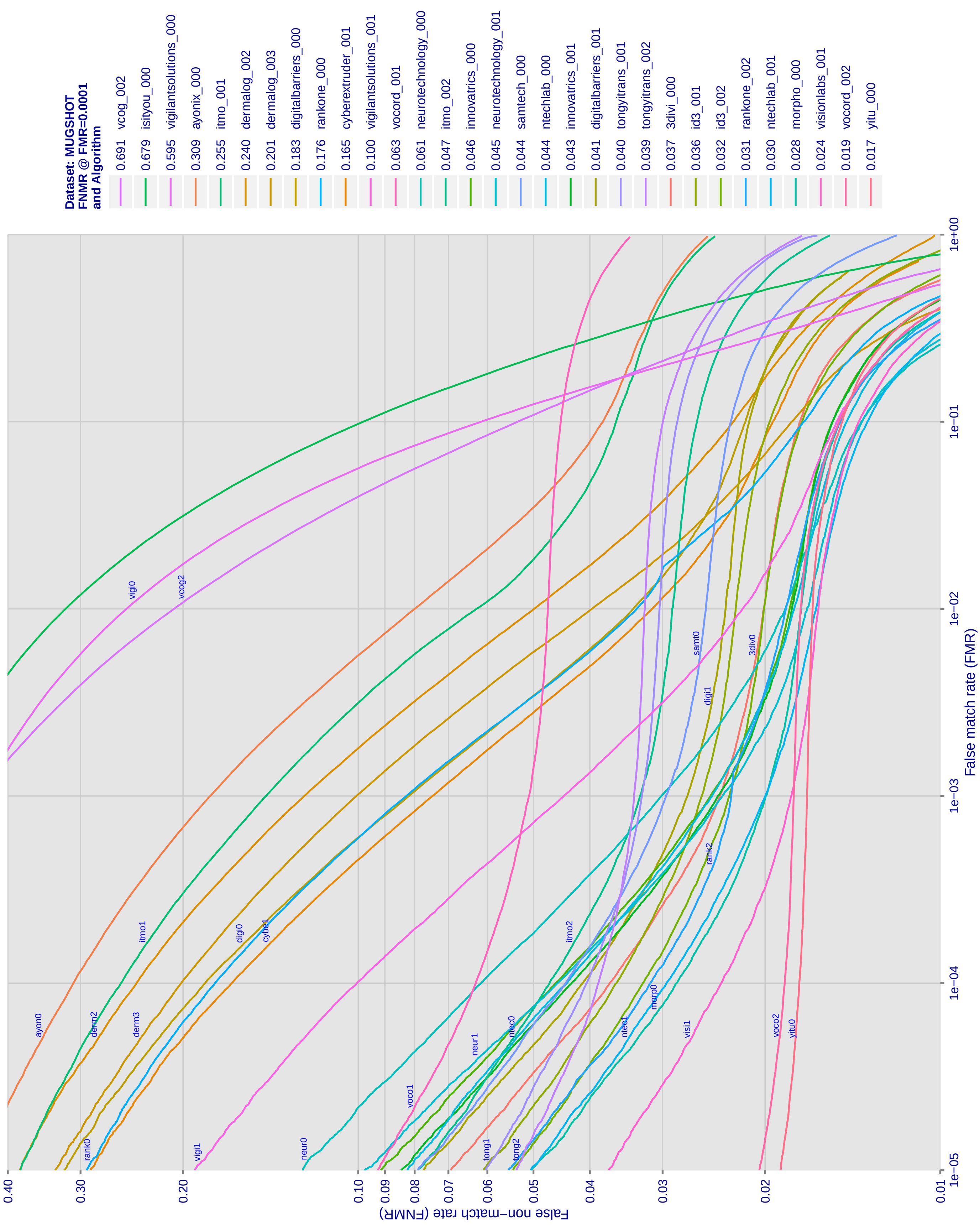


Figure 6: For the mugshot images, detection error tradeoff (DET) *characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show decades of* FMR.

FNMR(T)	"False non-match rate"
FMR(T)	"False match rate"

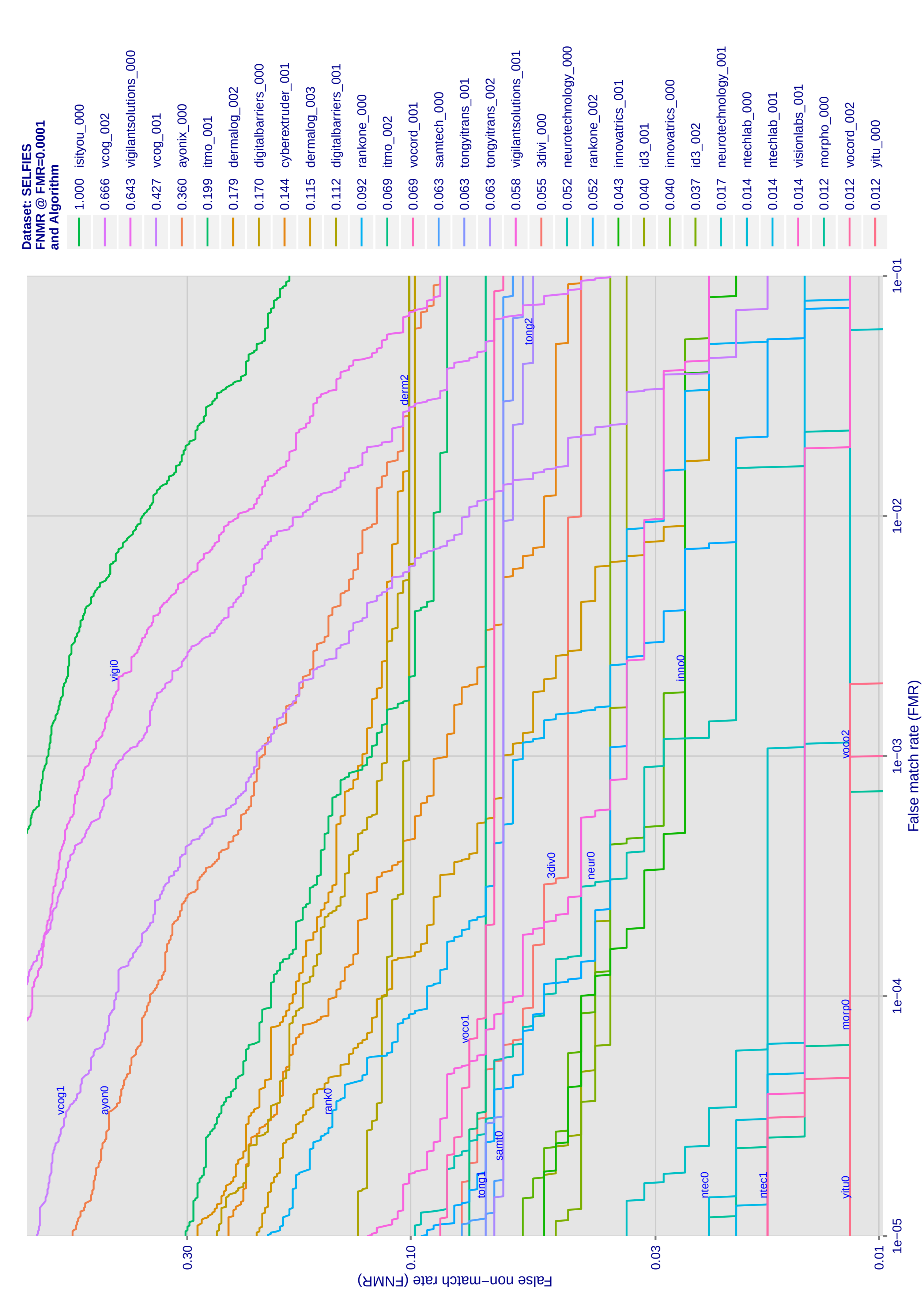


Figure 7: For the selfie-to-portrait comparisons, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show several decades of FMR. Caution: The FNMR values here are optimistic statements of accuracy because the image pairs were collected on the same day. This is known across biometrics to give better accuracy, and is operationally relevant only in special cases.

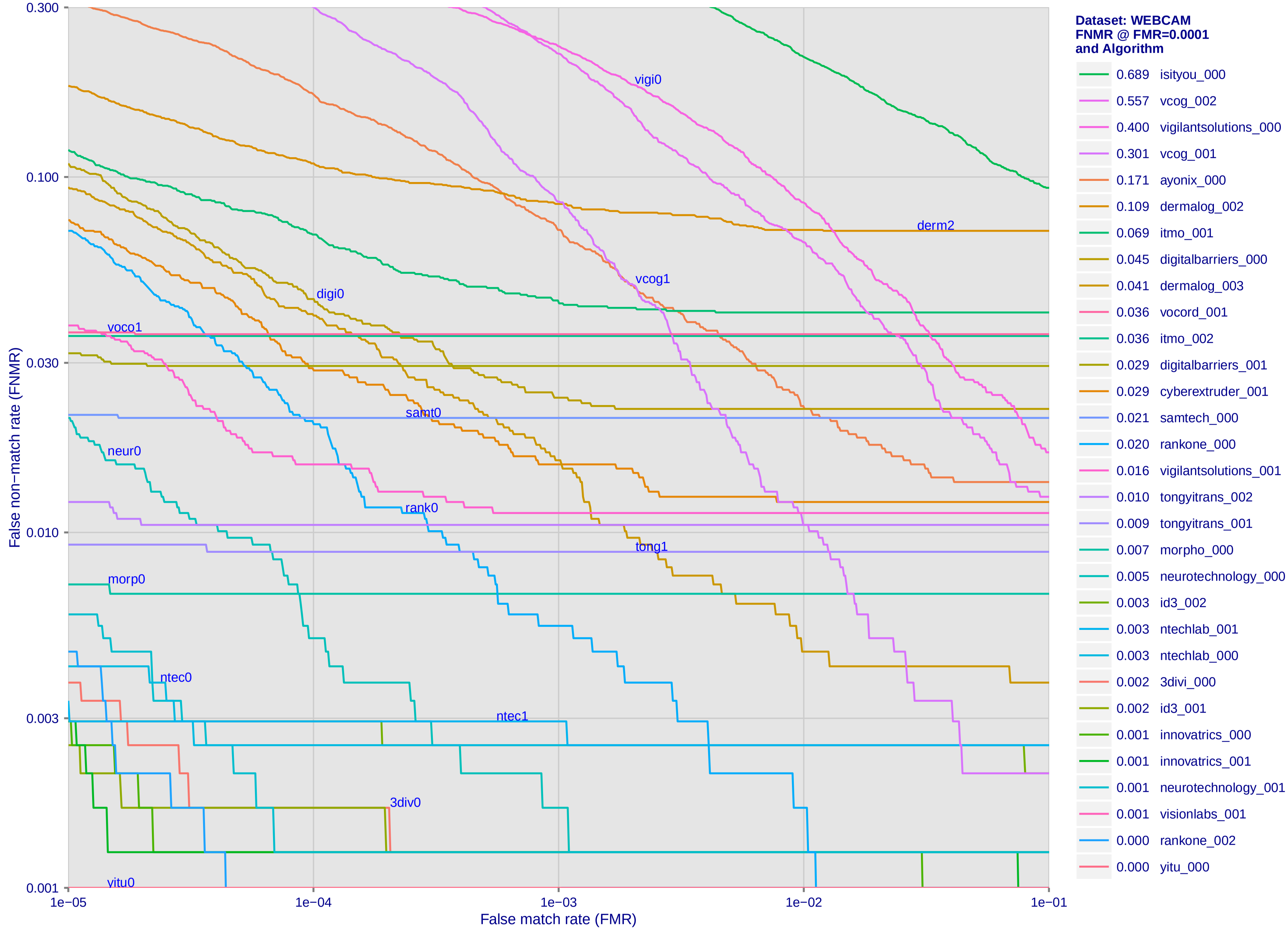


Figure 8: For the webcam-to-portrait comparisons, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show several decades of FMR. Caution: The FNMR values here are optimistic statements of accuracy because the image pairs were collected on the same day. This is known across biometrics to give better accuracy, and is operationally relevant only in special cases.

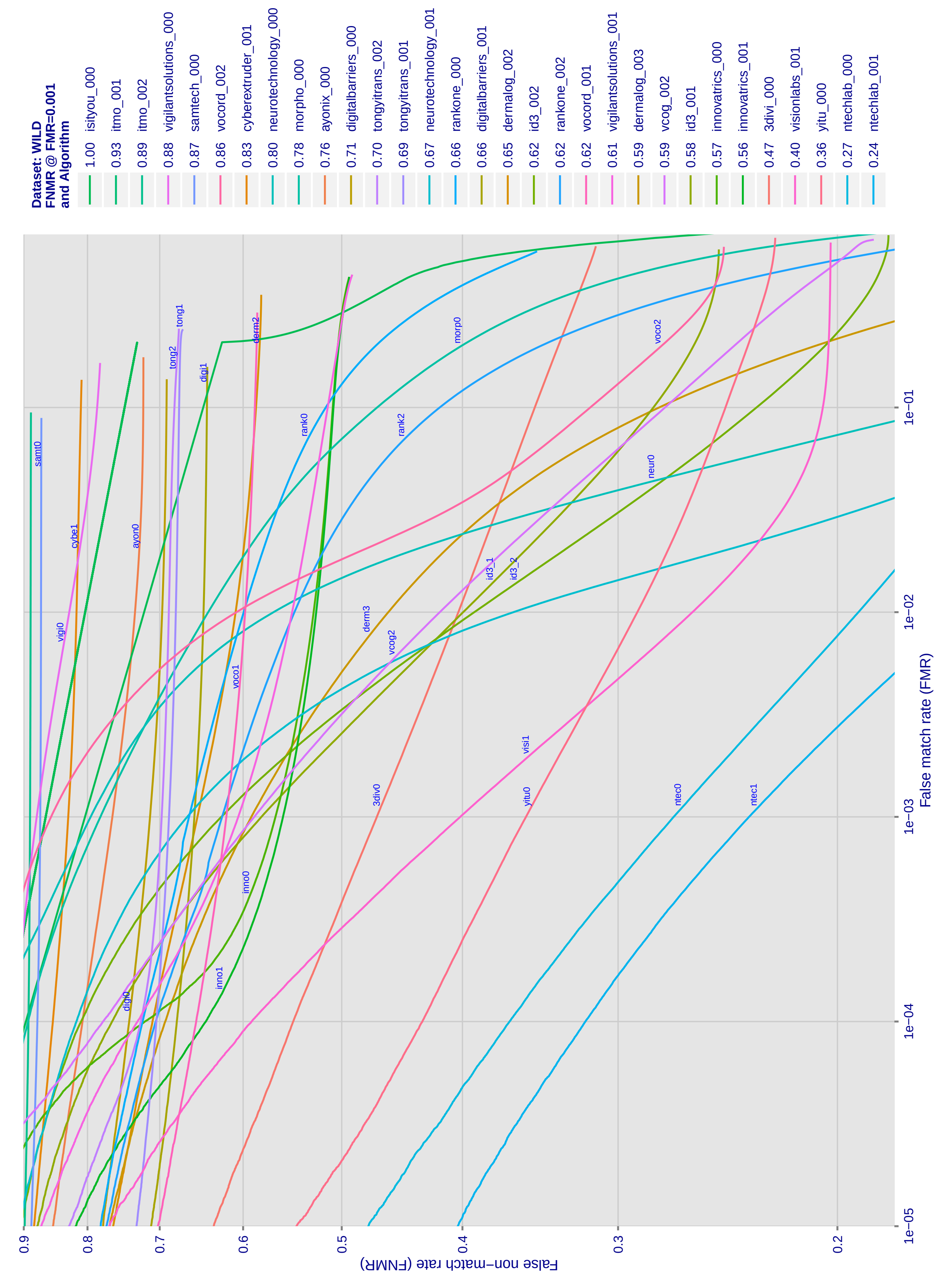


Figure 9: For the wild image comparisons, detection error tradeoff (DET) *characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show several decades of* FMR.

FNMR(T) "False non-match rate"
FMR(T) "False match rate"

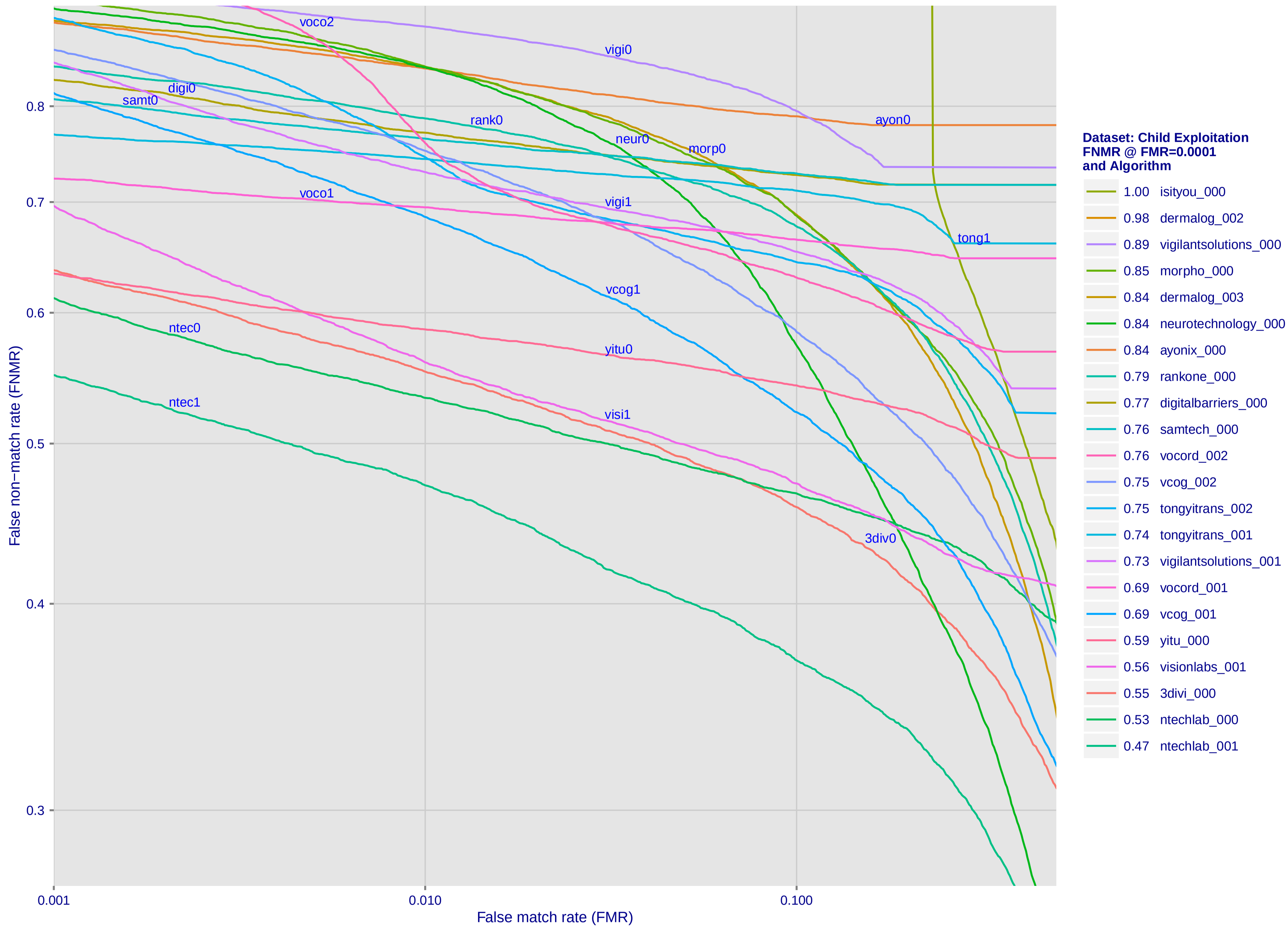


Figure 10: For child exploitation images, detection error tradeoff (DET) *characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of* FMR*. Accuracy is poor because many images have adverse quality characteristics, and because detection and enrollment fails.*

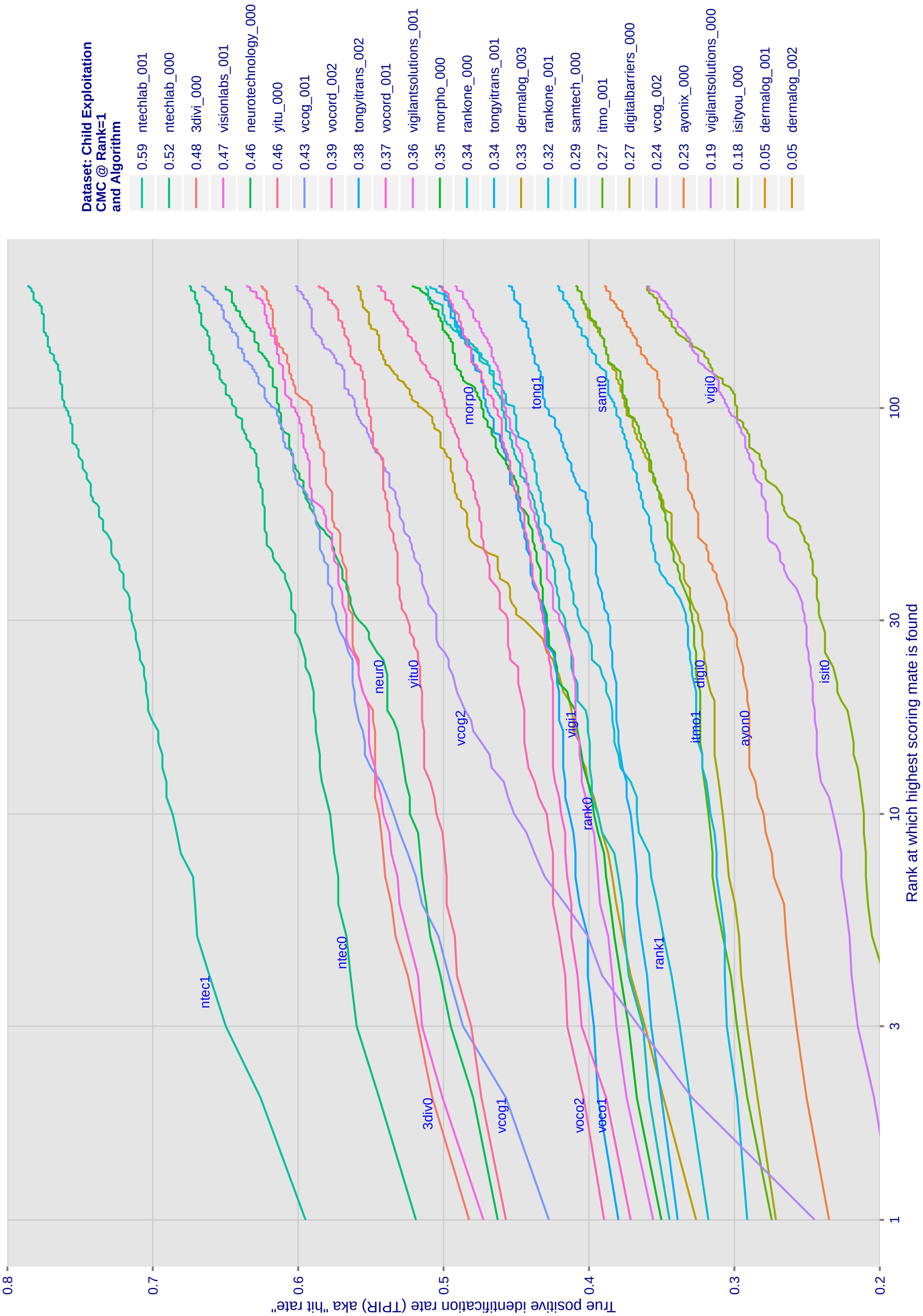


Figure 11: For child exploitation images, cumulative match characteristics (CMC) showing true positive identification rate vs. rank. This is simulation of a one-to-many search experiment - see discussion in section 4.2. The scales are logarithmic in order to show the effect of long candidate lists. Accuracy is poor but much improved relative to the 1:1 DETs of Fig. 10 because a search can succeed if any of a subject's several enrolled images matches the search image with a high score.

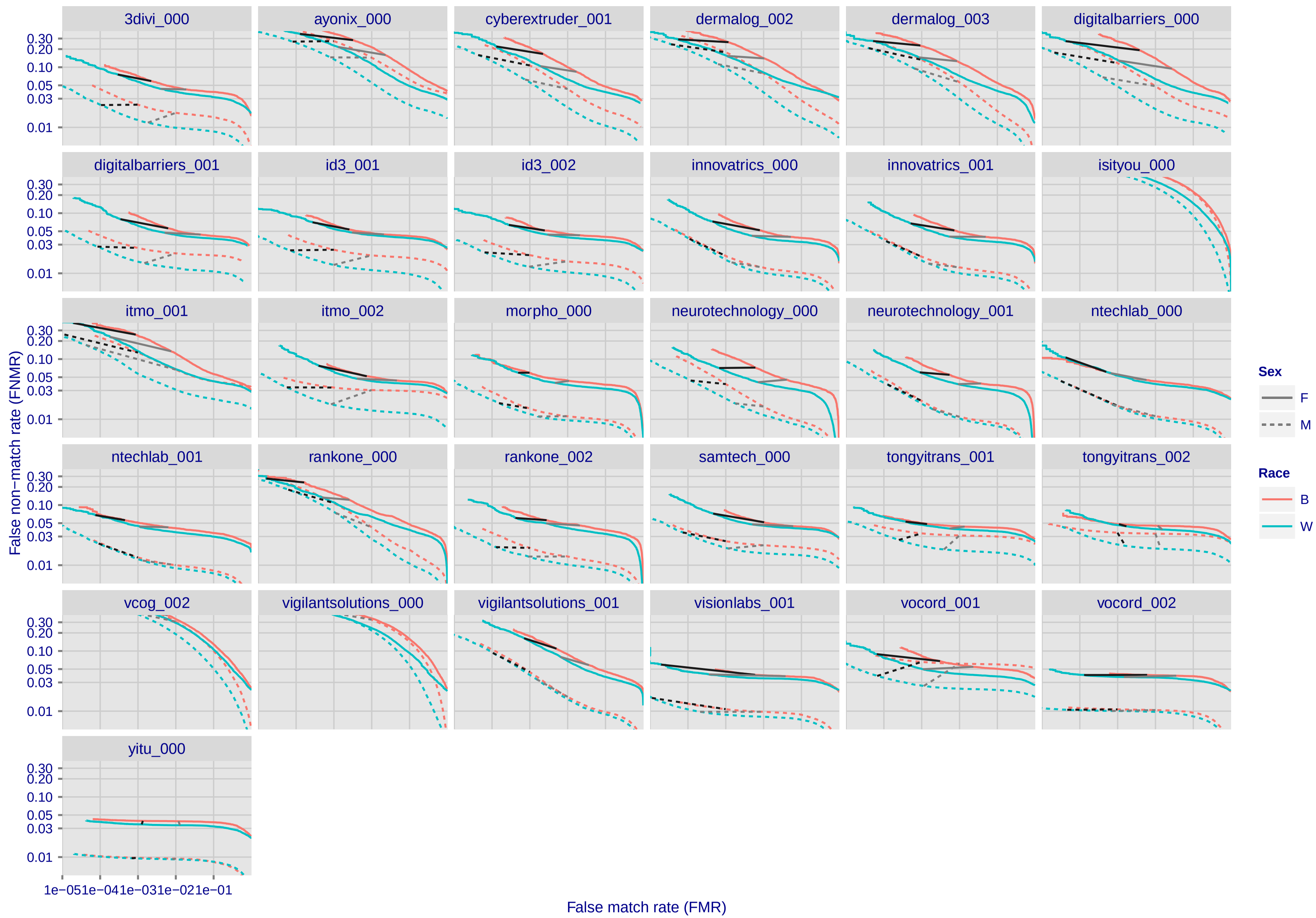


Figure 12: For the mugshot images, error tradeoff characteristics for white females, black females, black males and white males. The grey lines correspond to fixed thresholds, showing how both *FNMR and FMR vary at one operating threshold. Important: Many of the plots will naively be read as saying whites gives lower error rates than blacks because the blue traces lie beneath the red ones. However, this is misleading and incomplete: The grey lines show the traces are generally shifted horizontally. Thus for the dermalog-001 algorithm FNMR for whites is* higher than for blacks *at a fixed threshold* but, *at the same time, FMR is higher for blacks - see Figure 17. As access control systems almost always operate at a fixed threshold, the naive interpretation is incorrect.*

Figure 13: For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between $(\text{FMR},\text{FNMR})_{\text{MALE}}$ and $(\text{FMR},\text{FNMR})_{\text{FEMALE}}$ showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds to female. The six operating thresholds are selected to give the nominal false match rates given in the legend, and are computed over all impostor pairs regardless of age, sex, and place of birth. The plotted FMR values are broadly an order of magnitude larger than the nominal rates because FMR is computed over demographically-matched impostor pairs i.e individuals of the same sex, from the same geographic region (see section 4.6.1), and the same age group (see section 4.6.2).

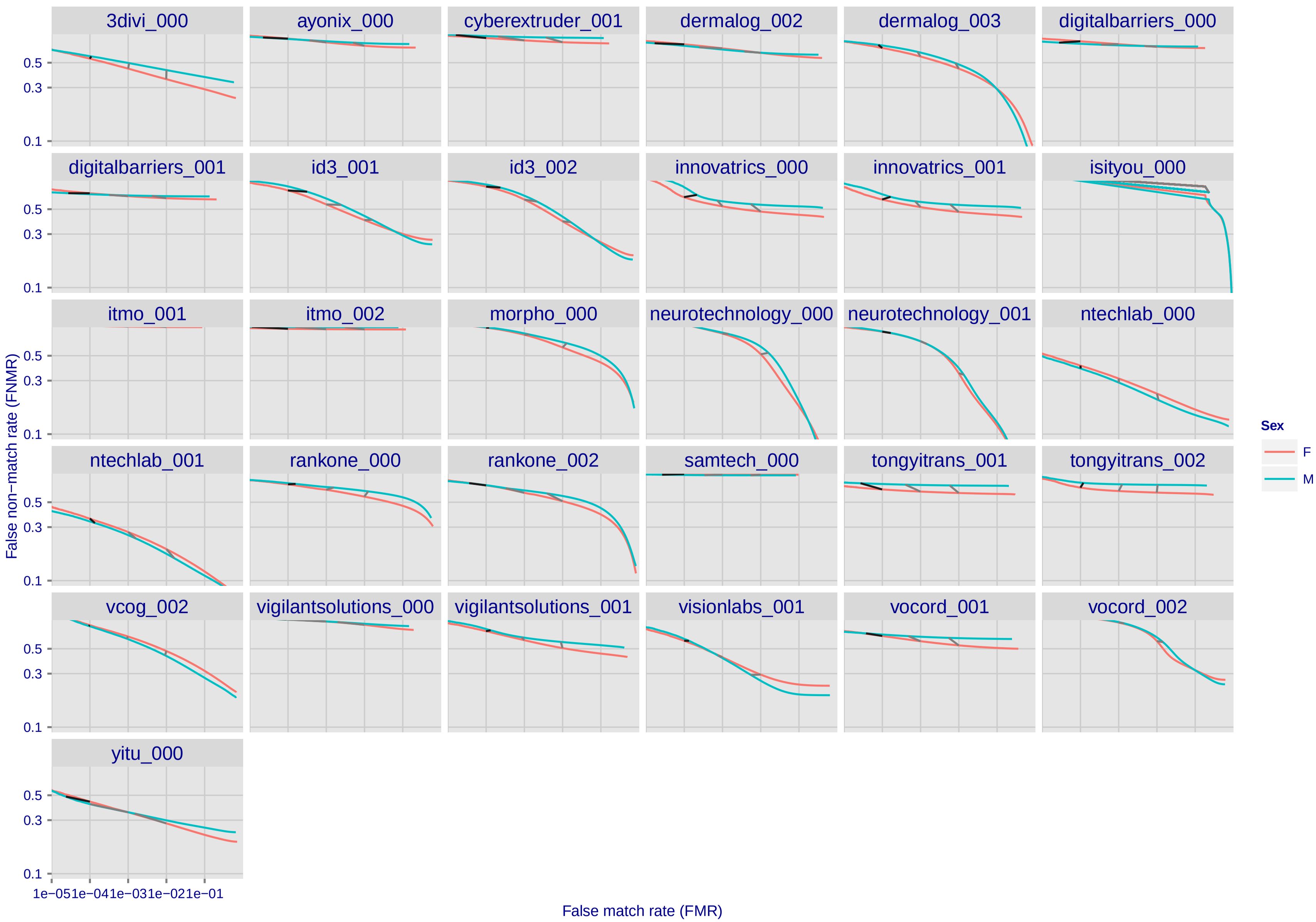


Figure 14: For the wild image comparisons, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. Error rates are higher here than in the generic wild DET (Fig 9) becuase the impostor pairs here are same-sex only. The scales are logarithmic in order to show several decades of FMR.

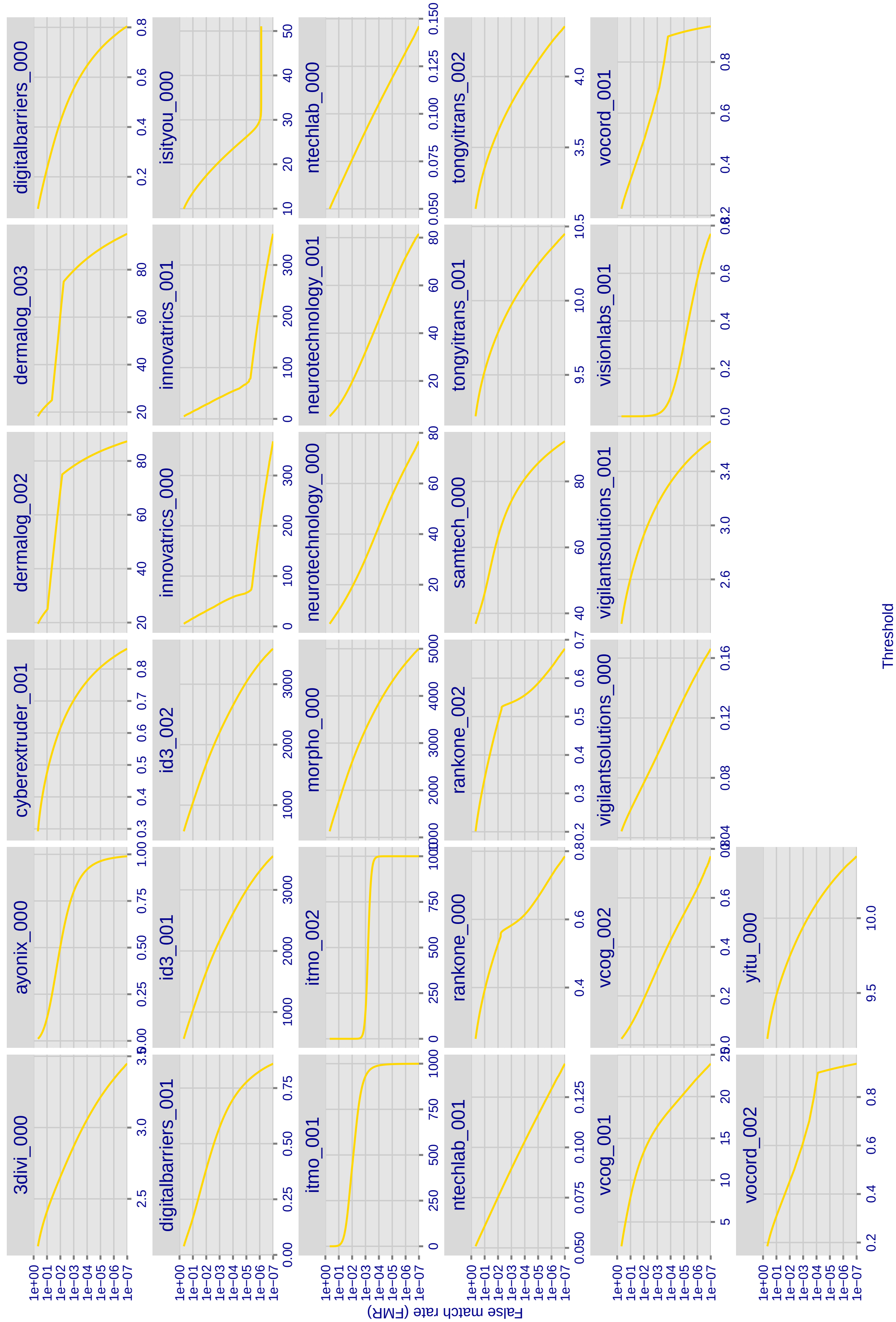


Figure 15: *For the visa images, the false match calibration curves show false match rate vs. threshold. These curves apply to zero-effort impostors. As shown later (sec. 4.6), FMR is higher for demographic-matched impostors.*

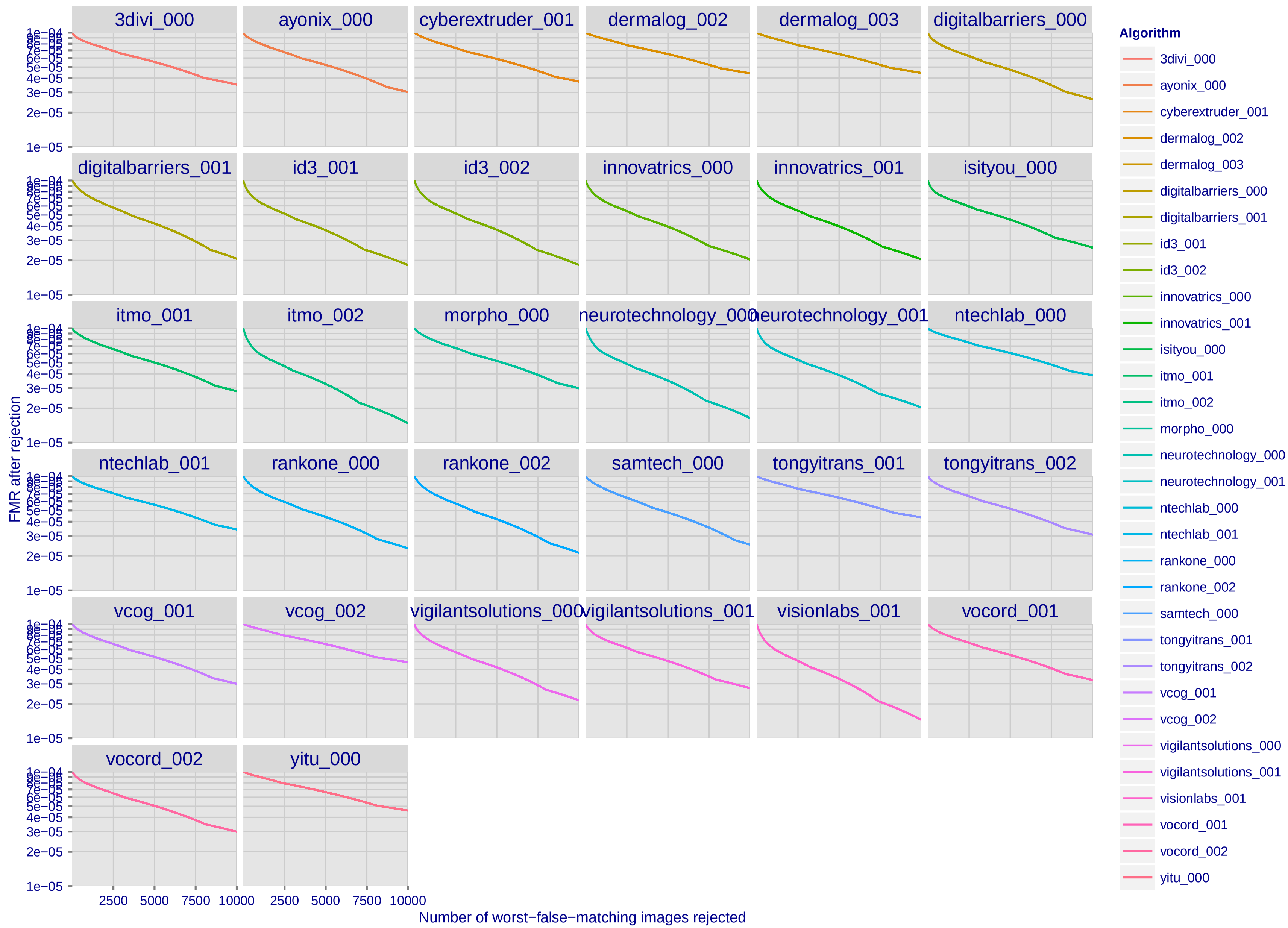


Figure 16: For the visa images, the curves show how false matches are concentrated in certain images. Specifically each line plots FMR(k) with k the number of images rejected in decreasing order of how many false matches that image was involved in. FMR$(0) = 10^{-4}$. In terms of the biometric zoo, the most "wolf-ish" images are rejected first i.e. those enrollment or verification images involved in false matches. A flat response is considered superior. A steeply descending response indicates that certain kinds of images false match against others. A hypothetical example would be if images of men with particular mustaches falsely match others.

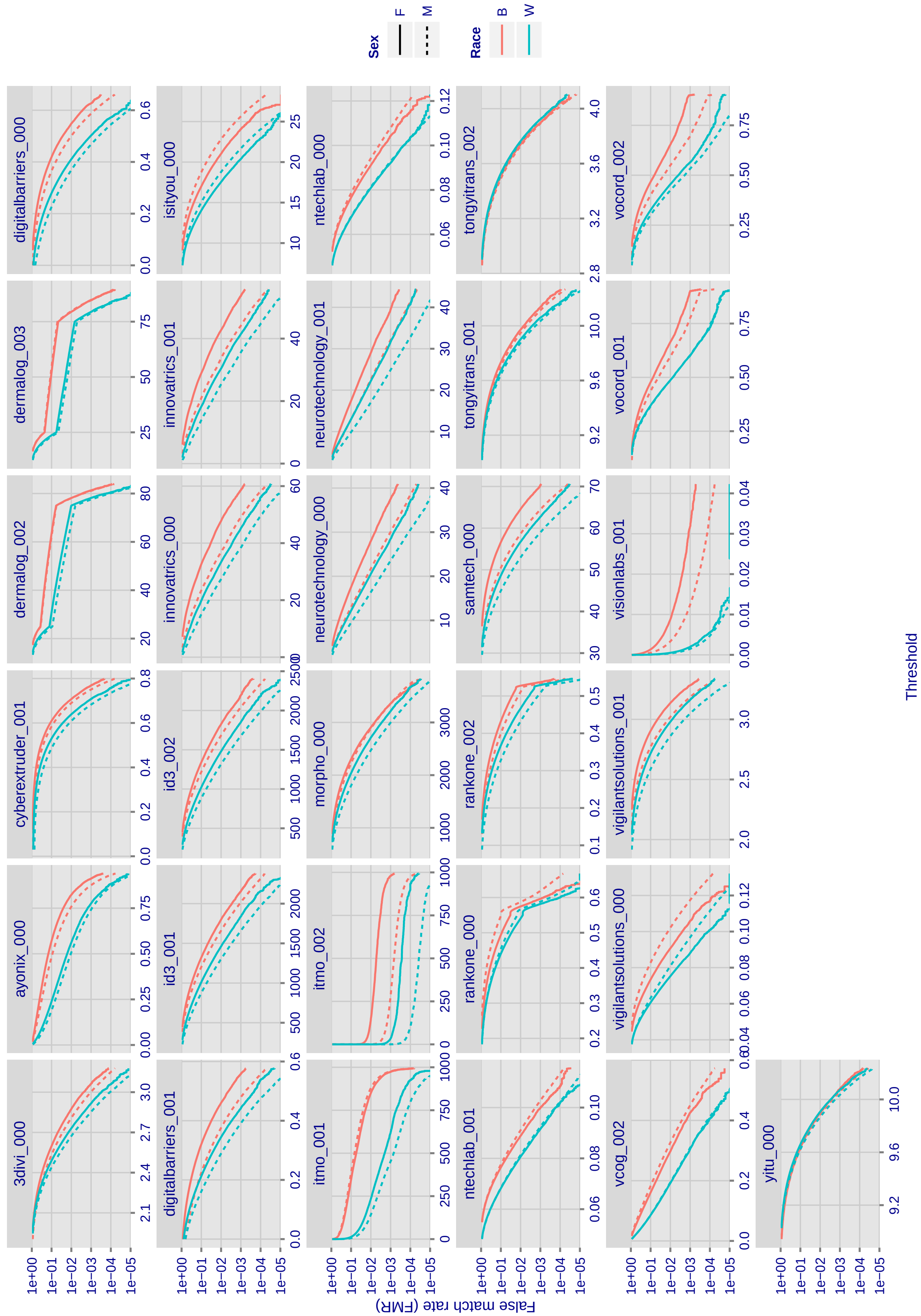


Figure 17: For the mugshot images, the false match calibration curves show false match rate vs. threshold. Separate curves appear for white females, black females, black males and white males.

4.5 Genuine distribution stability

4.5.1 Effect of birth place on the genuine distribution

Background: Both skin tone and bone structure vary geographically. Prior studies have reported variations in FNMR and FMR.

Goal: To measure false non-match rate (FNMR) variation with country of birth.

Methods: Thresholds are determined that give FMR = $\{0.001, 0.0001\}$ over the entire impostor set. Then FNMR is measured over 1000 bootstrap replications of the genuine scores. Only those countries with at least 140 individuals are included in the analysis.

Results: Figure 18 shows FNMR by country of birth for the two thresholds.

Caveats: The results may not relate to subject-specific properties. Instead they could reflect image-specific quality differences, which could occur due to collection protocol or software processing variations.

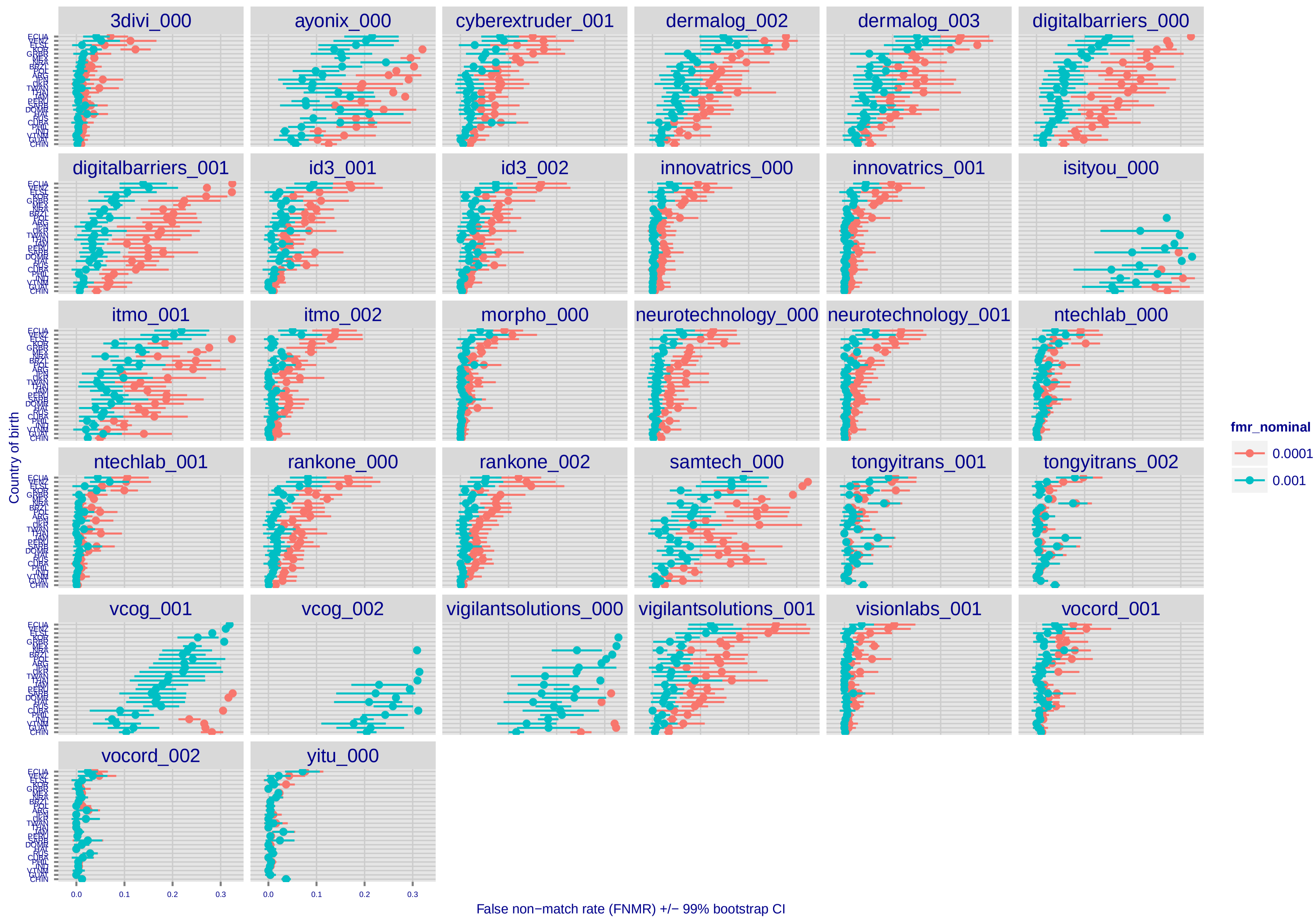


Figure 18: For the visa images, the dots show FNMR by country of birth for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all $O(10^{10})$ impostor scores. The figures shows an order of magnitude variation in FNMR across country of birth; these effects are due to quality variations. The least accurate countries vary by algorithm.

4.5.2 Effect of age on genuine subjects

Background: Faces change appearance throughout life. Face recognition algorithms have previously beeen reported to give better accuracy on older individuals (See NIST IR 8009).

Goal: To quantify false non-match rates (FNMR) as a function of age. We do not aim to quantify ageing effects here as the separation between two samples is limited to just a few years.

Methods: Using the visa images, thresholds are determined that give FMR = 0.001 and 0.0001 over the entire impostor set. Then FNMR is measured over 1000 bootstrap replications of the genuine scores.

Results: For the visa images, Figure 19 shows how false non-match rates for genuine users, as a function of age group.

The notable aspects are:

- ▷ Younger subjects give considerably higher FNMR. This is likely due to rapid growth and change in facial appearance.
- ▷ FNMR trends down throughout life. The last bin, AGE $>$ 72, contains fewer than 140 mated pairs, and may be affected by small sample size.

Caveats: None.

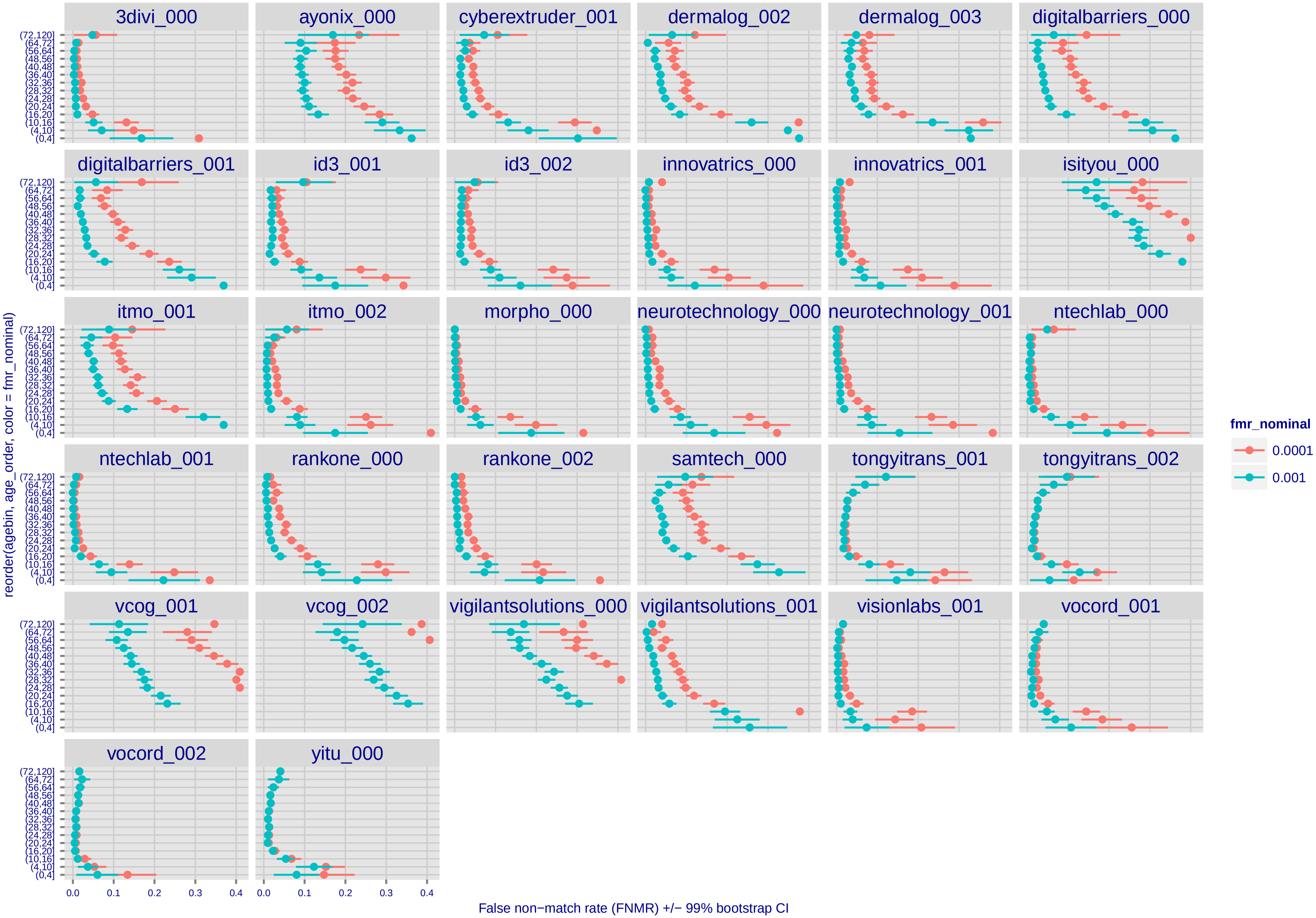


Figure 19: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all $O(10^{10})$ impostor scores. Given a pair of face images taken at different times, we assign a false non-match to the bin that is the arithmetic average of the subject's ages. This plot shows only the effect of age, not ageing. The number of comparisons in each bin is generally in the thousands. However the FNMR for the first and last bins are each computed over fewer than 150 comparisons.

4.6 Impostor distribution stability

4.6.1 Effect of birth place on the impostor distribution

Background: Facial appearance varies geographically, both in terms of skin tone, cranio-facial structure and size. This section addresses whether false match rates vary intra- and inter-regionally.

Goals:

▷ To show the effect of birth region of the impostor and enrollee on false match rates.

▷ To determine whether some algorithms give better impostor distribution stability.

Methods:

▷ For the visa images, NIST defined 10 regions: Sub-Saharan Africa, South Asia, Polynesia, North Africa, Middle East, Europe, East Asia, Central and South America, Central Asia, and the Caribbean.

▷ For the visa images, NIST mapped each country of birth to a region. There is some arbitrariness to this. For example, Egypt could reasonably be assigned to the Middle East instead of North Africa. An alternative methodology could, for example, assign the Philippines to *both* Polynesia and East Asia.

▷ FMR is computed for cases where all face images of impostors born in region r_2 are compared with enrolled face images of persons born in region r_1.

$$\text{FMR}(r_1, r_2, T) = \frac{\sum_{i=1}^{N_{r_1,r_2}} H(s_i - T)}{N_{r_1,r_2}} \quad (5)$$

where the same threshold, T, is used in all cells, and H is the unit step function. The threshold is set to give FMR(T) = 0.001 over the entire set of visa image impostor comparisons.

▷ This analysis is then repeated by country-pair, but only for those country pairs where both have at least 1000 images available. The countries[1] appear in the axes of graphs that follow.

▷ The mean number of impostor scores in any cross-region bin is 33 million. The smallest number of impostor scores in any bin is 135000, for Central Asia - North Africa. While these counts are large enough to support reasonable significance, the number of individual faces is much smaller, O($N^{0.5}$).

▷ The numbers of impostor scores in any cross-country bin is shown in Figure 91.

Results: Subsequent figures show heatmaps that use color to represent the base-10 logarithm of the false match rate. Red colors indicate high (bad) false match rates. Dark colors indicate benign false match rates. There are two series of graphs corresponding to aggregated geographical regions, and to countries. The notable observations are:

▷ The on-diagonal elements correspond to within-region impostors. FMR is generally above the nominal value of FMR = 0.001. Particularly there is usually higher FMR in, Sub-Saharan Africa, South Asia, and the Caribbean. Europe and Central Asia, on the other hand, usually give FMR closer to the nominal value.

▷ The off-diagonal elements correspond to across-region impostors. The highest FMR is produced between the Caribbean and Sub-Saharan Africa.

▷ Algorithms vary.

[1] These are Argentina, Australia, Brazil, Chile, China, Costa Rica, Cuba, Czech Republic, Dominican Republic, Ecuador, Egypt, El Salvador, Germany, Ghana, Great Britain, Greece, Guatamala, Haiti, Hong Kong, Honduras, Indonesia, India, Israel, Jamaica, Japan, Kenya, Korea, Lebanon, Mexico, Malaysia, Nepal, Nigeria, Peru, Philippines, Pakistan, Poland, Romania, Russia, South Africa, Saudi Arabia, Thailand, Trinidad, Turkey, Taiwan, Ukraine, Venezuela, and Vietnam.

▷ We computed the same quantities for a global FMR = 0.0001. The effects are similar.

Caveats:

▷ The effects of variable impostor rates on one-to-many identification systems may well differ from what's implied by these one-to-one verification results. Two reasons for this are a) the enrollment galleries are usually imbalanced across countries of birth, age and sex; b) one-to-many identification algorithms often implement techniques aimed at stabilizing the impostor distribution. Further research is necessary.

▷ In principle, the effects seen in this subsection could be due to differences in the image capture process. We consider this unlikely since the effects are maintained across geography - e.g. Caribbean vs. Africa, or Japan vs. China.

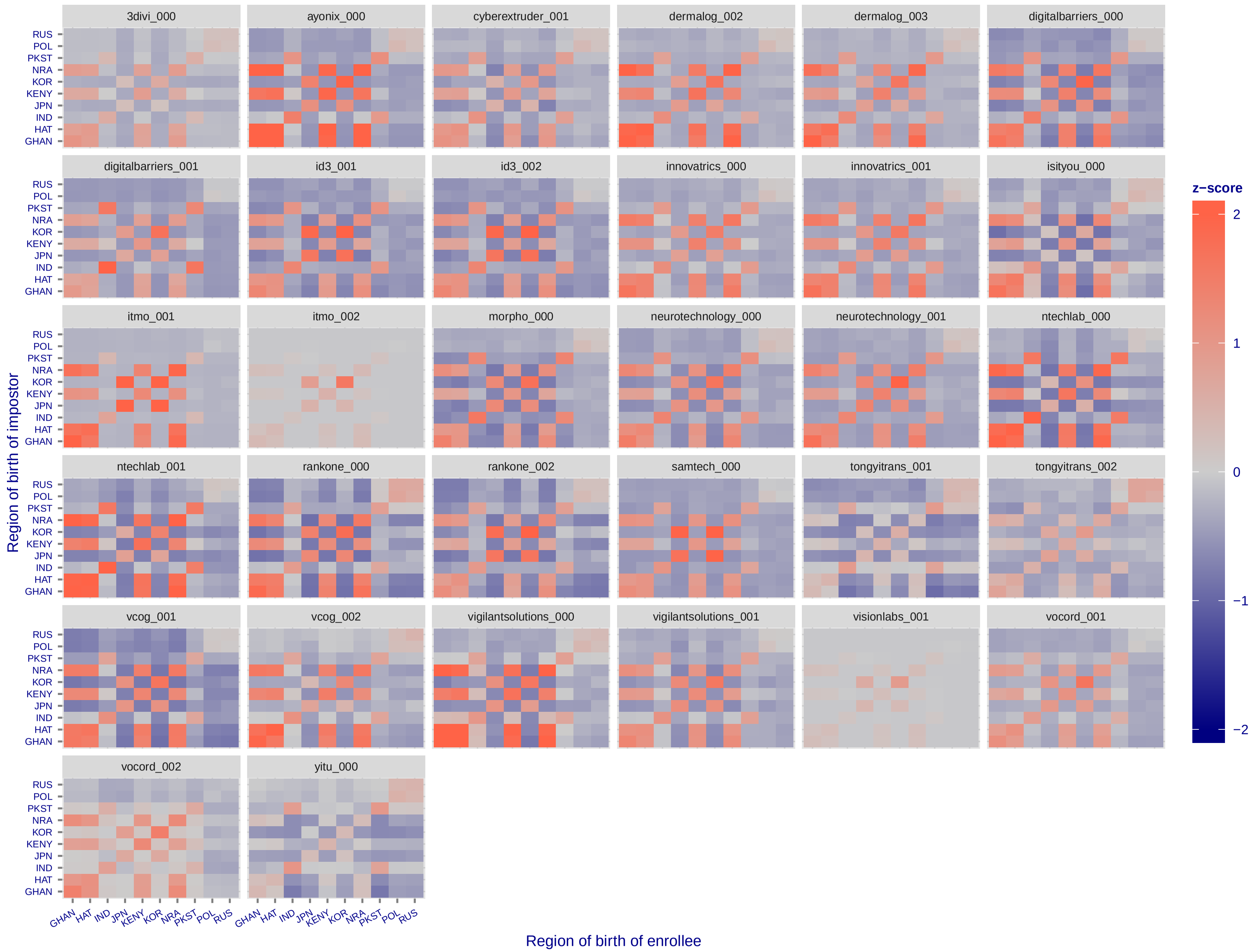


Figure 20: For visa images, the heatmap shows how the mean of the impostor distribution for the country pair (a,b) is shifted relative to the mean of the global impostor distribution, expressed as a number of standard deviations of the global impostor distribution. This statistic is designed to show shifts in the entire impostor distribution, not just tail effects that manifest as the anomalously high (or low) false match rates that appear in the subsequent figures. The countries are chosen to show that skin tone alone does not explain impostor distribution shifts. The reduced shift in Asian populations with the Yitu and TongYiTrans algorithms, is accompanied by positive shifts in the European populations. This reversal relative to most other algorithms, may derive from use of nationally weighted training sets. The Visionlabs algorithm appears most insensitive to country effects. The figure is computed from same-sex and same-age impostor pairs.

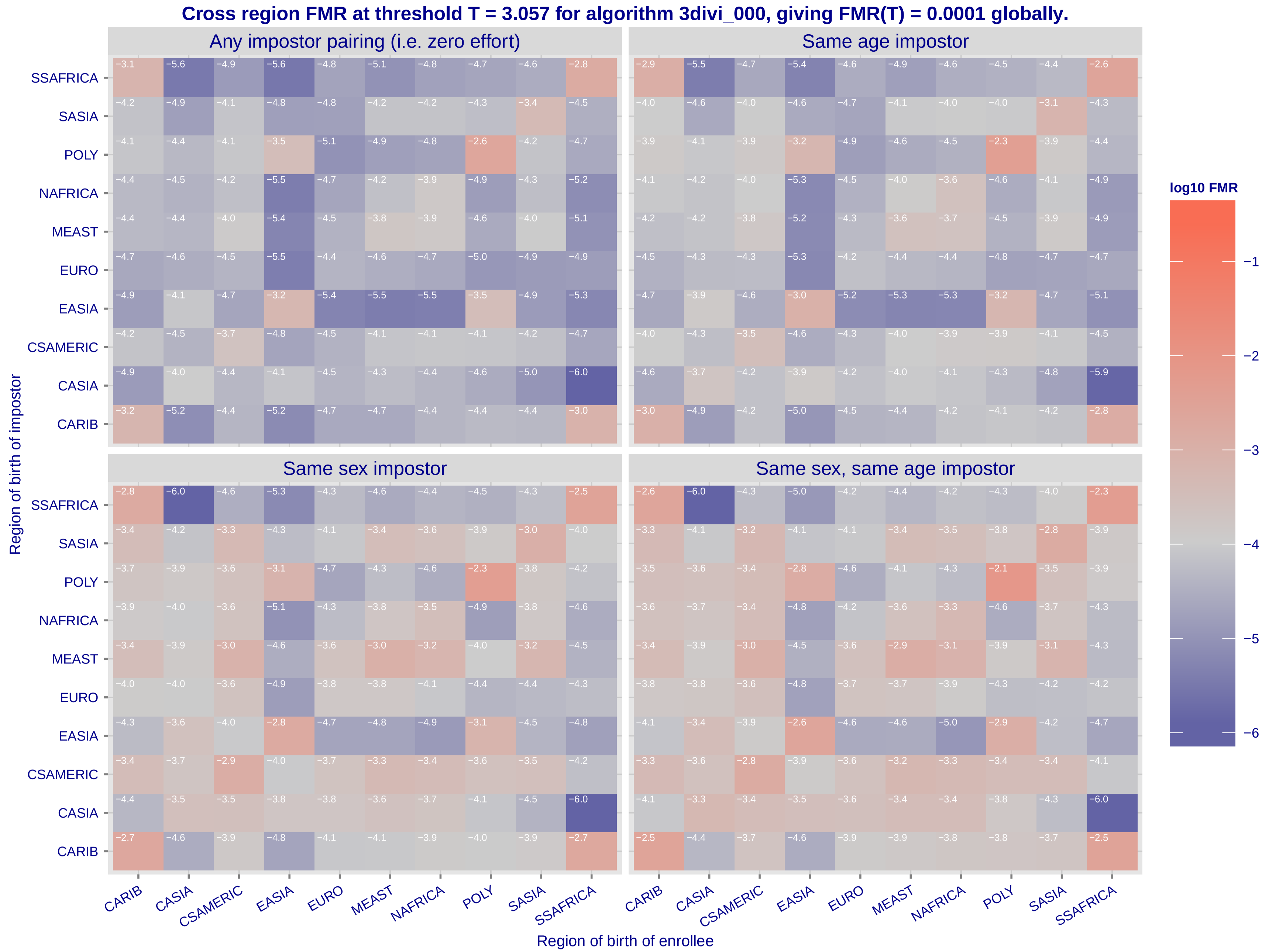


Figure 21: For algorithm 3divi-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

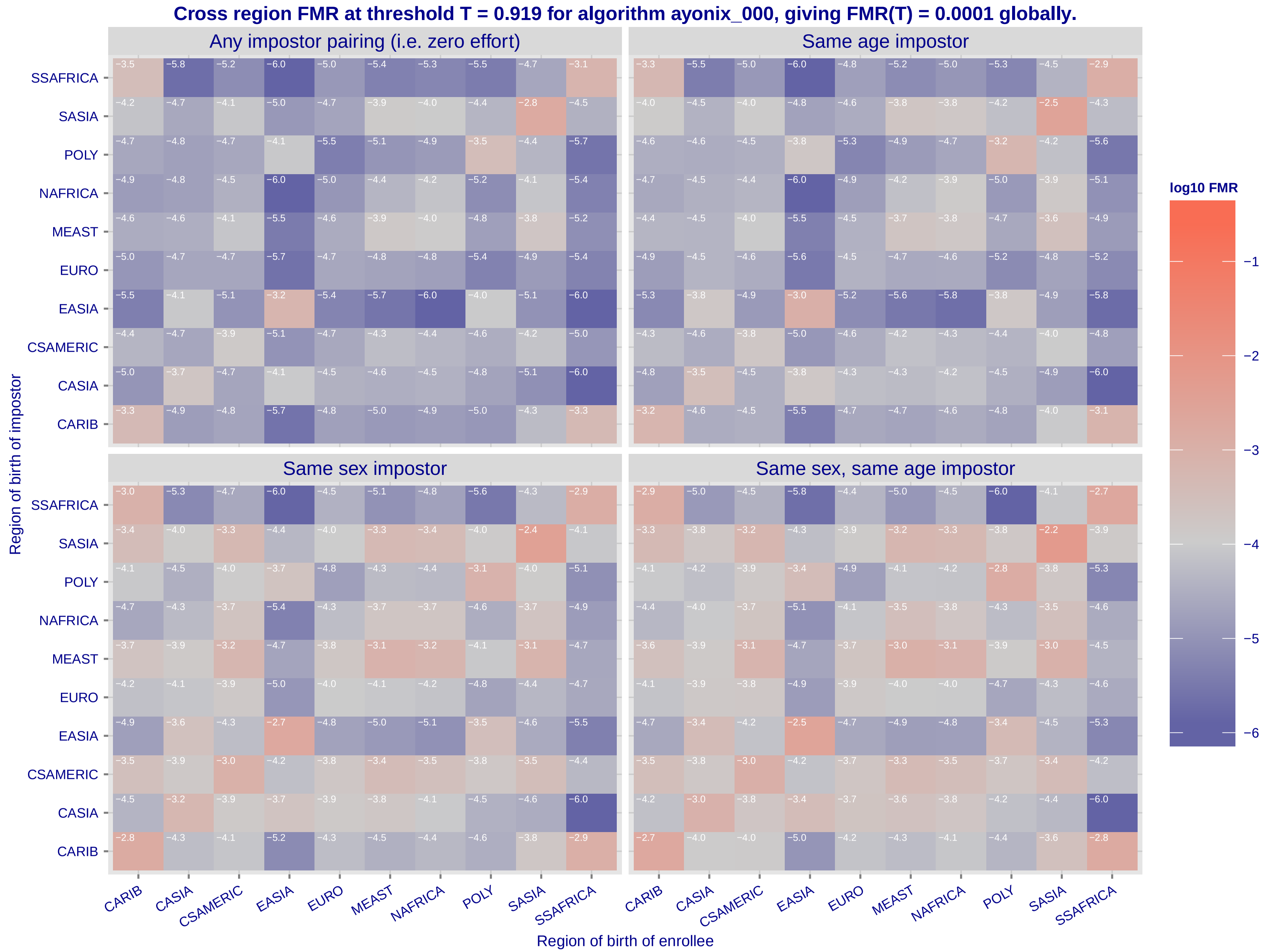


Figure 22: For algorithm ayonix-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

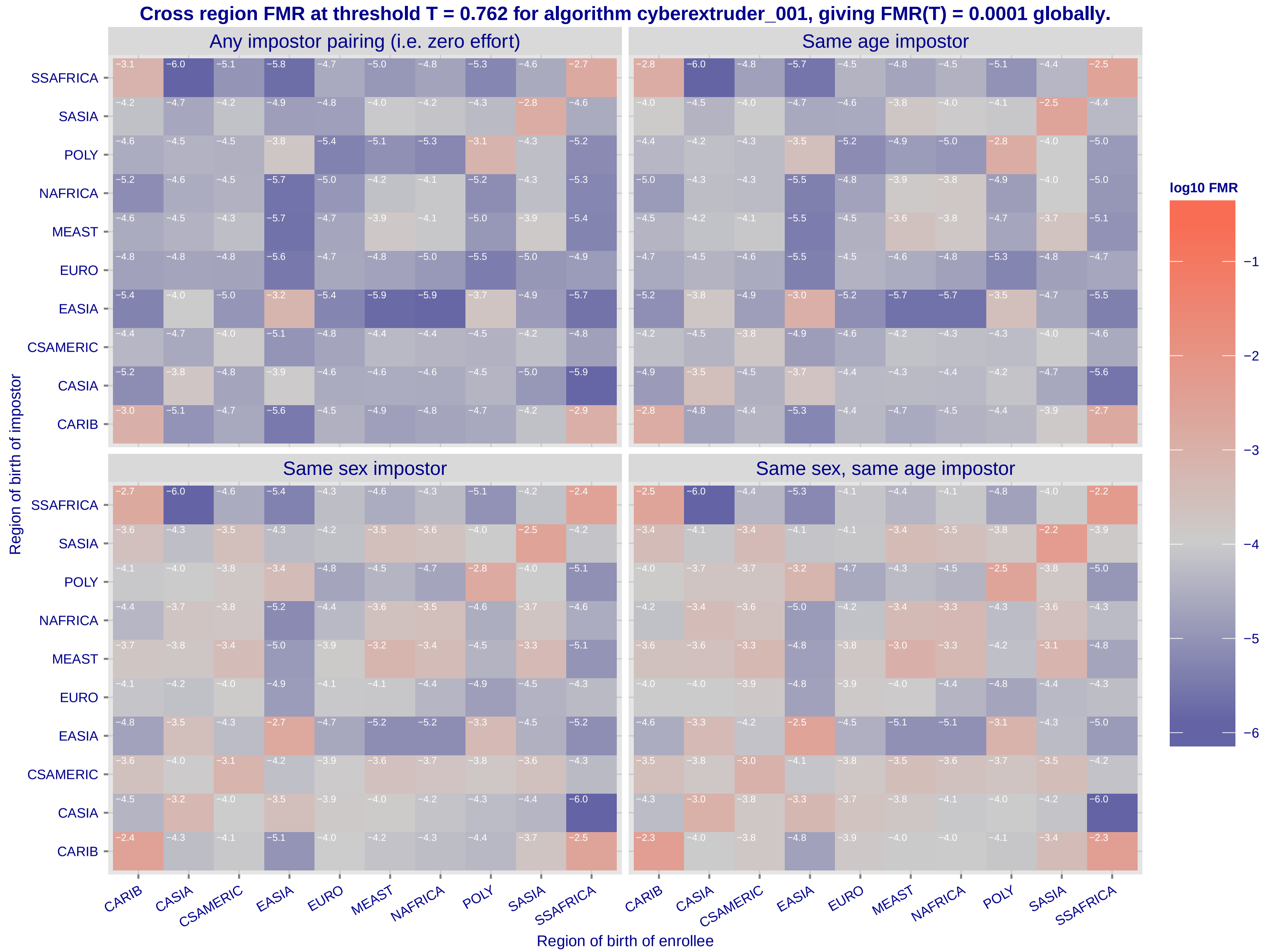


Figure 23: For algorithm cyberextruder-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

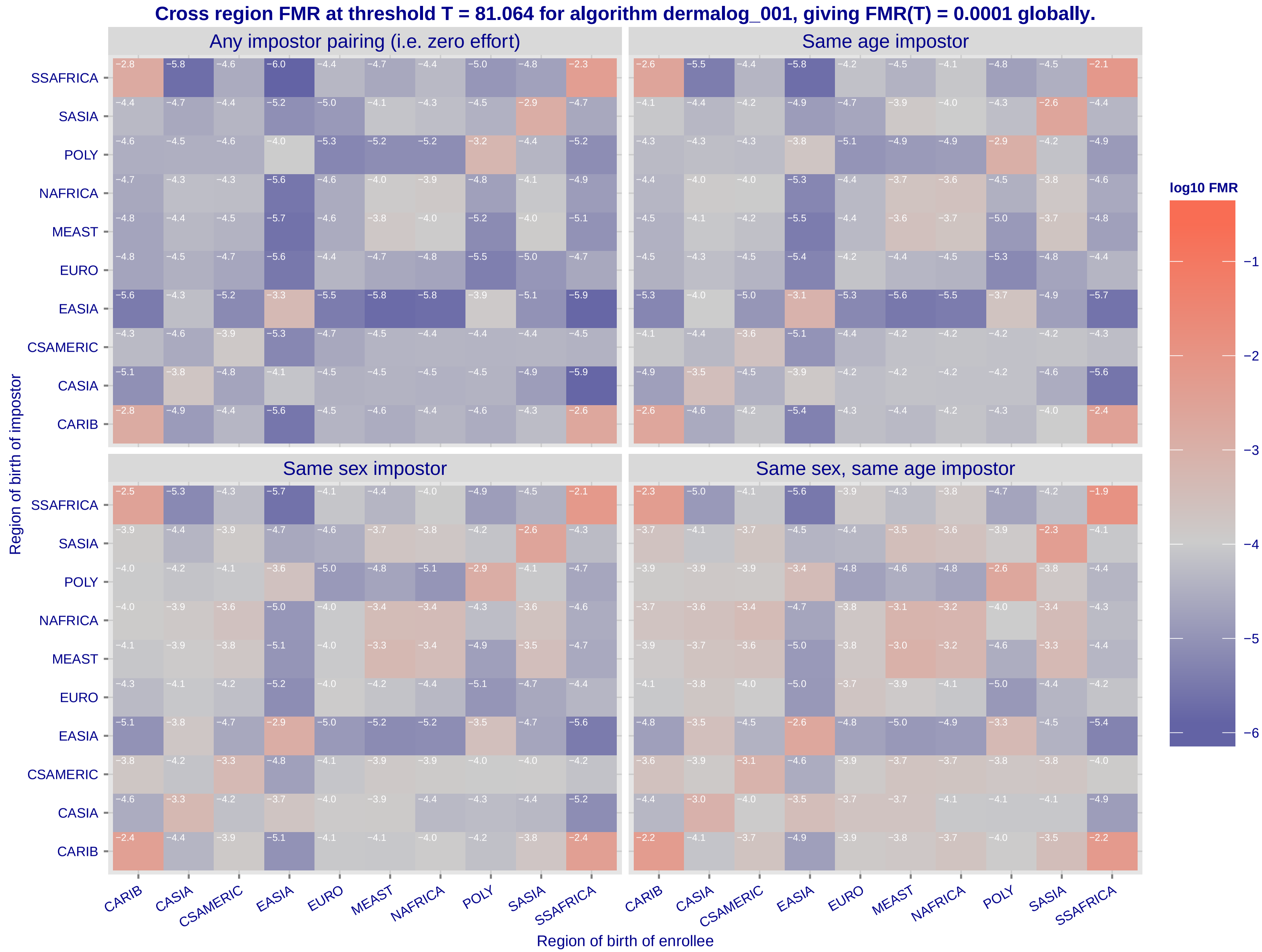


Figure 24: For algorithm dermalog-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

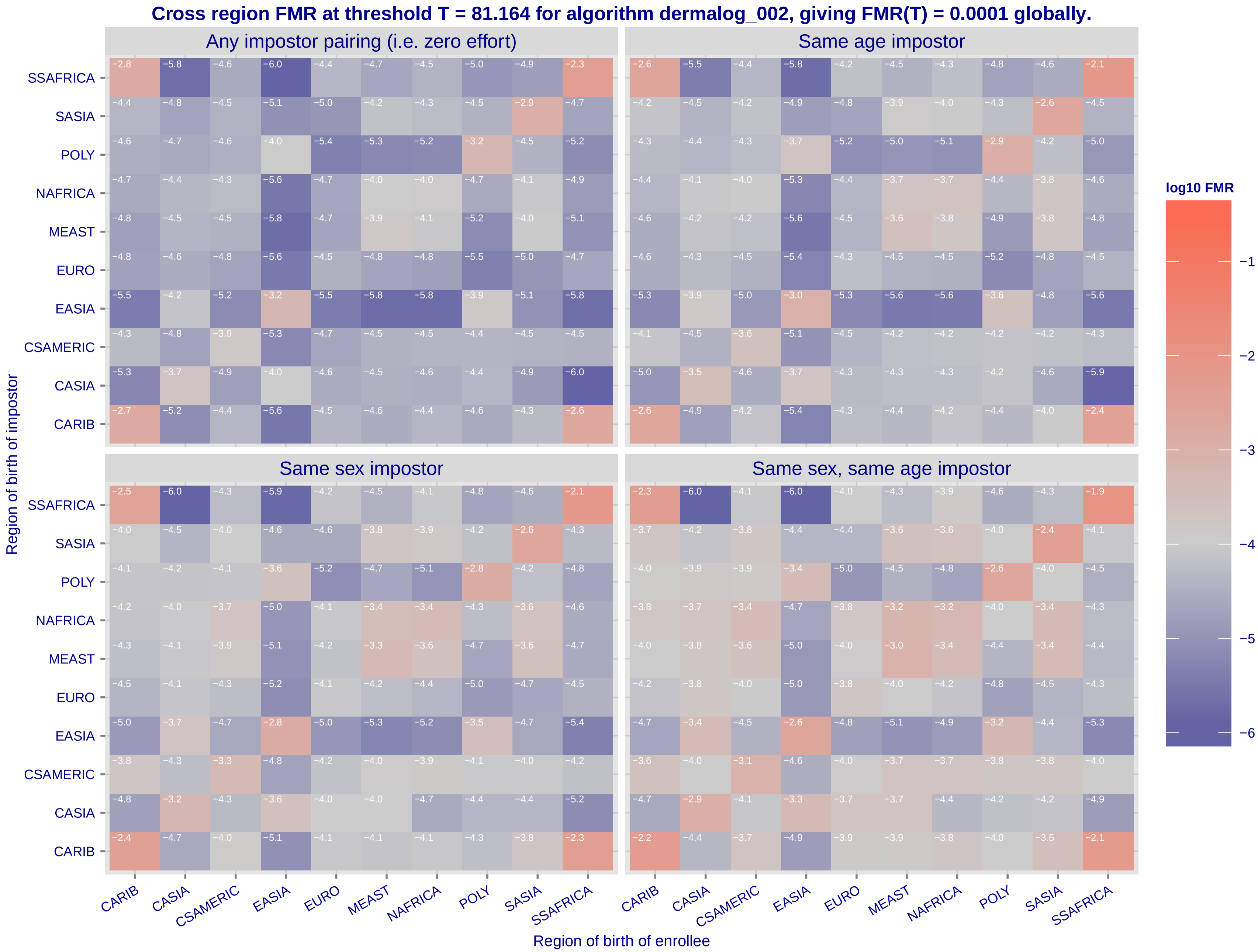


Figure 25: For algorithm dermalog-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

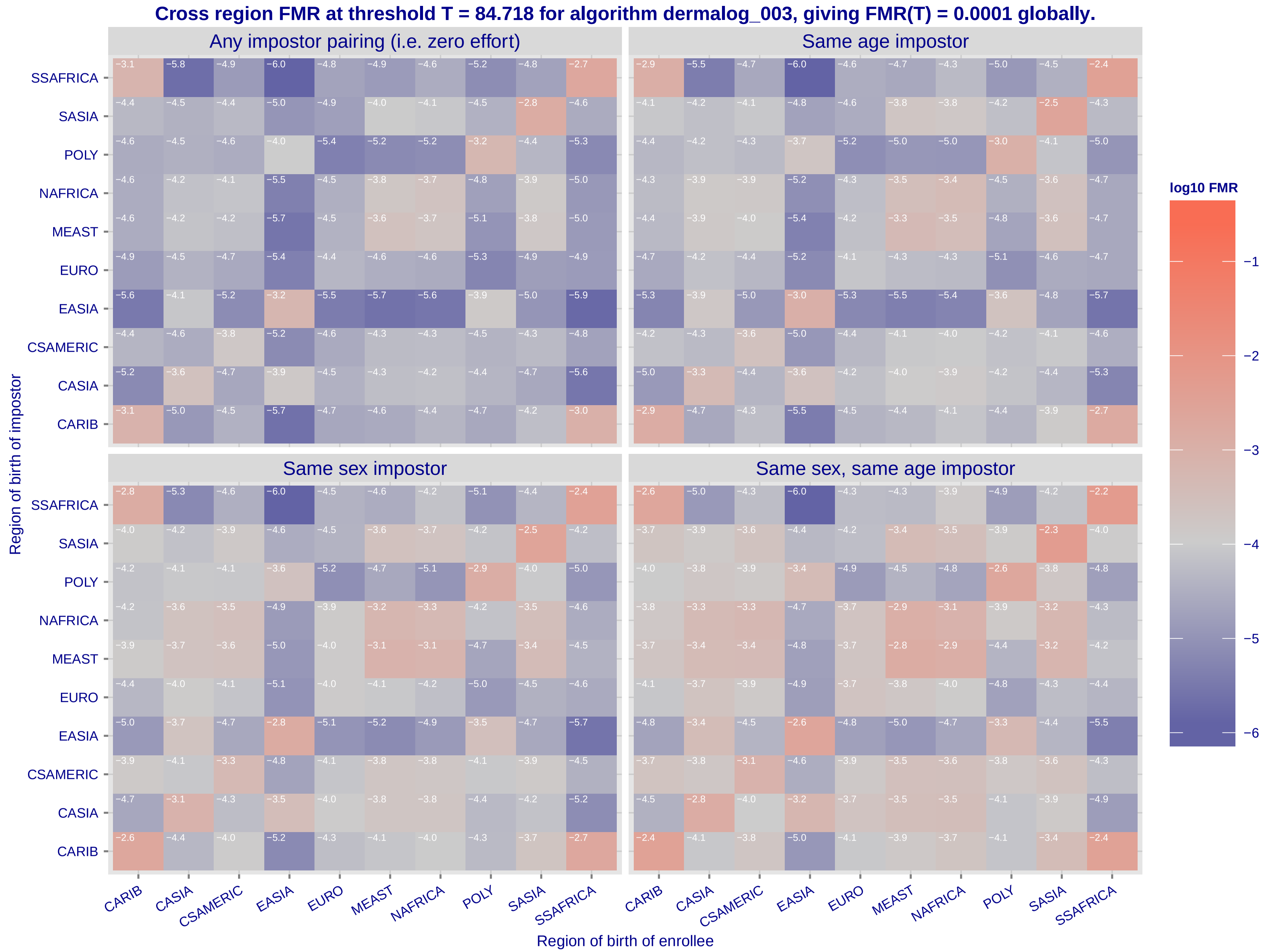


Figure 26: For algorithm dermalog-003 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

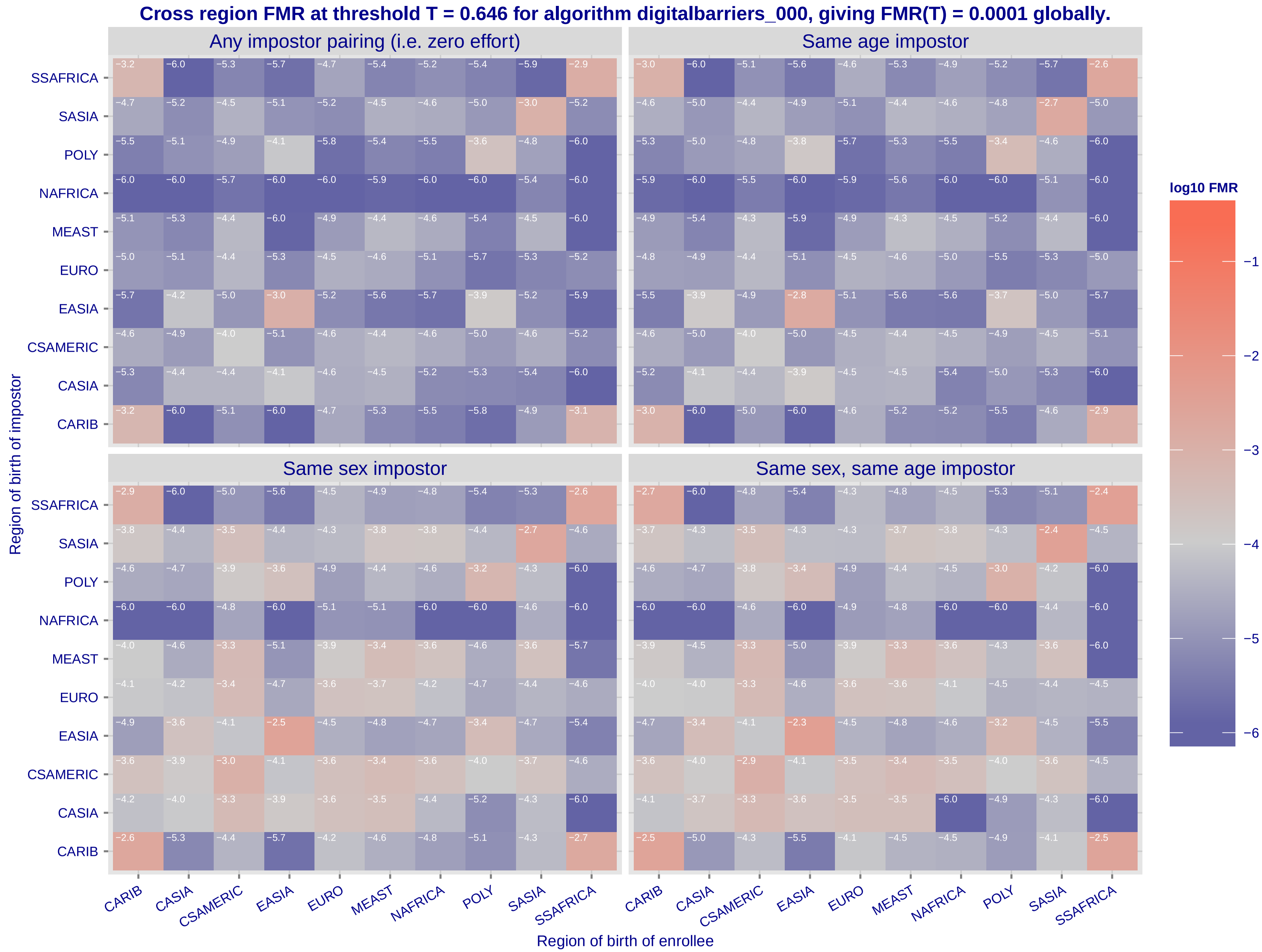


Figure 27: For algorithm digitalbarriers-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

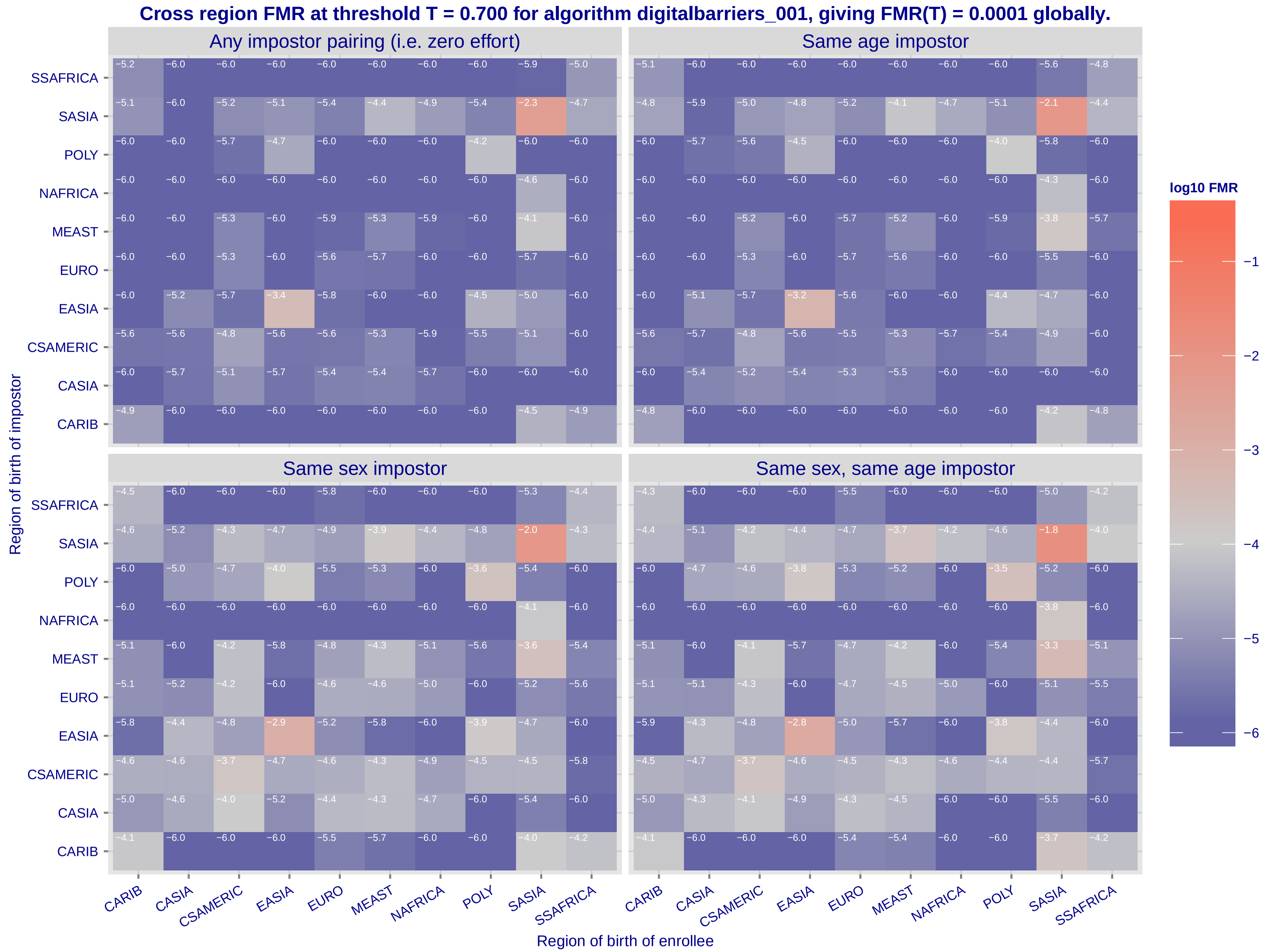


Figure 28: *For algorithm digitalbarriers-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.*

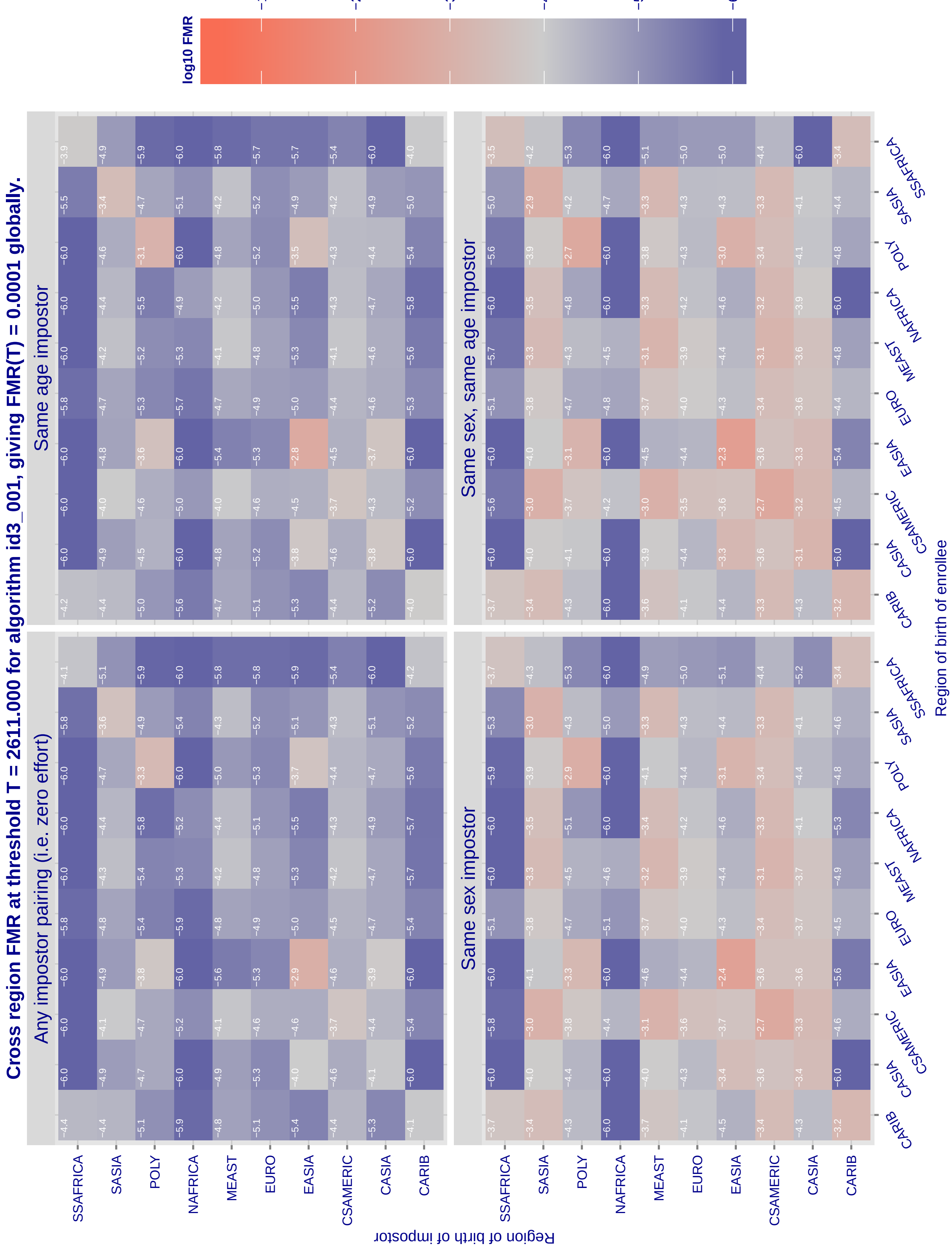


Figure 29: For algorithm id3-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all O(10^{10}) impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

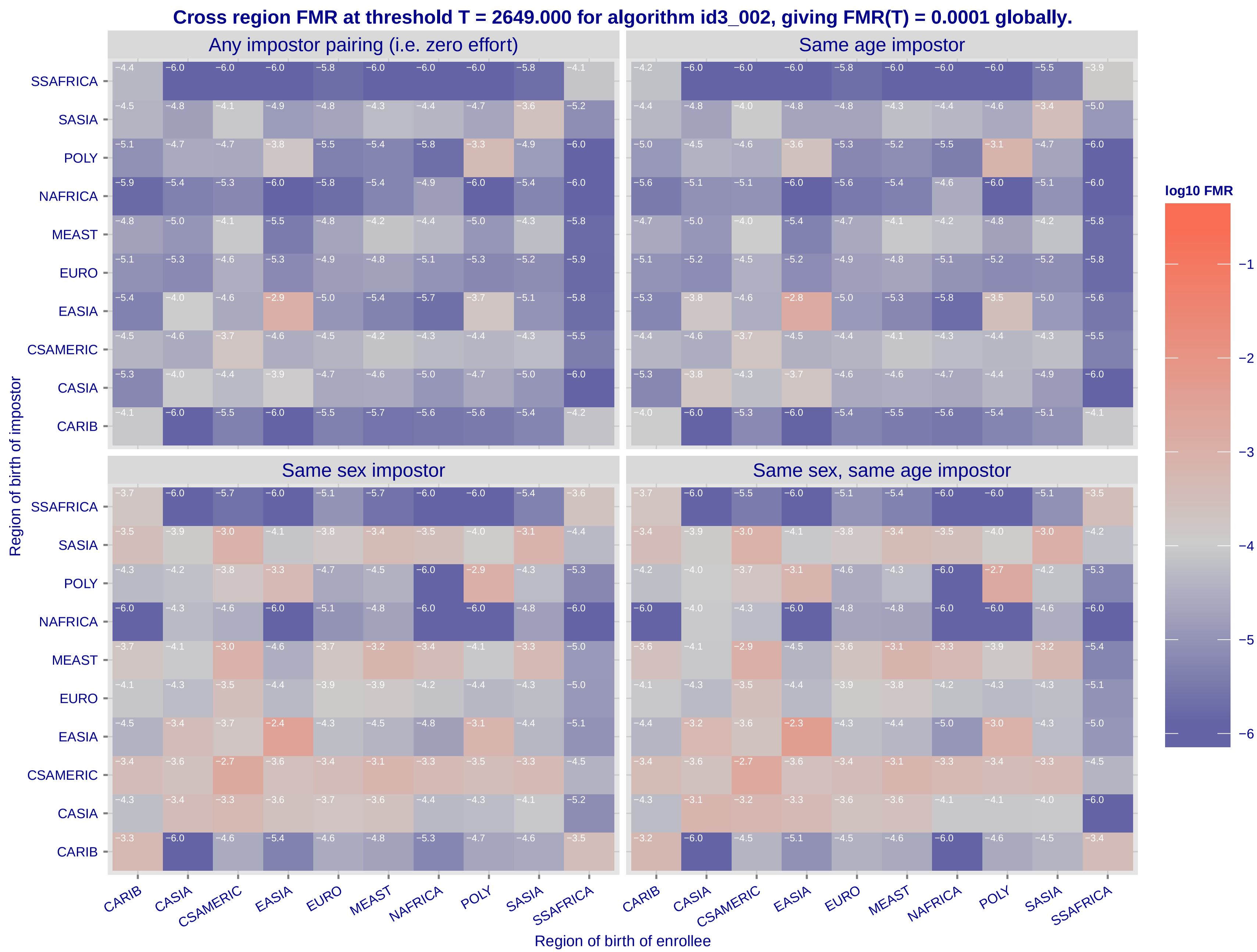


Figure 30: For algorithm id3-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

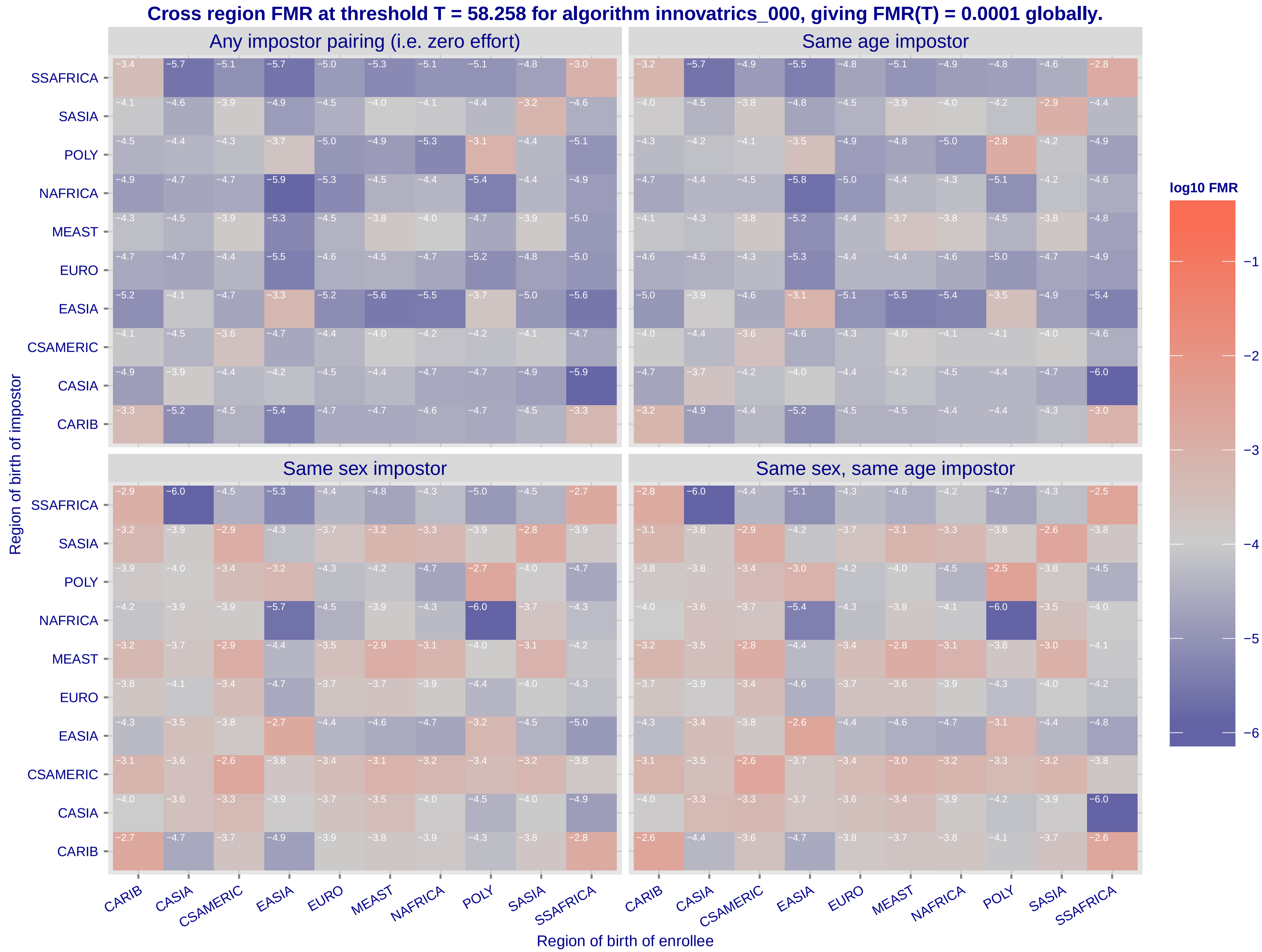


Figure 31: For algorithm innovatrics-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

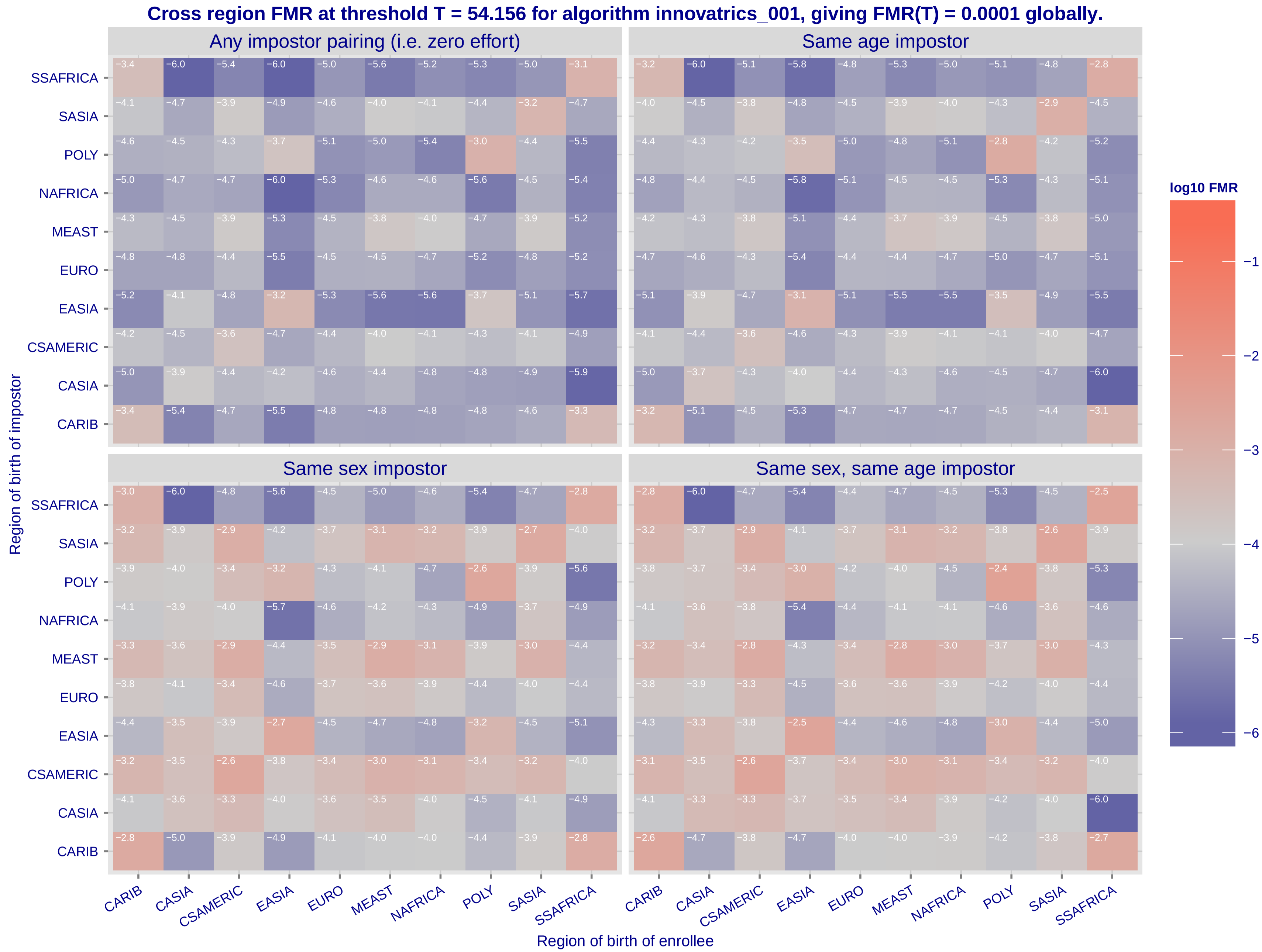


Figure 32: For algorithm innovatrics-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

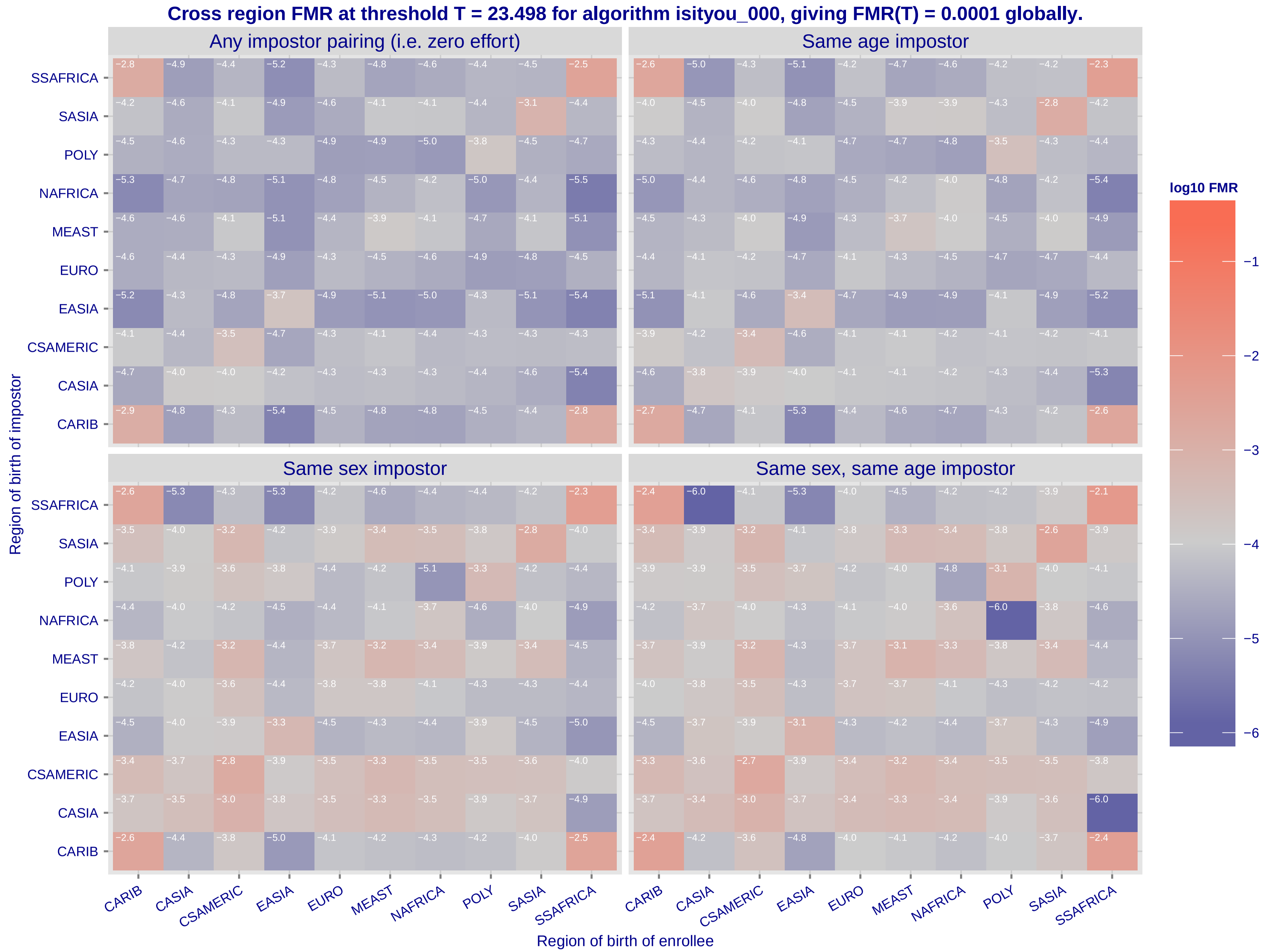


Figure 33: For algorithm isityou-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

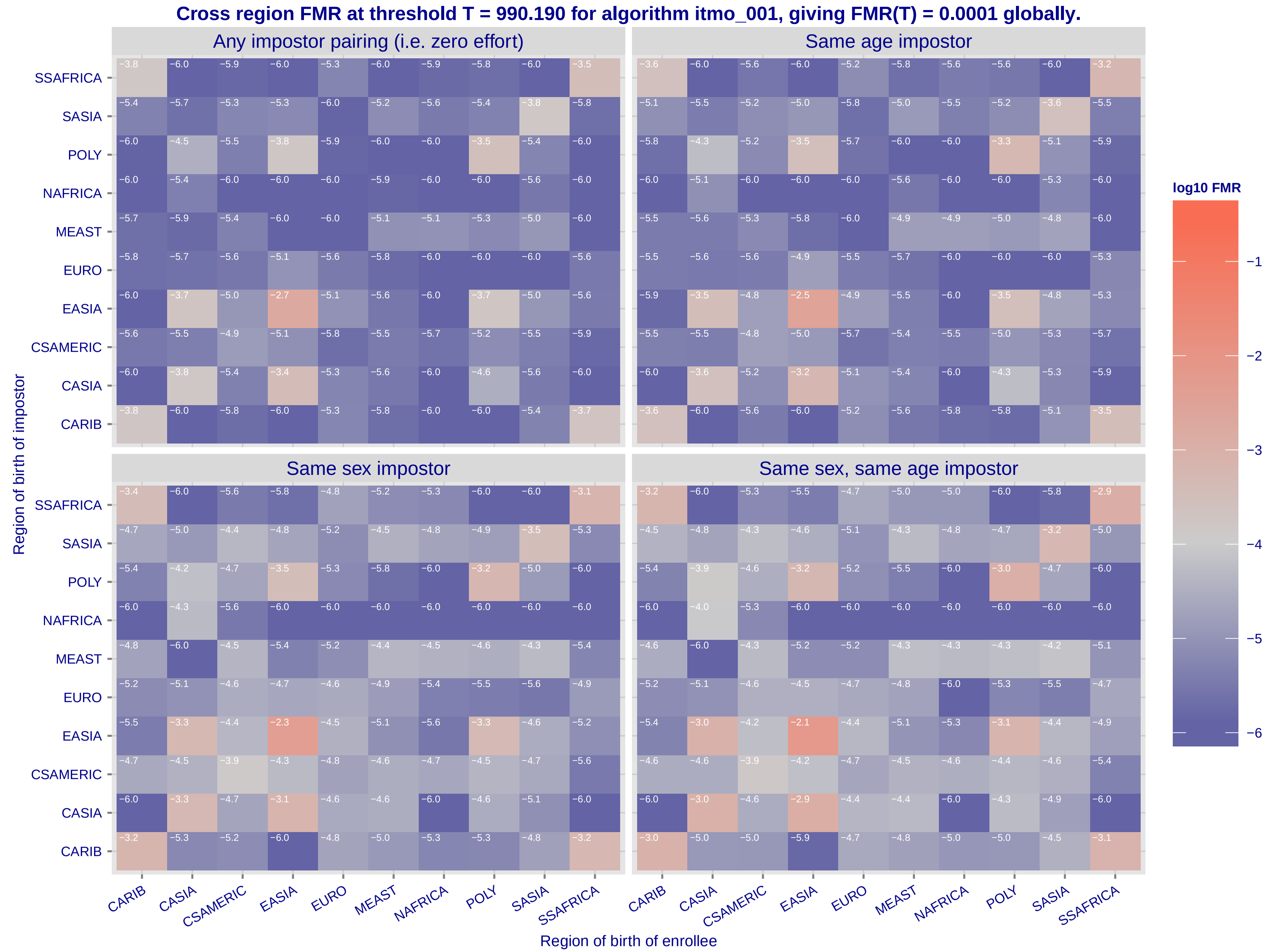


Figure 34: For algorithm itmo-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

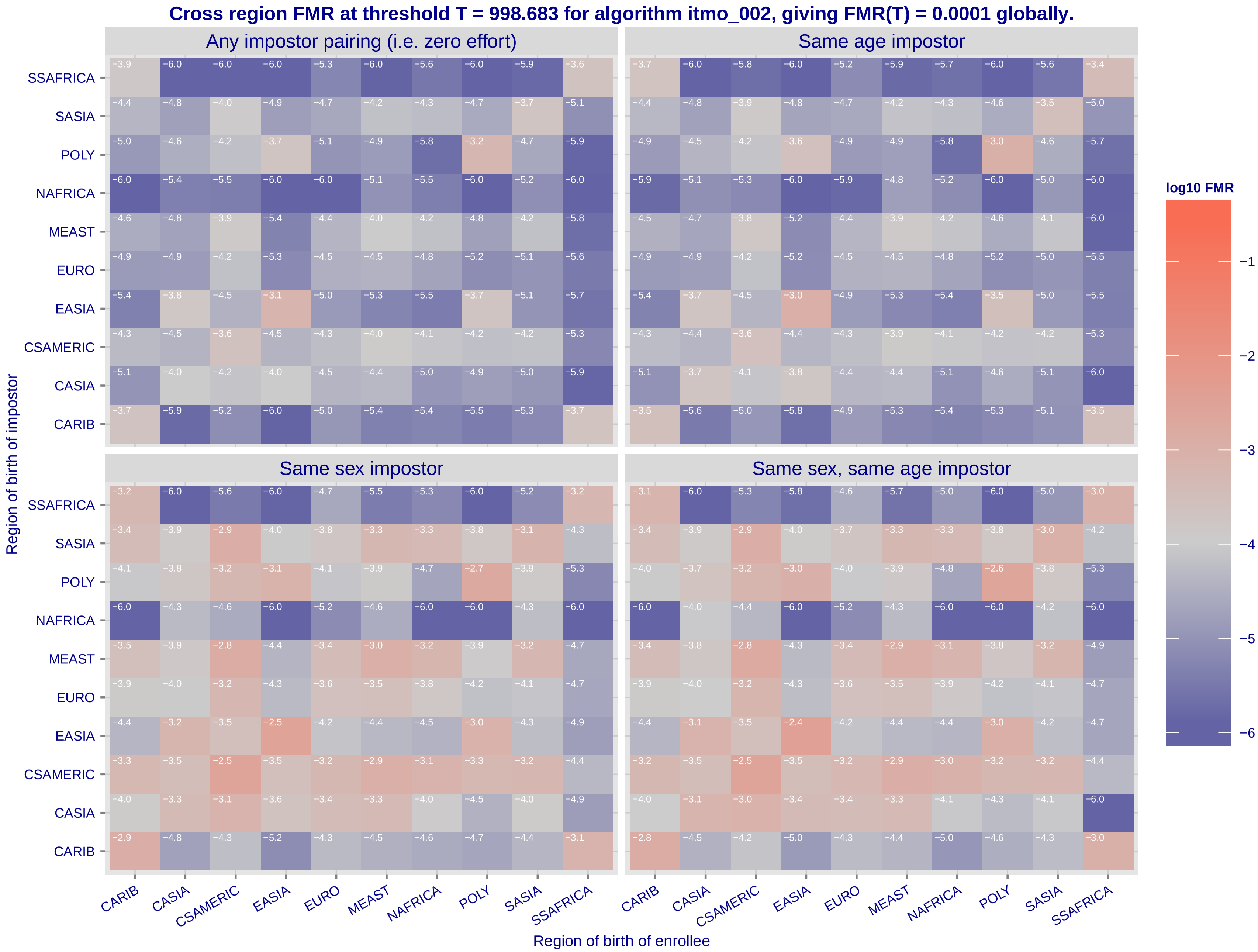


Figure 35: For algorithm itmo-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

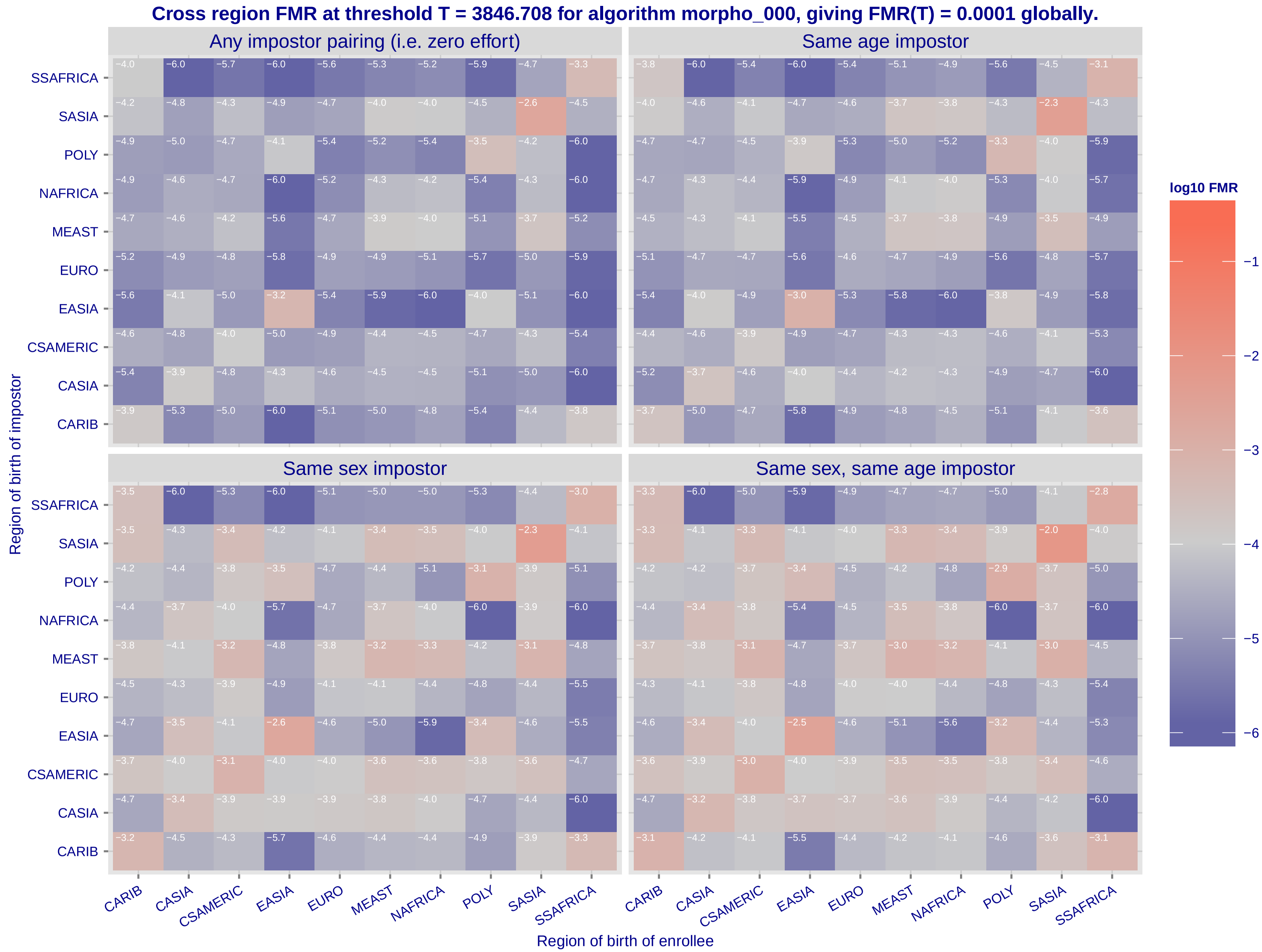


Figure 36: For algorithm morpho-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

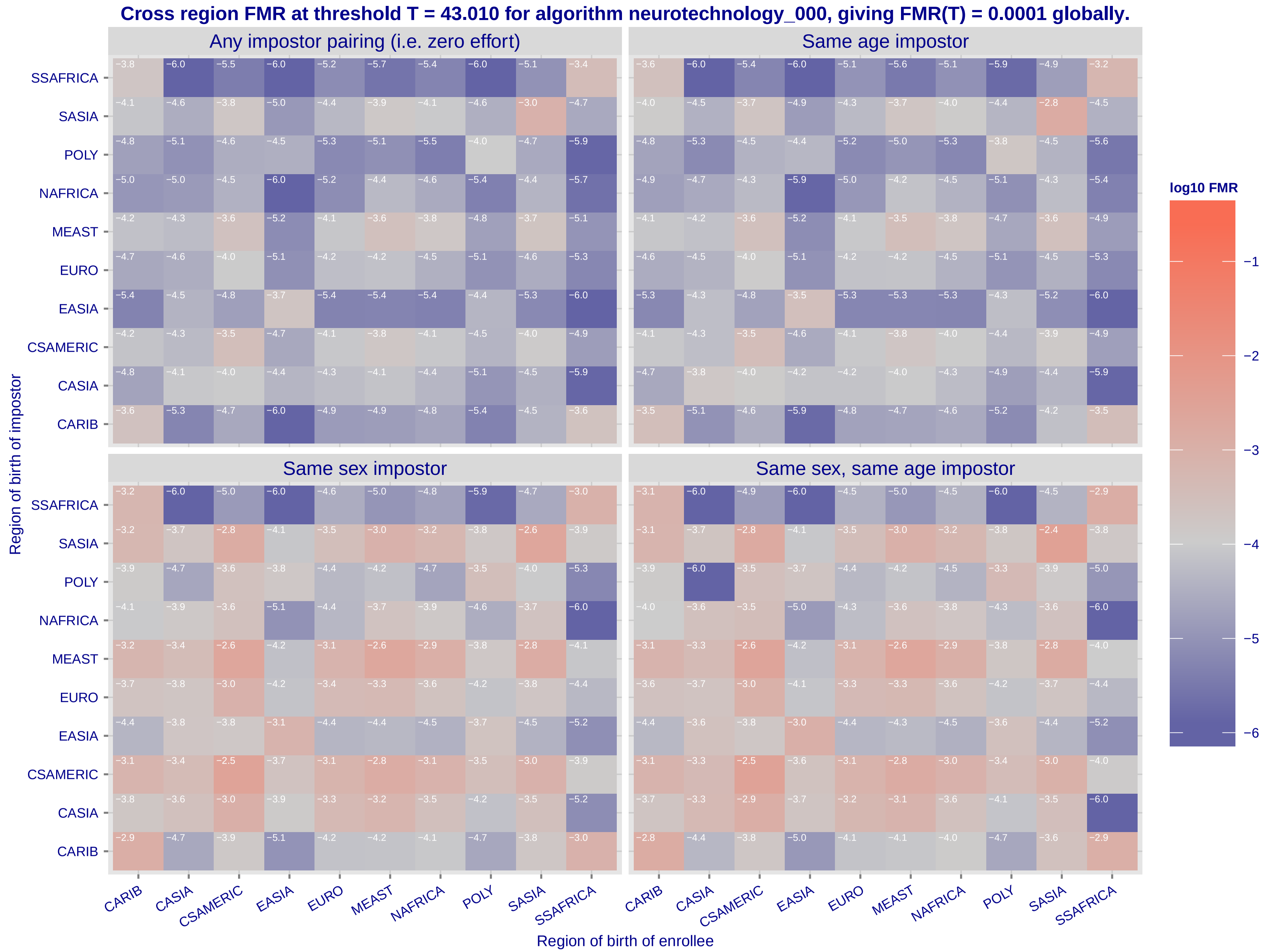


Figure 37: For algorithm neurotechnology-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

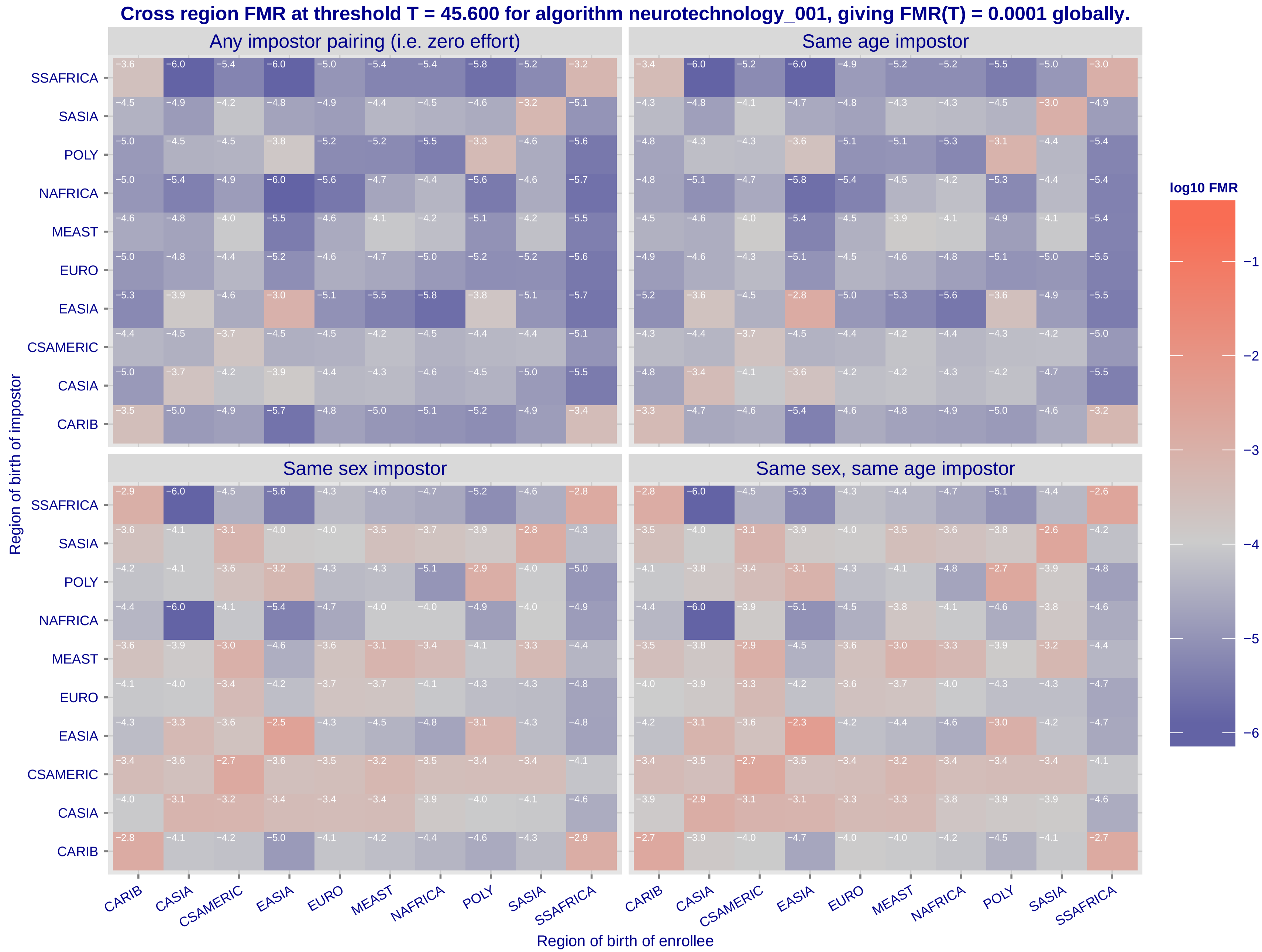


Figure 38: For algorithm neurotechnology-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

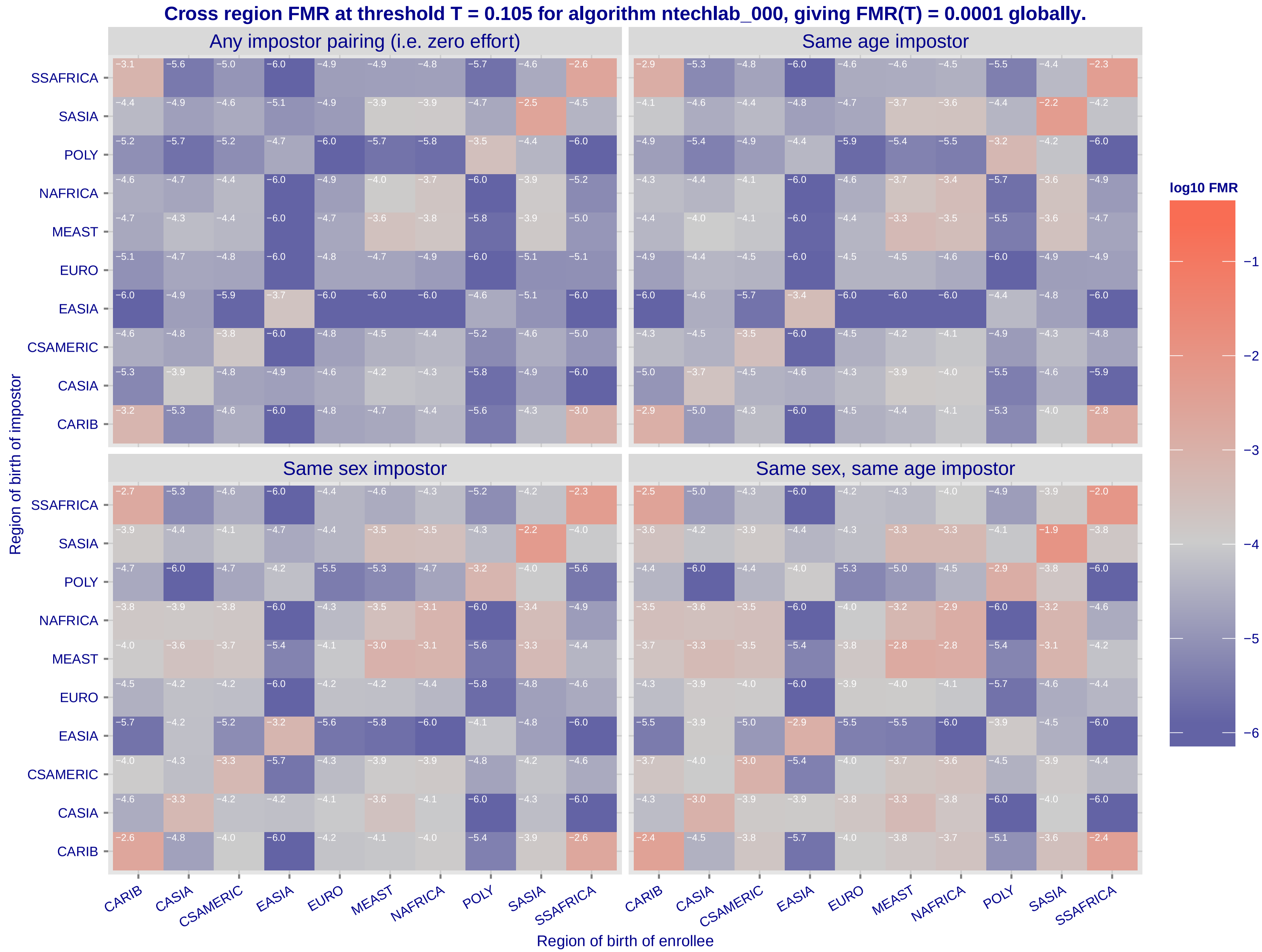


Figure 39: For algorithm ntechlab-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

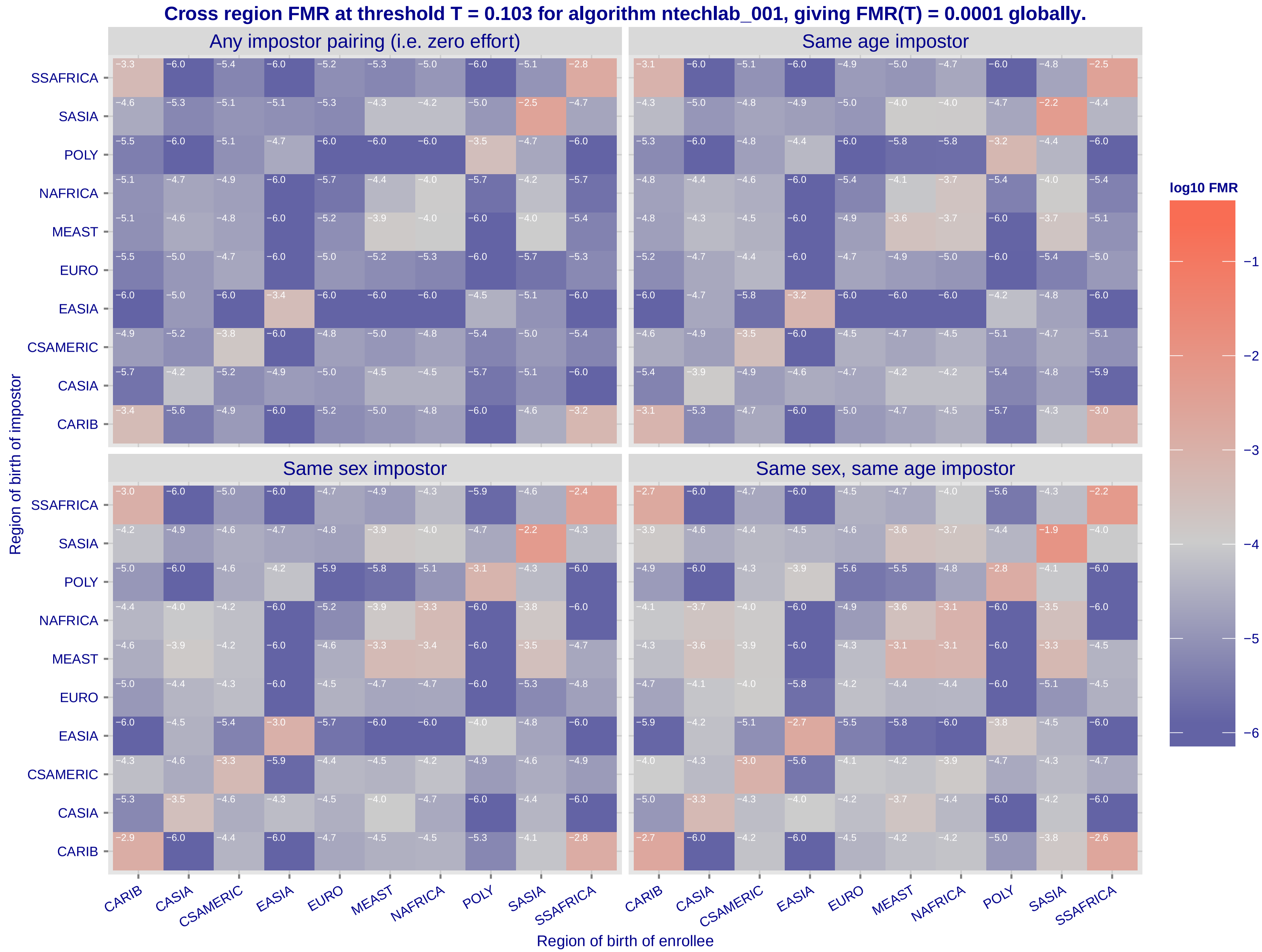


Figure 40: For algorithm ntechlab-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

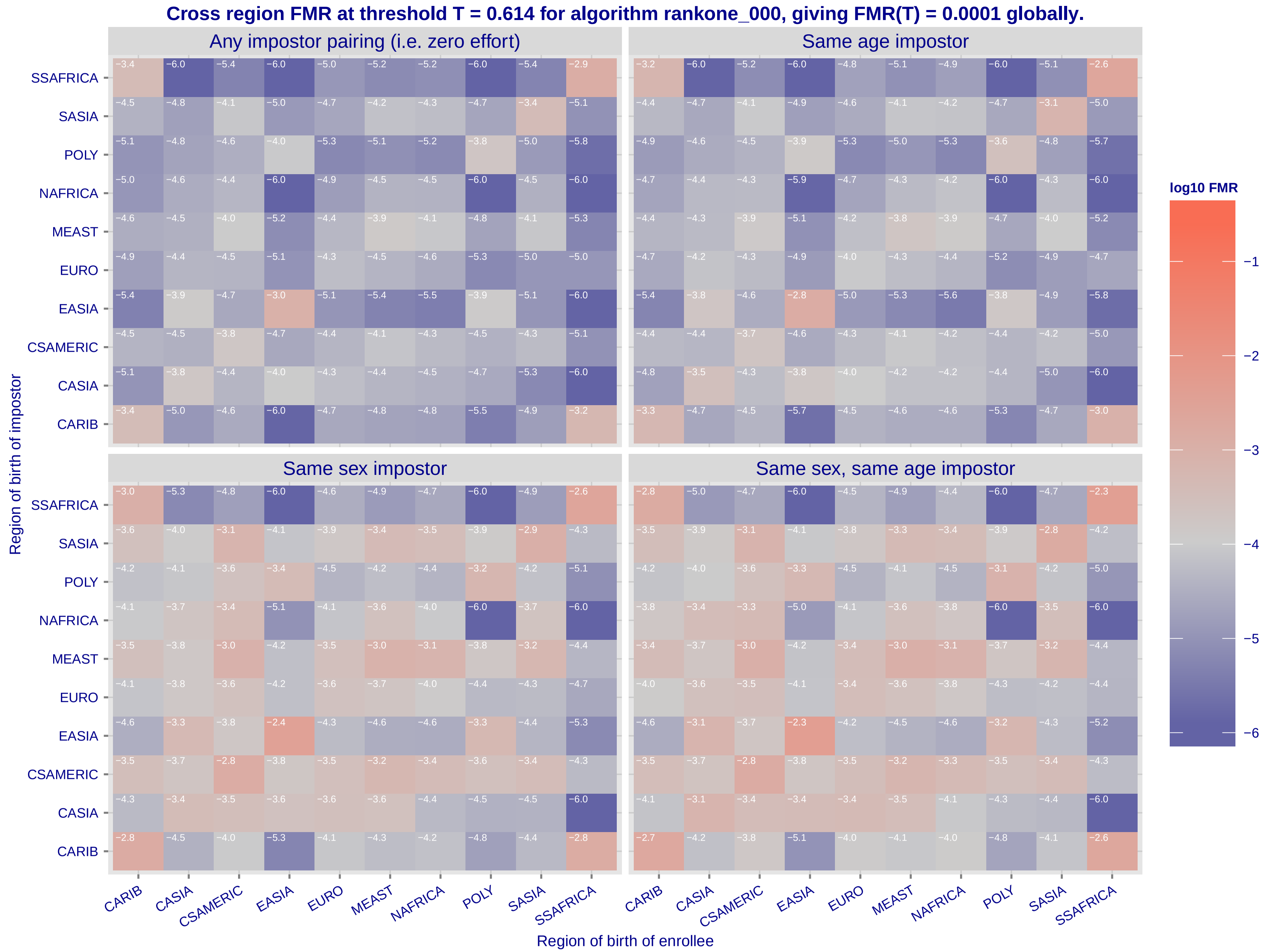


Figure 41: For algorithm rankone-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

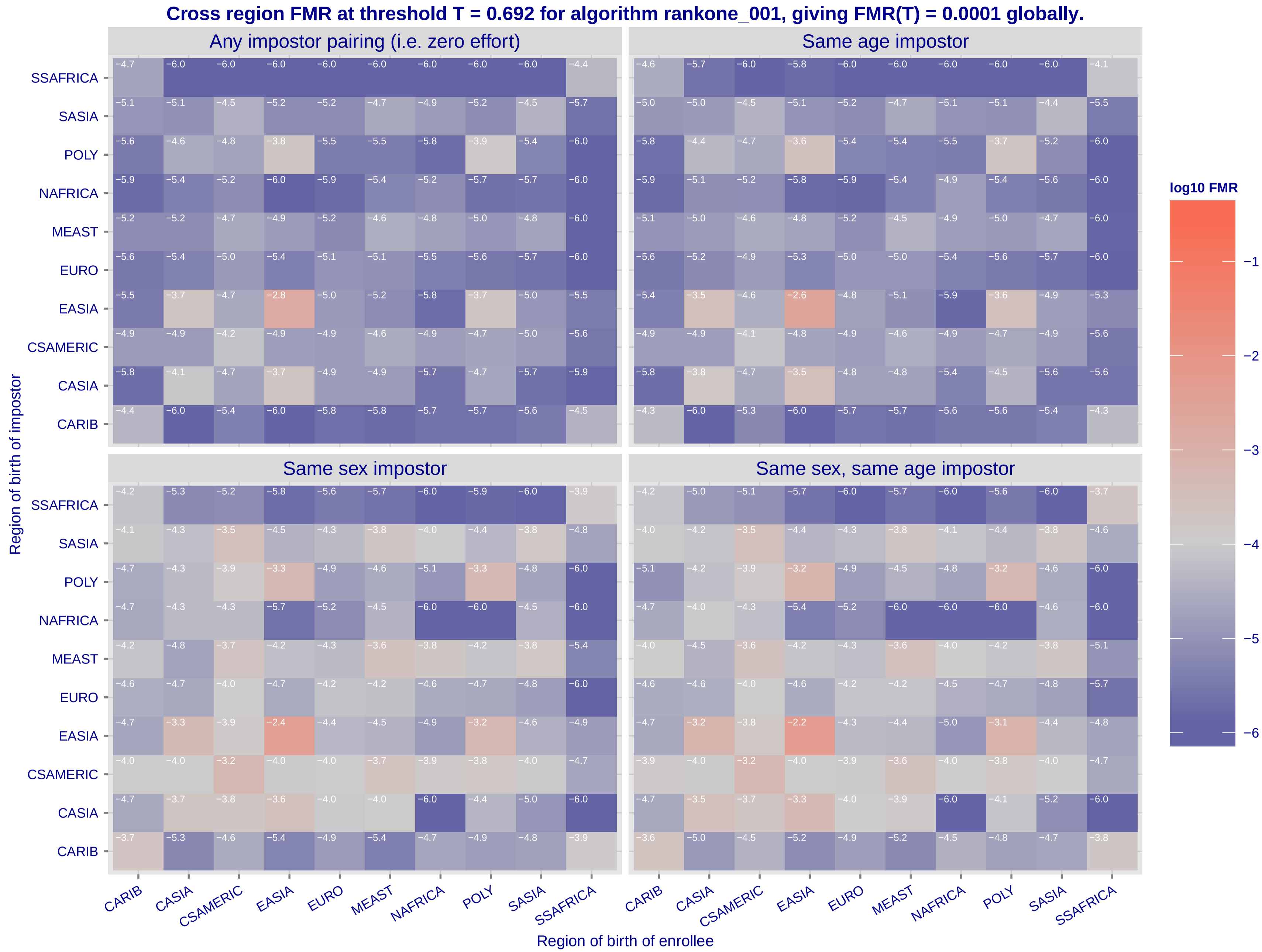


Figure 42: For algorithm rankone-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

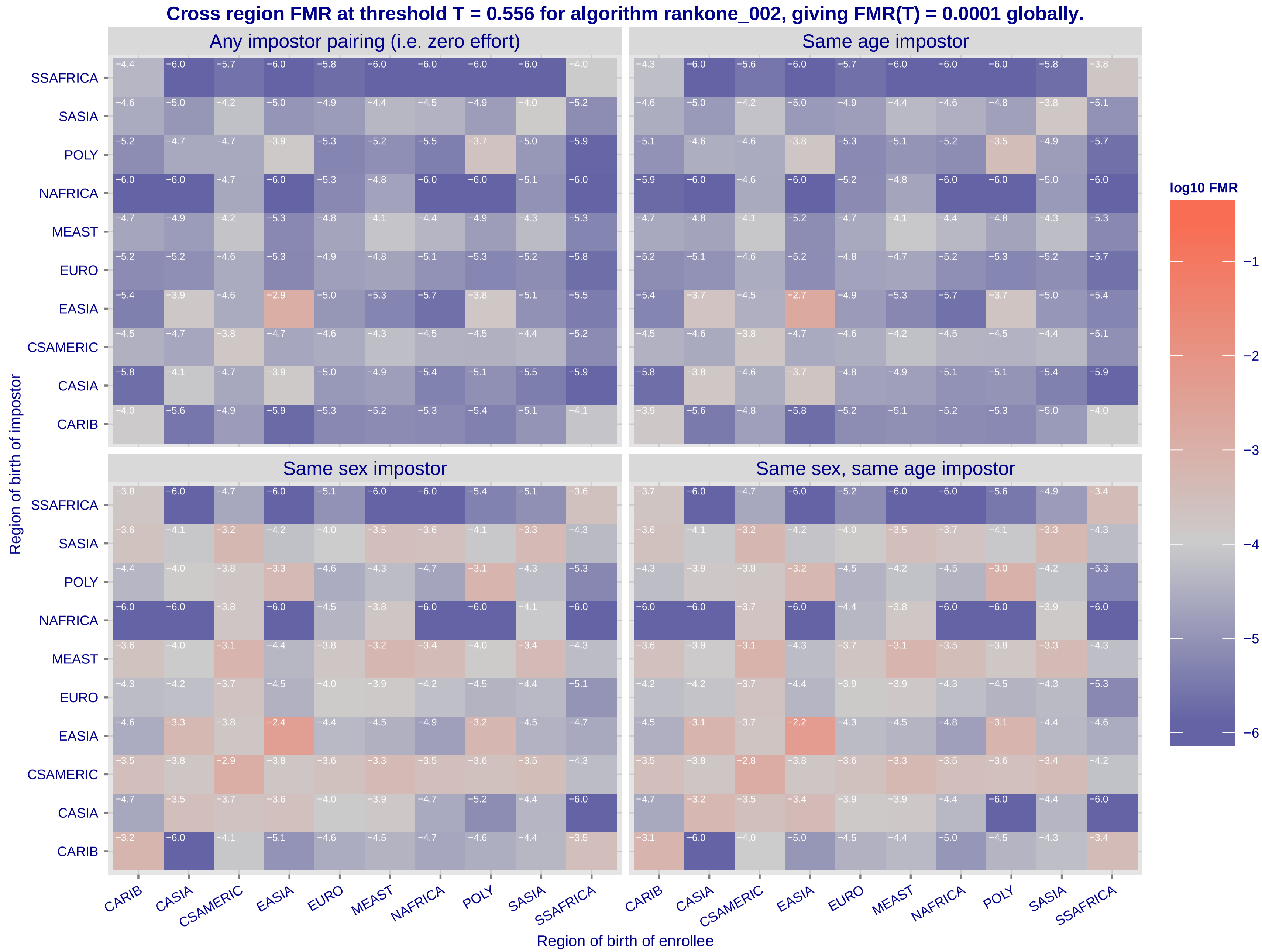


Figure 43: For algorithm rankone-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

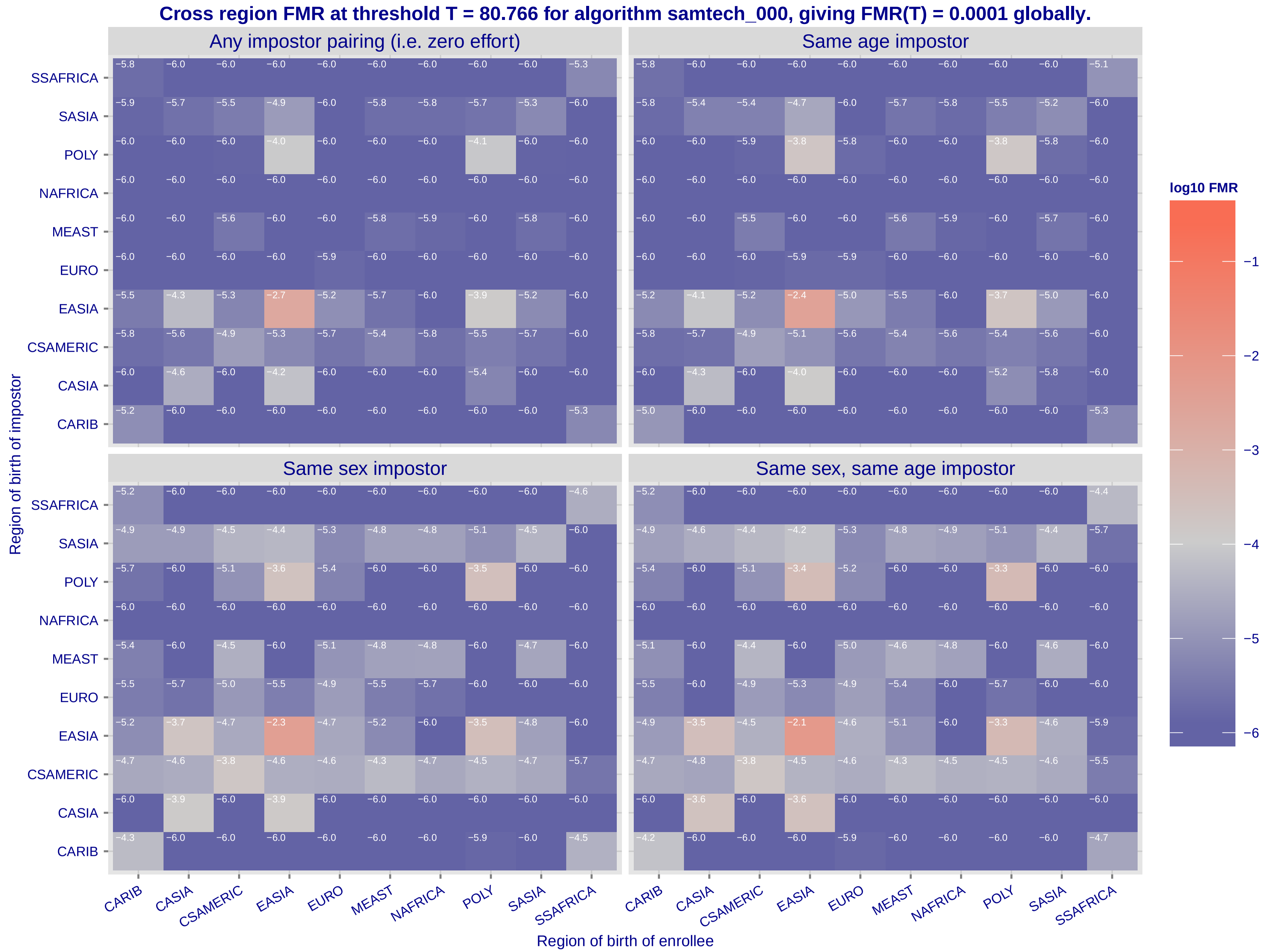


Figure 44: For algorithm samtech-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all O(10^{10}) impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

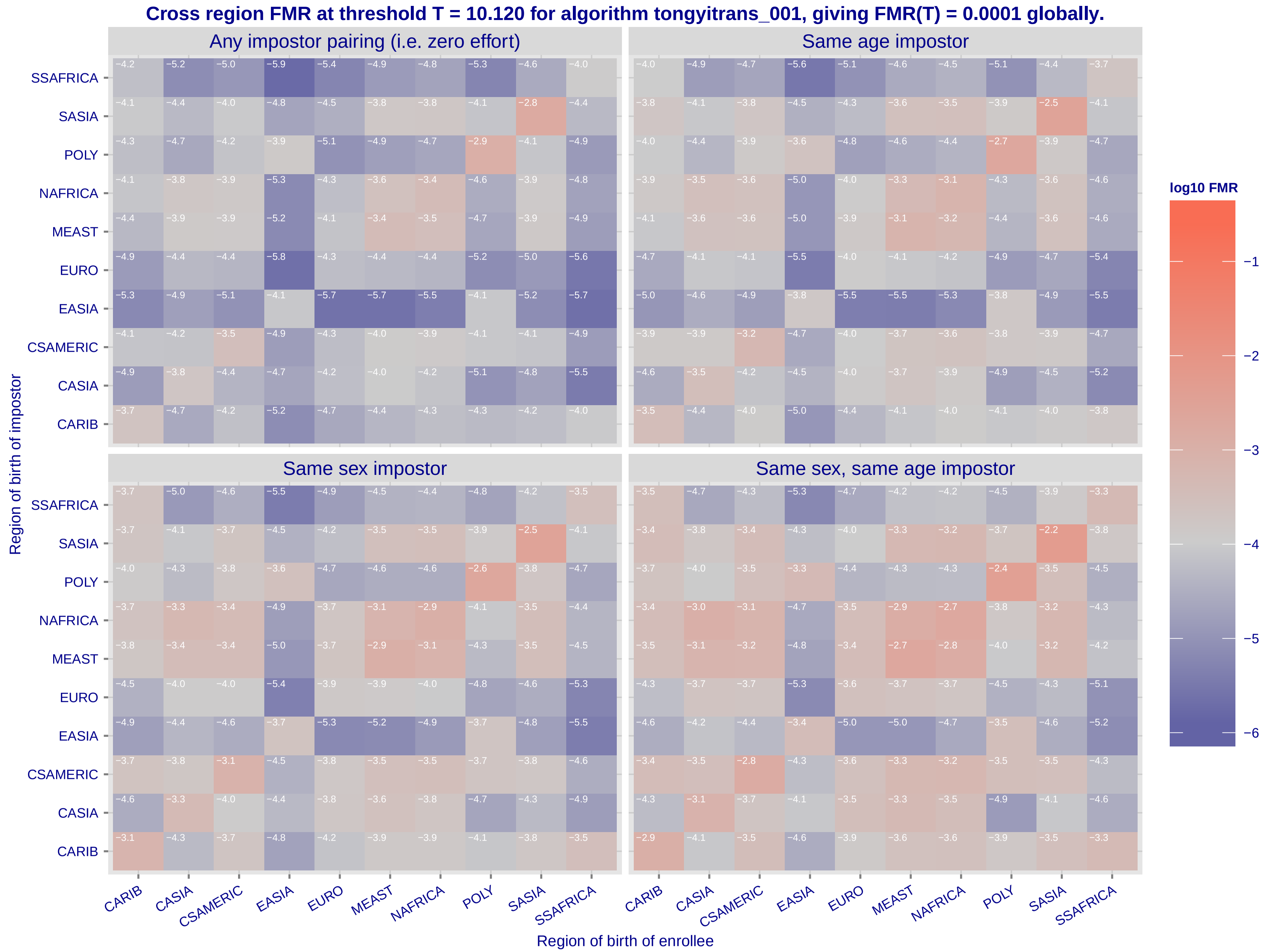


Figure 45: For algorithm tongyitrans-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

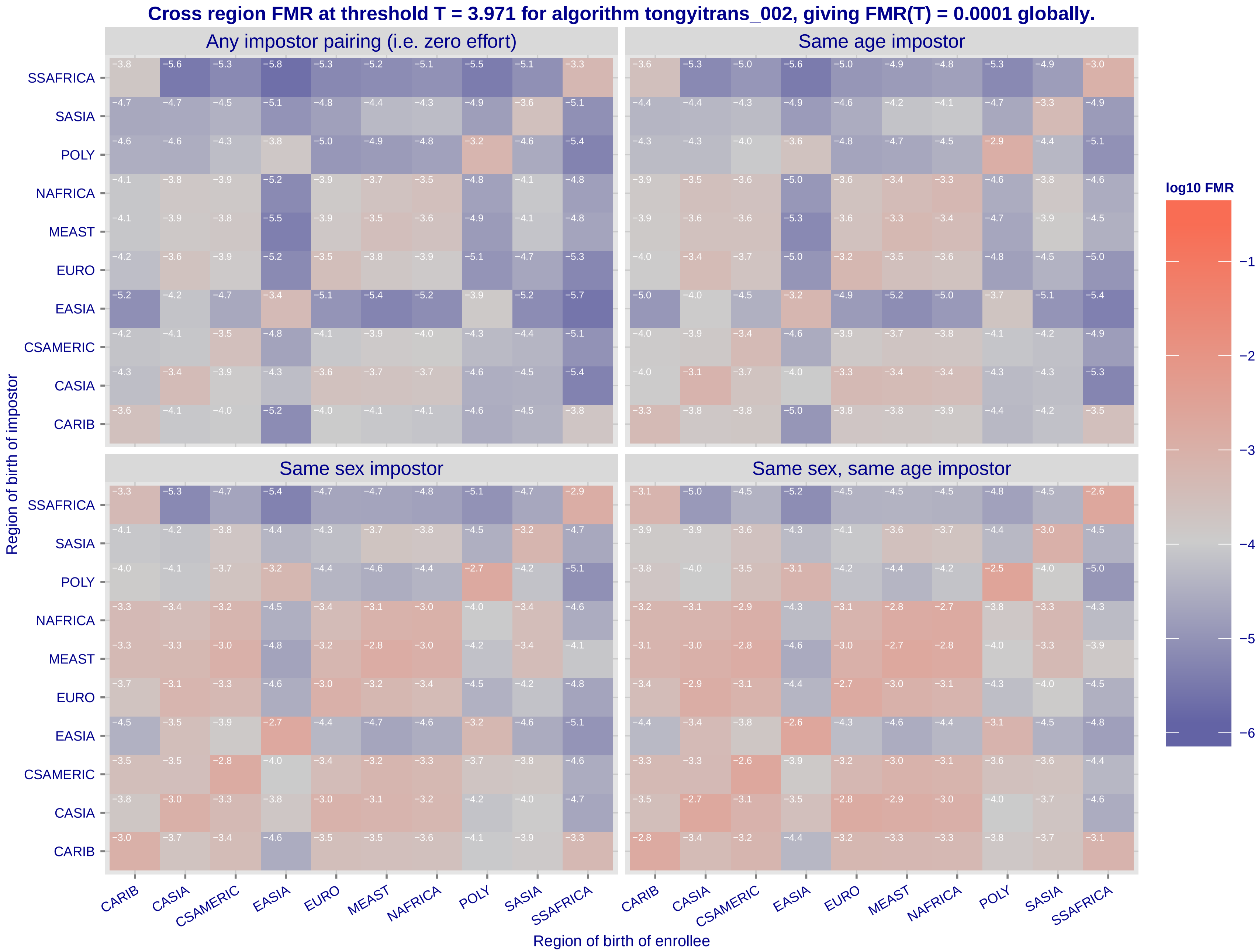


Figure 46: *For algorithm tongyitrans-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.*

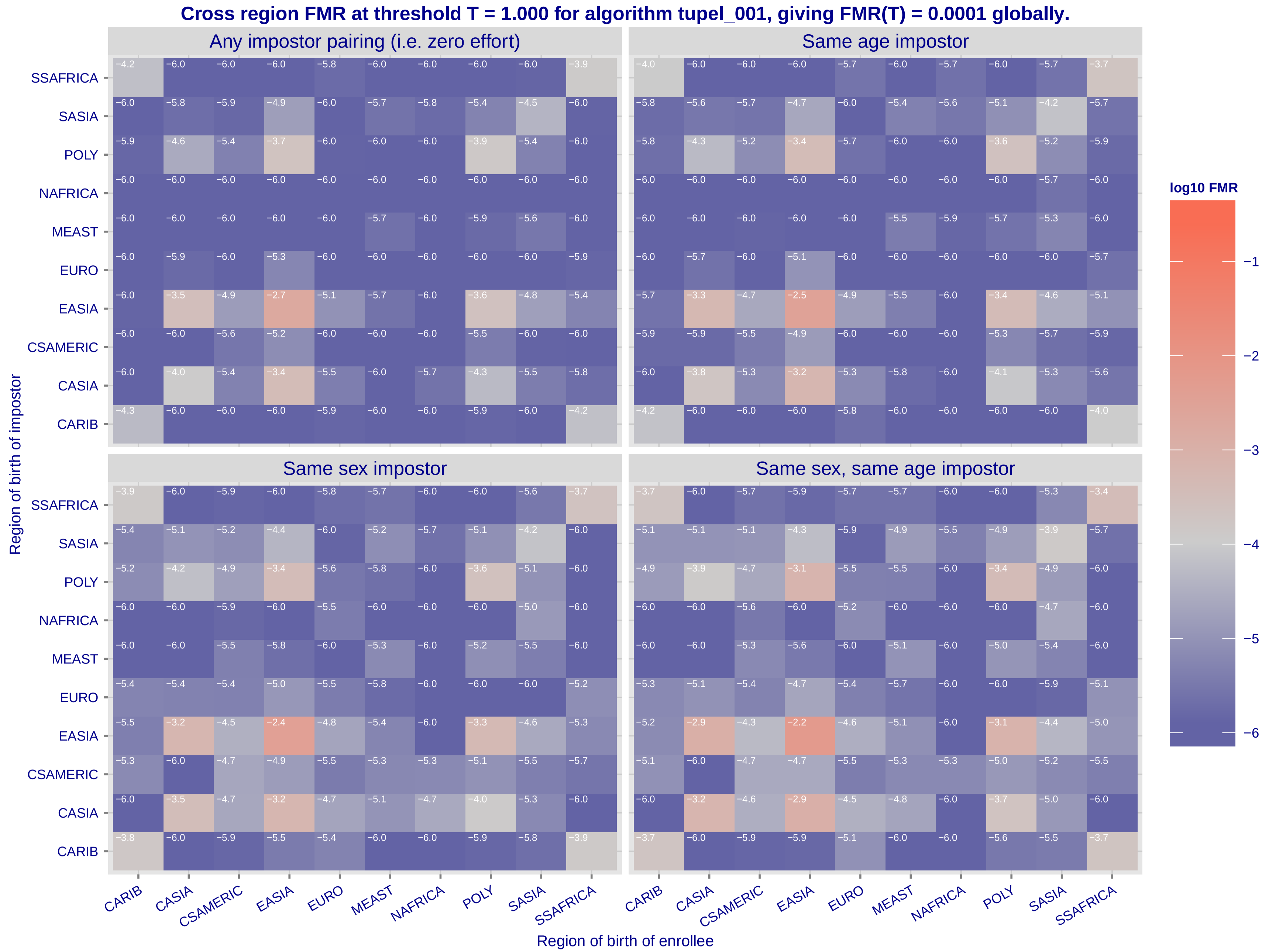


Figure 47: For algorithm tupel-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

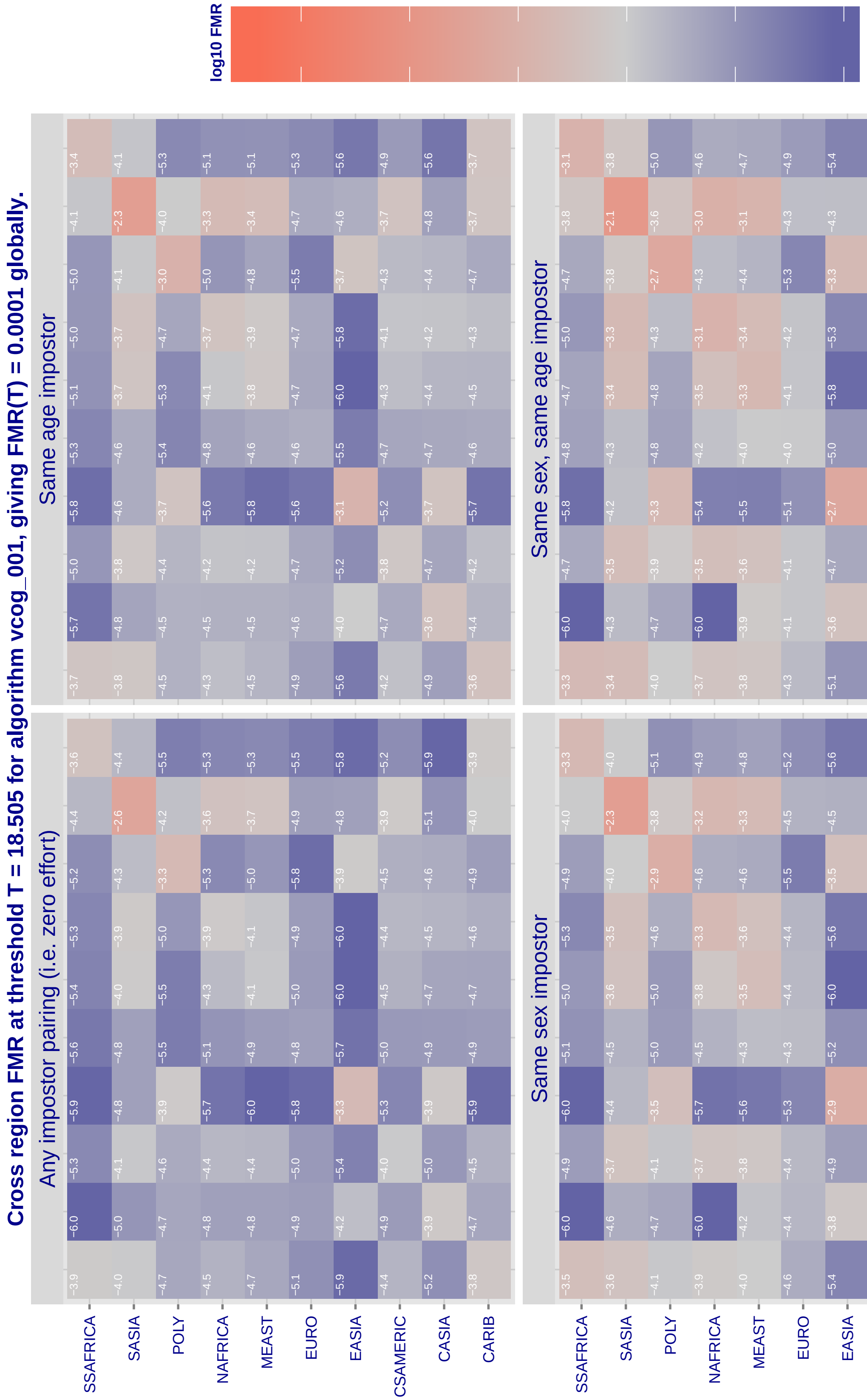


Figure 48: For algorithm vcog-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all O(10^{10}) impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

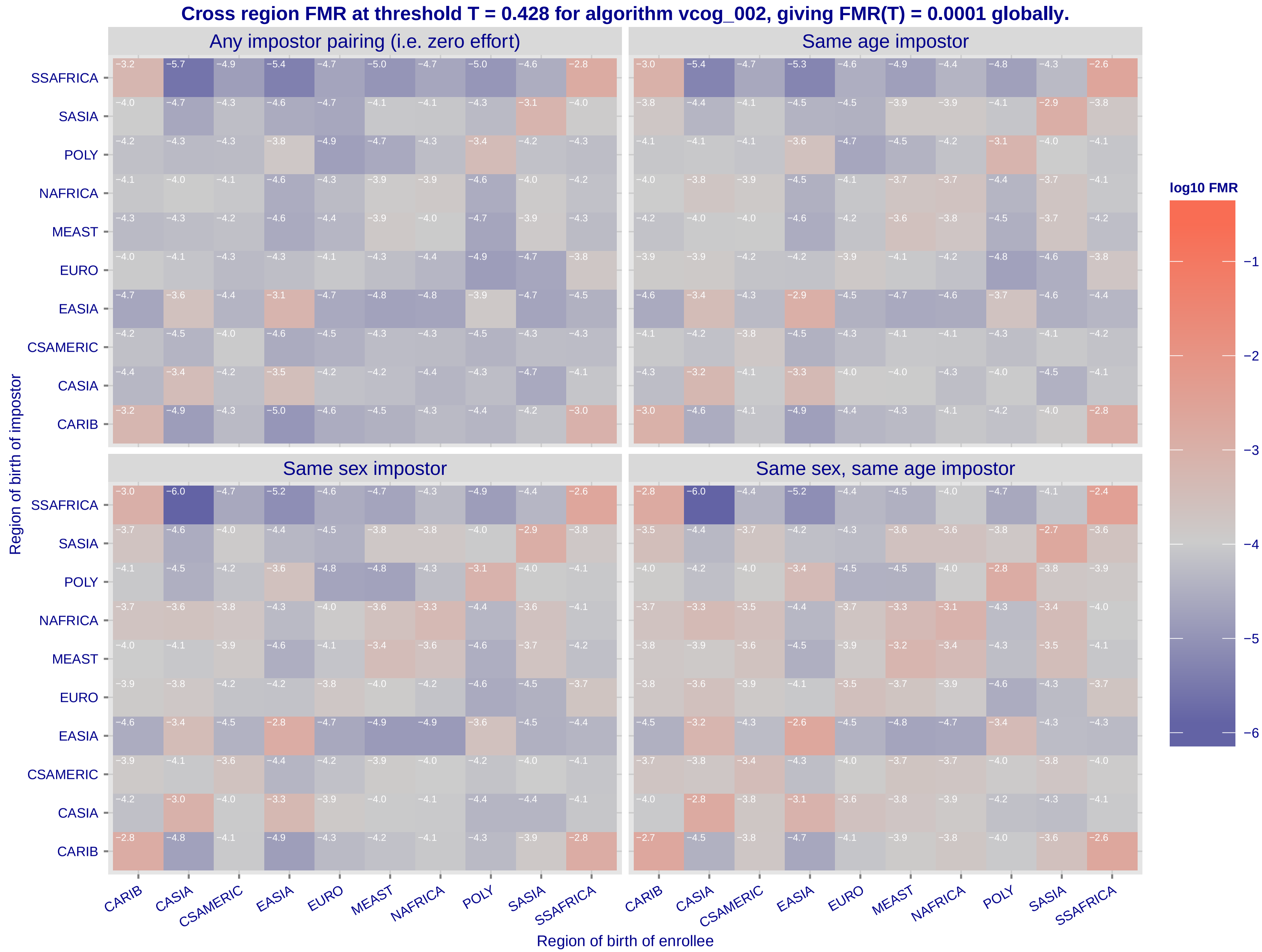


Figure 49: For algorithm vcog-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

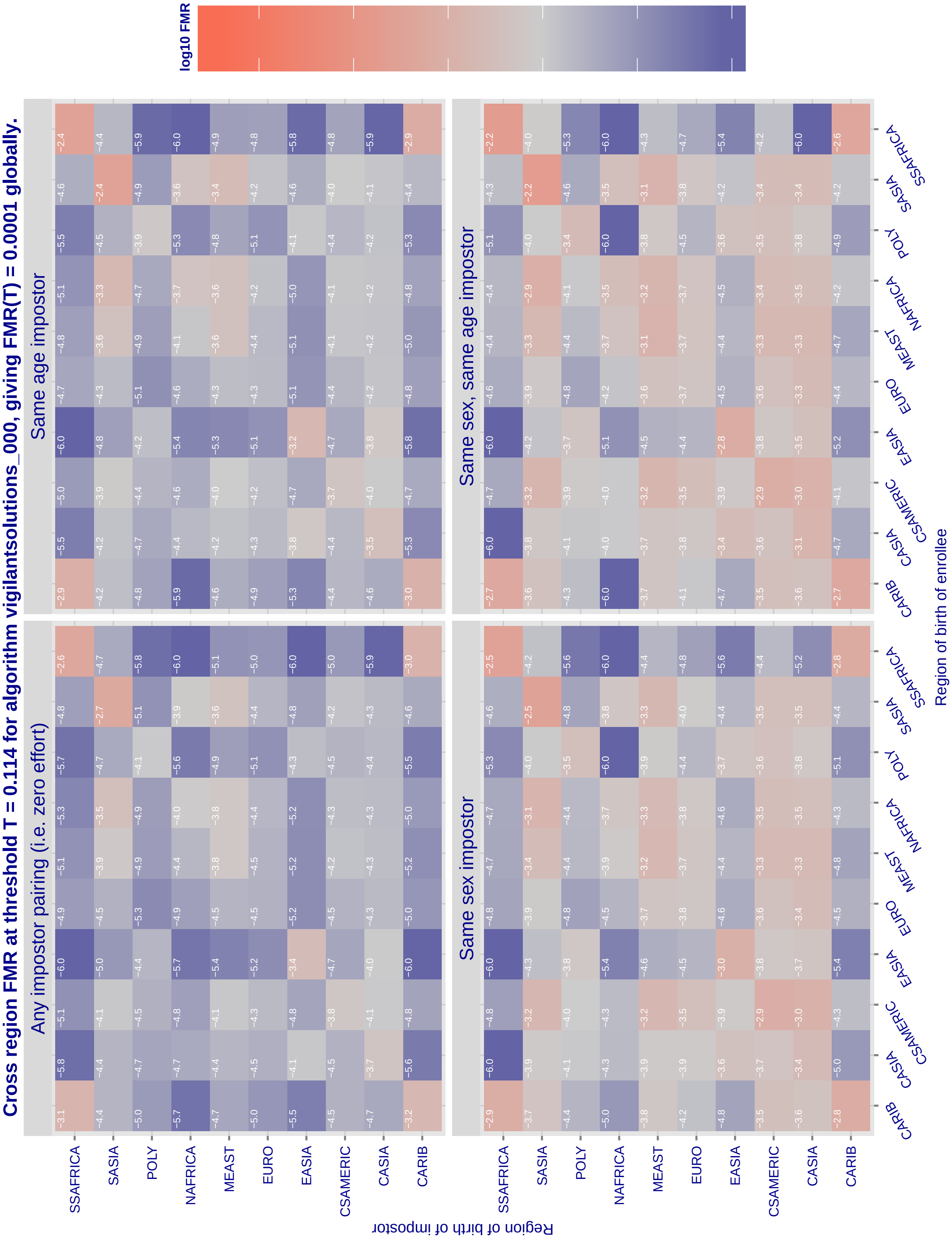


Figure 50: For algorithm vigilantsolutions-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

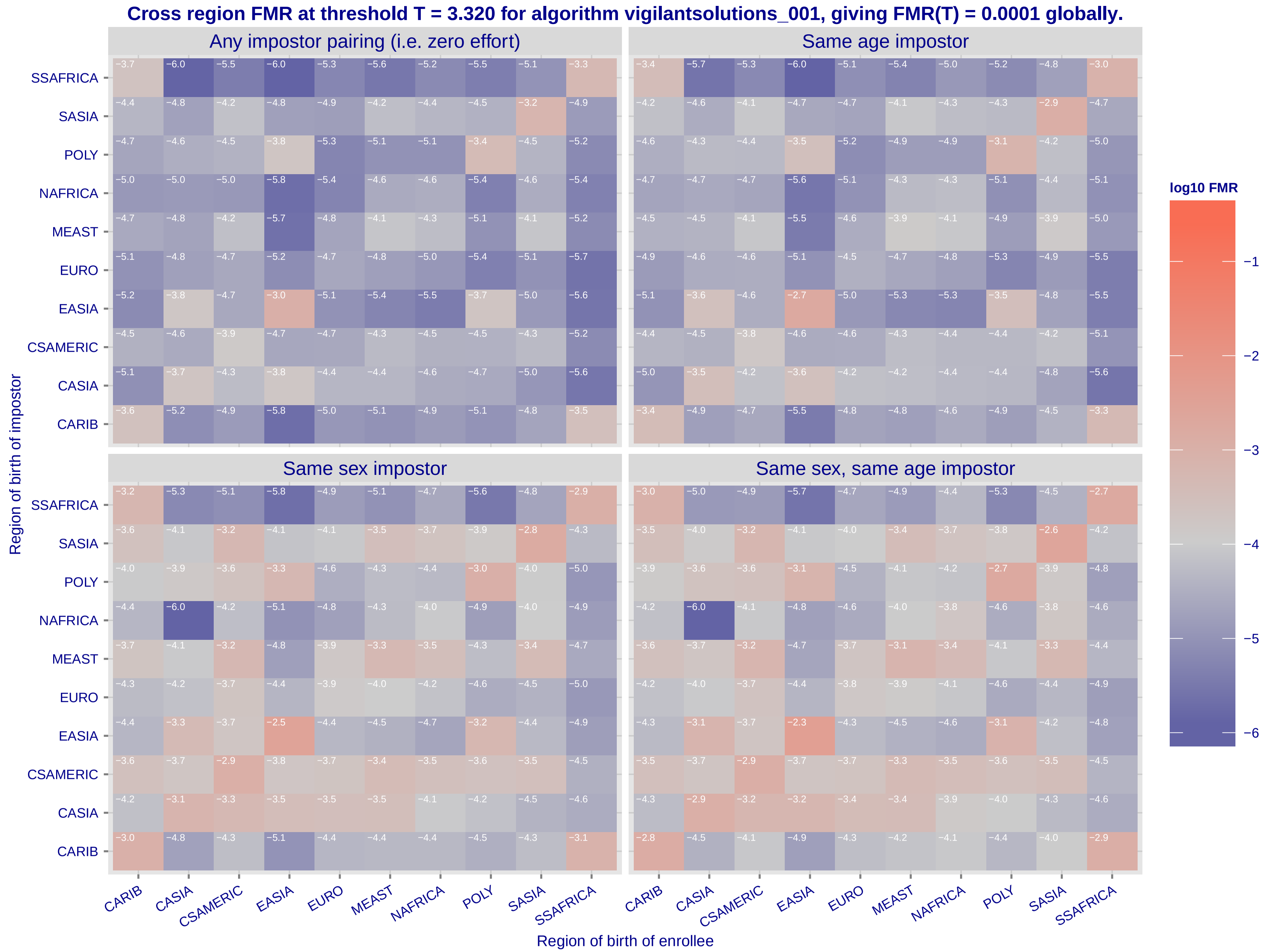


Figure 51: For algorithm vigilantsolutions-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

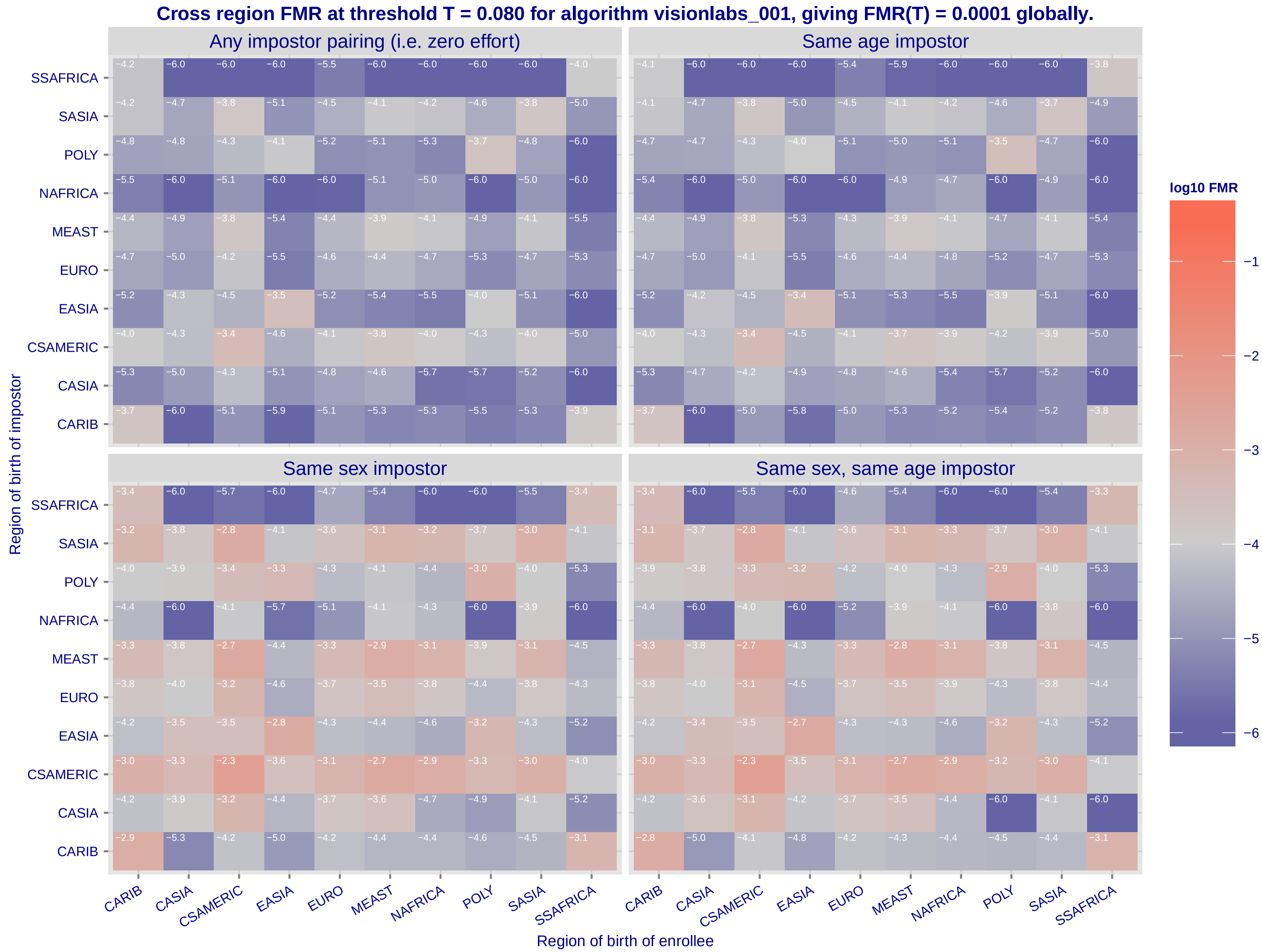


Figure 52: For algorithm visionlabs-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all O(10^{10}) impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

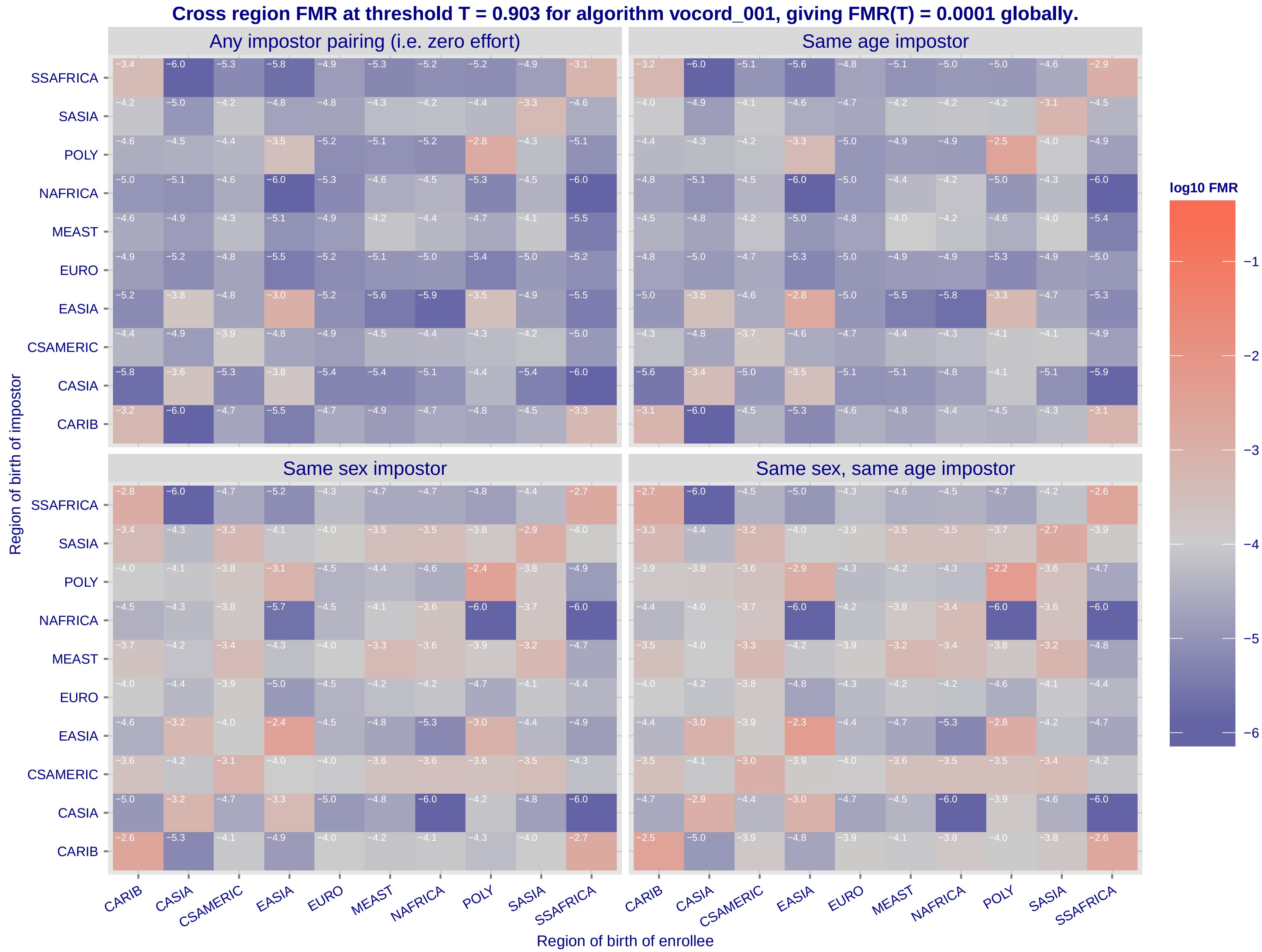


Figure 53: For algorithm vocord-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

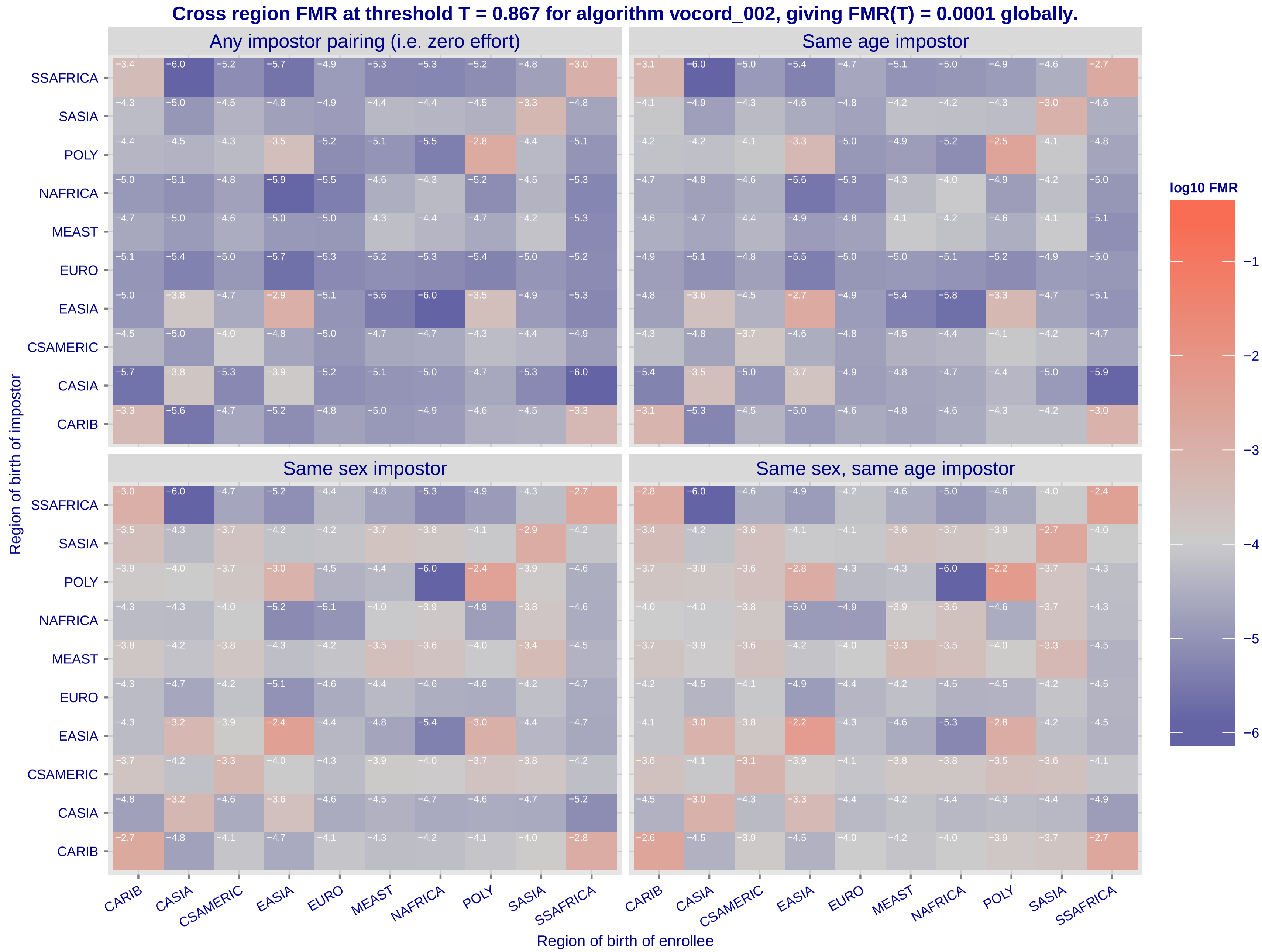


Figure 54: For algorithm vocord-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

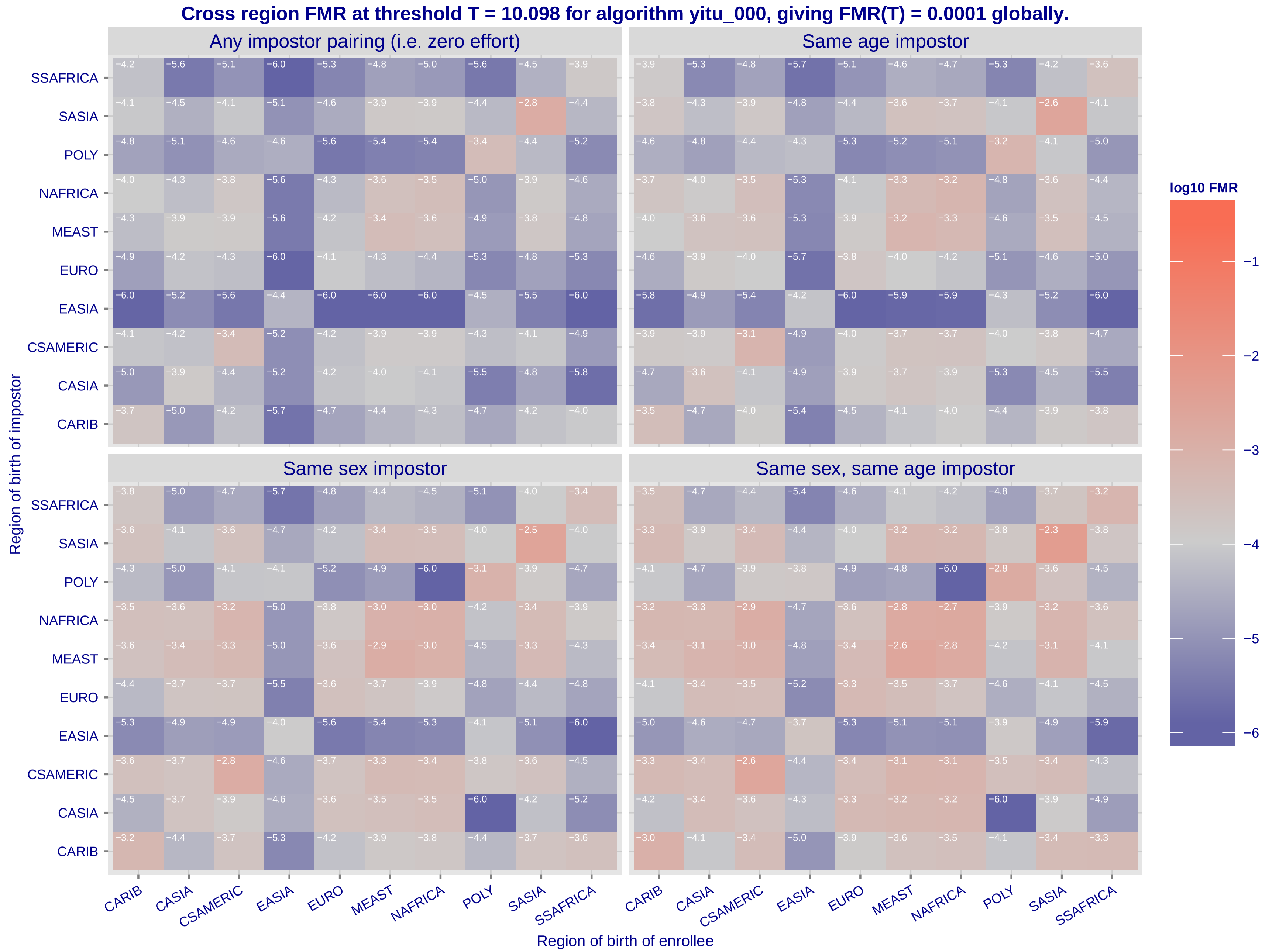


Figure 55: For algorithm yitu-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

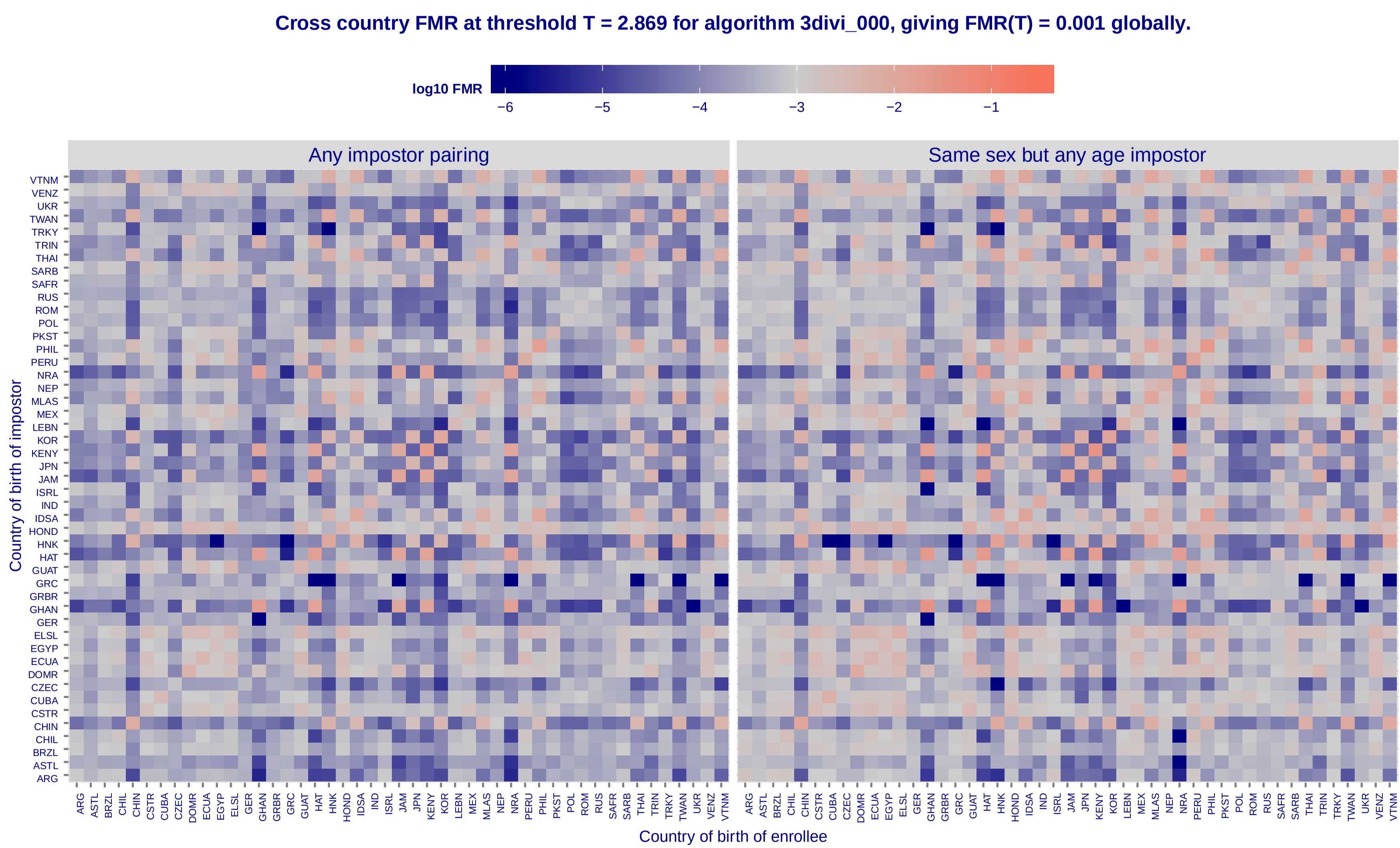


Figure 56: For algorithm 3divi-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

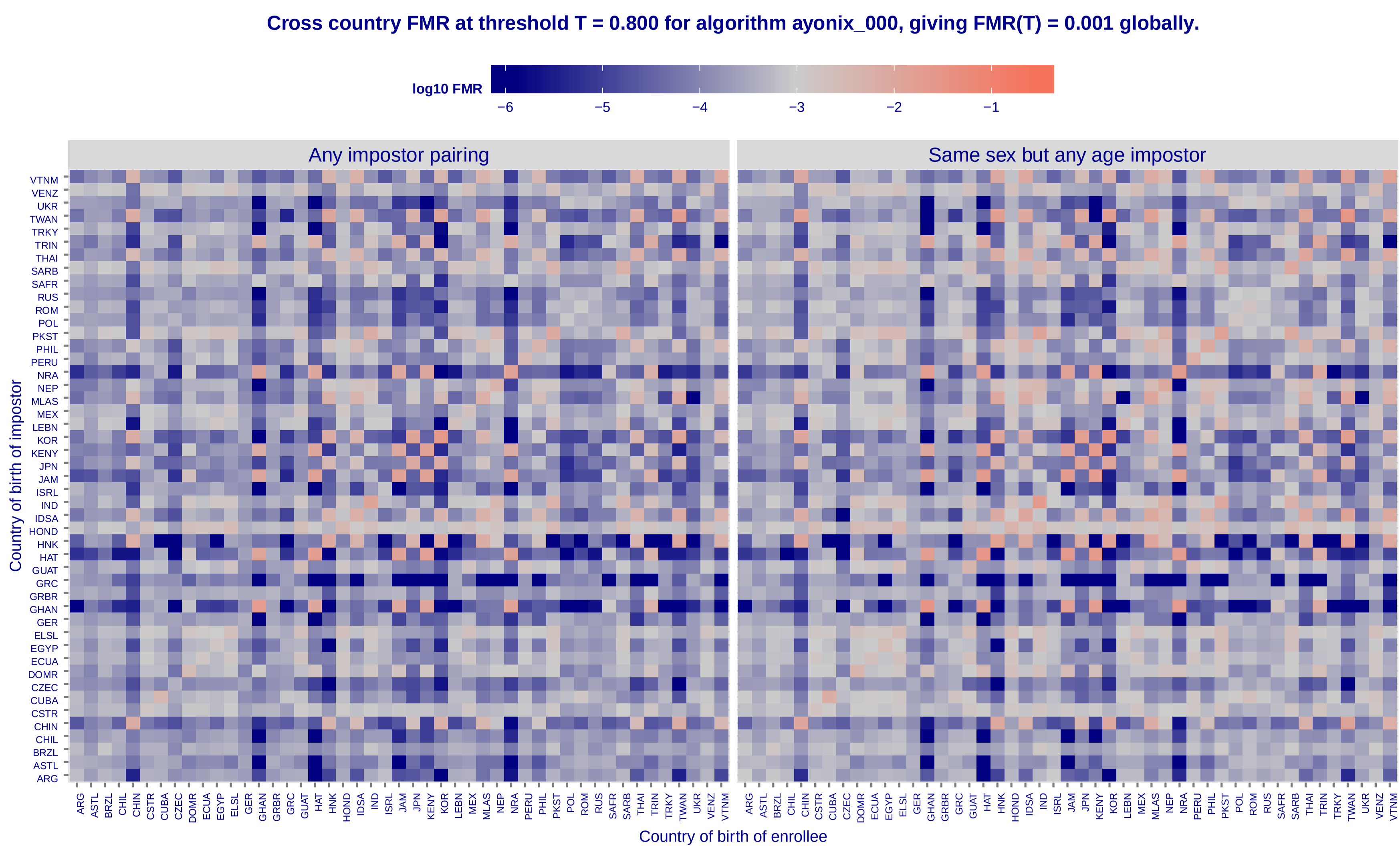


Figure 57: For algorithm ayonix-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

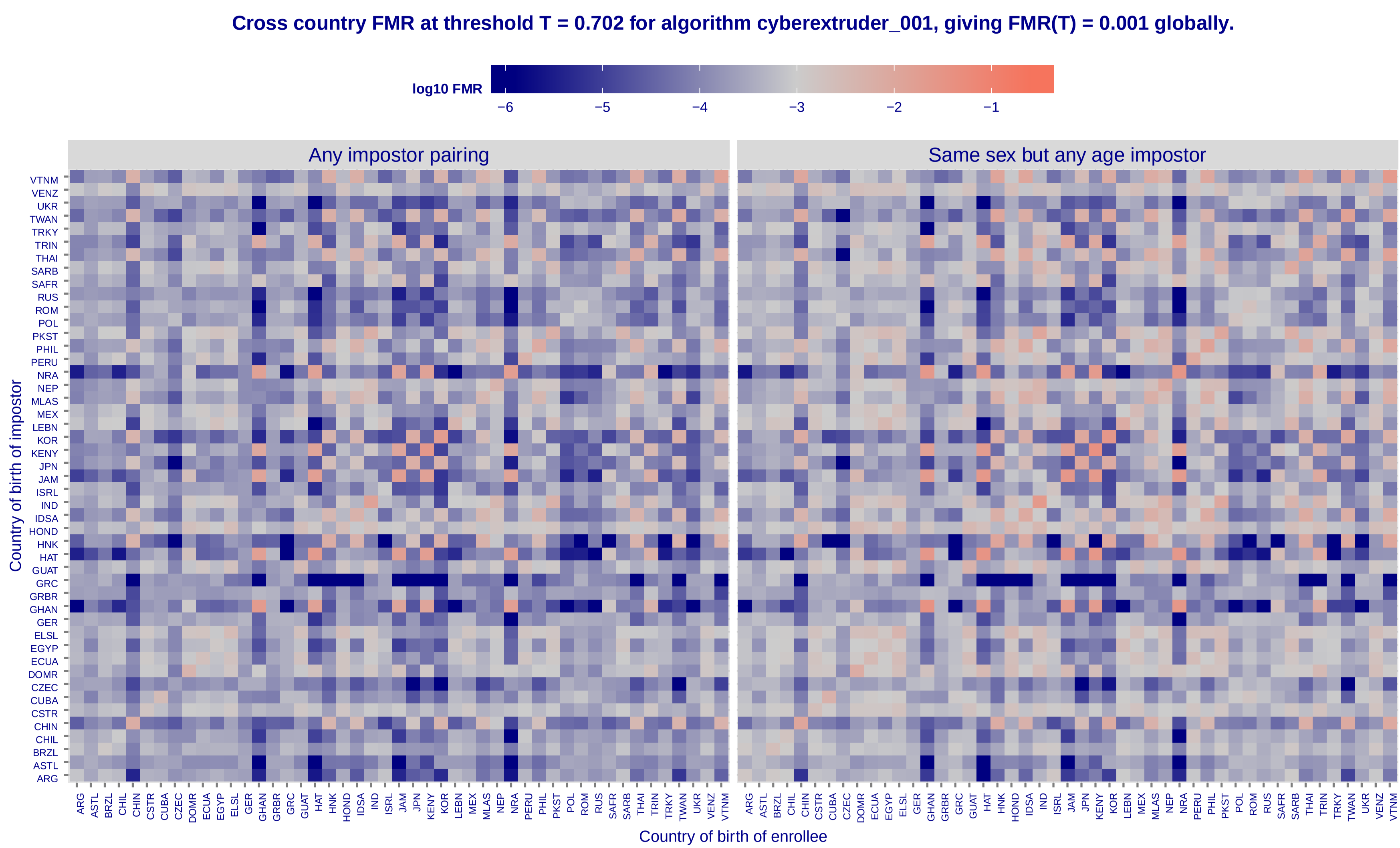


Figure 58: For algorithm cyberextruder-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all O(10^{10}) impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

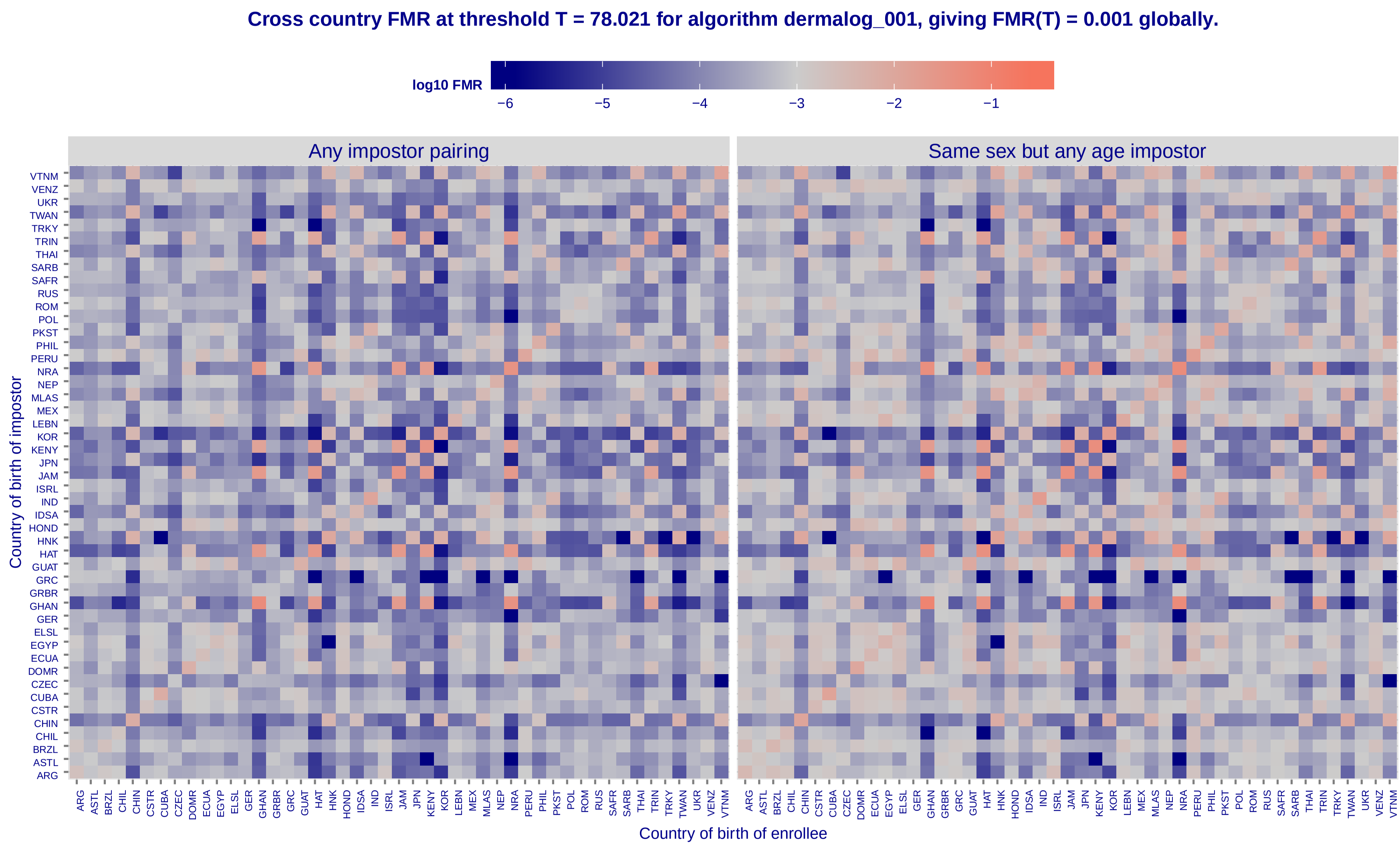


Figure 59: For algorithm dermalog-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

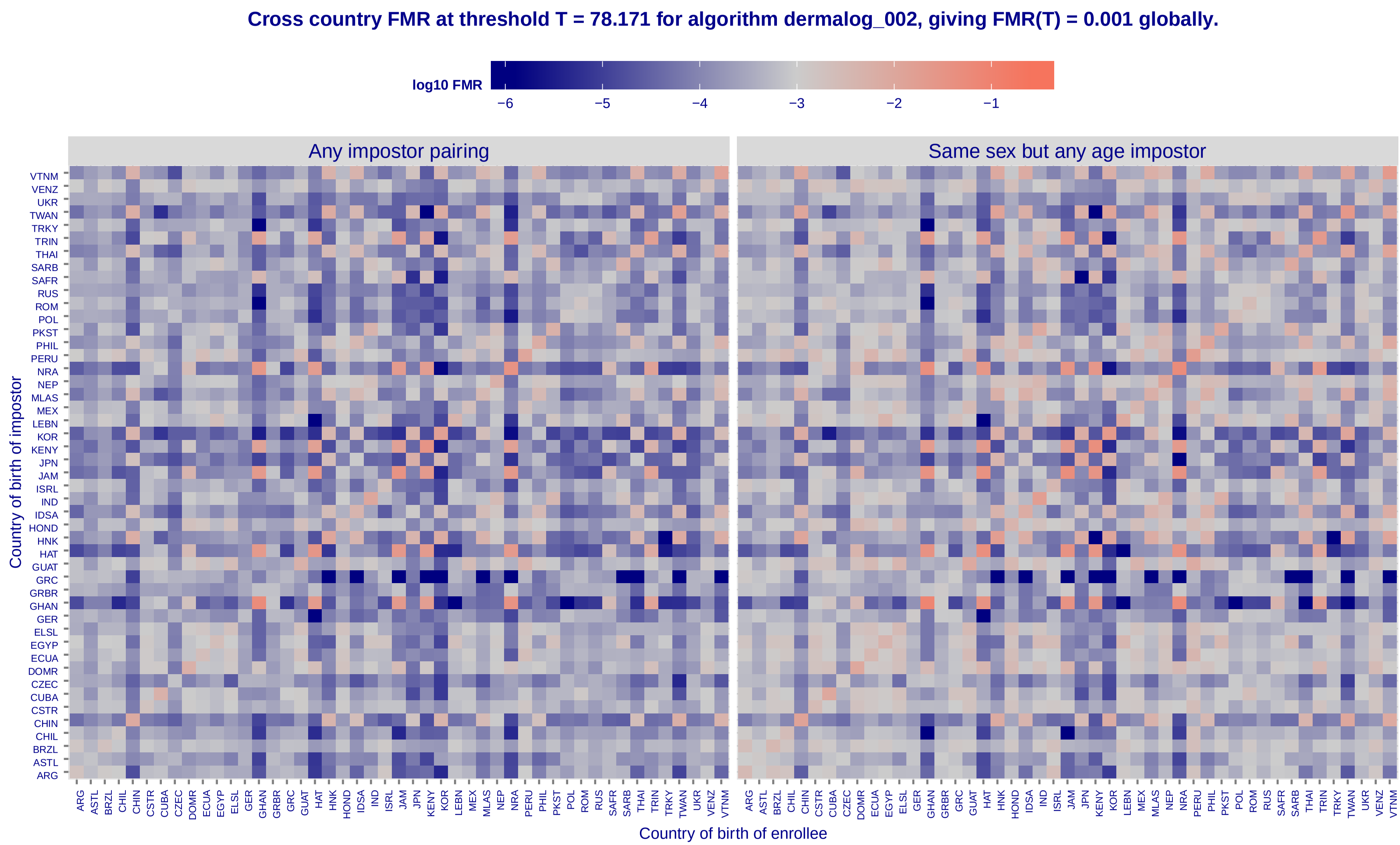


Figure 60: For algorithm dermalog-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all O(10^{10}) impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

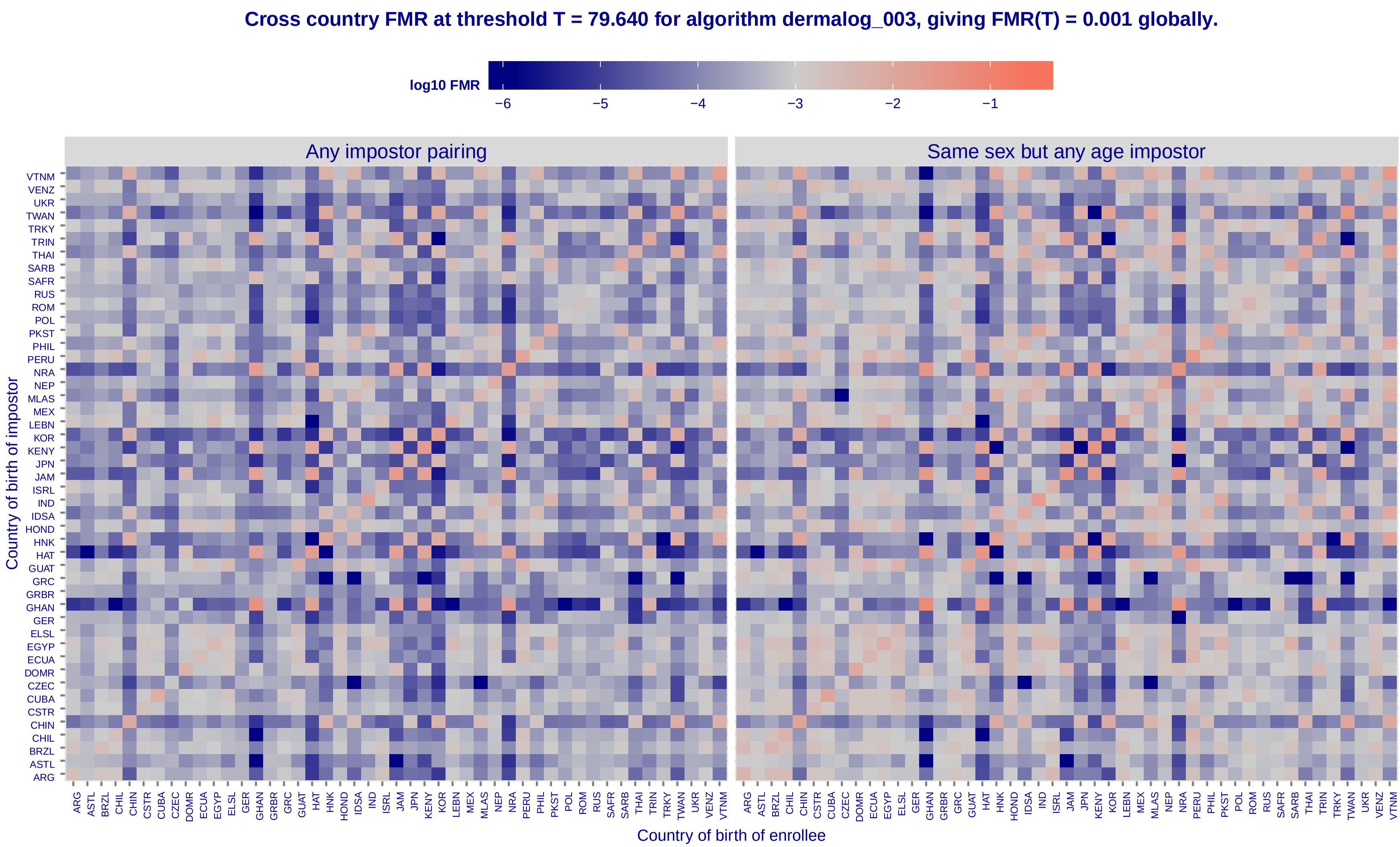


Figure 61: For algorithm dermalog-003 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

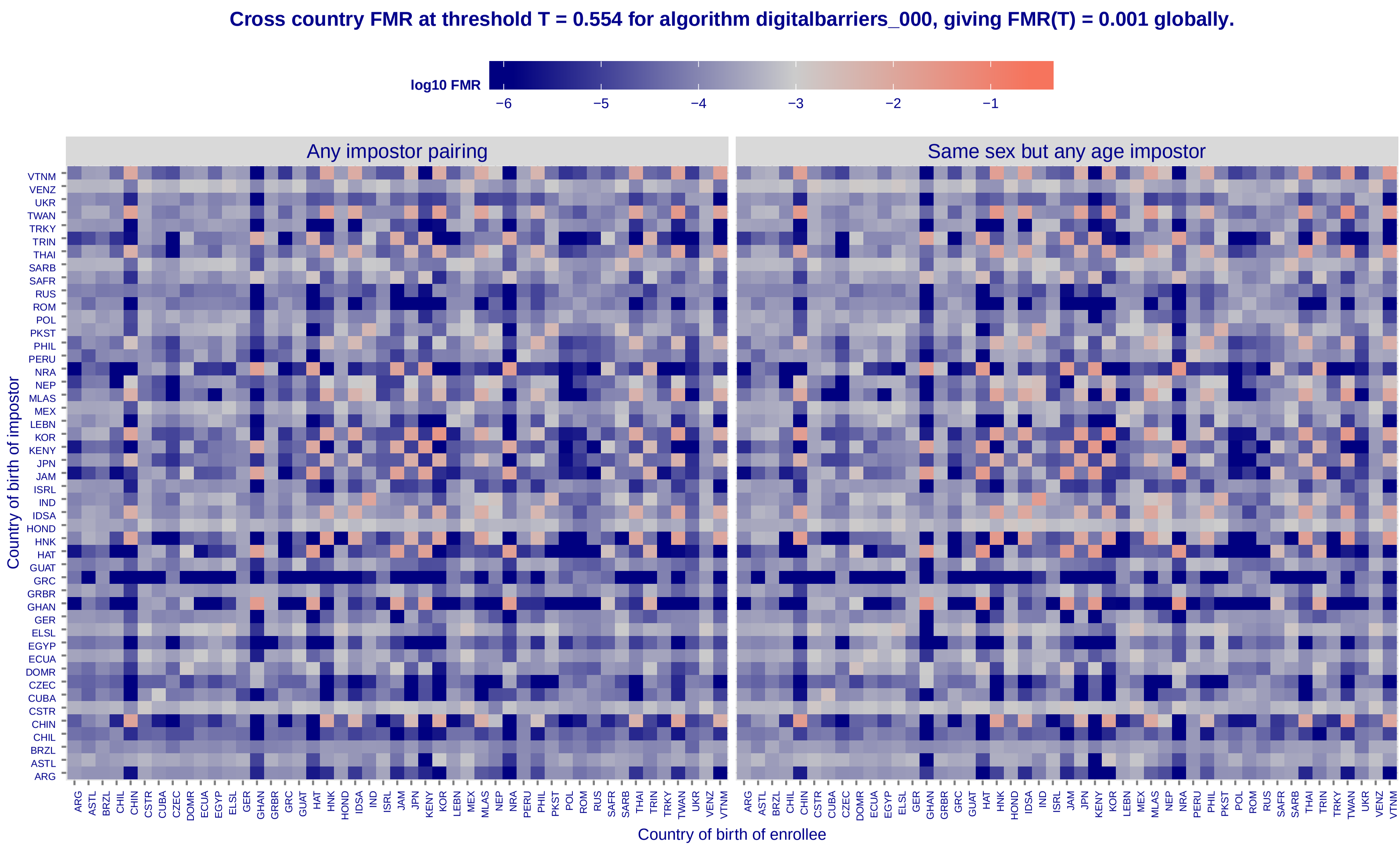


Figure 62: For algorithm digitalbarriers-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

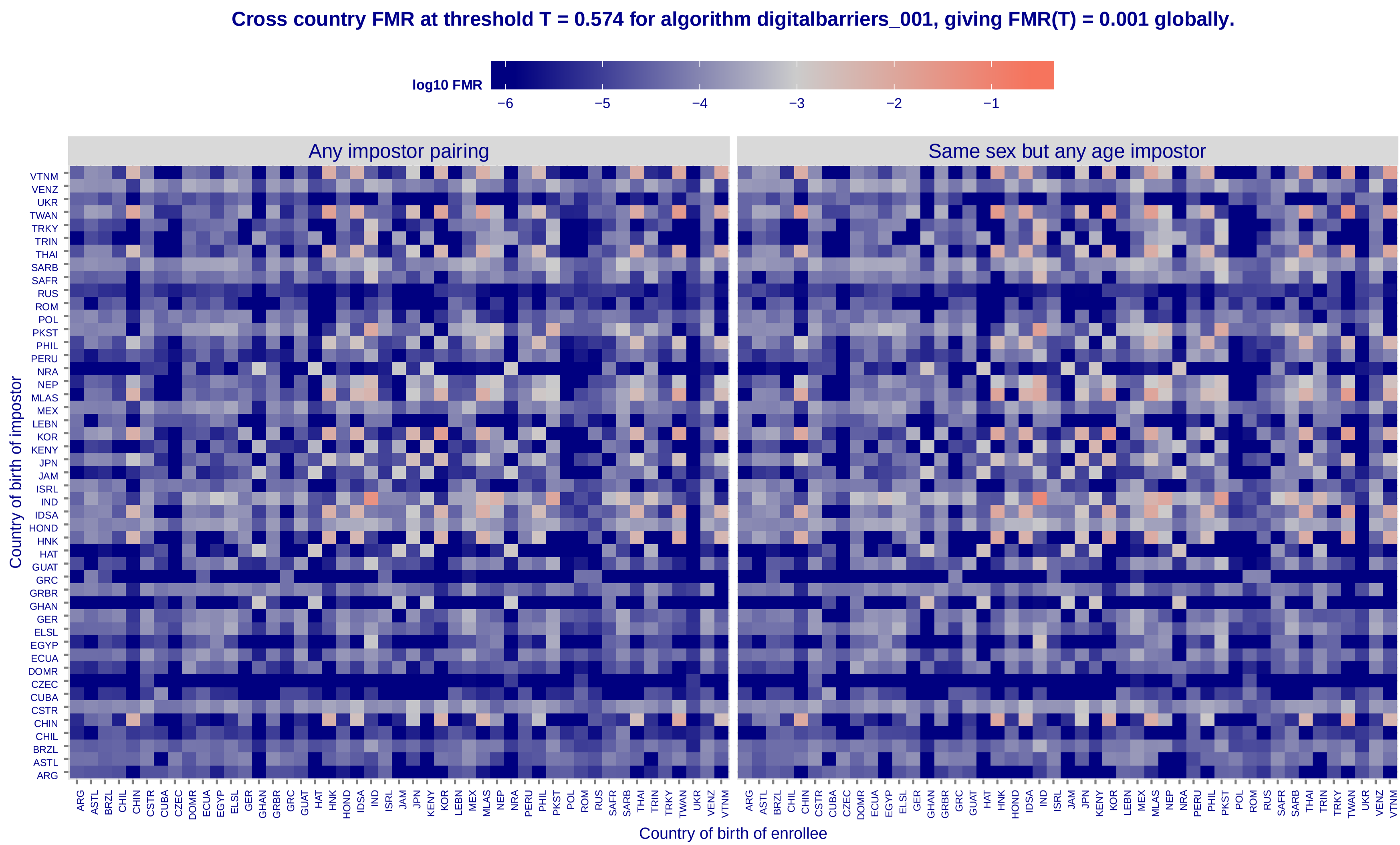


Figure 63: For algorithm digitalbarriers-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

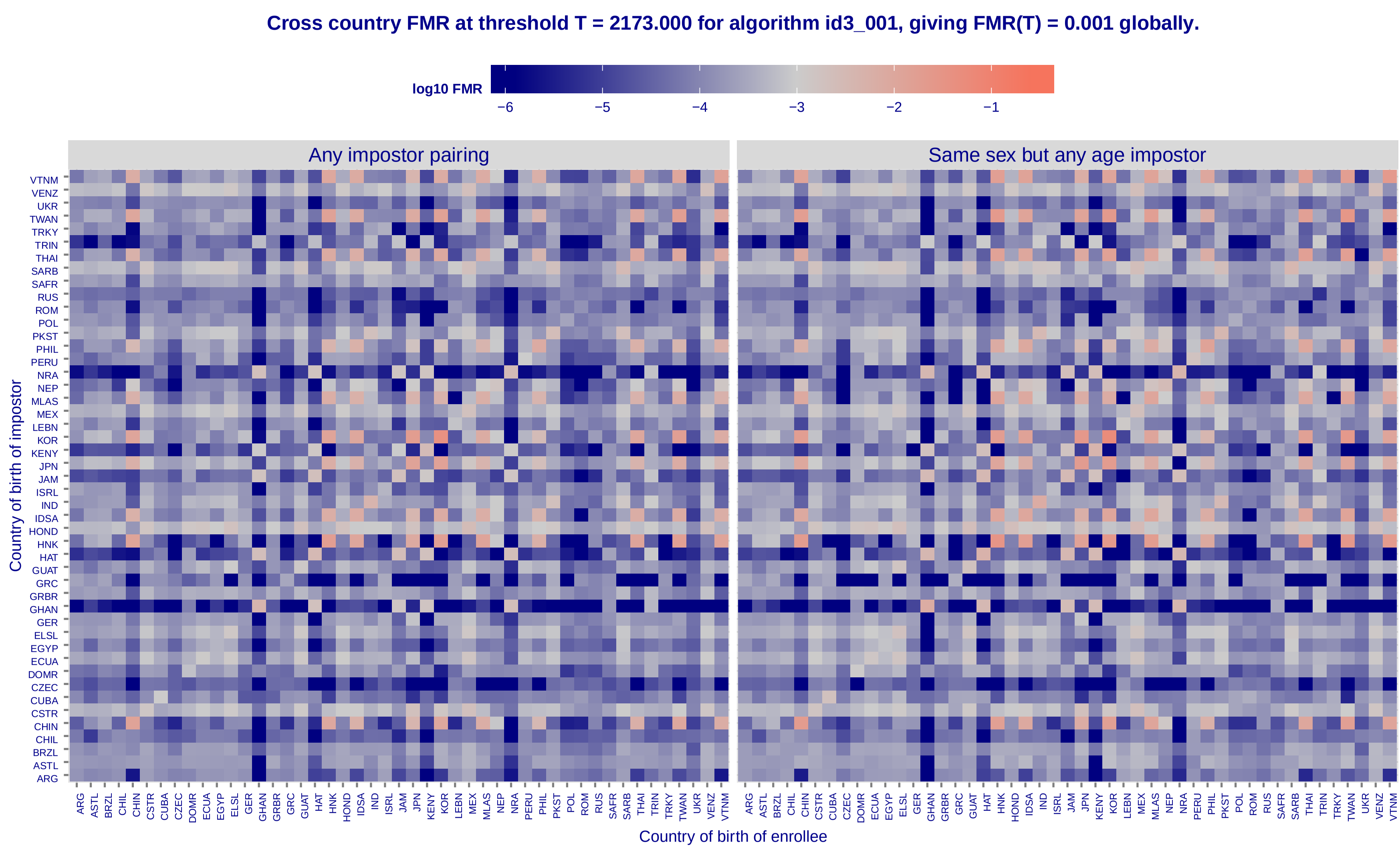


Figure 64: For algorithm id3-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

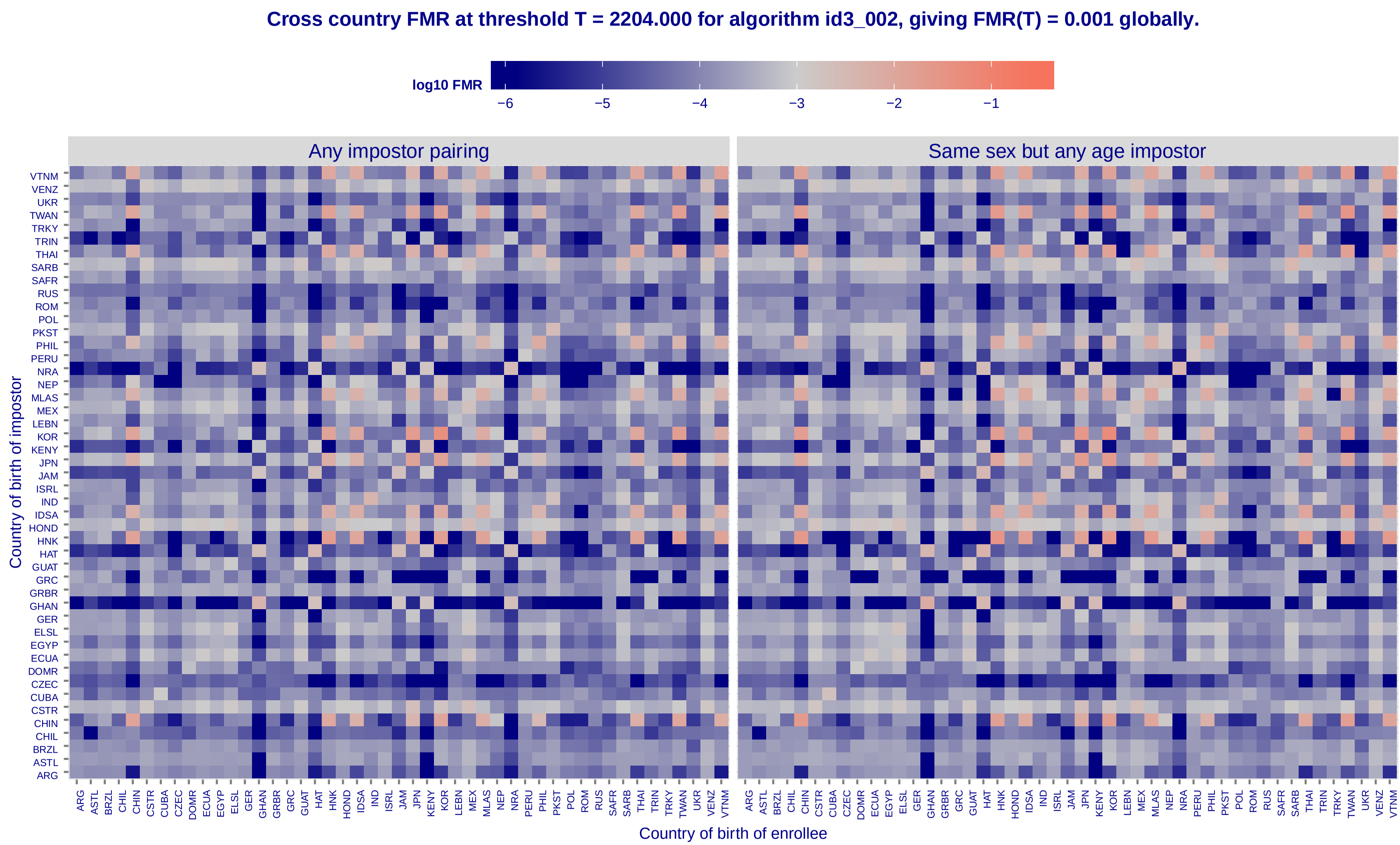


Figure 65: For algorithm id3-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

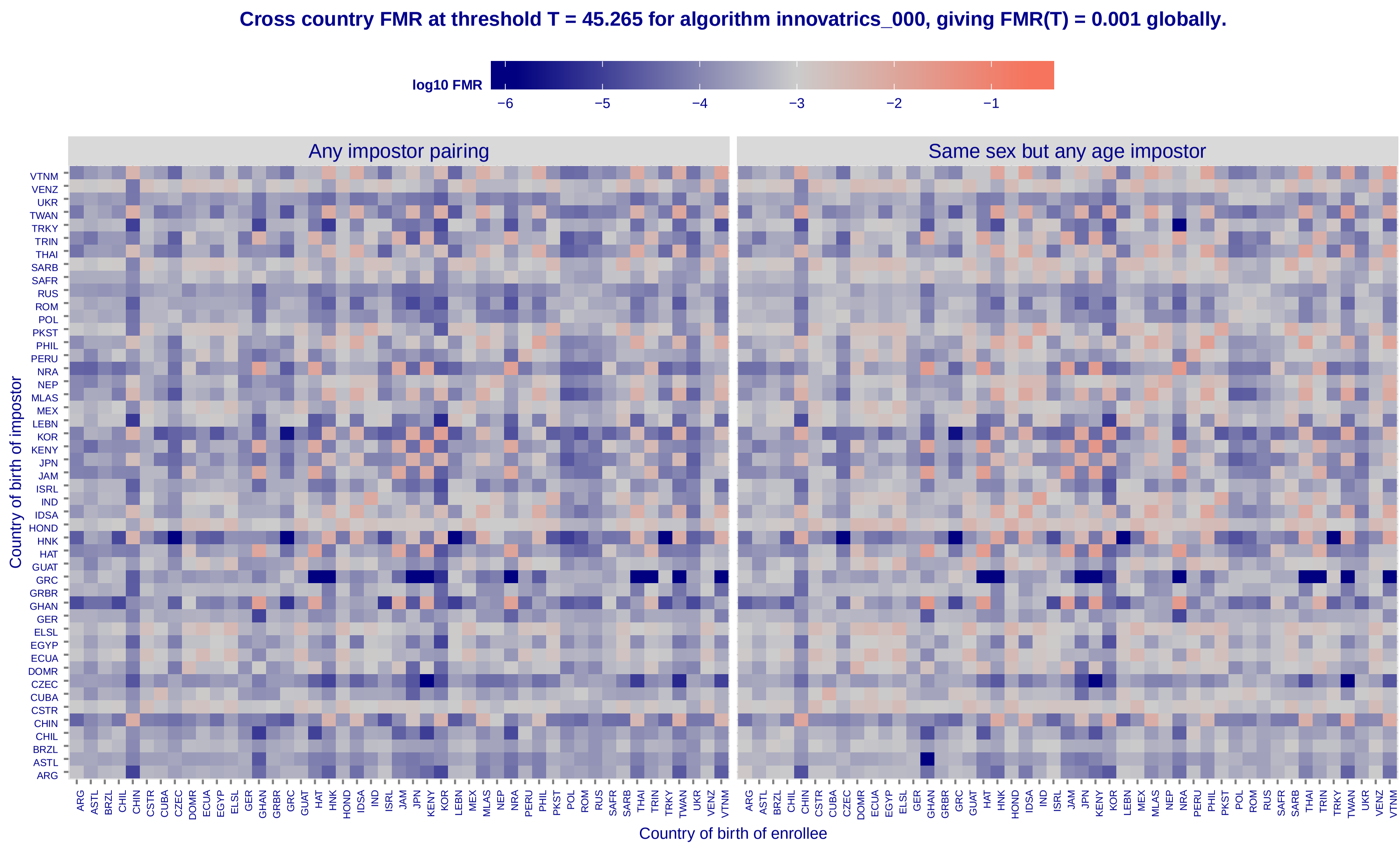


Figure 66: For algorithm innovatrics-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

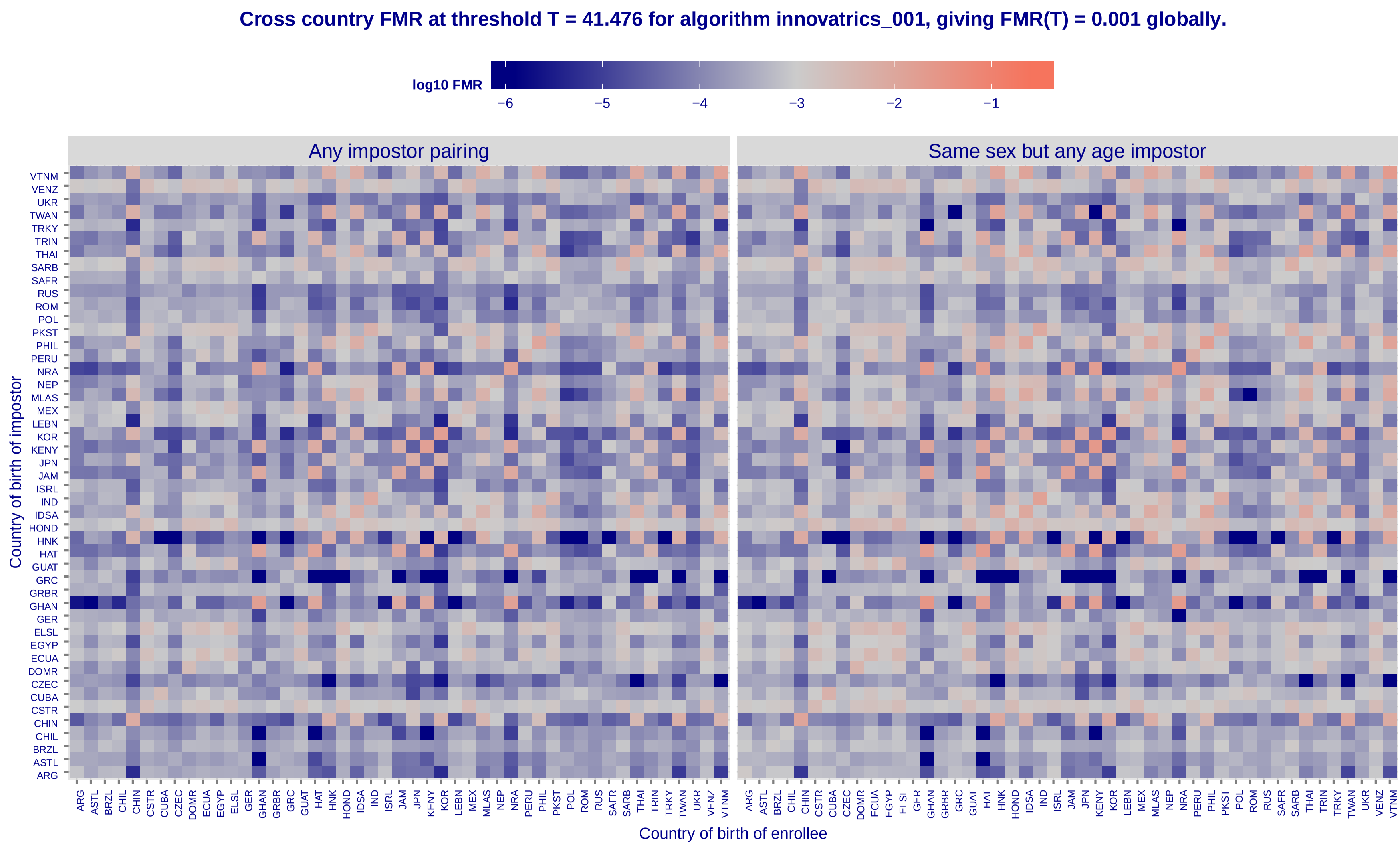


Figure 67: For algorithm innovatrics-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all O(10^{10}) impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

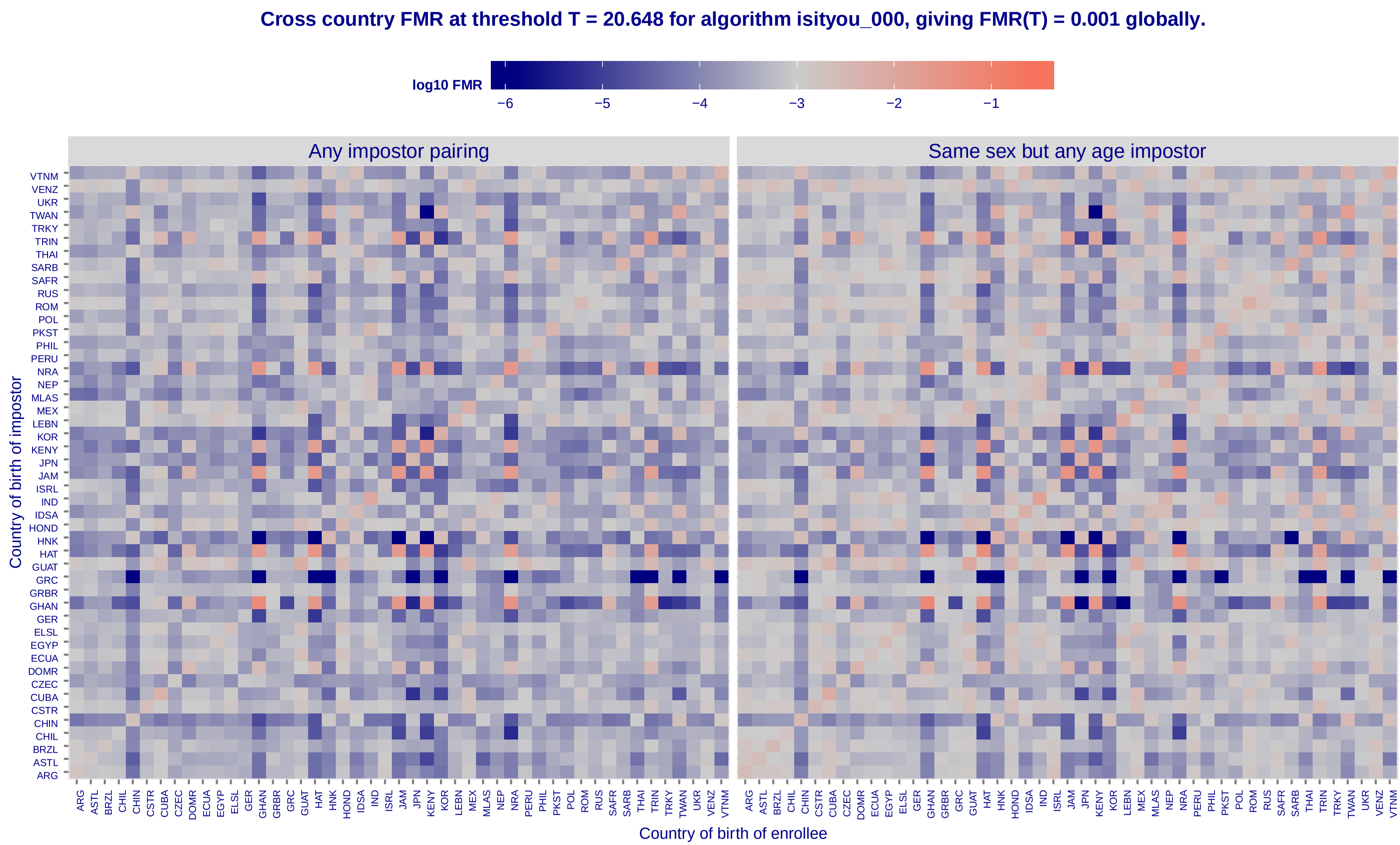


Figure 68: For algorithm isityou-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

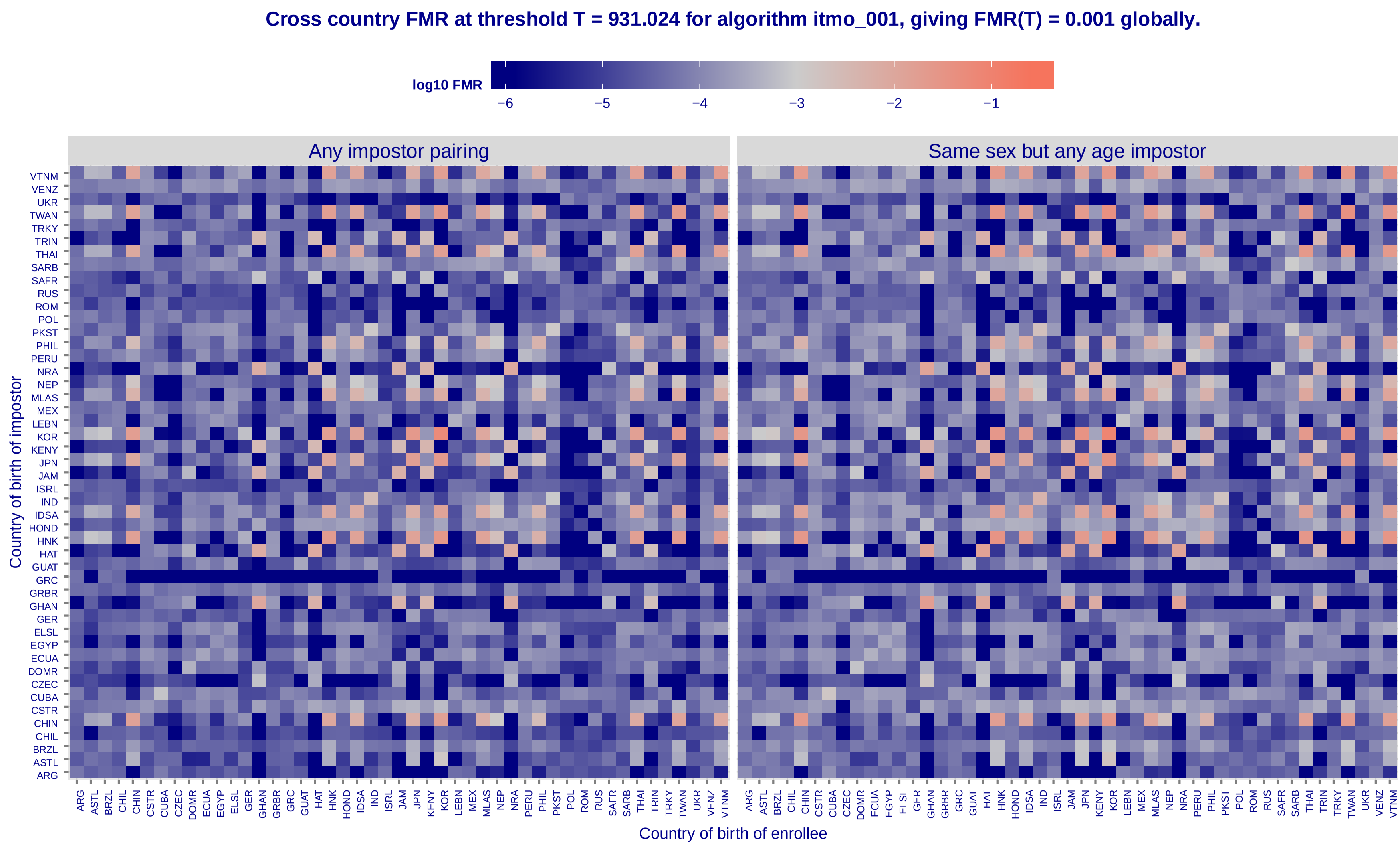


Figure 69: For algorithm itmo-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all O(10^{10}) impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

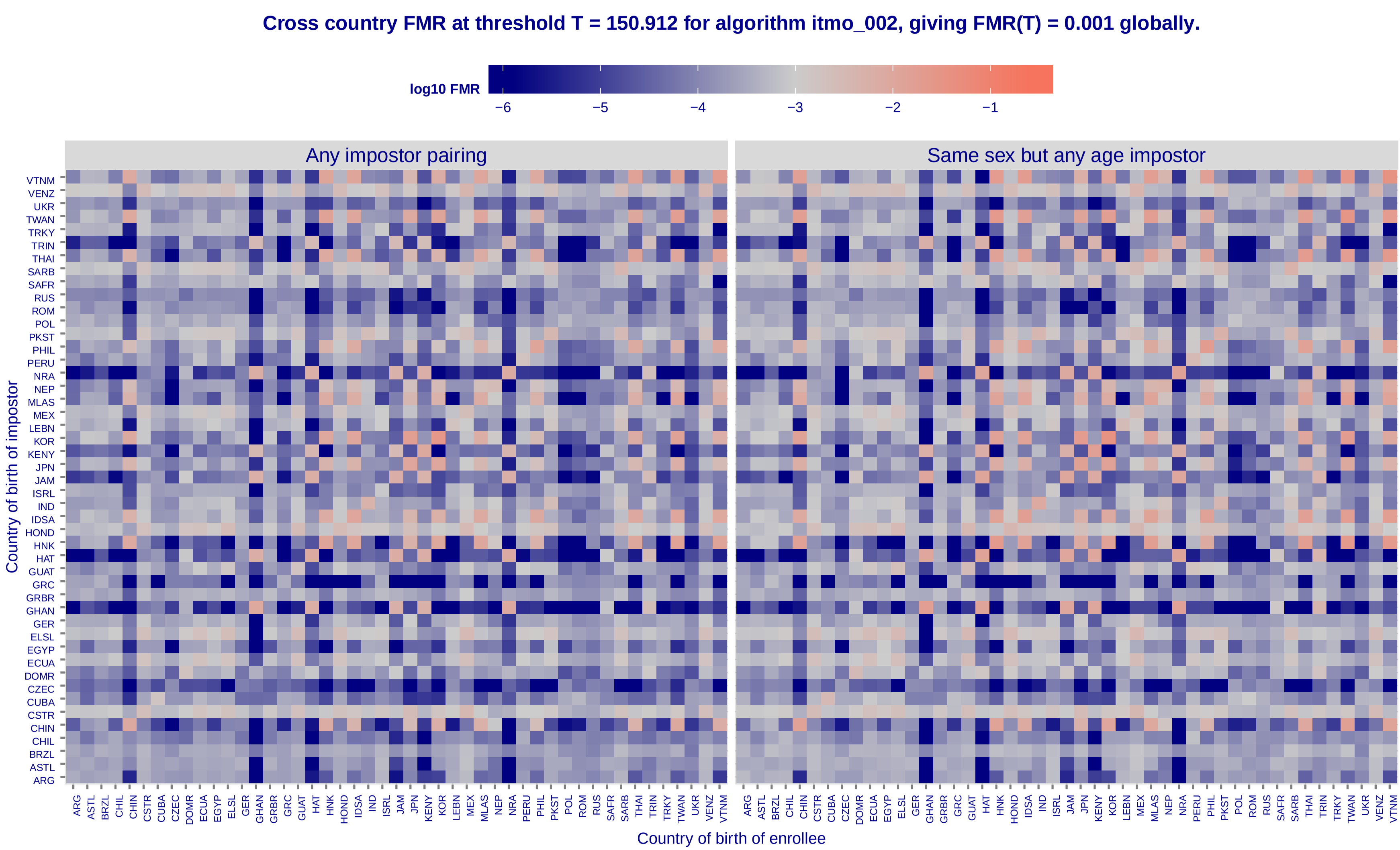


Figure 70: For algorithm itmo-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

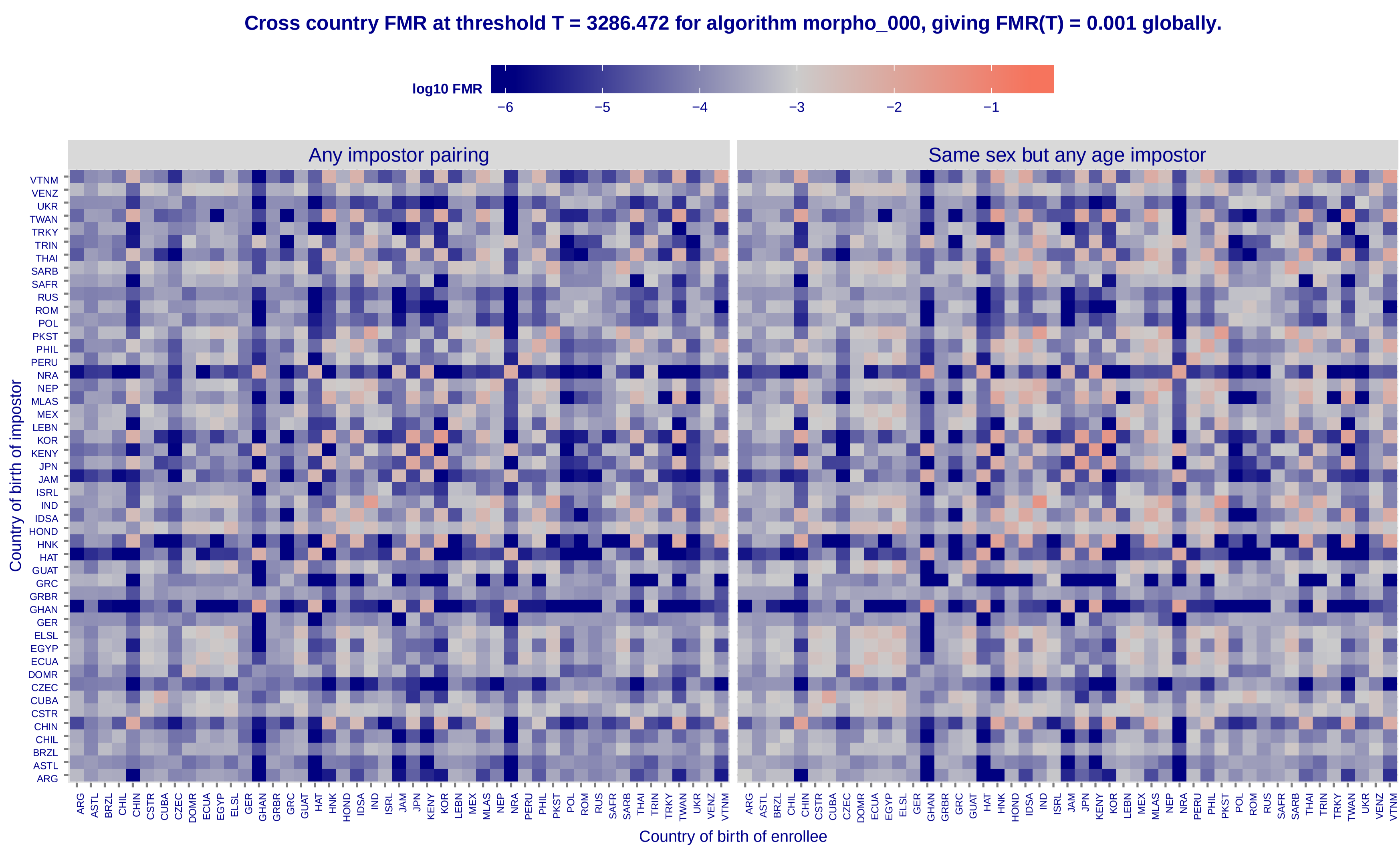


Figure 71: For algorithm morpho-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all O(10^{10}) impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

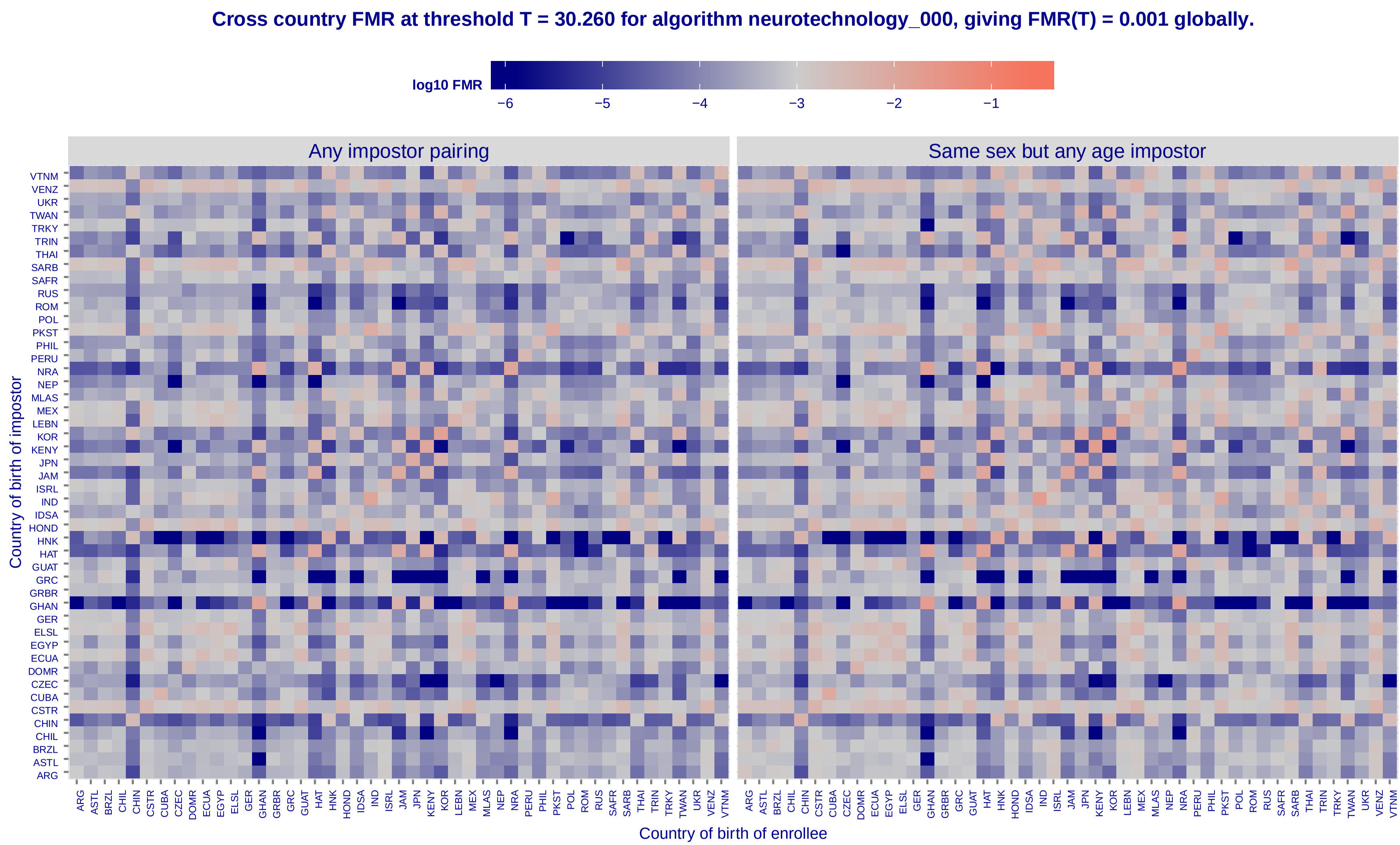


Figure 72: For algorithm neurotechnology-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

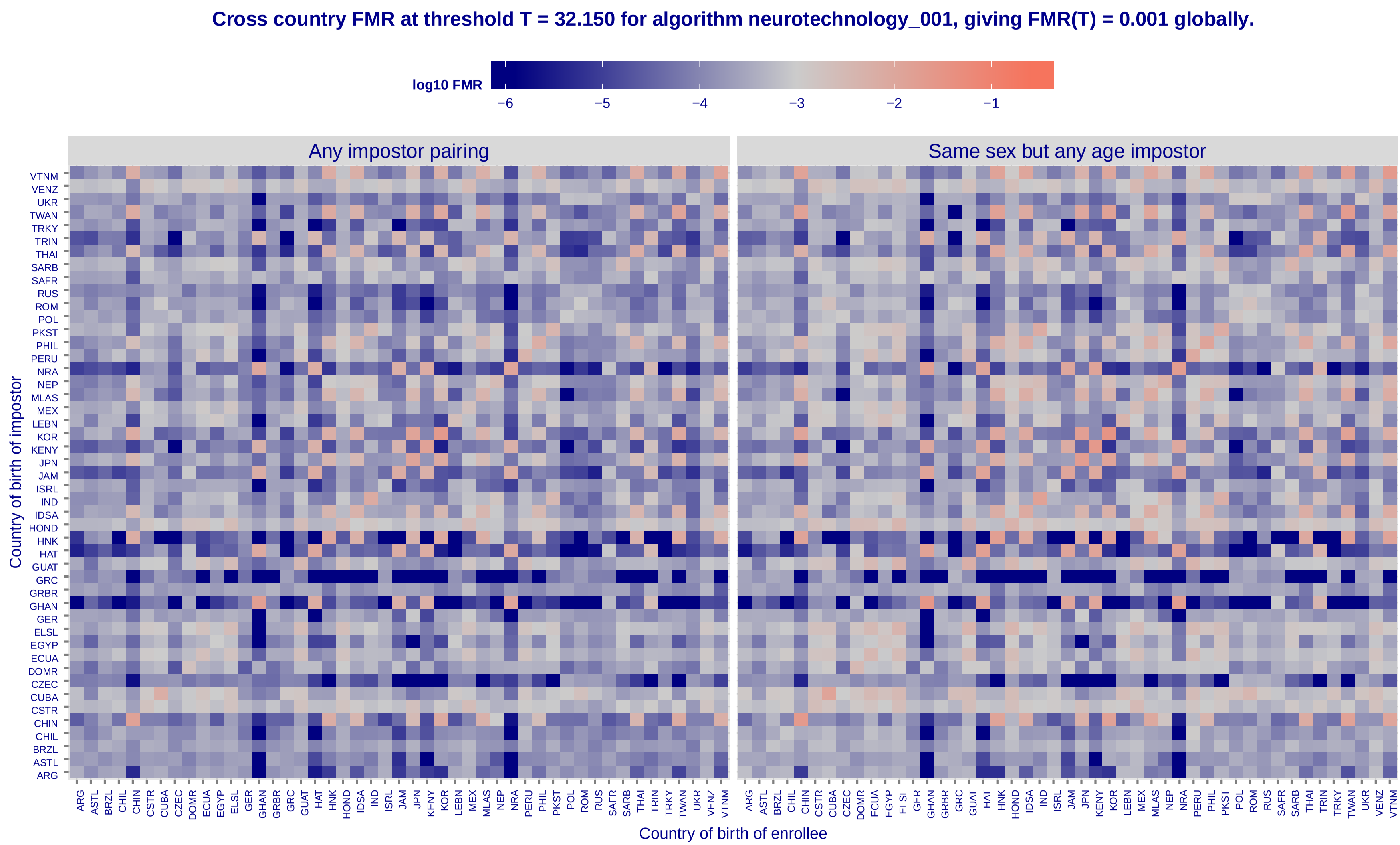


Figure 73: For algorithm neurotechnology-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

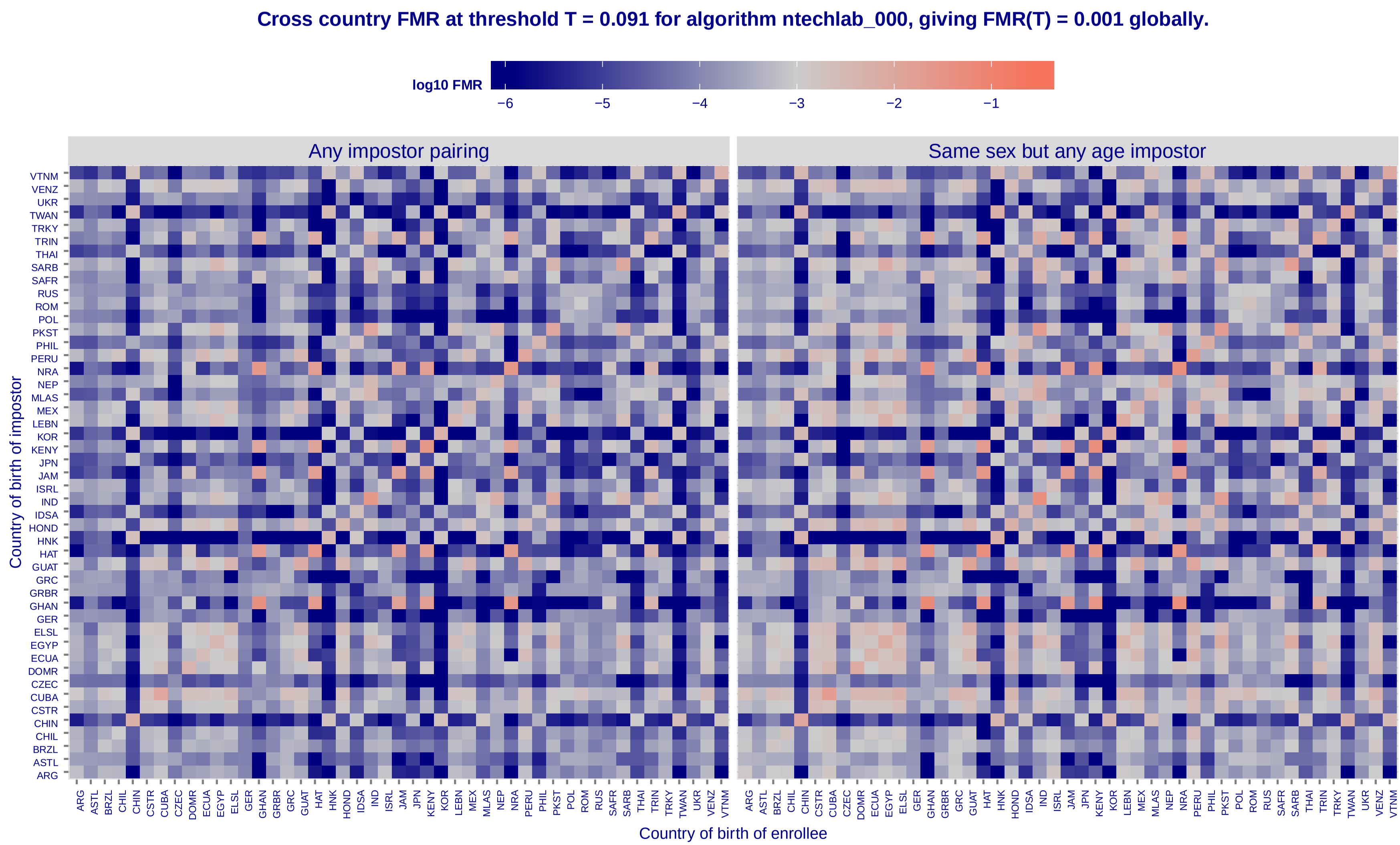


Figure 74: For algorithm ntechlab-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all O(10^{10}) impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

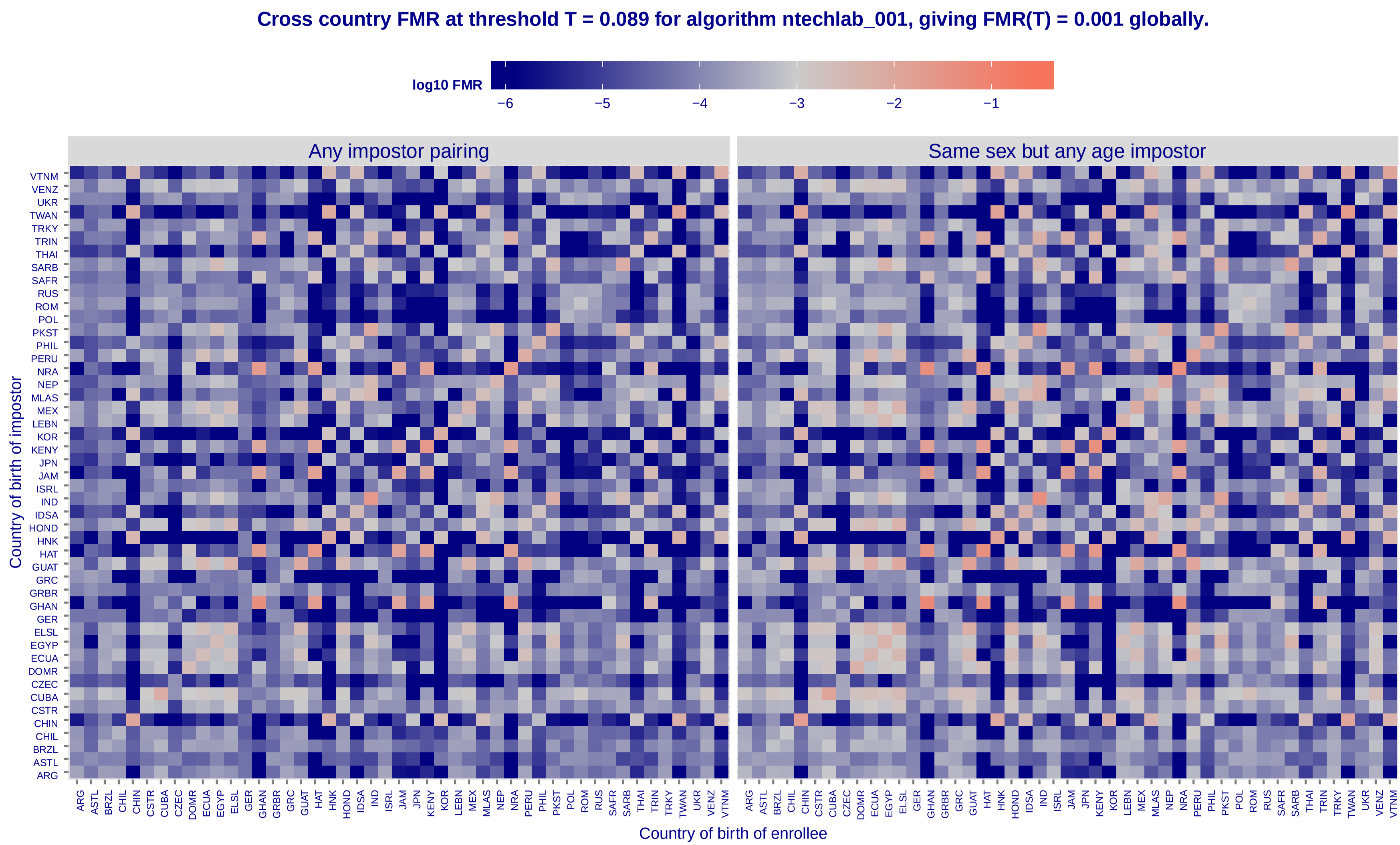


Figure 75: For algorithm ntechlab-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

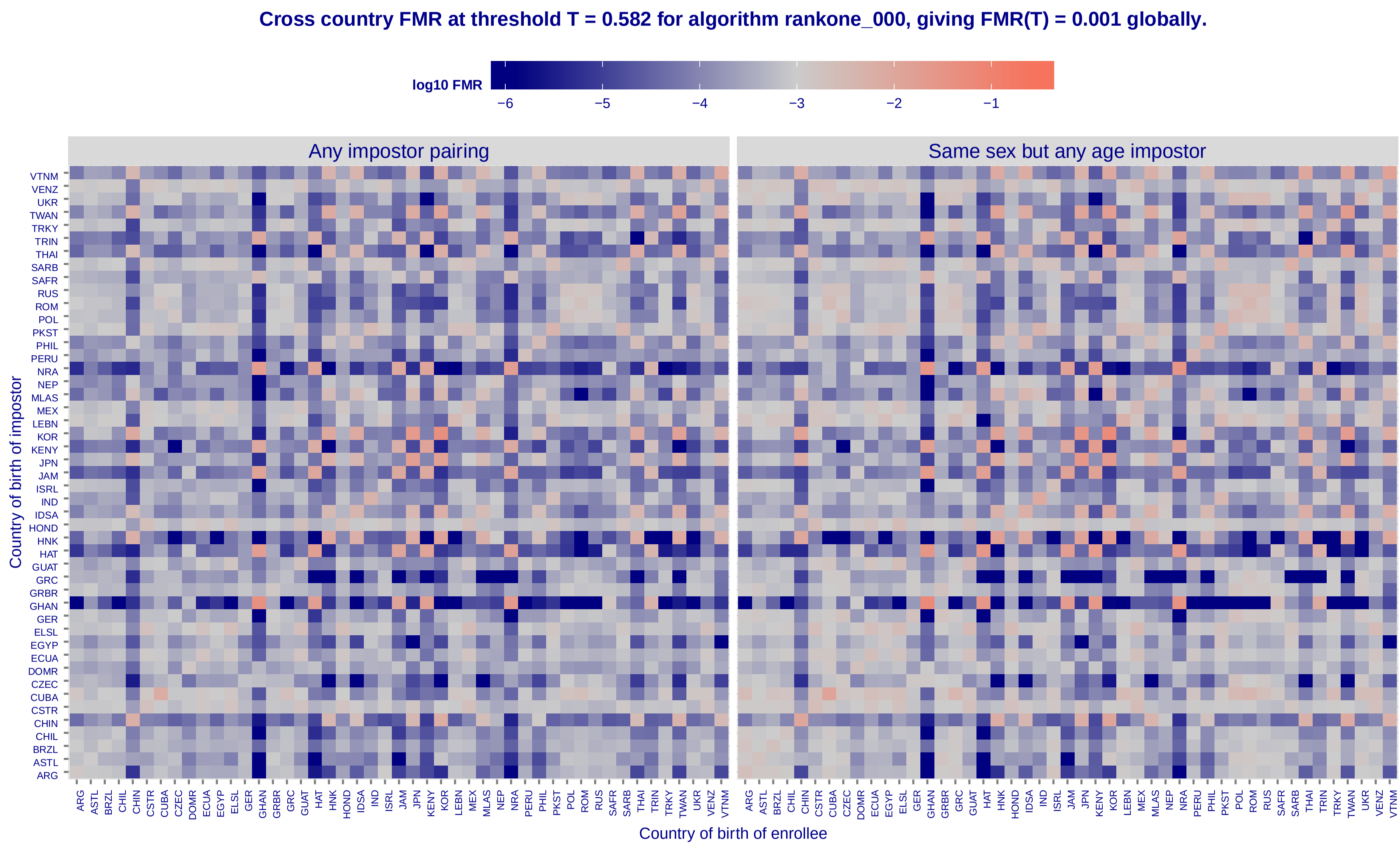


Figure 76: For algorithm rankone-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

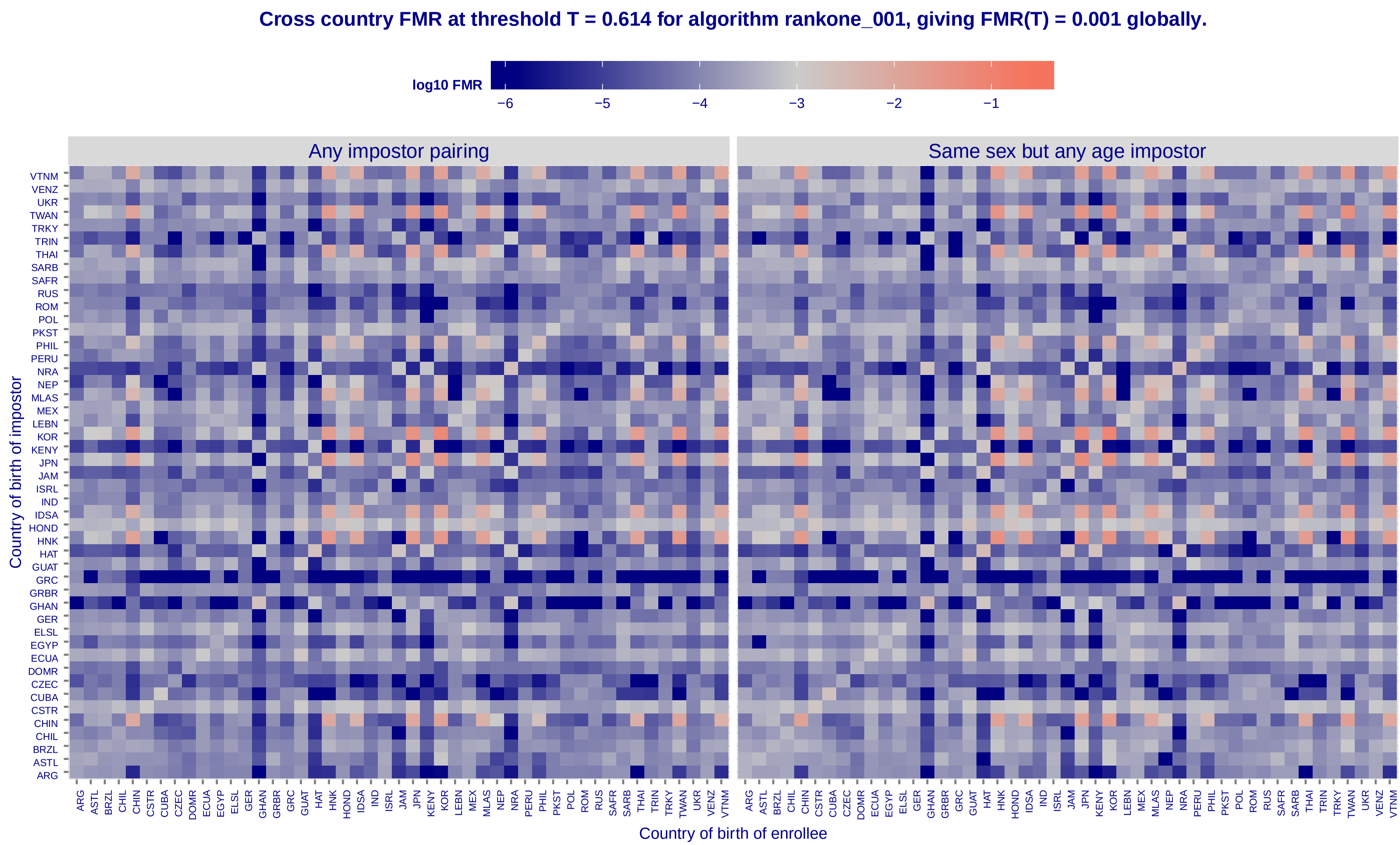


Figure 77: For algorithm rankone-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

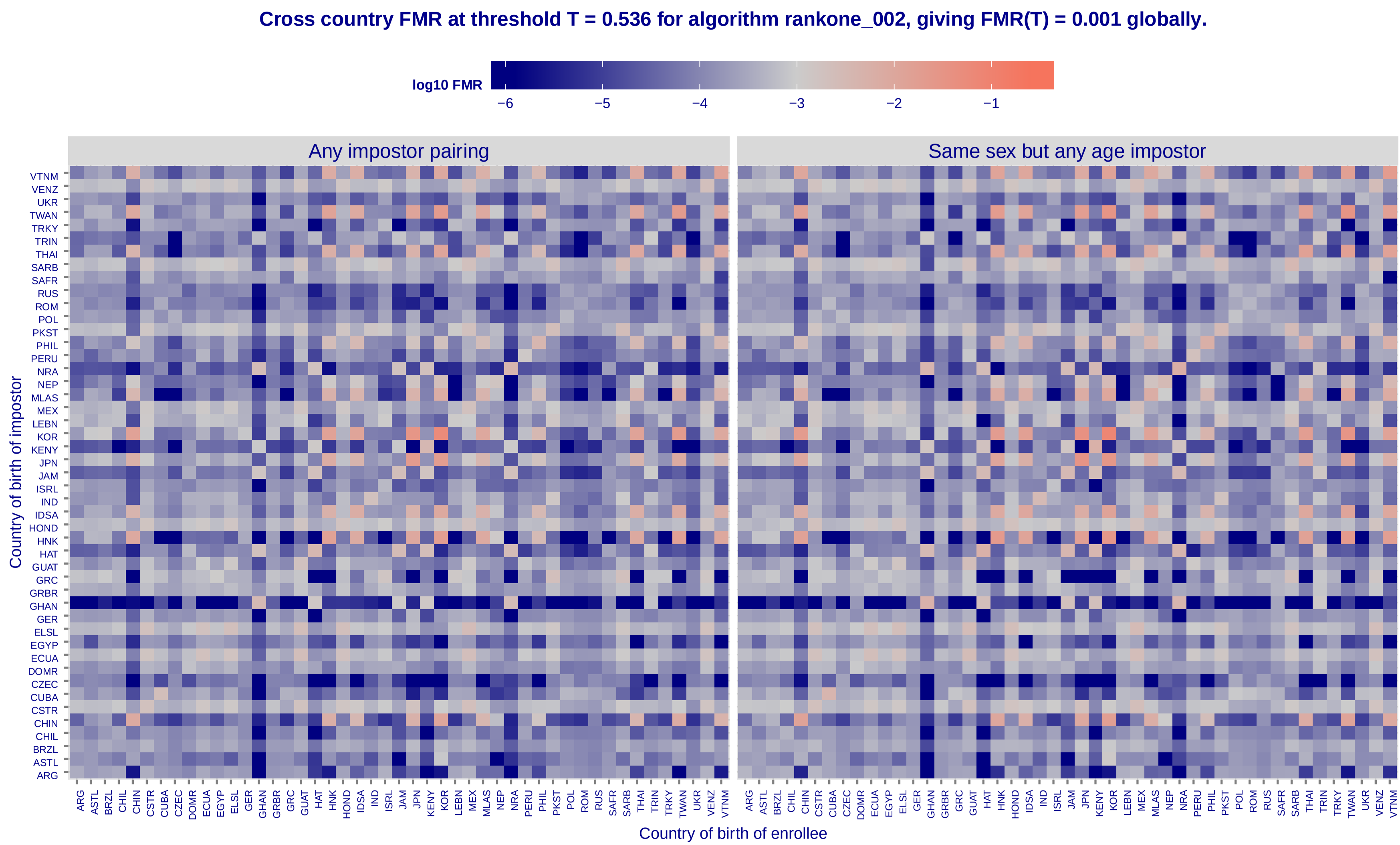


Figure 78: For algorithm rankone-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

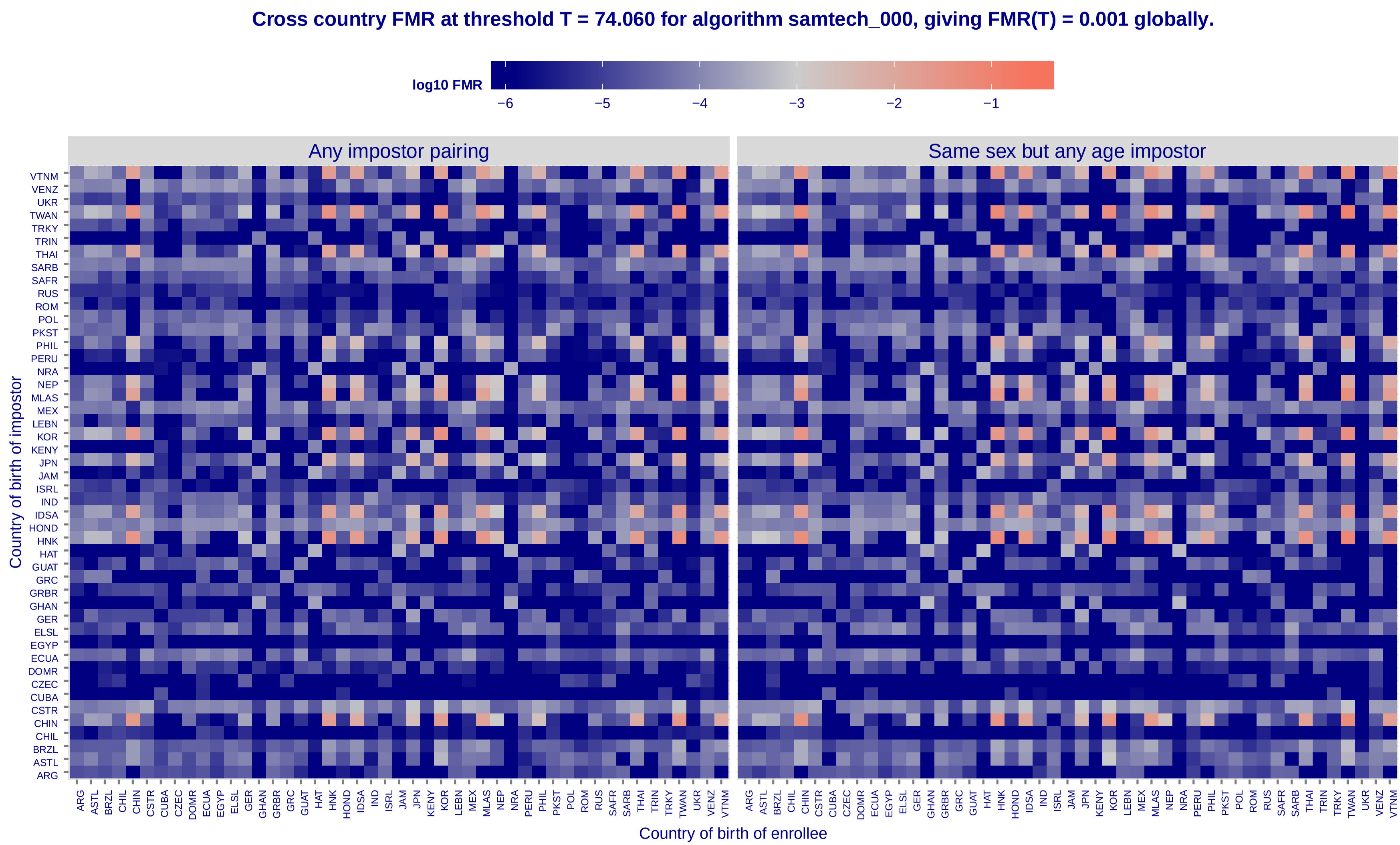


Figure 79: For algorithm samtech-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

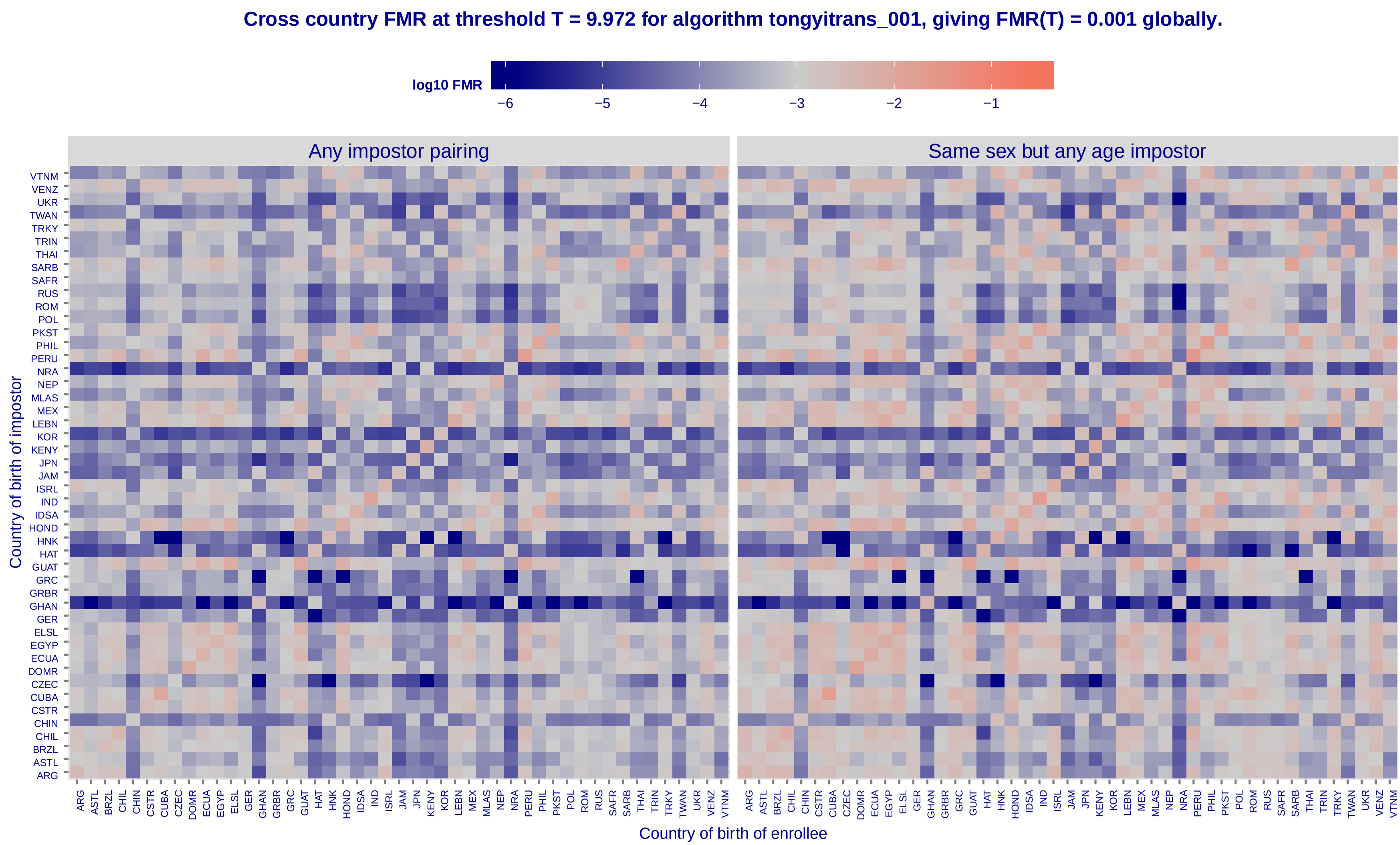


Figure 80: For algorithm tongyitrans-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

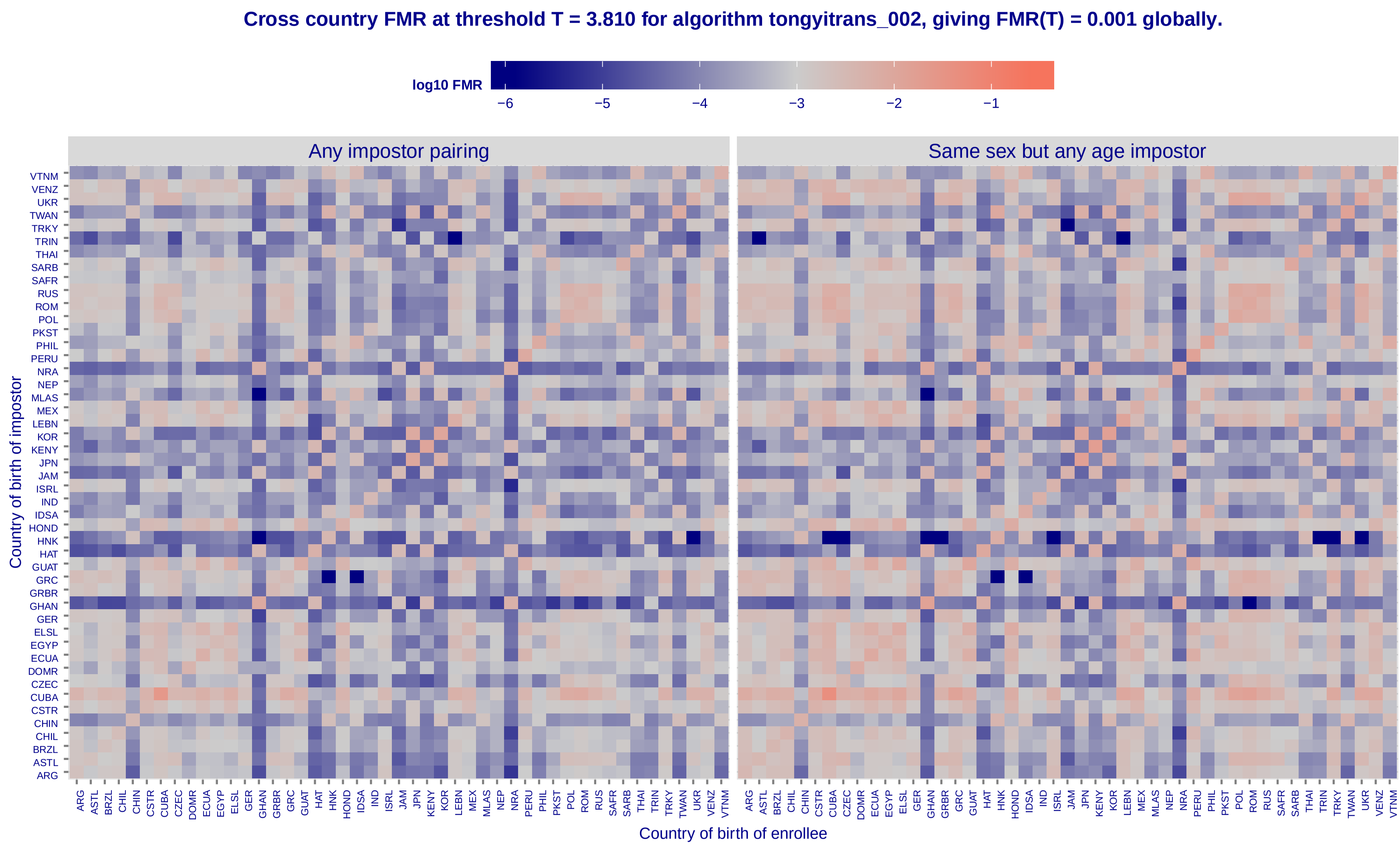


Figure 81: For algorithm tongyitrans-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

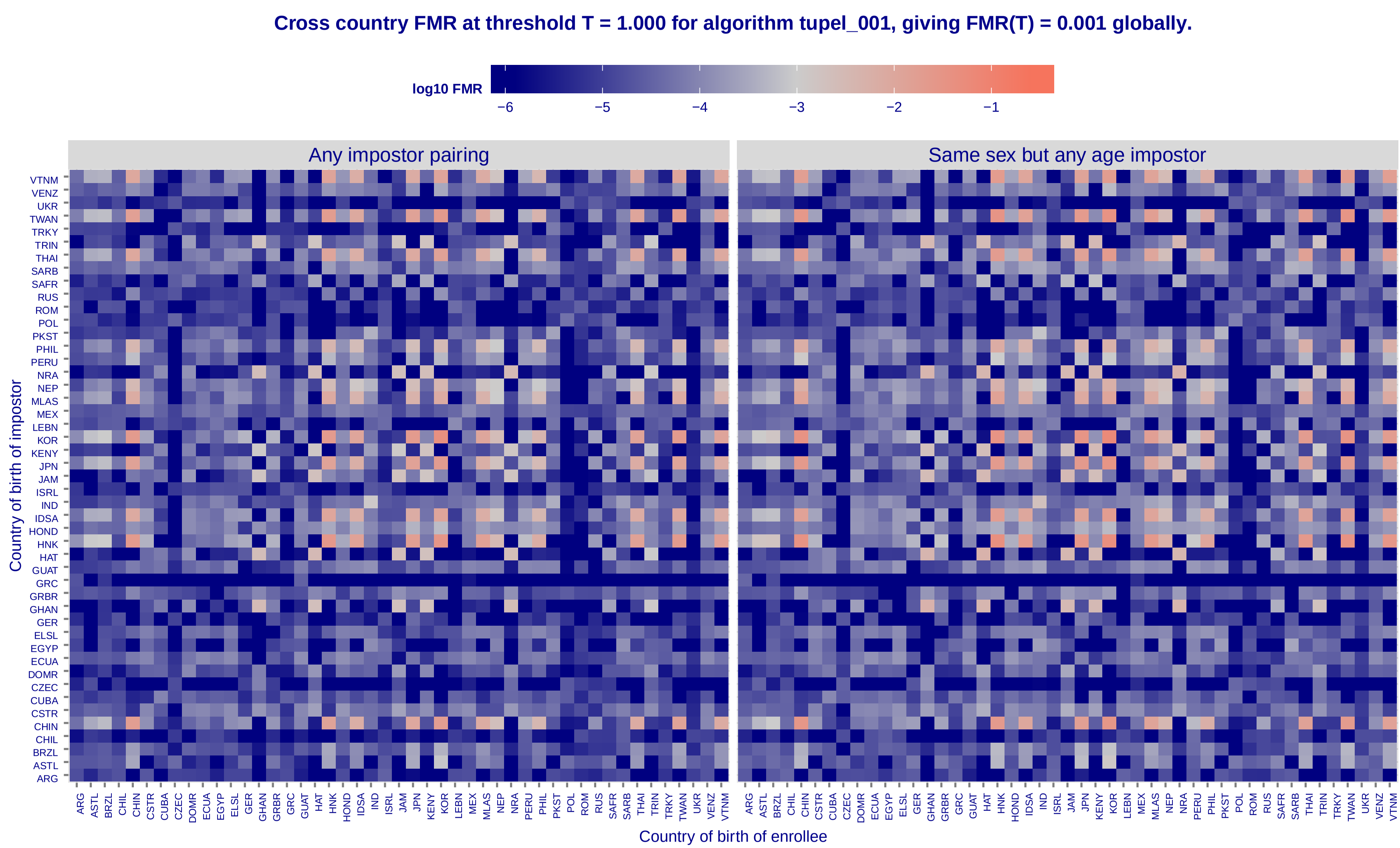


Figure 82: For algorithm tupel-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

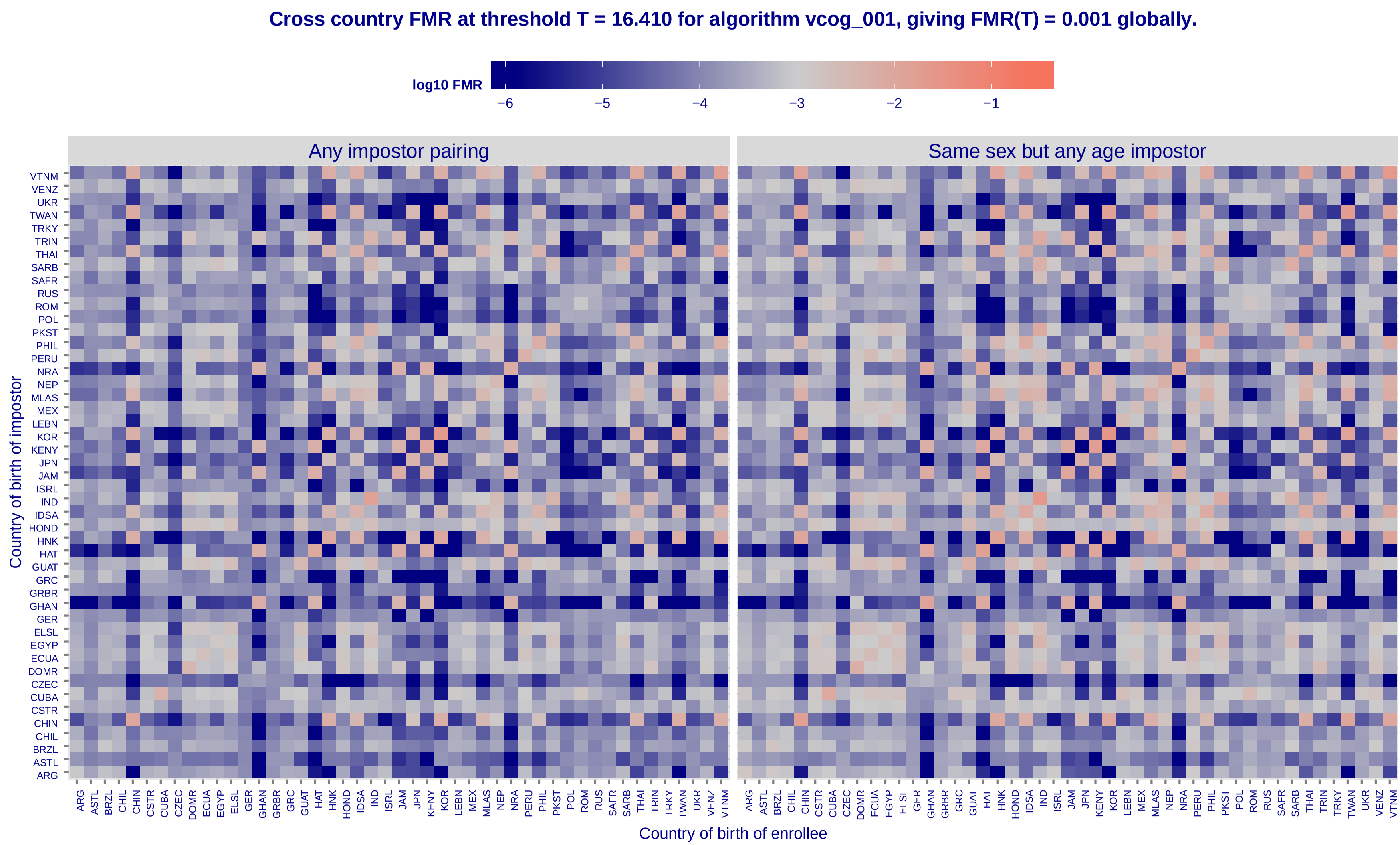


Figure 83: For algorithm vcog-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all O(10^{10}) impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

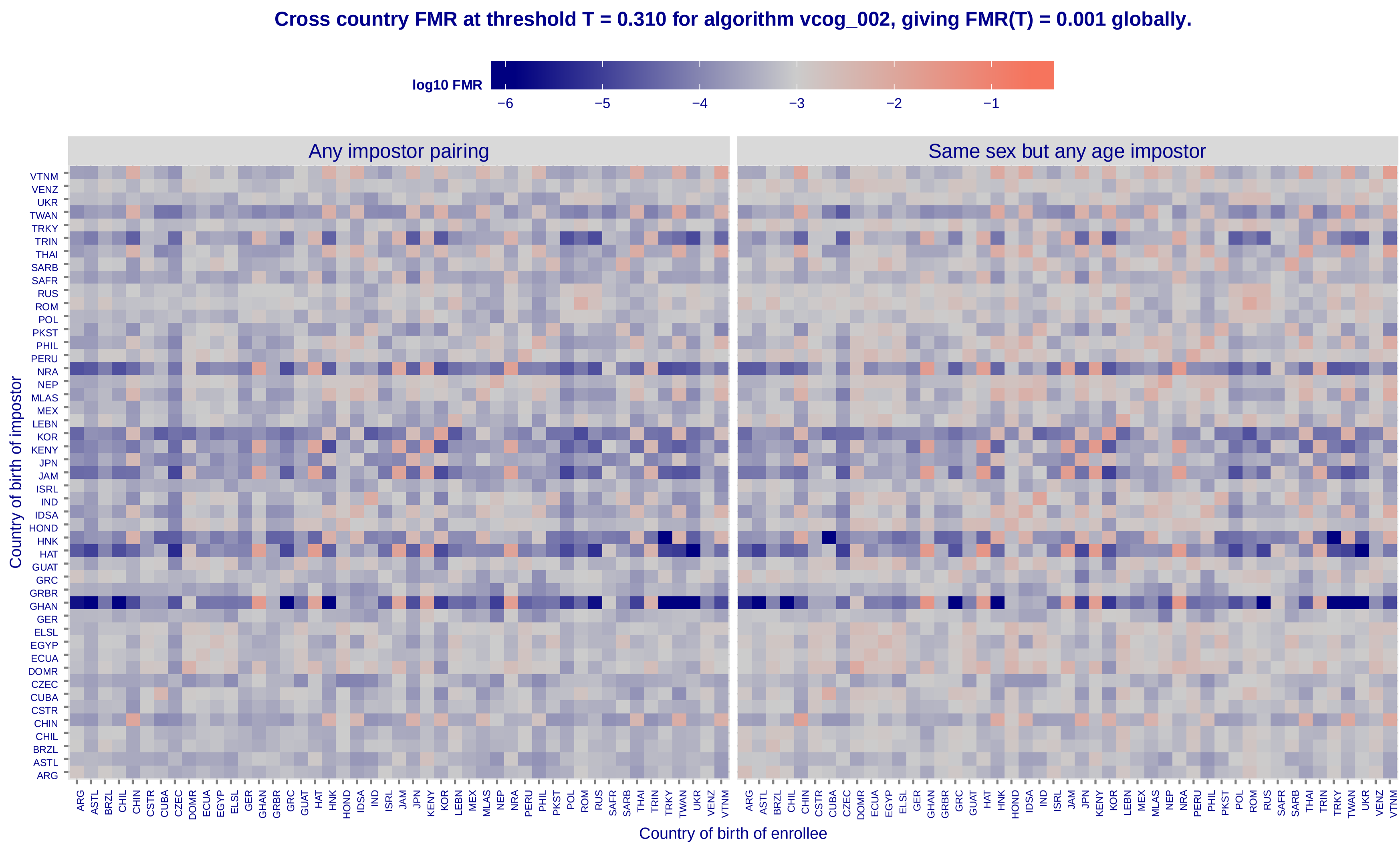


Figure 84: For algorithm vcog-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all O(10^{10}) impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

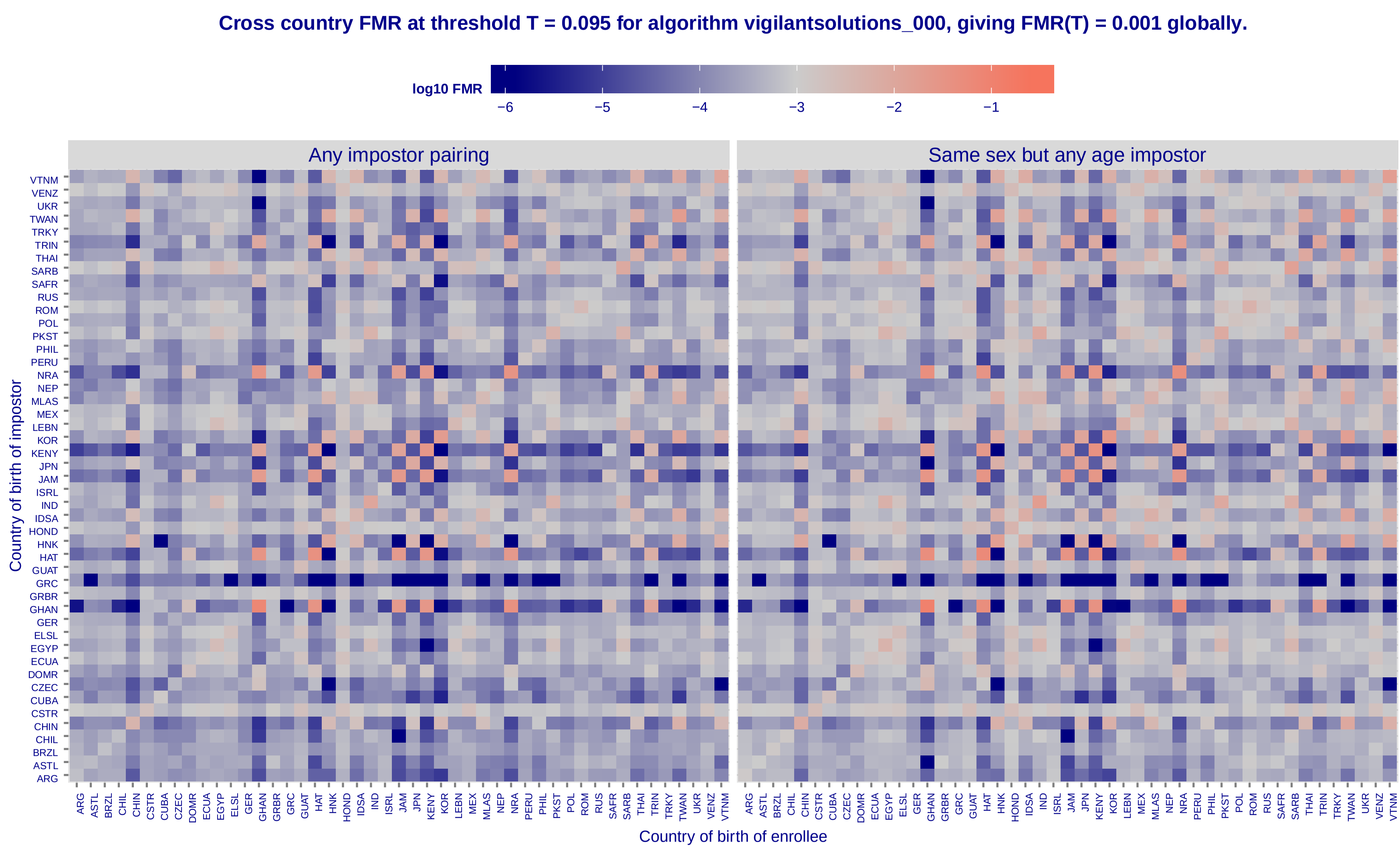


Figure 85: For algorithm vigilantsolutions-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

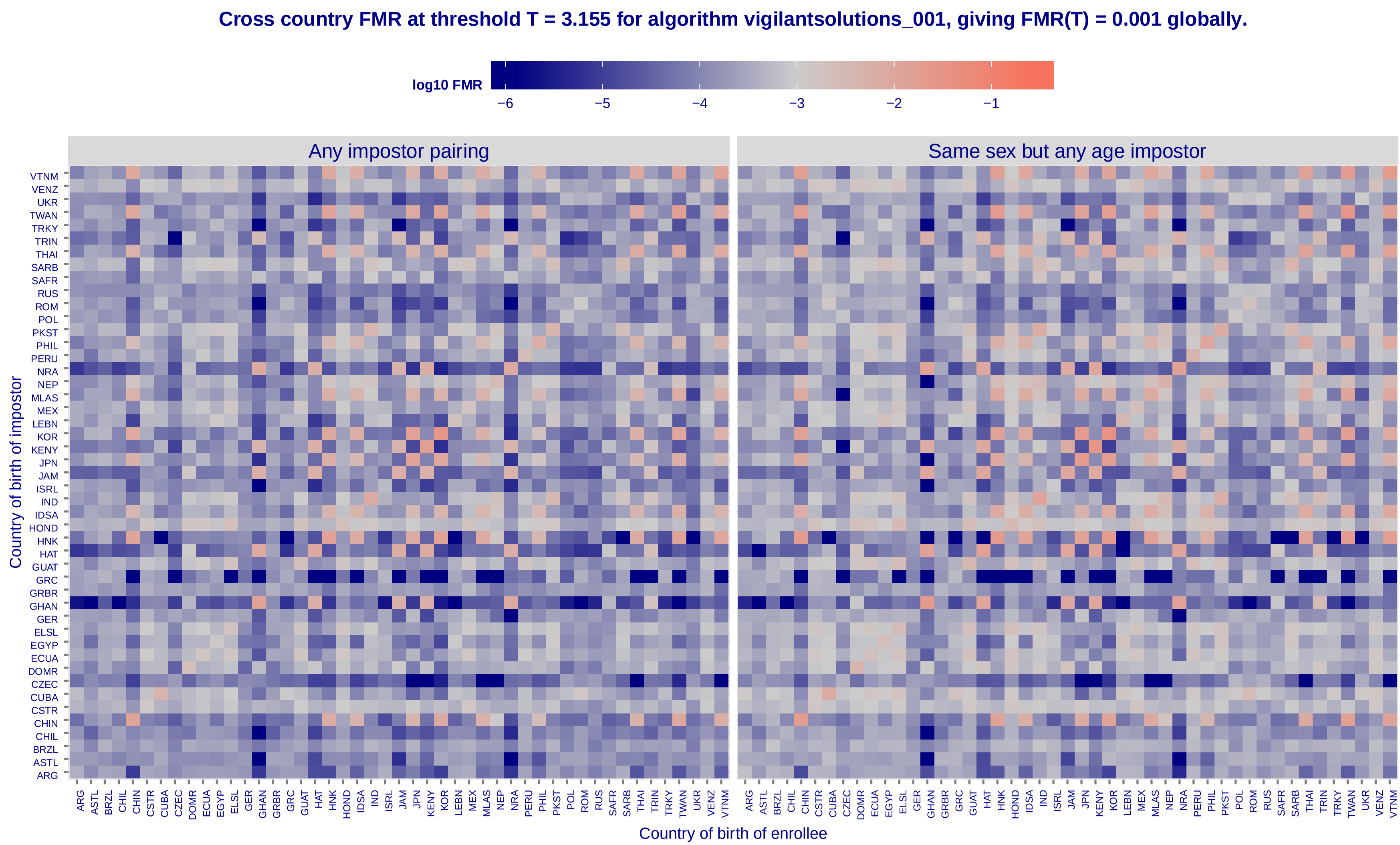


Figure 86: For algorithm vigilantsolutions-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

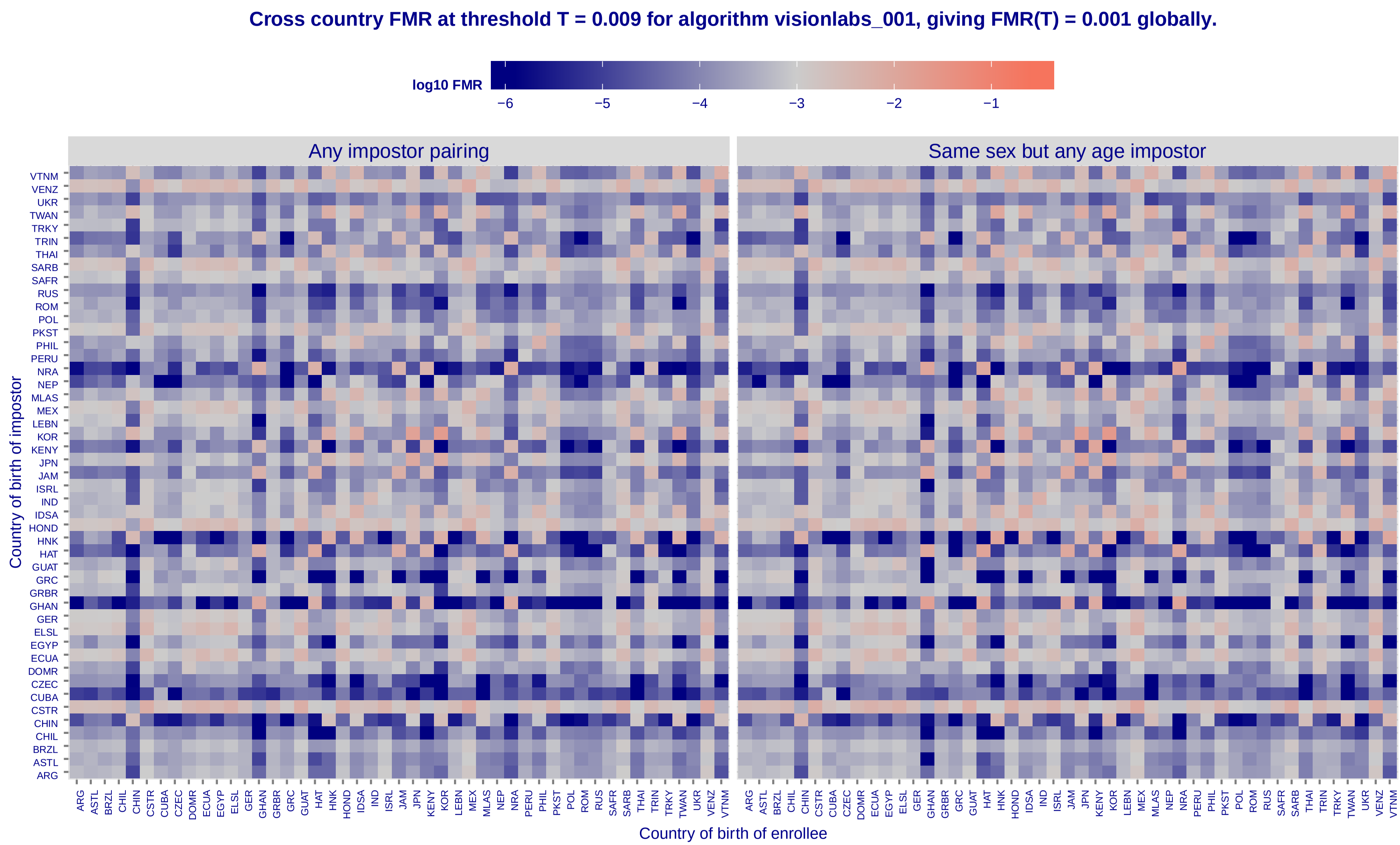


Figure 87: For algorithm visionlabs-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

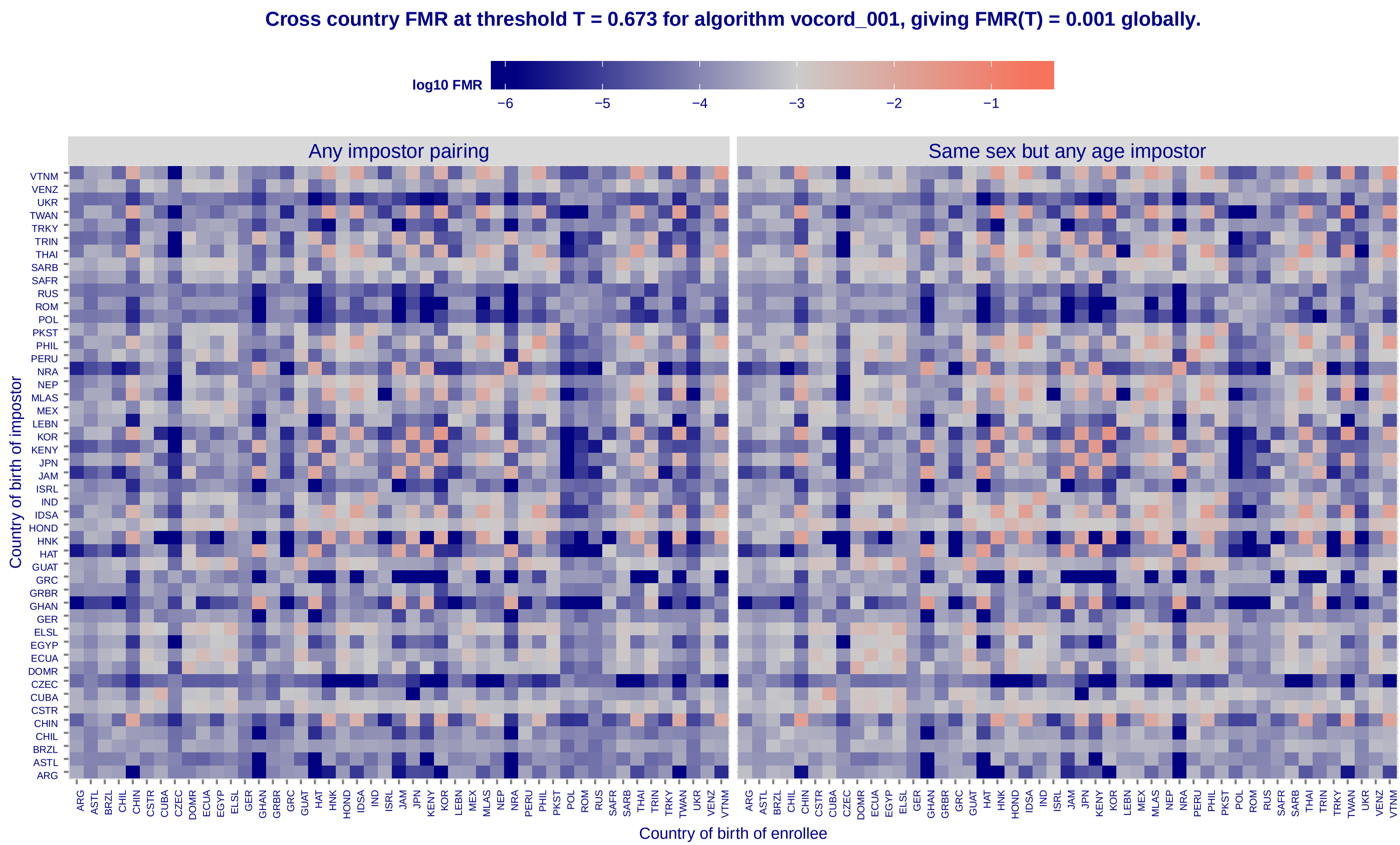


Figure 88: For algorithm vocord-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

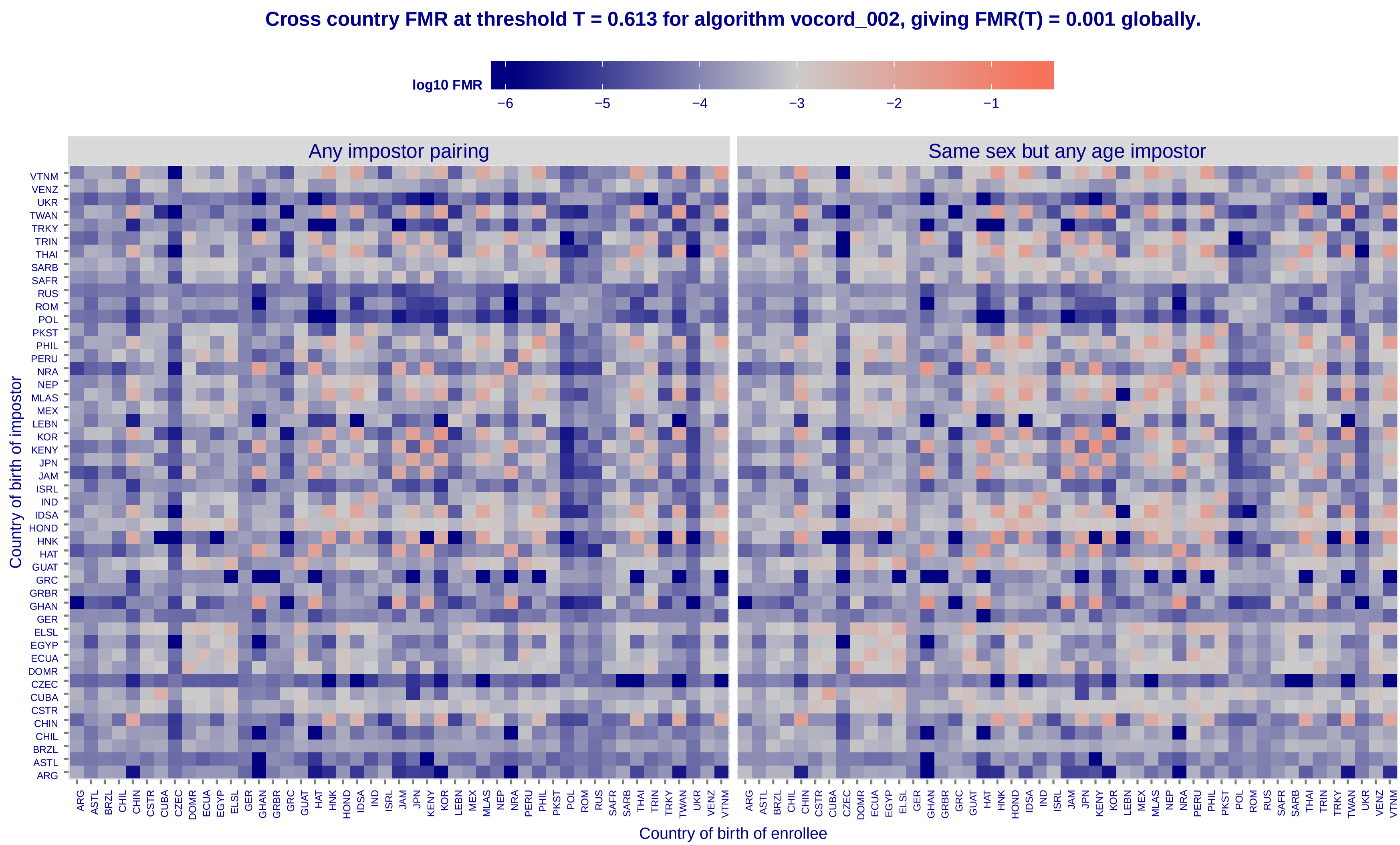


Figure 89: For algorithm vocord-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all $O(10^{10})$ impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

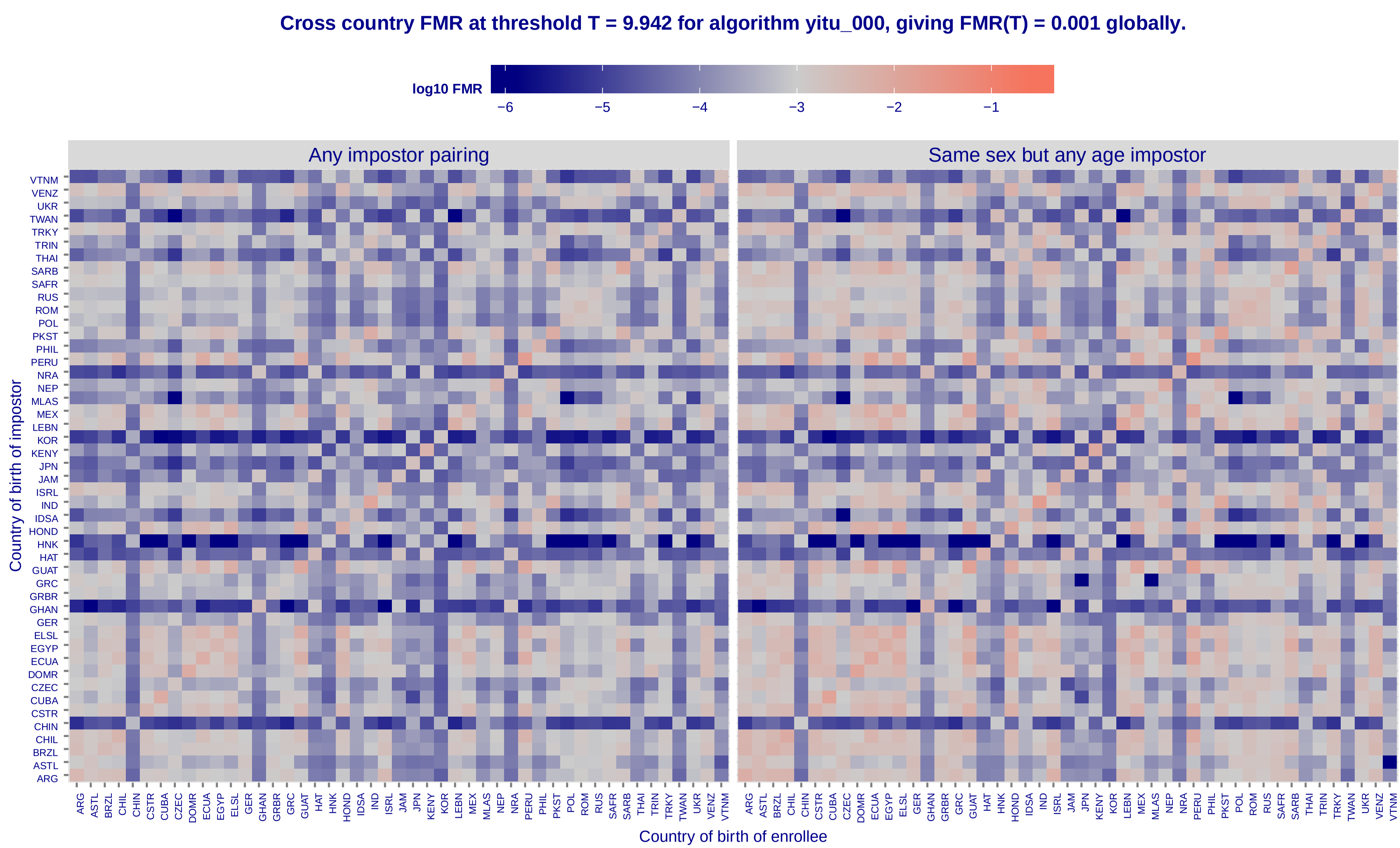


Figure 90: For algorithm yitu-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR in the plot title, computed over all O(10^{10}) impostor comparisons. If text appears in each box it give the same quantity as that coded by the color. Grey indicates FMR is at the intended FMR target level. Light red colors present a security vulnerability to, for example, a passport gate. Each +1 increase in log10 FMR corresponds to a factor of 10 increase in FMR. The matrix is not quite symmetric because images in the enrollment and verification sets are different.

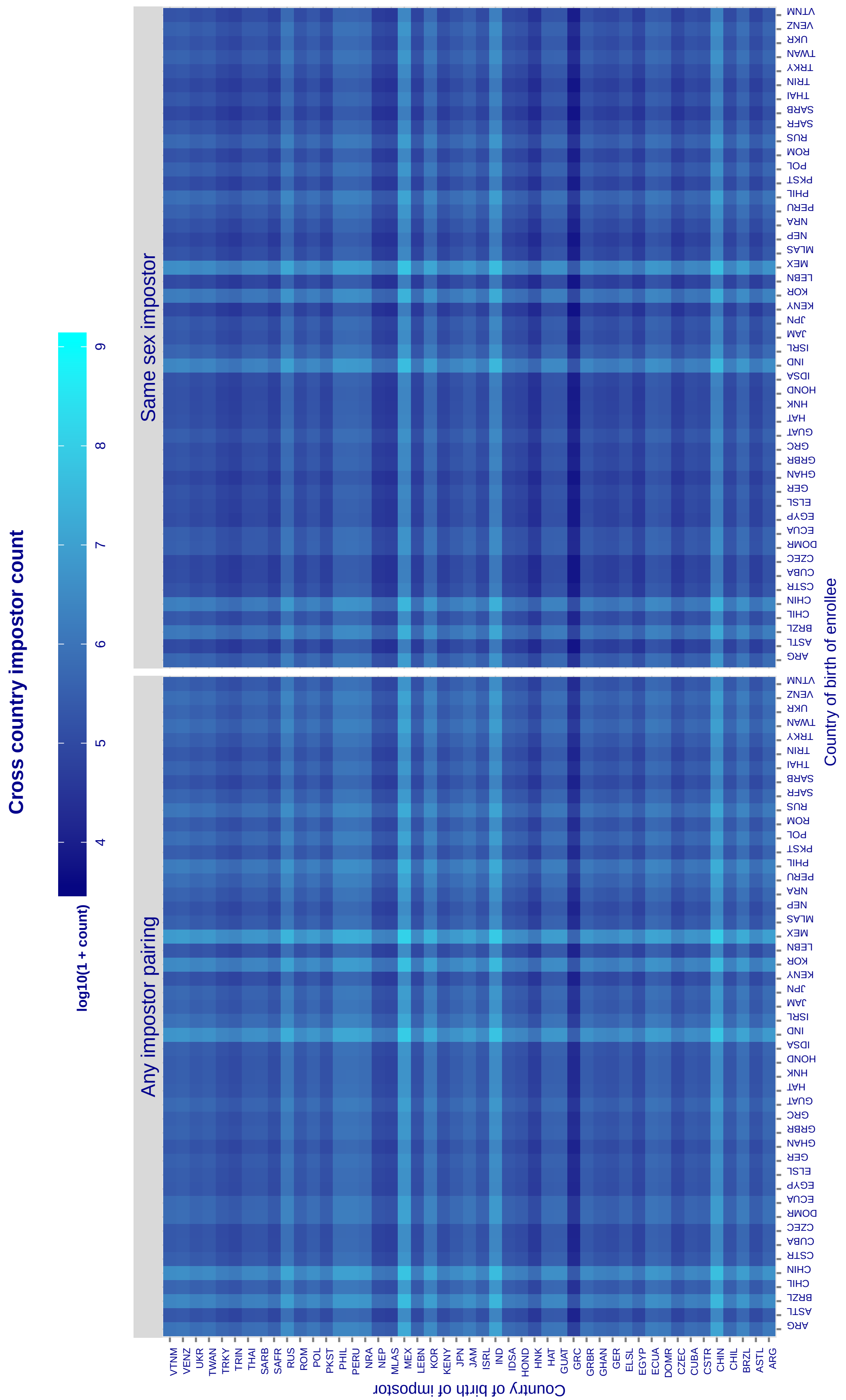


Figure 91: For visa images, the heatmap shows the count of impostor comparisons of faces from different individuals who were born in the given country pair.

4.6.2 Effect of age on impostors

Background: This section shows the effect of age on the impostor distribution. The ideal behaviour is that the age of the enrollee and the impostor would not affect impostor scores. This would support FMR stability over sub-populations.

Goals:

- ▷ To show the effect of relative ages of the impostor and enrollee on false match rates.
- ▷ To determine whether some algorithms have better impostor distribution stability.

Methods:

- ▷ Define 14 age group bins, spanning 0 to over 100 years old.
- ▷ Compute FMR over all impostor comparisons for which the subjects in the enrollee and impostor images have ages in two bins.
- ▷ Compute FMR over all impostor comparisons for which the subjects are additionally of the same sex, and born in the same geographic region.

Results:

The notable aspects are:

- ▷ Diagonal dominance: Impostors are more likely to be matched against their same age group.
- ▷ Same sex and same region impostors are more successful. On the diagonal, an impostor is more likely to succeed by posing as someone of the same sex. If $\Delta \log_{10} \text{FMR} = 0.2$, then same-sex same-region FMR exceeds the all-pairs FMR by factor of $10^{0.2} = 1.6$.
- ▷ Young children impostors give elevated FMR against young children. Older adult impostor give elevated FMR against older adults. These effects are quite large, for example if $\Delta \log_{10} \text{FMR} = 1.0$ larger than a 32 year old, then these groups have higher FMR by a factor of $10^1 = 10$. This would imply an FMR above 0.01 for a nominal (global) FMR = 0.001.
- ▷ Algorithms vary.
- ▷ We computed the same quantities for a global FMR = 0.0001. The effects are similar.

Note the calculations in this section include impostors paired across all countries of birth.

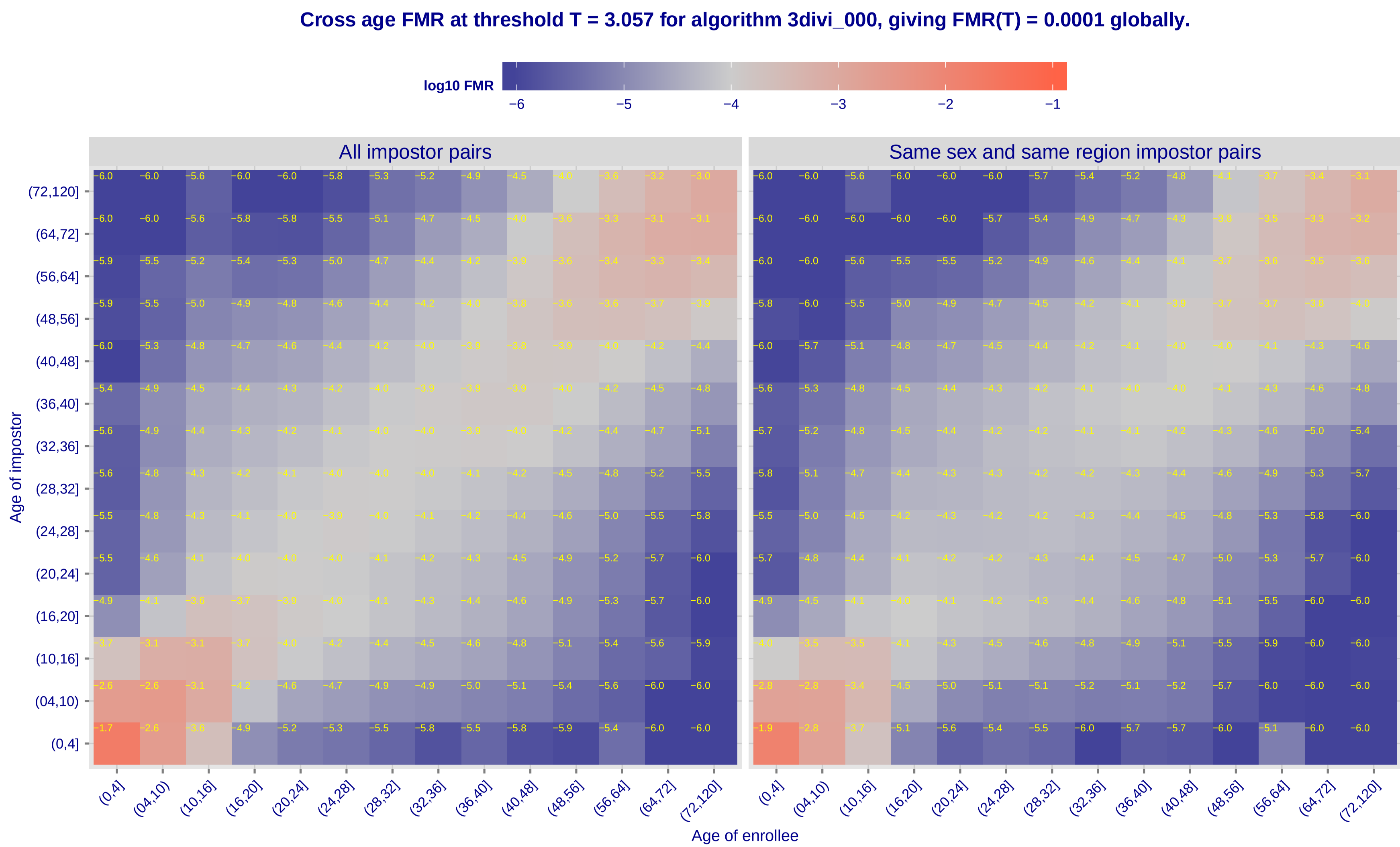


Figure 92: For algorithm 3divi-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

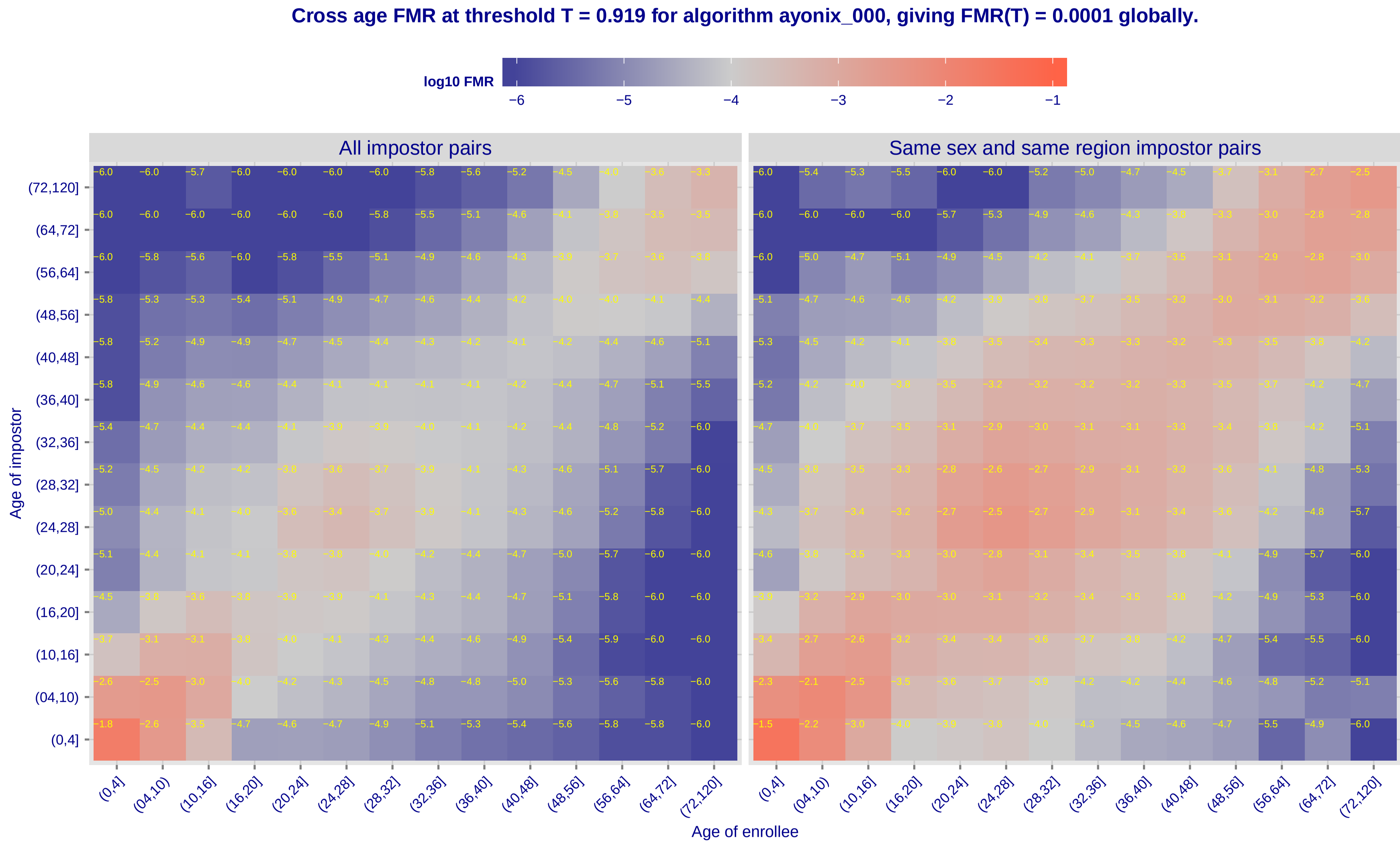


Figure 93: For algorithm ayonix-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

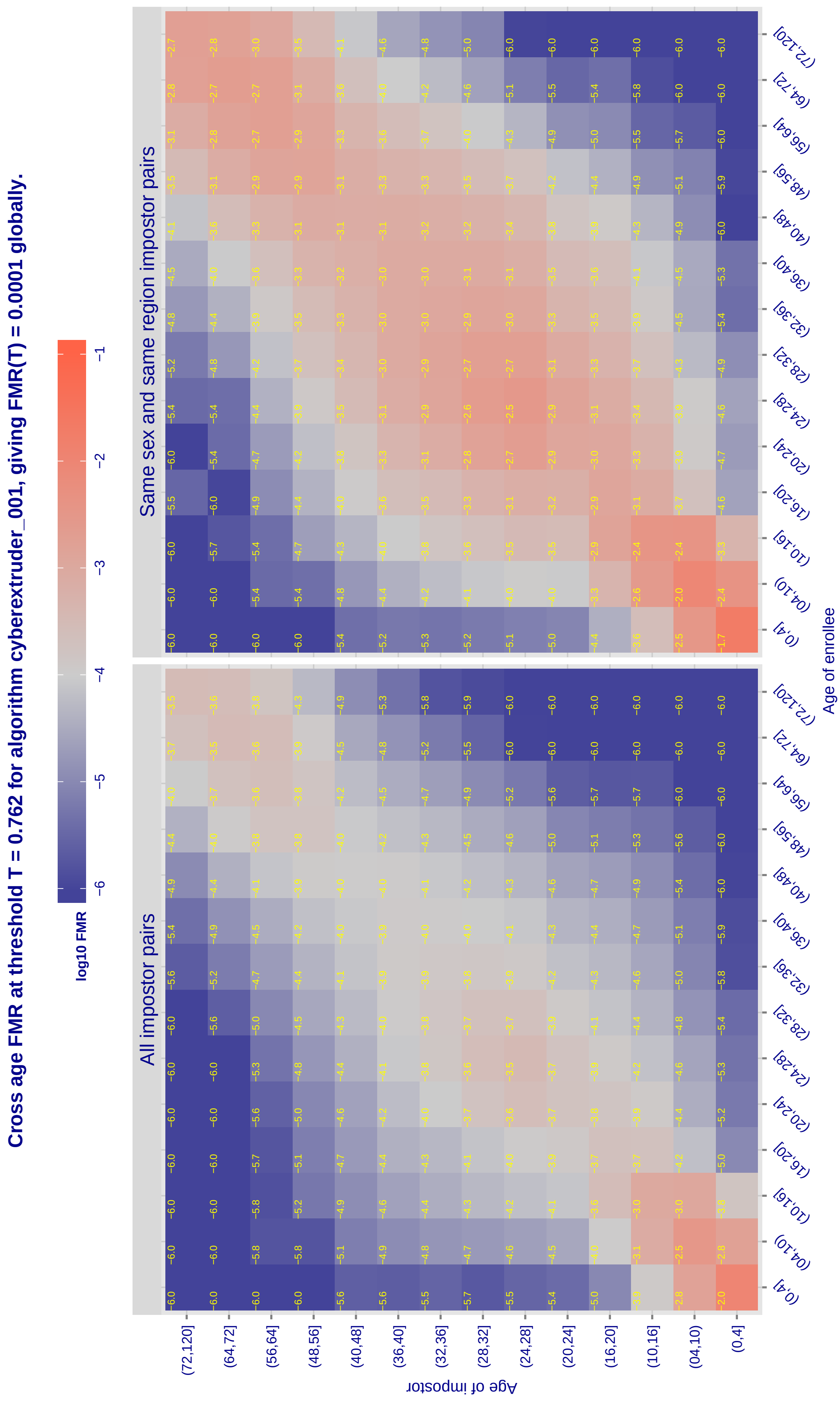


Figure 94: For algorithm cyberextruder-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

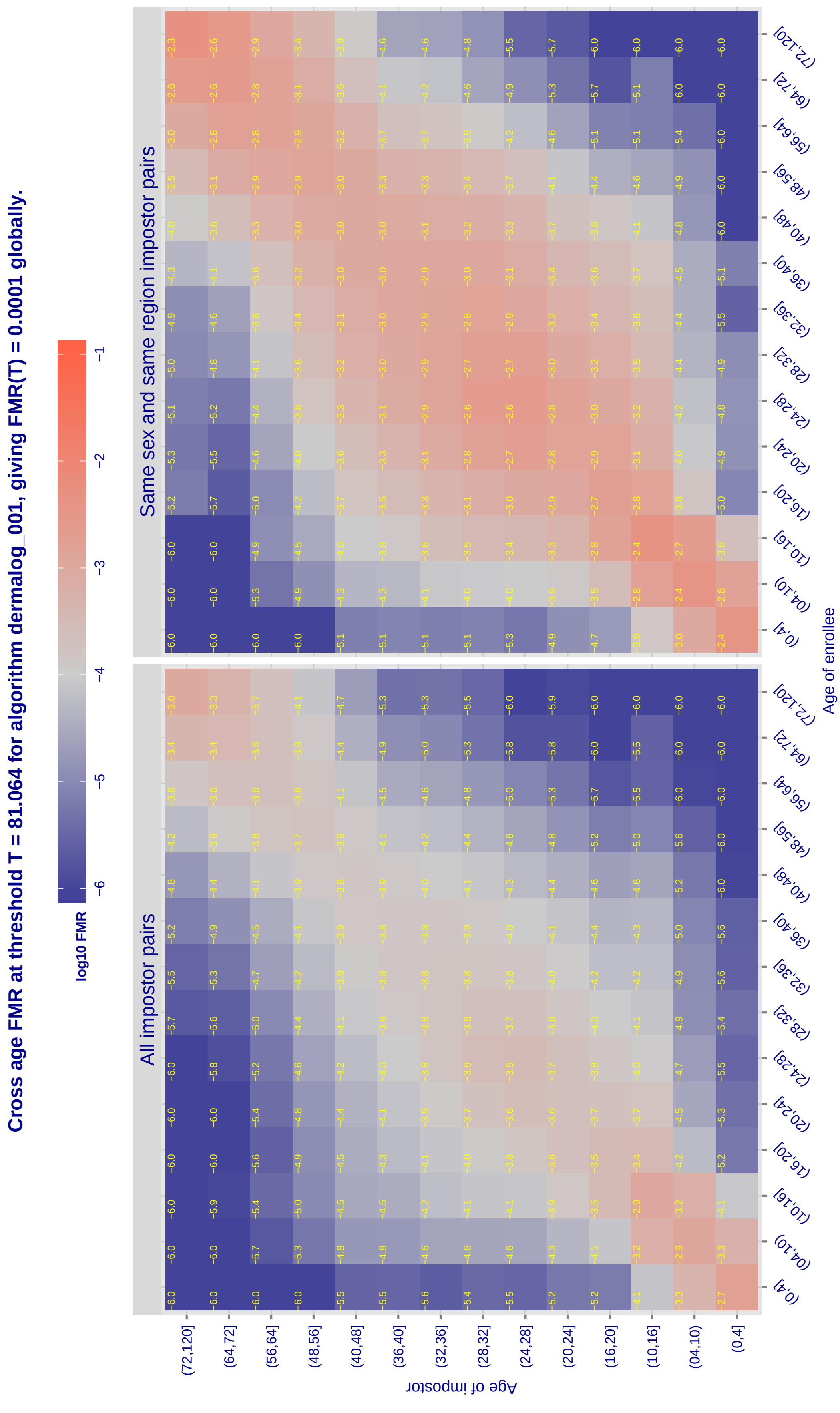


Figure 95: For algorithm dermalog-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

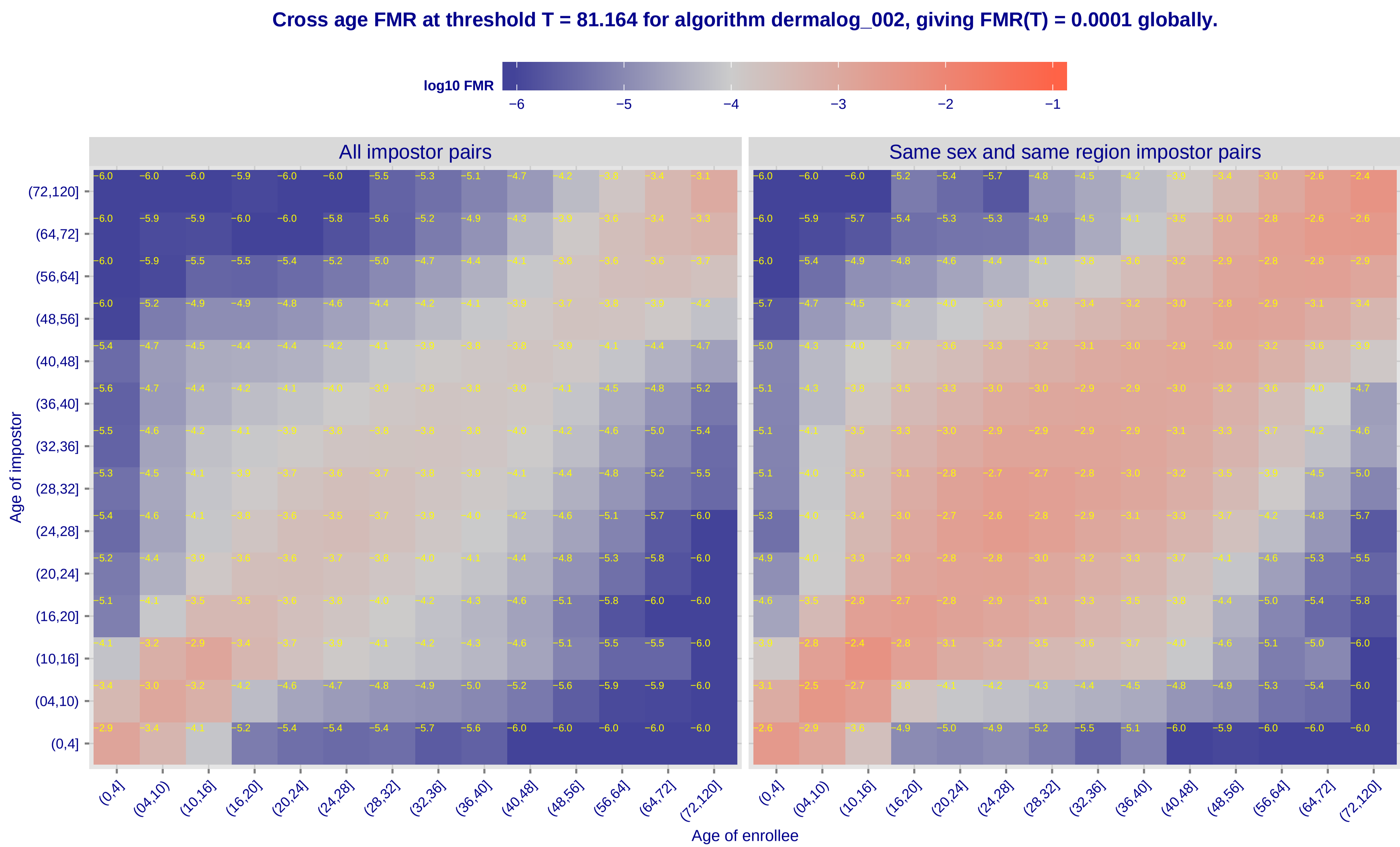


Figure 96: For algorithm dermalog-002 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

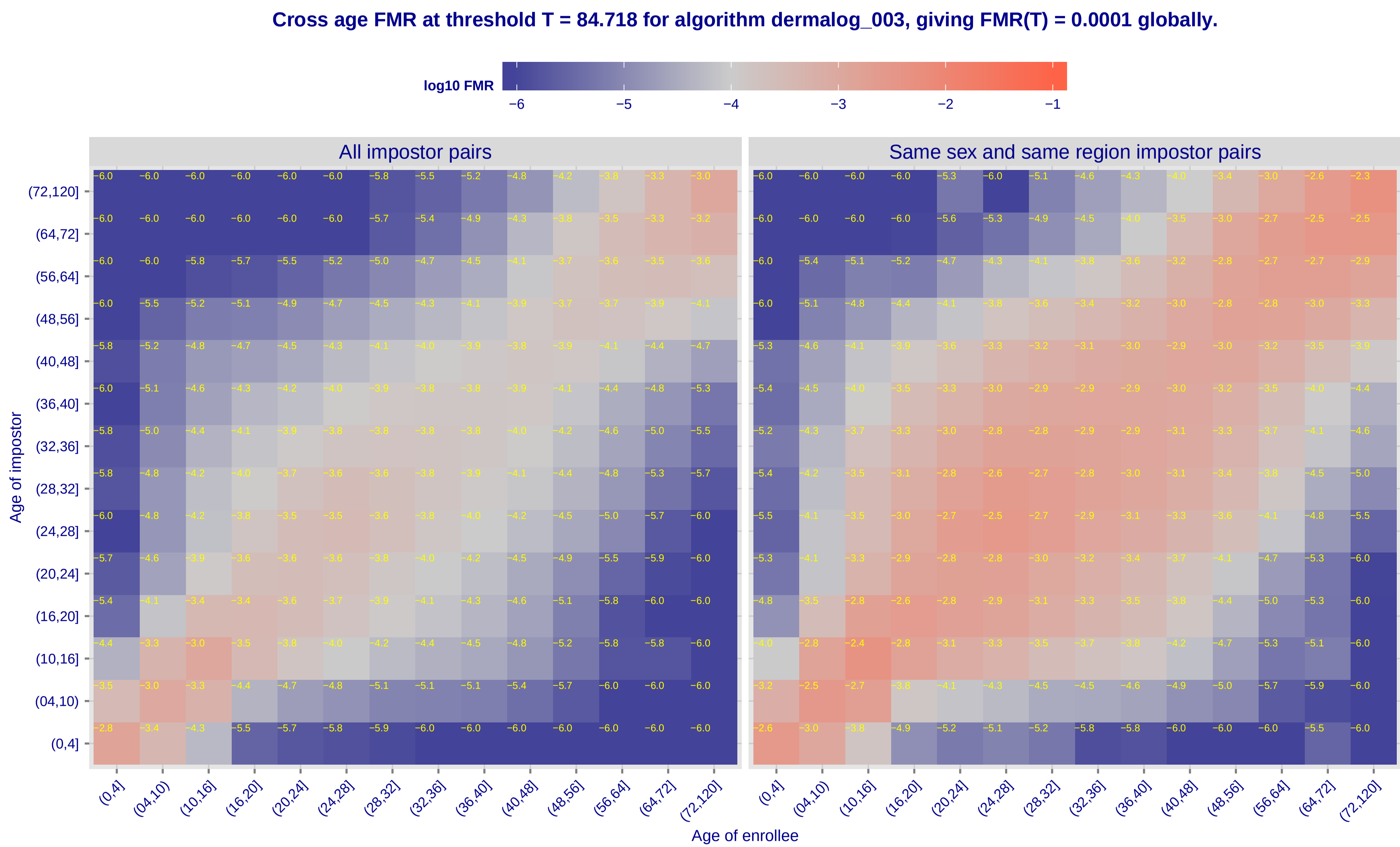


Figure 97: For algorithm dermalog-003 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

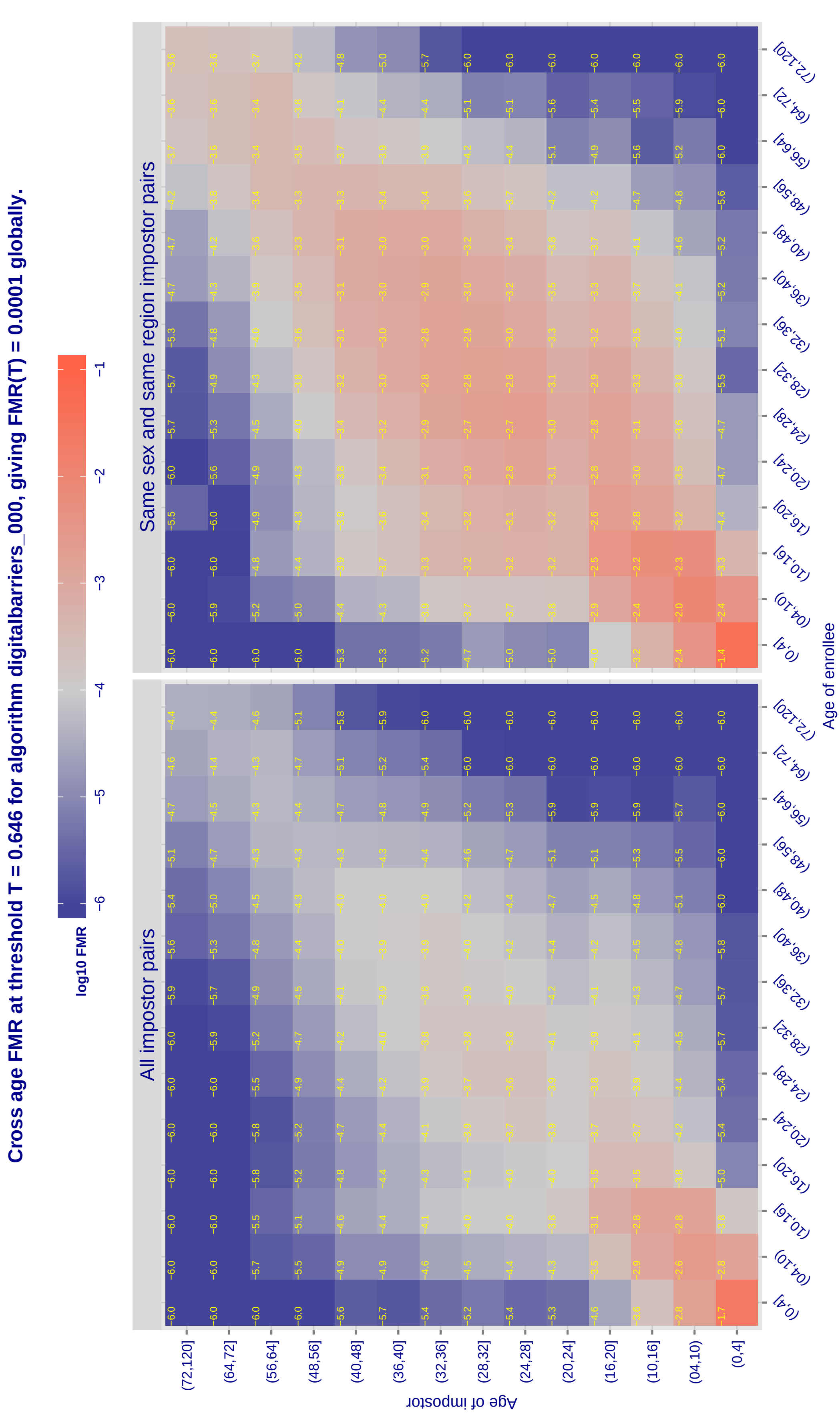


Figure 98: For algorithm digitalbarriers-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

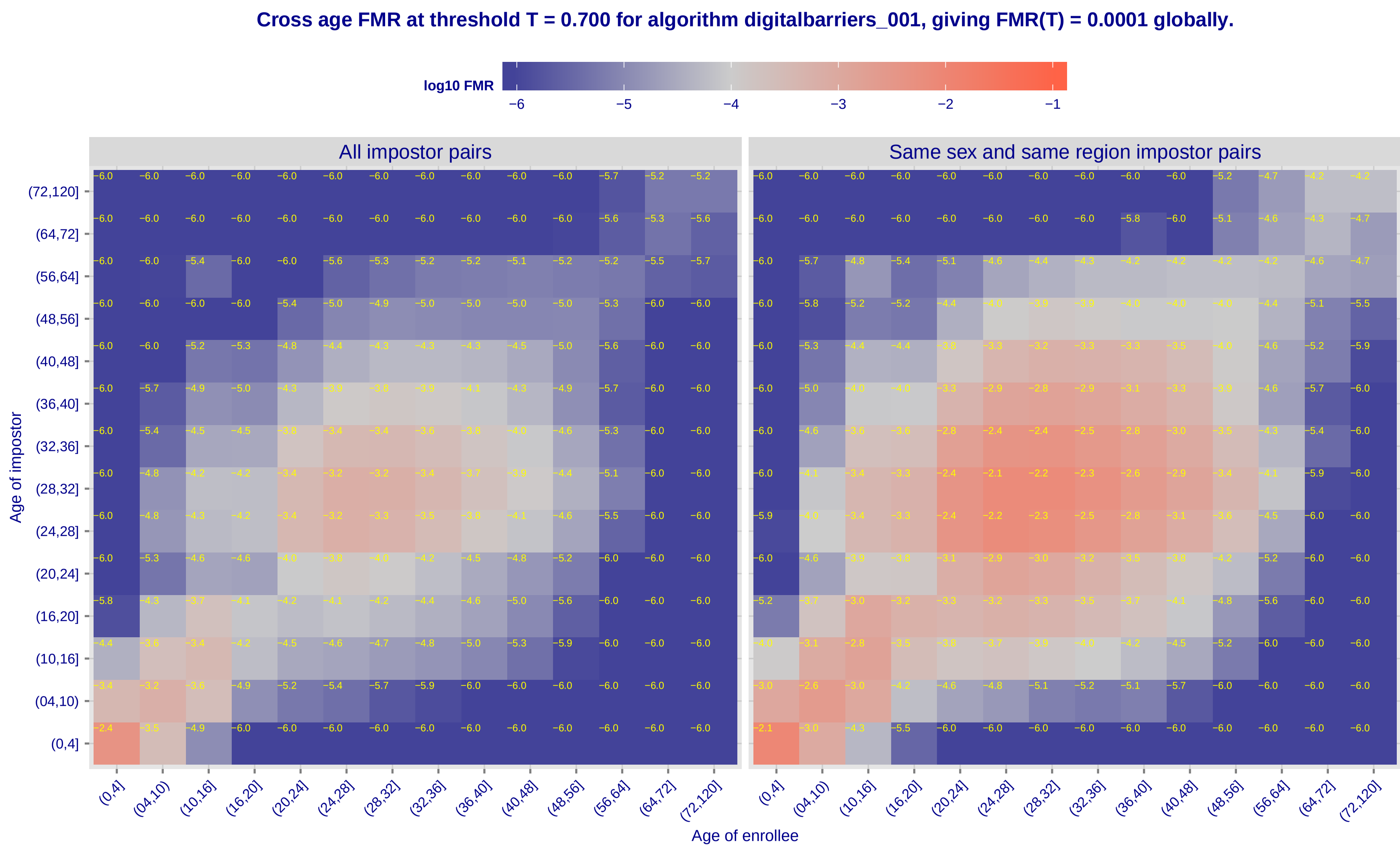


Figure 99: For algorithm digitalbarriers-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

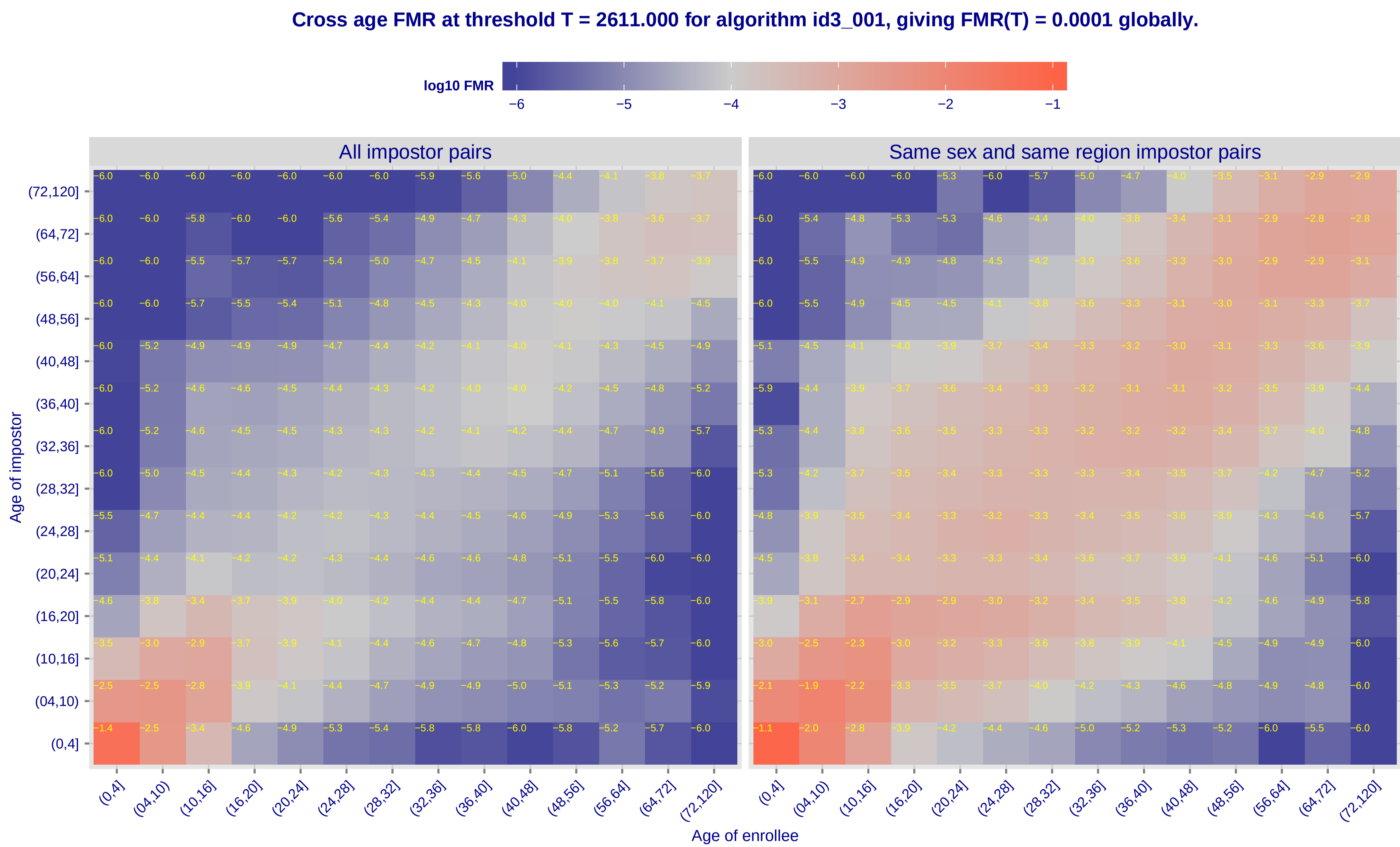


Figure 100: For algorithm id3-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

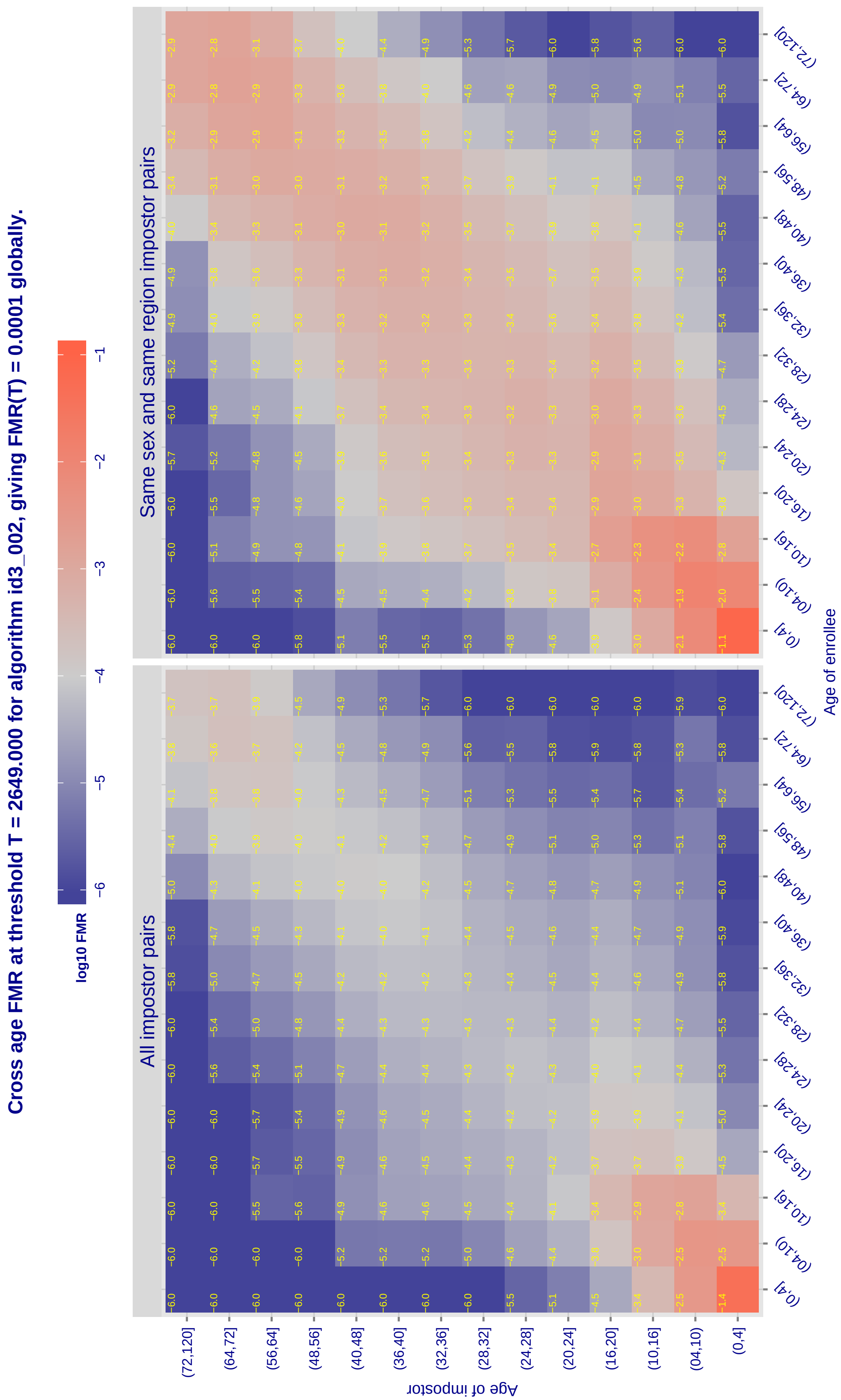


Figure 101: For algorithm id3-002 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

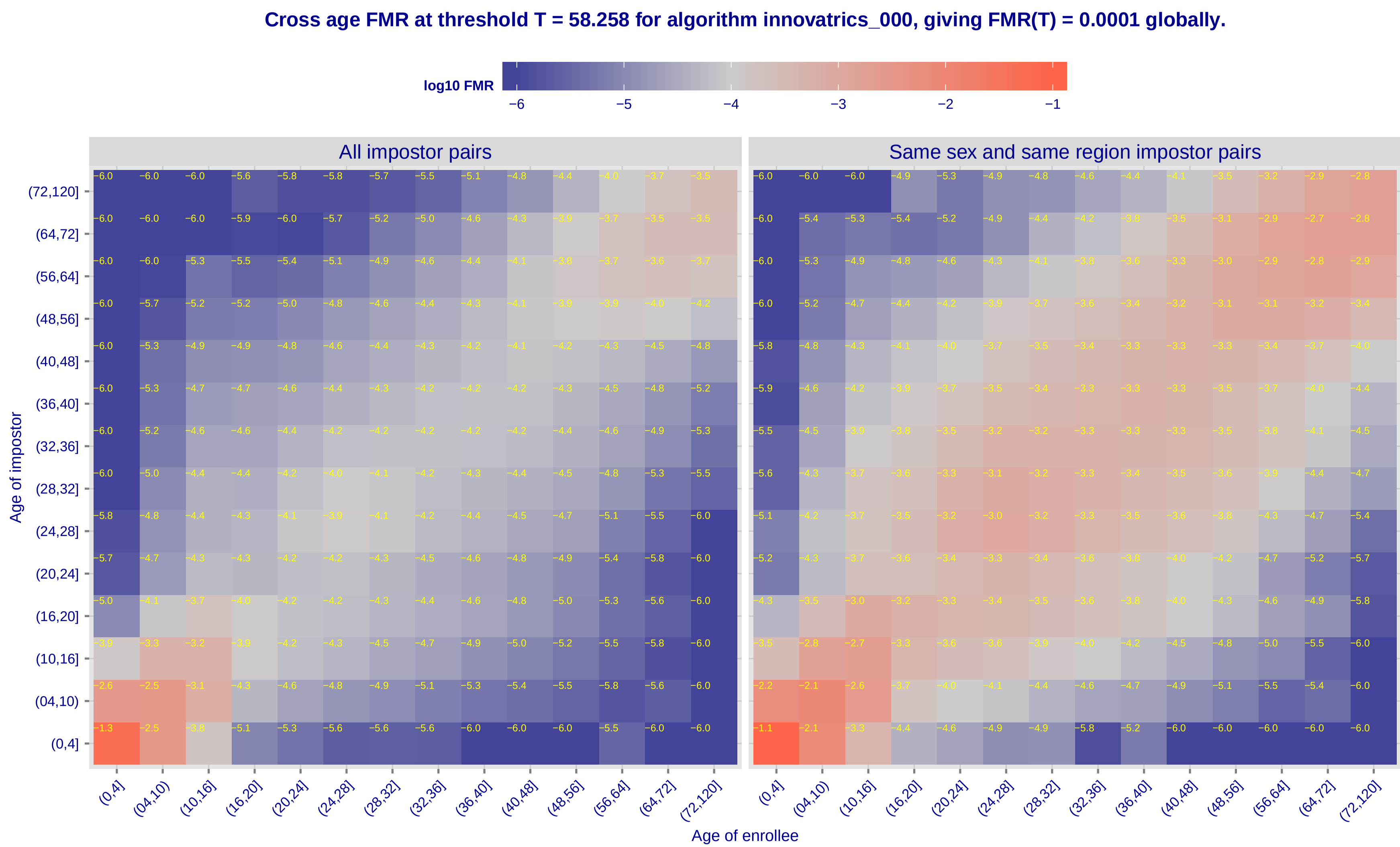


Figure 102: For algorithm innovatrics-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

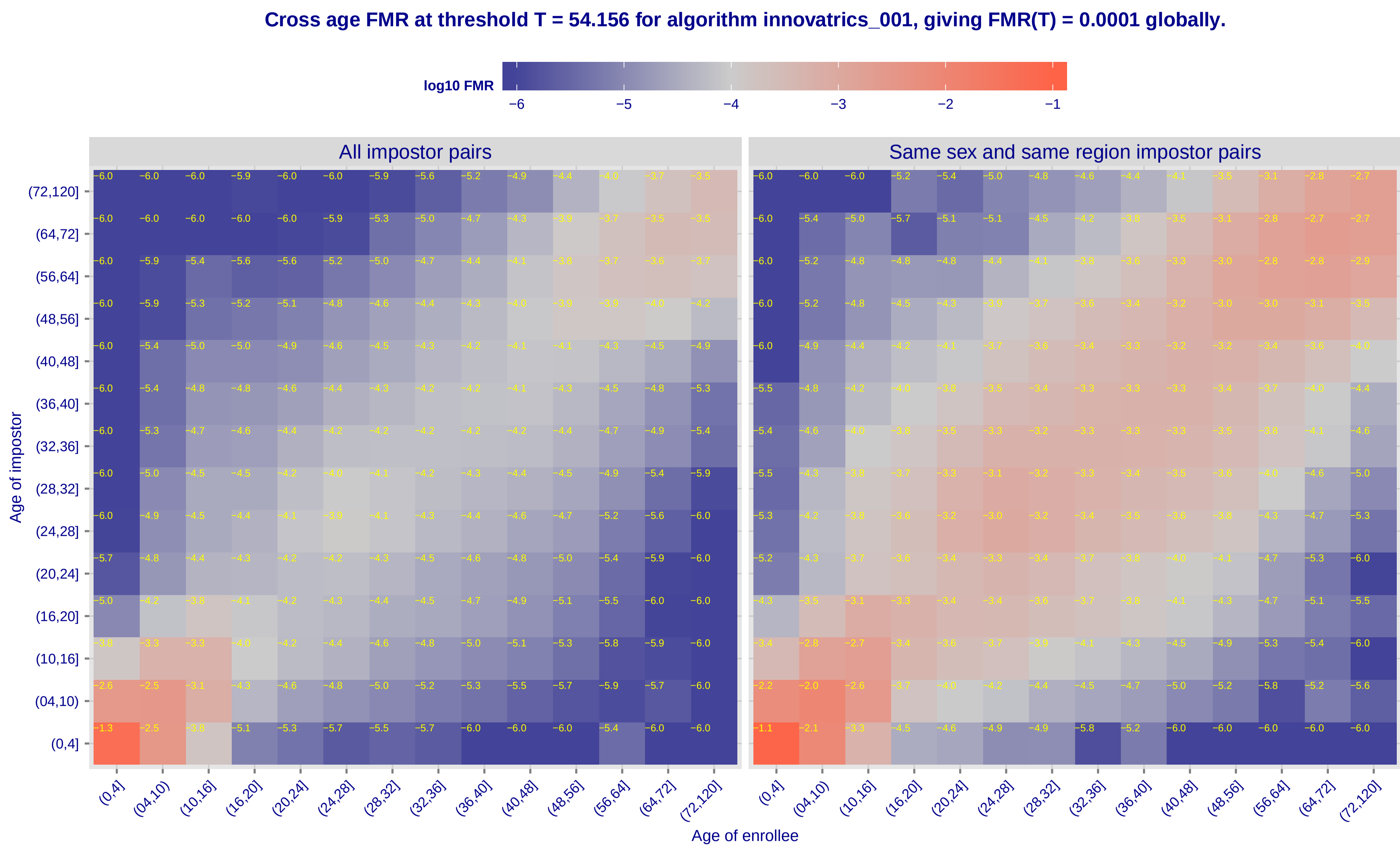


Figure 103: For algorithm innovatrics-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

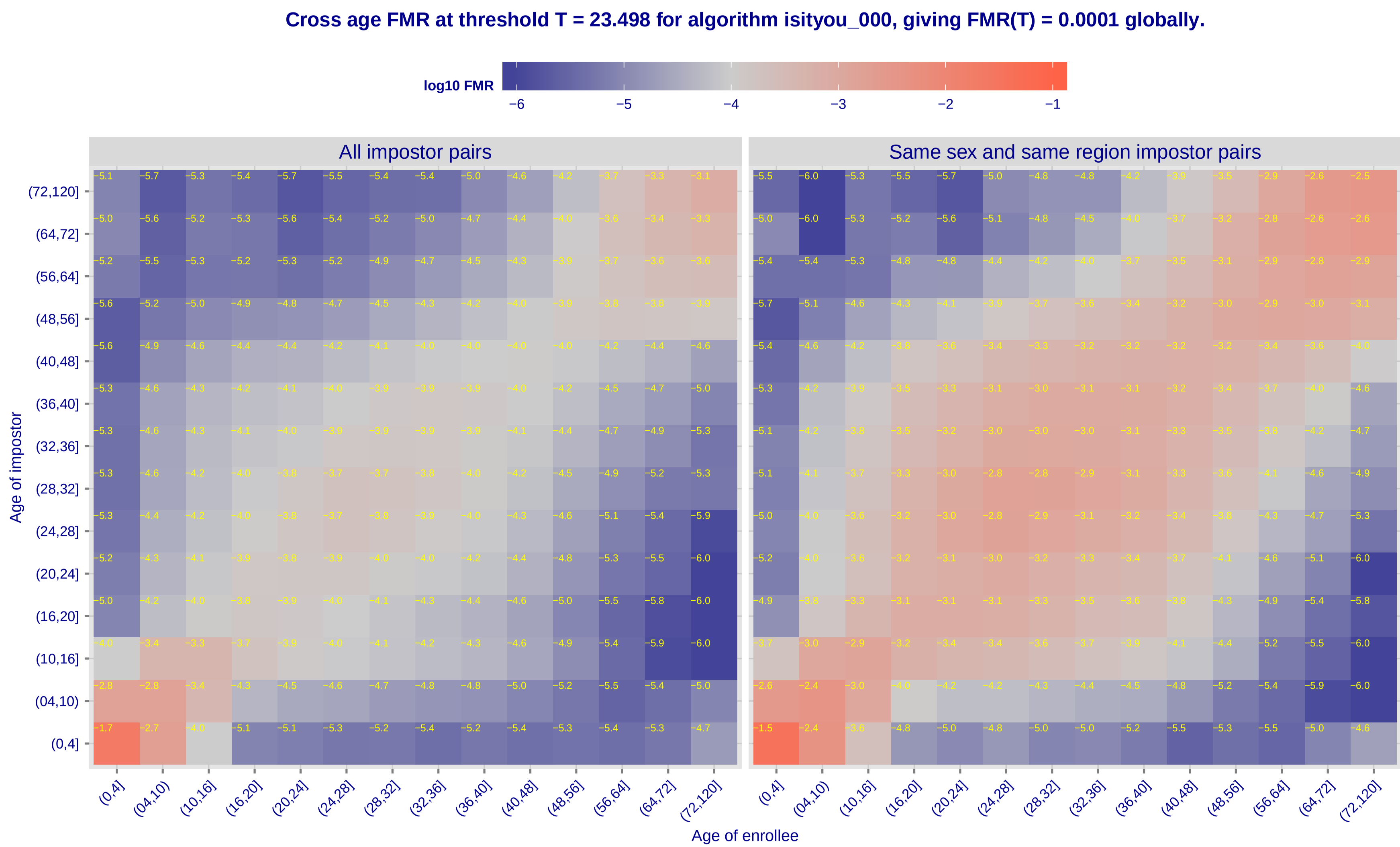


Figure 104: For algorithm isityou-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

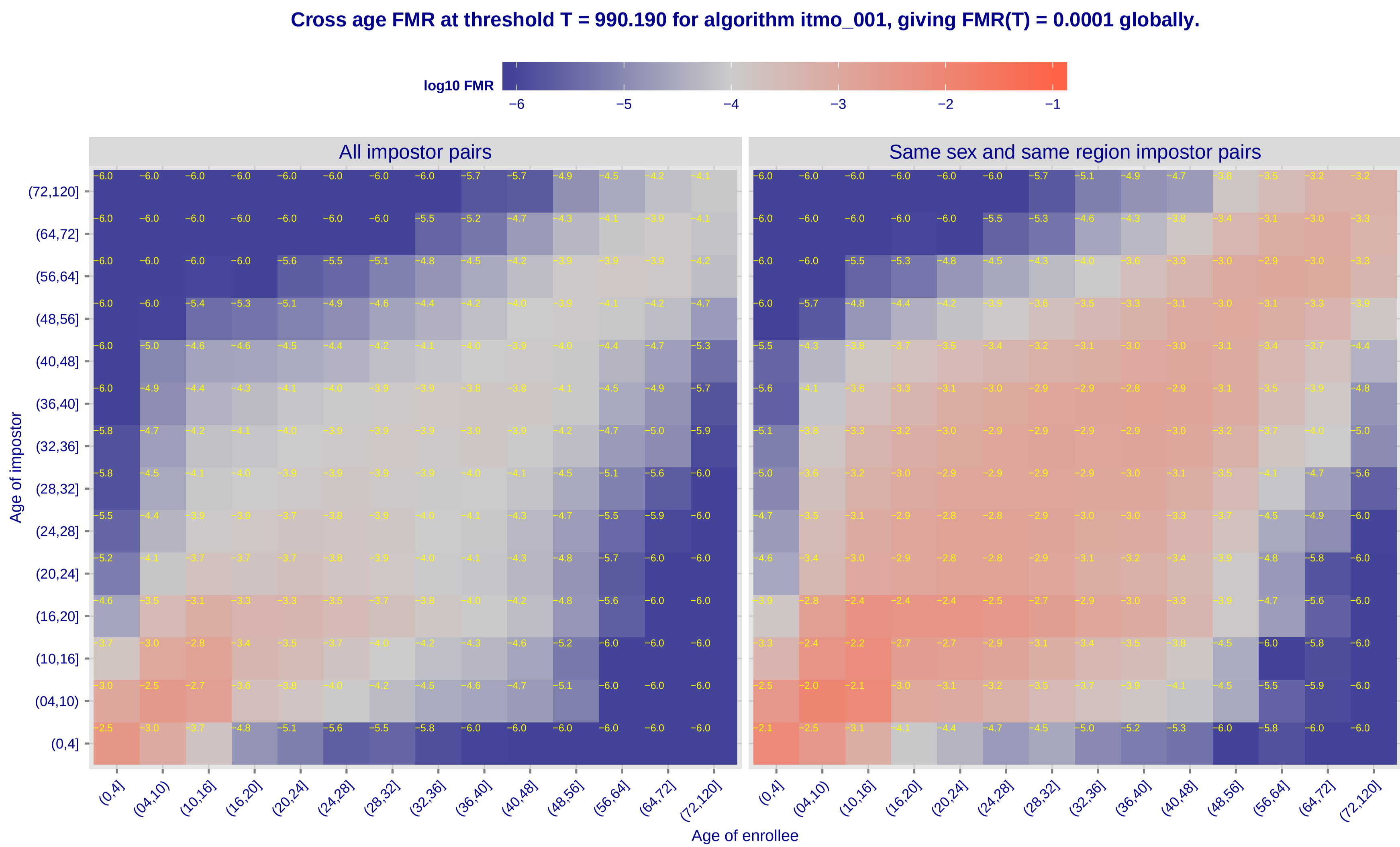


Figure 105: For algorithm itmo-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

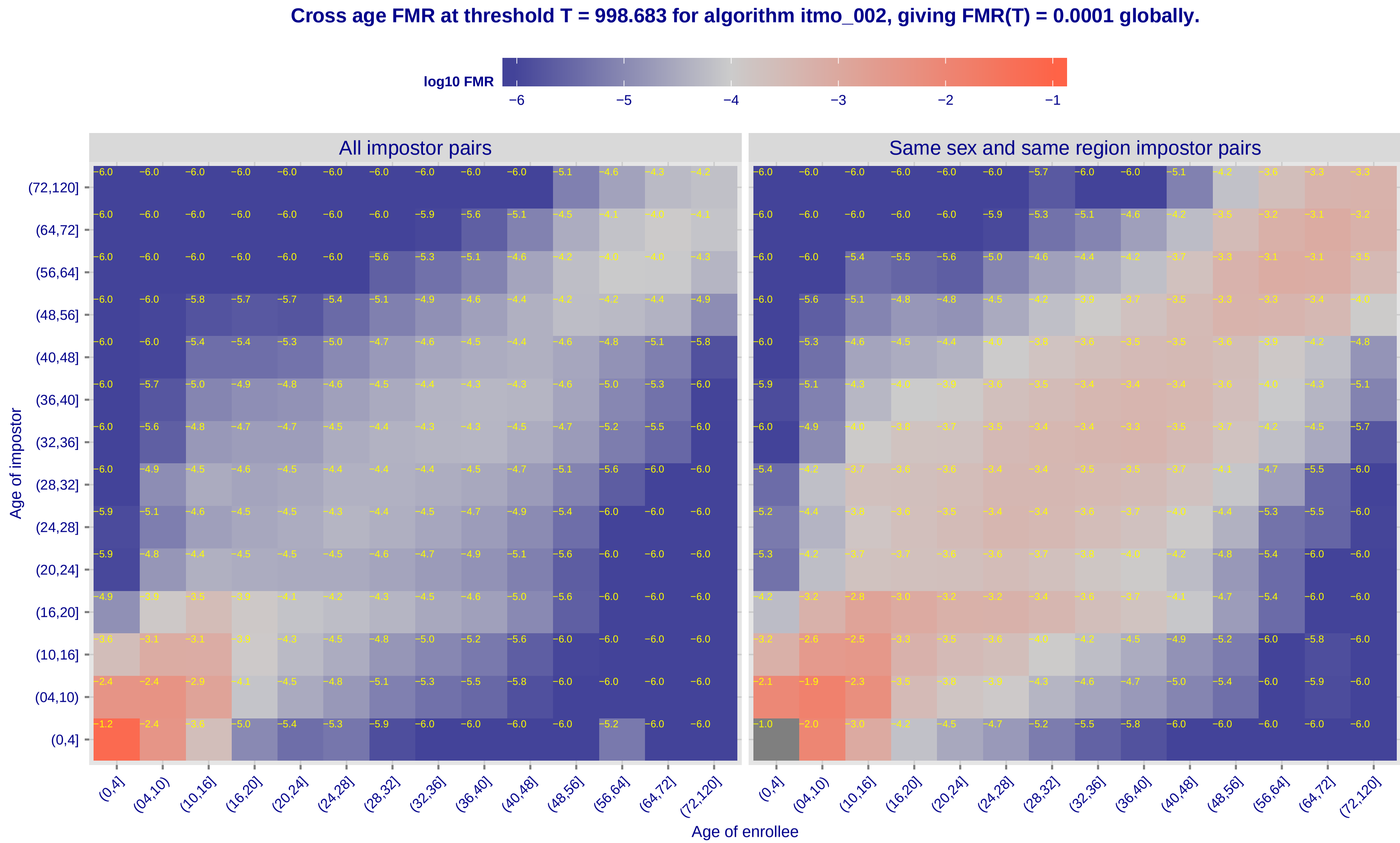


Figure 106: For algorithm itmo-002 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

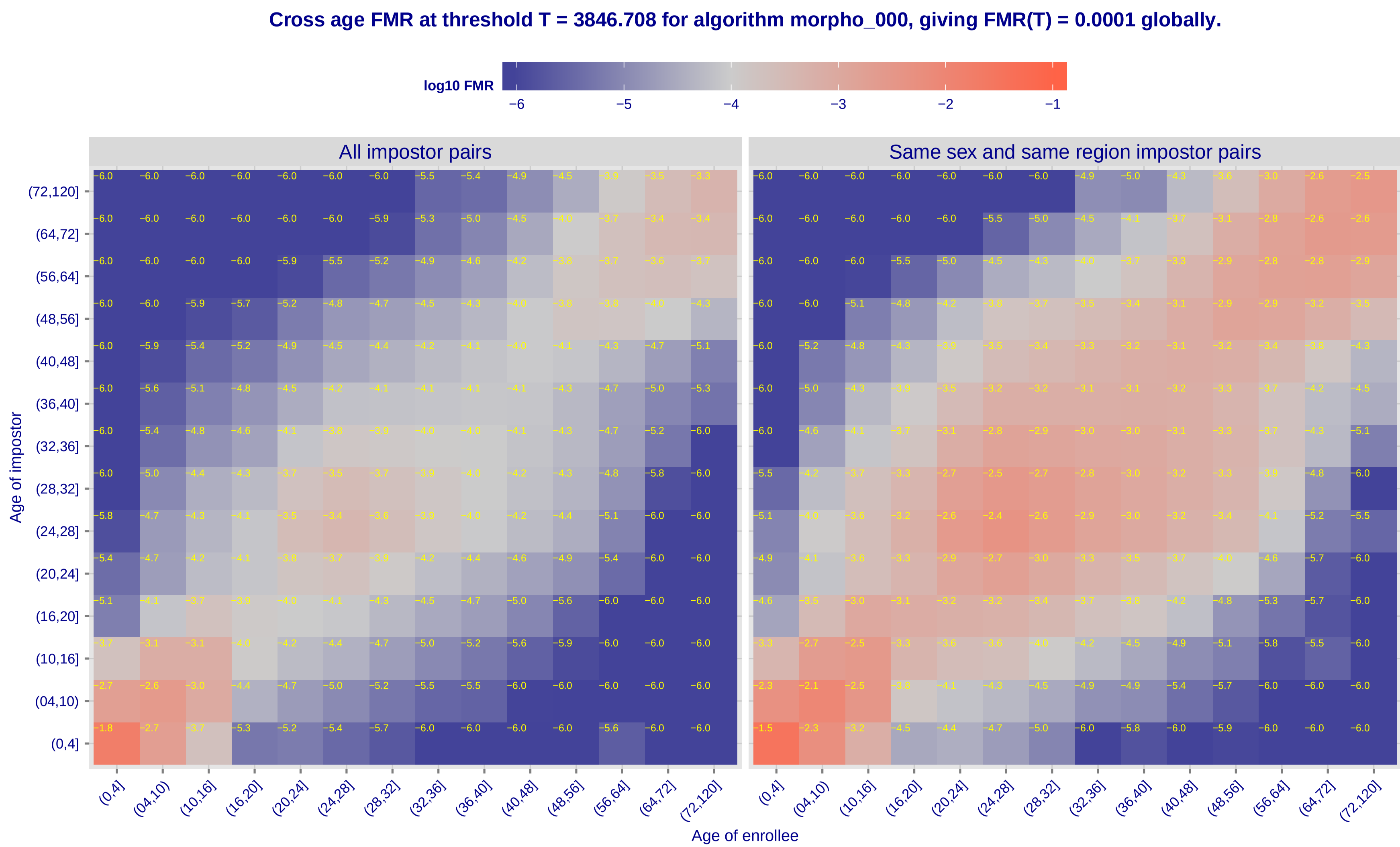


Figure 107: For algorithm morpho-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

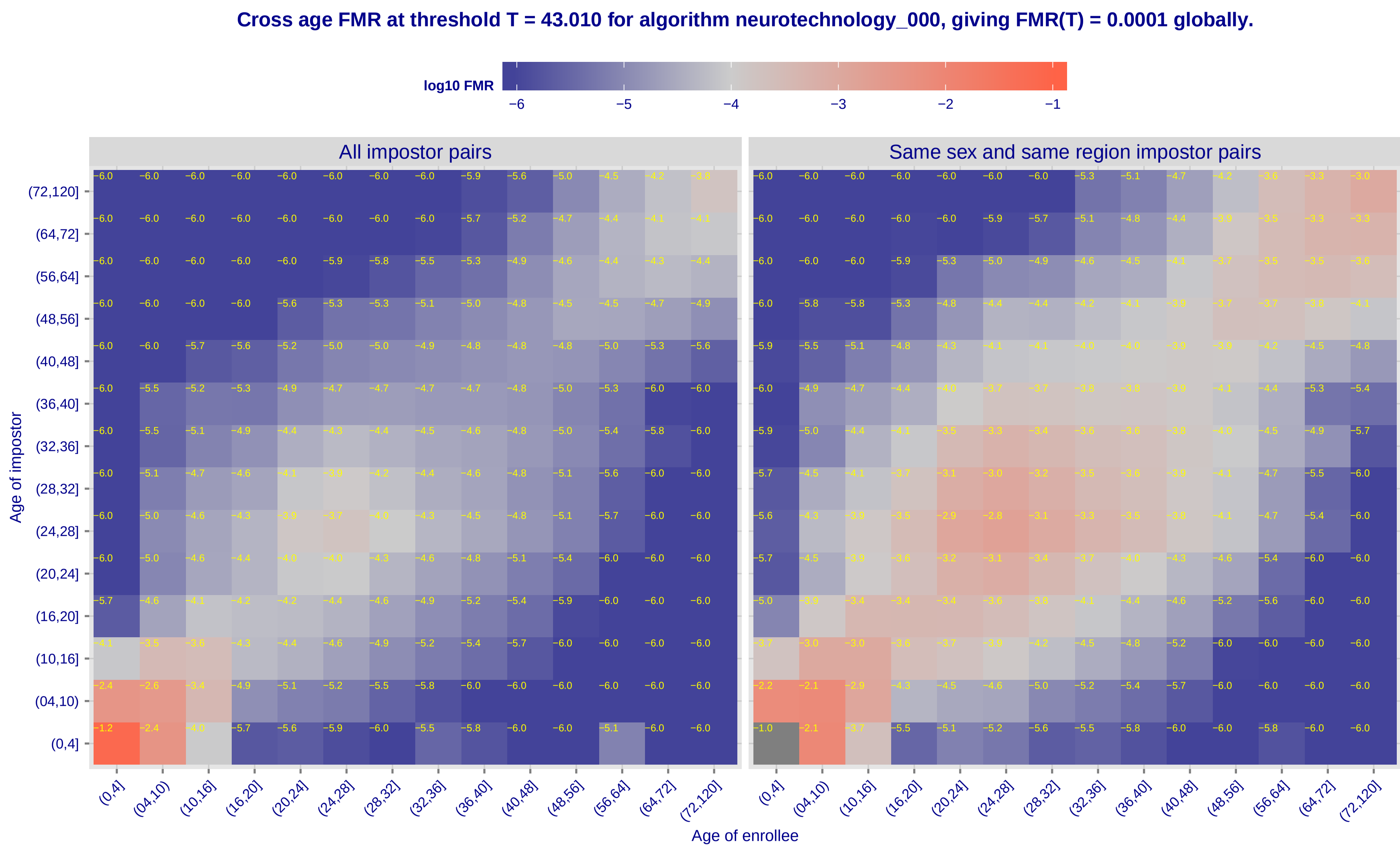


Figure 108: For algorithm neurotechnology-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

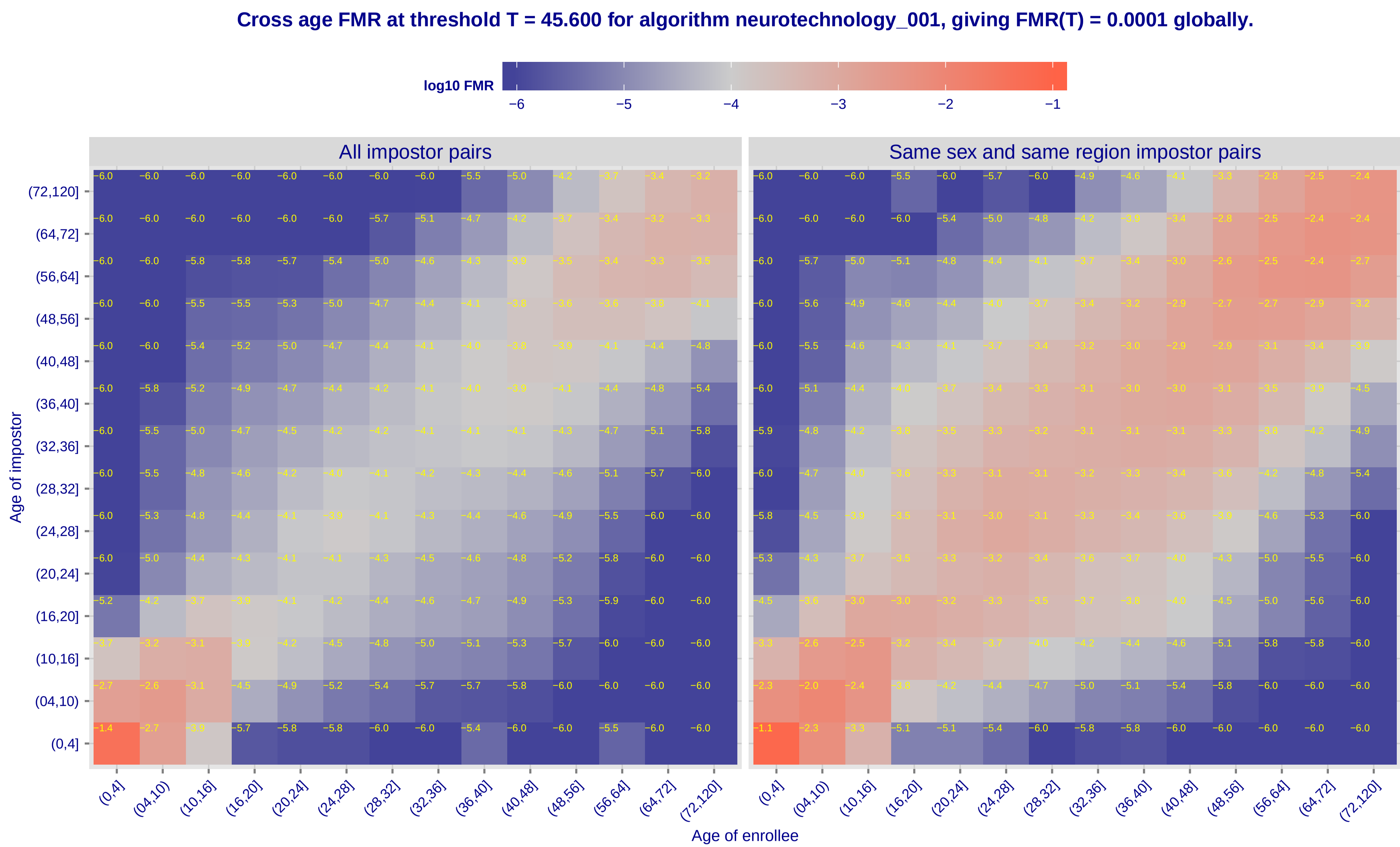


Figure 109: For algorithm neurotechnology-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

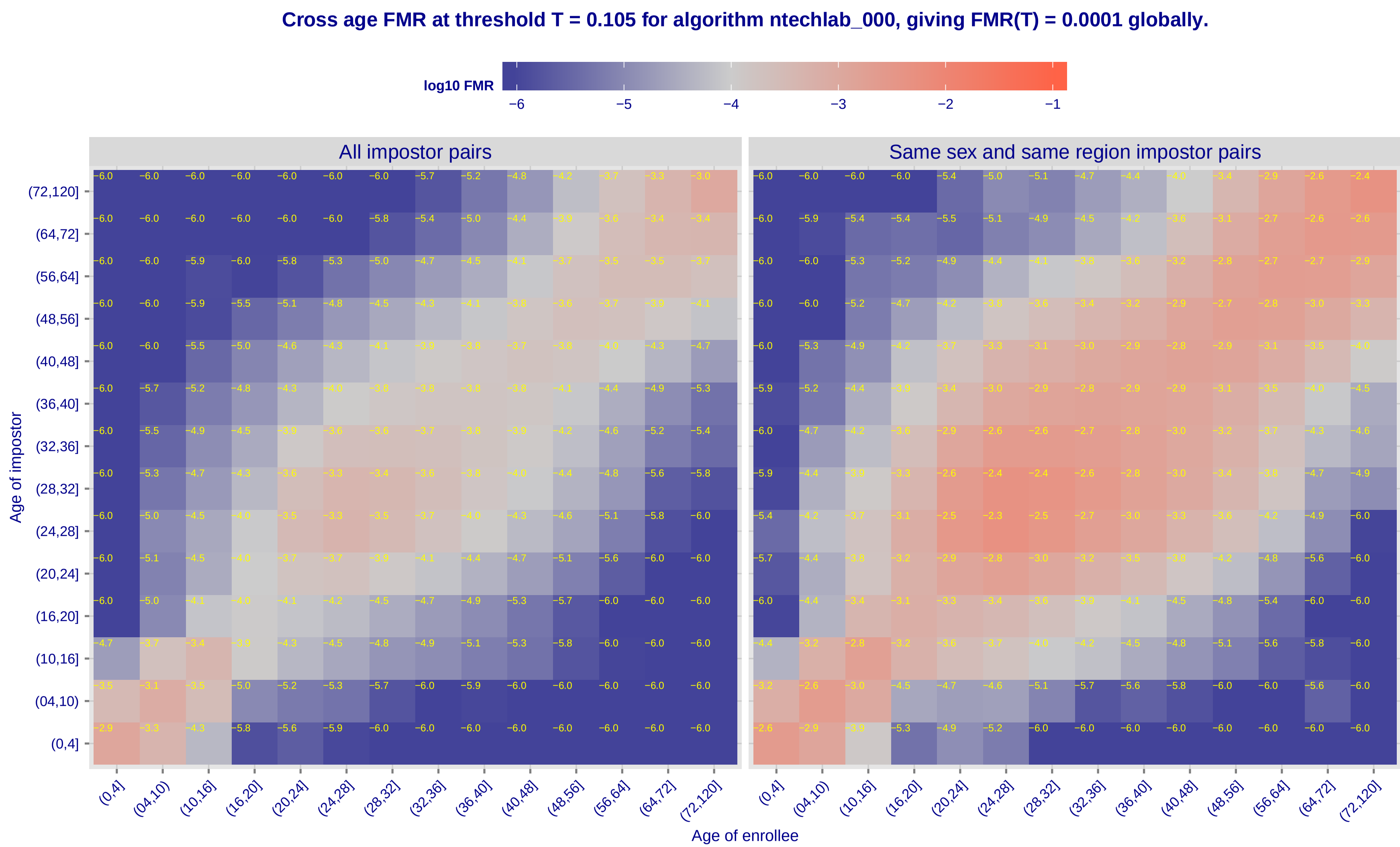


Figure 110: For algorithm ntechlab-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

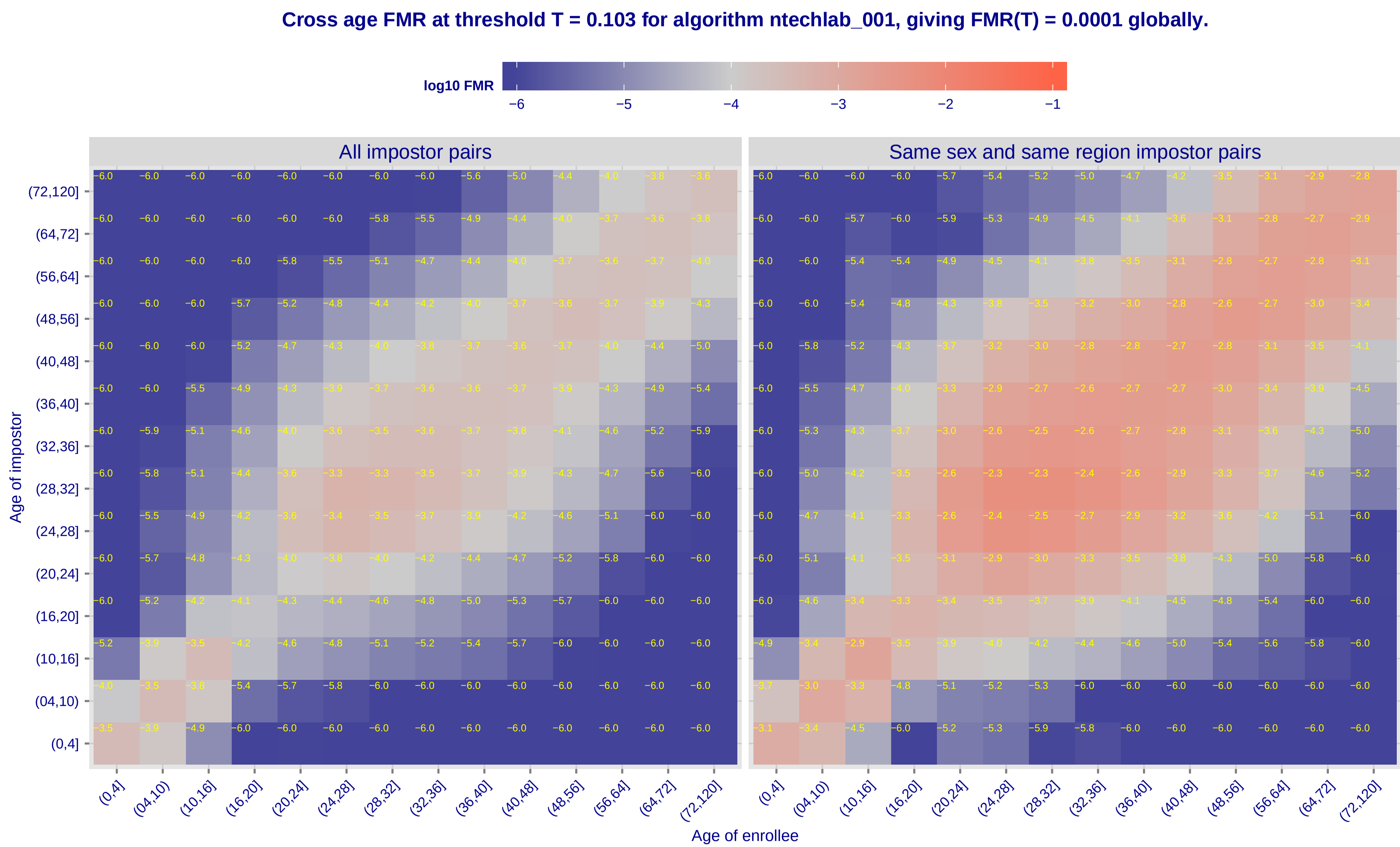


Figure 111: For algorithm ntechlab-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

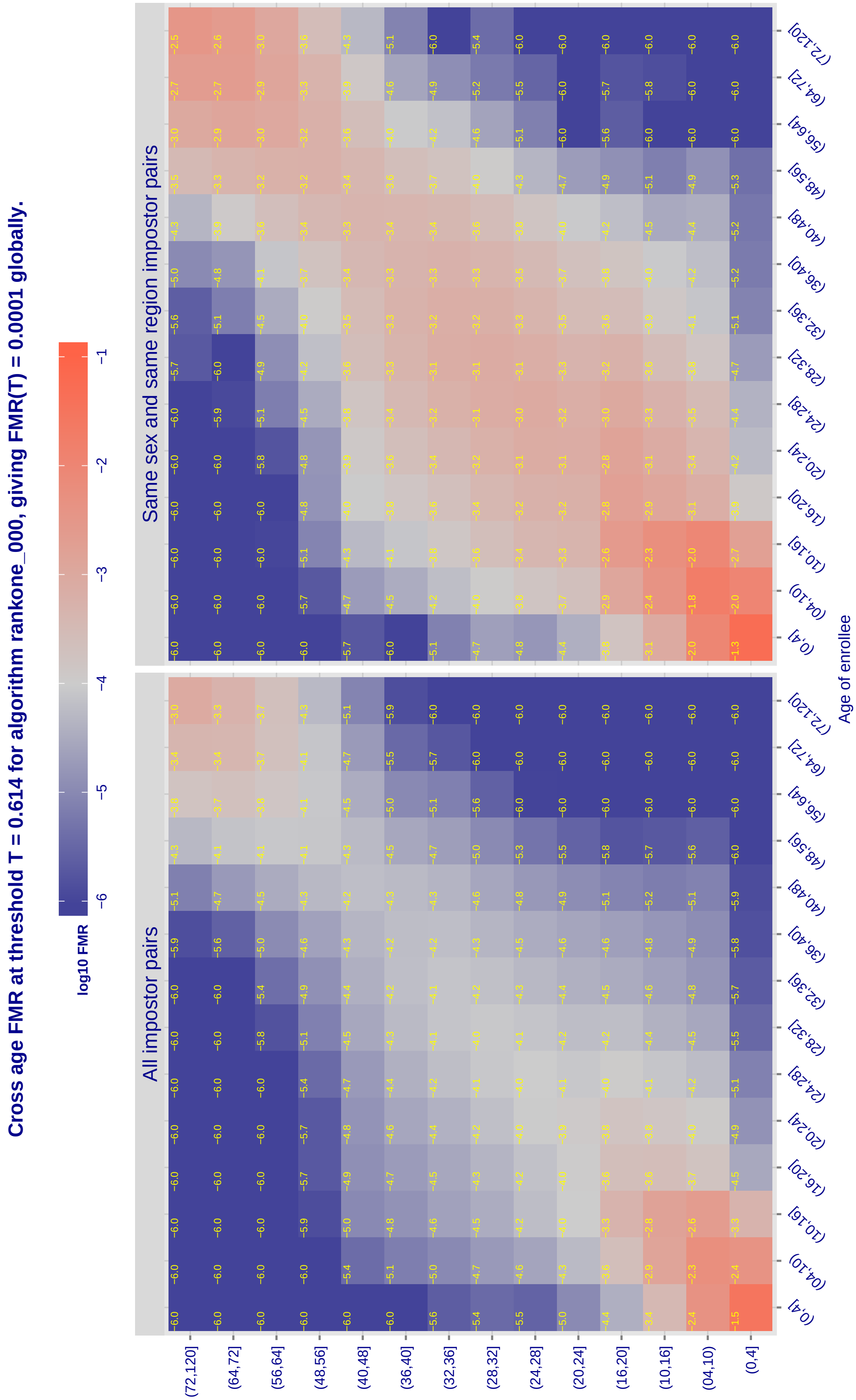


Figure 112: For algorithm rankone-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

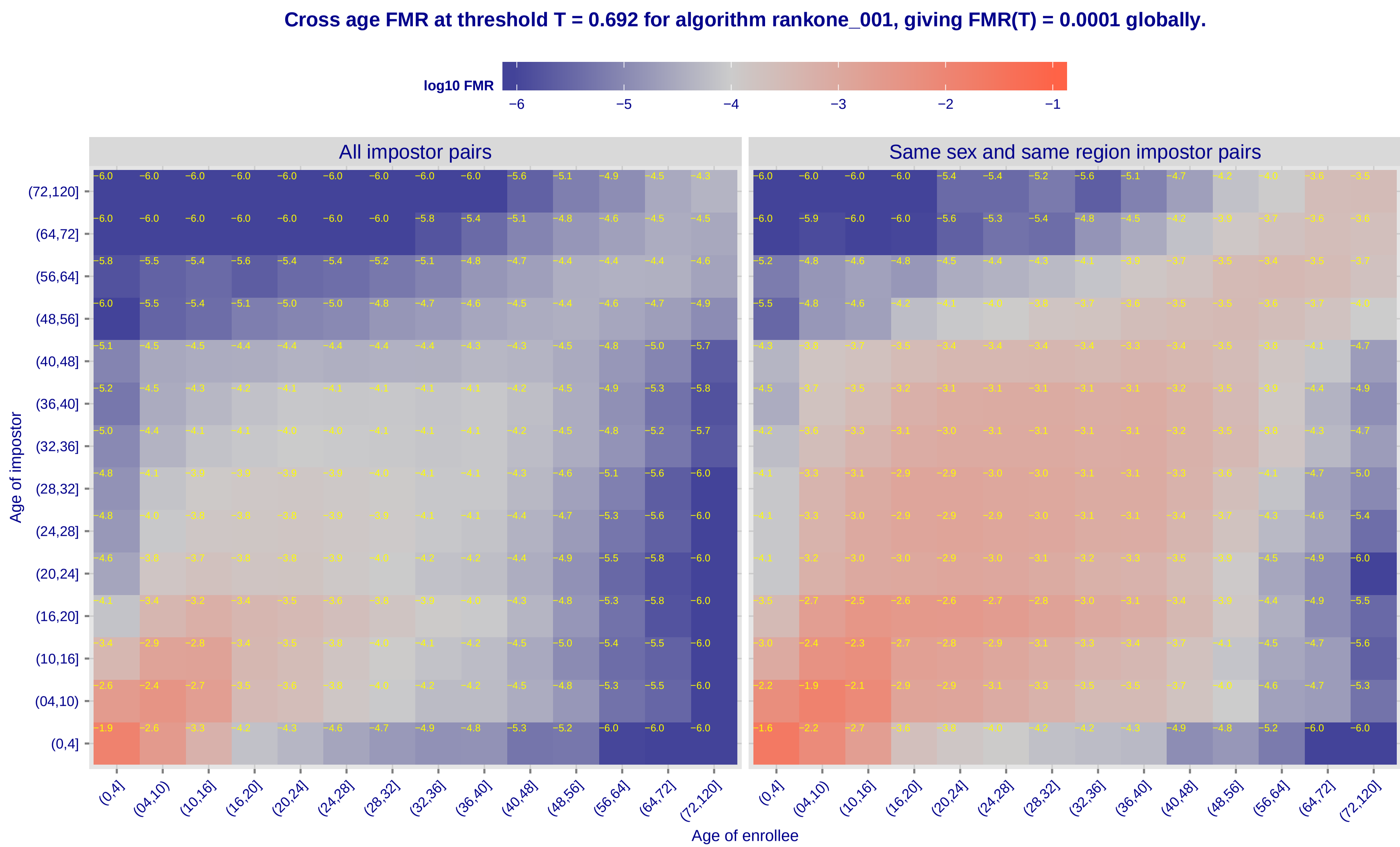


Figure 113: *For algorithm rankone-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.*

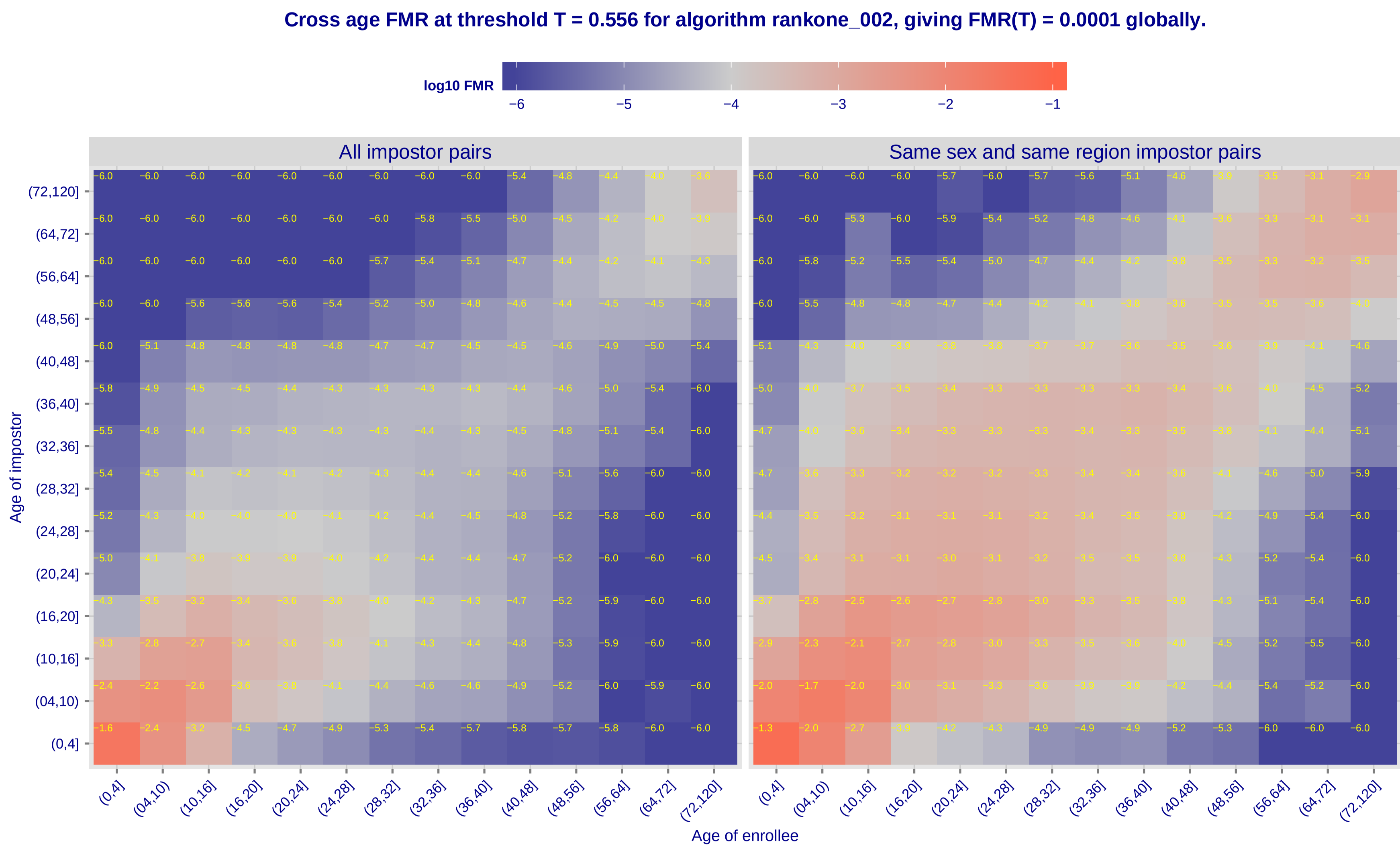


Figure 114: For algorithm rankone-002 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

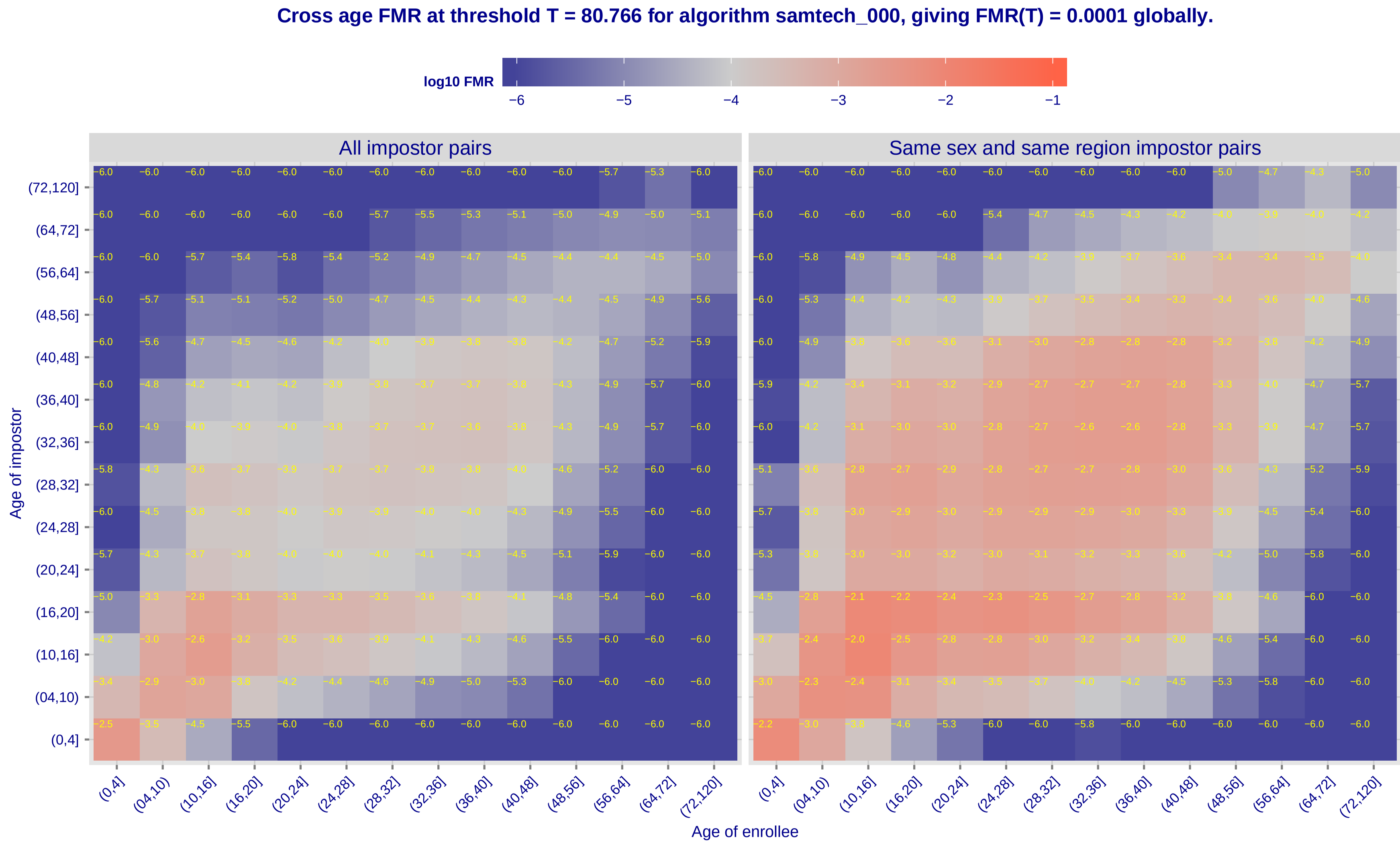


Figure 115: For algorithm samtech-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

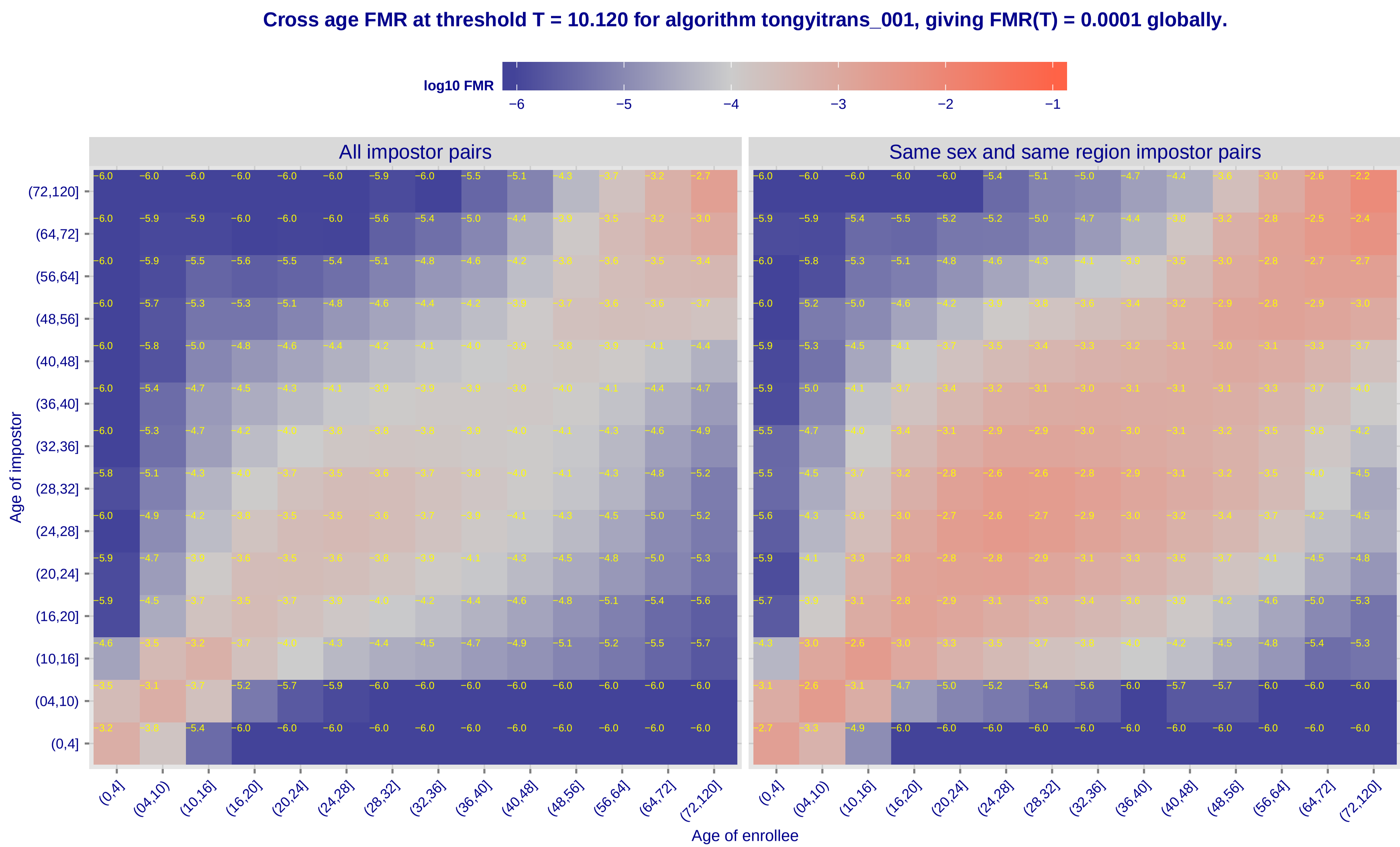


Figure 116: For algorithm tongyitrans-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

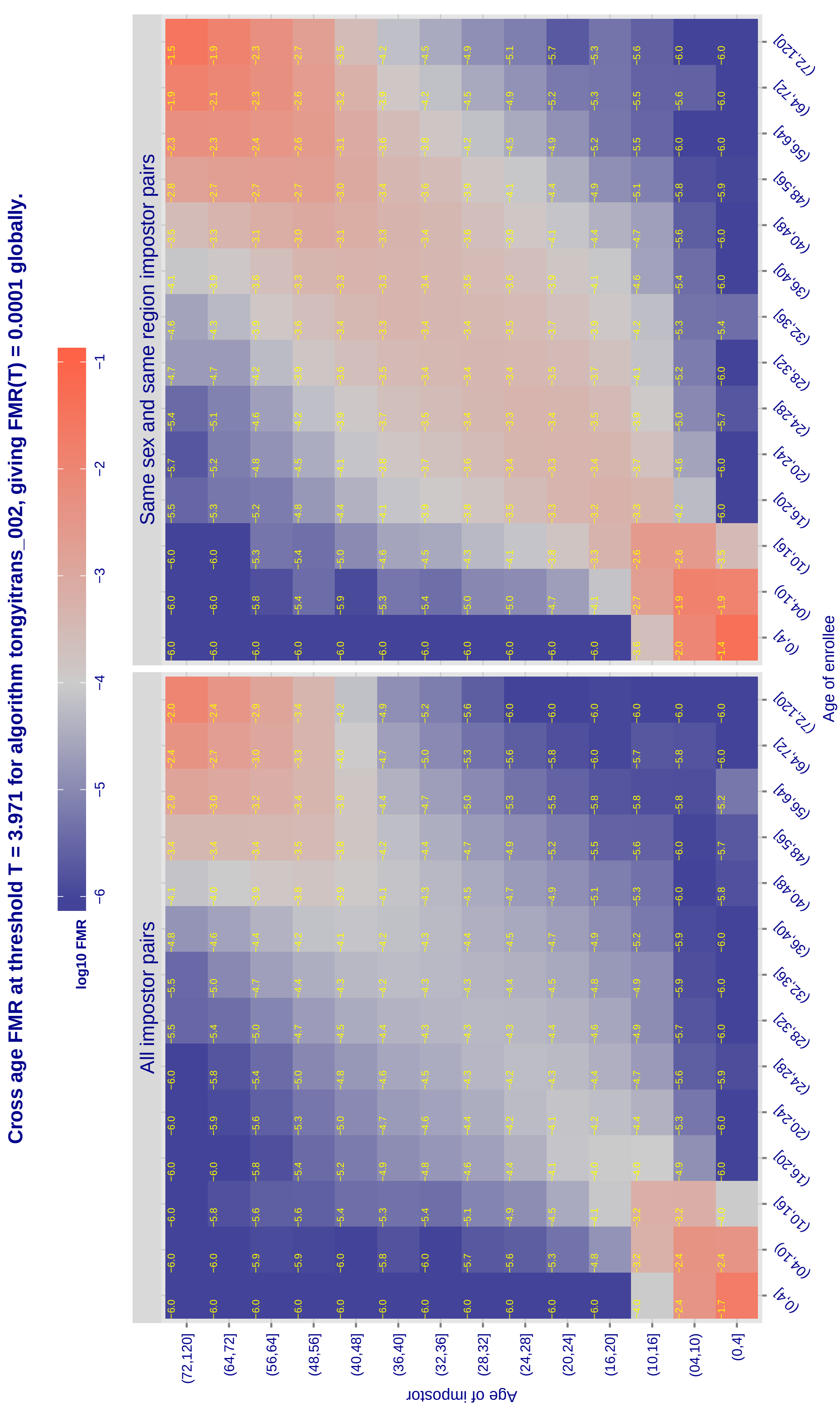


Figure 117: For algorithm tongyitrans-002 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

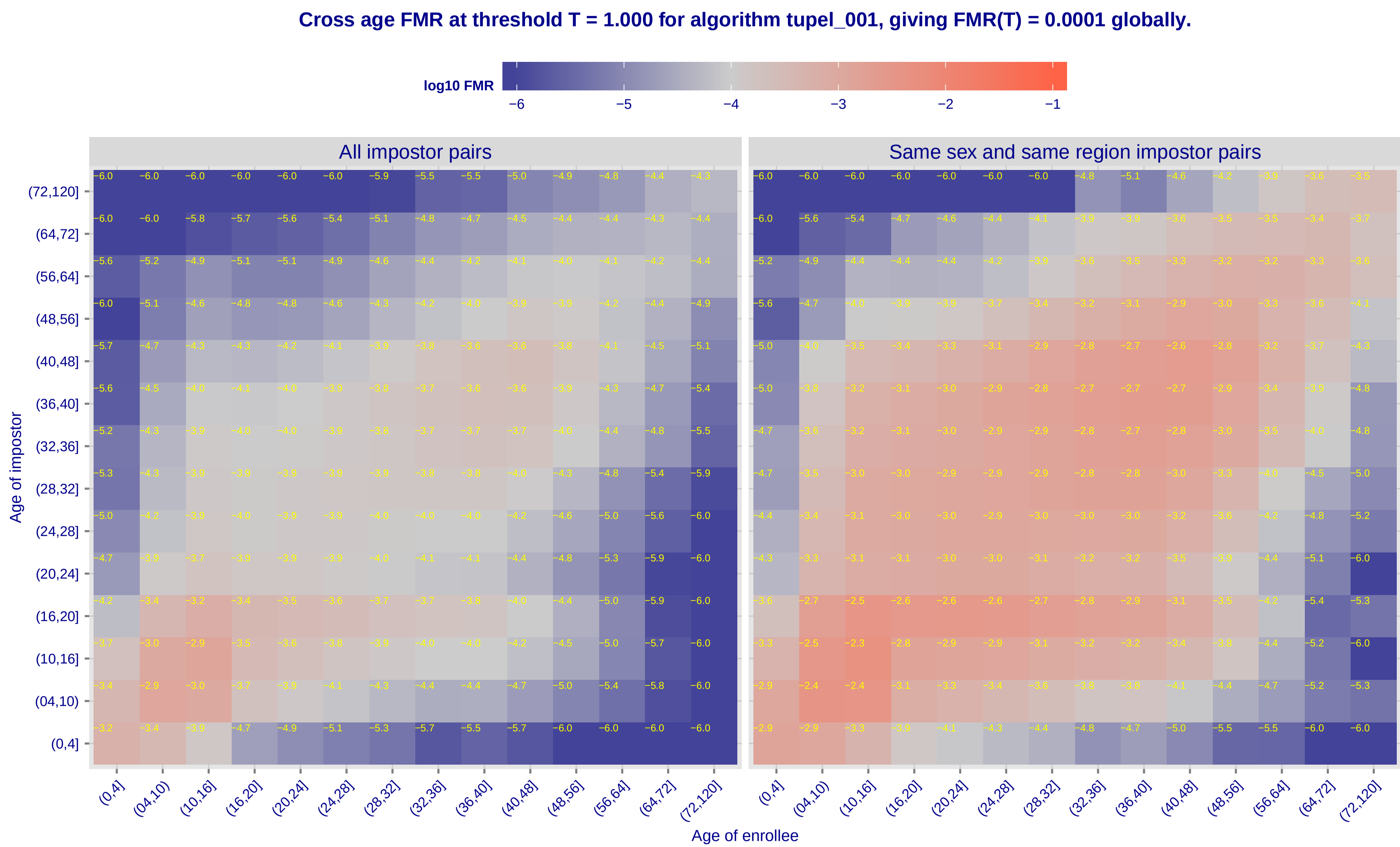


Figure 118: For algorithm tupel-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

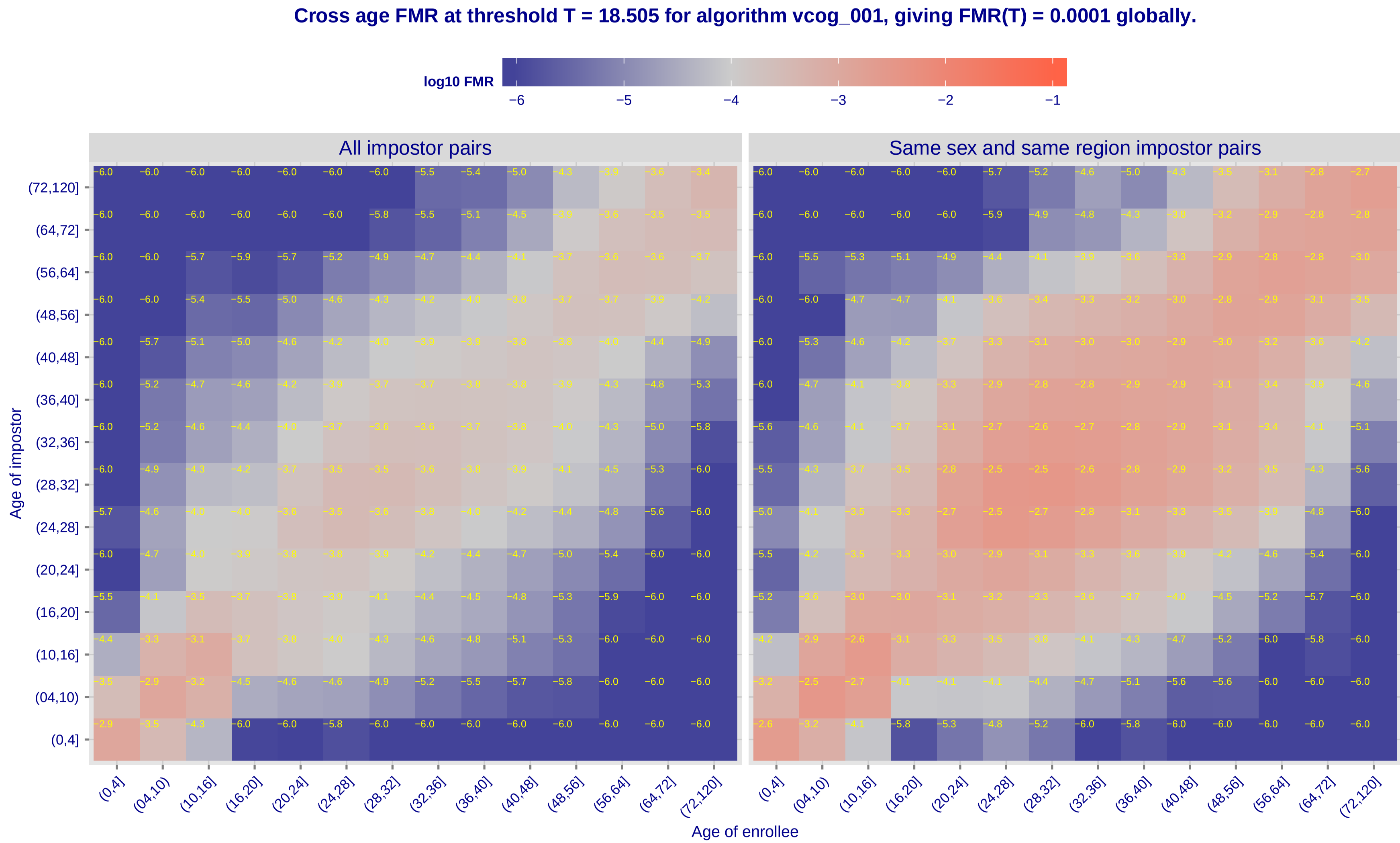


Figure 119: For algorithm vcog-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

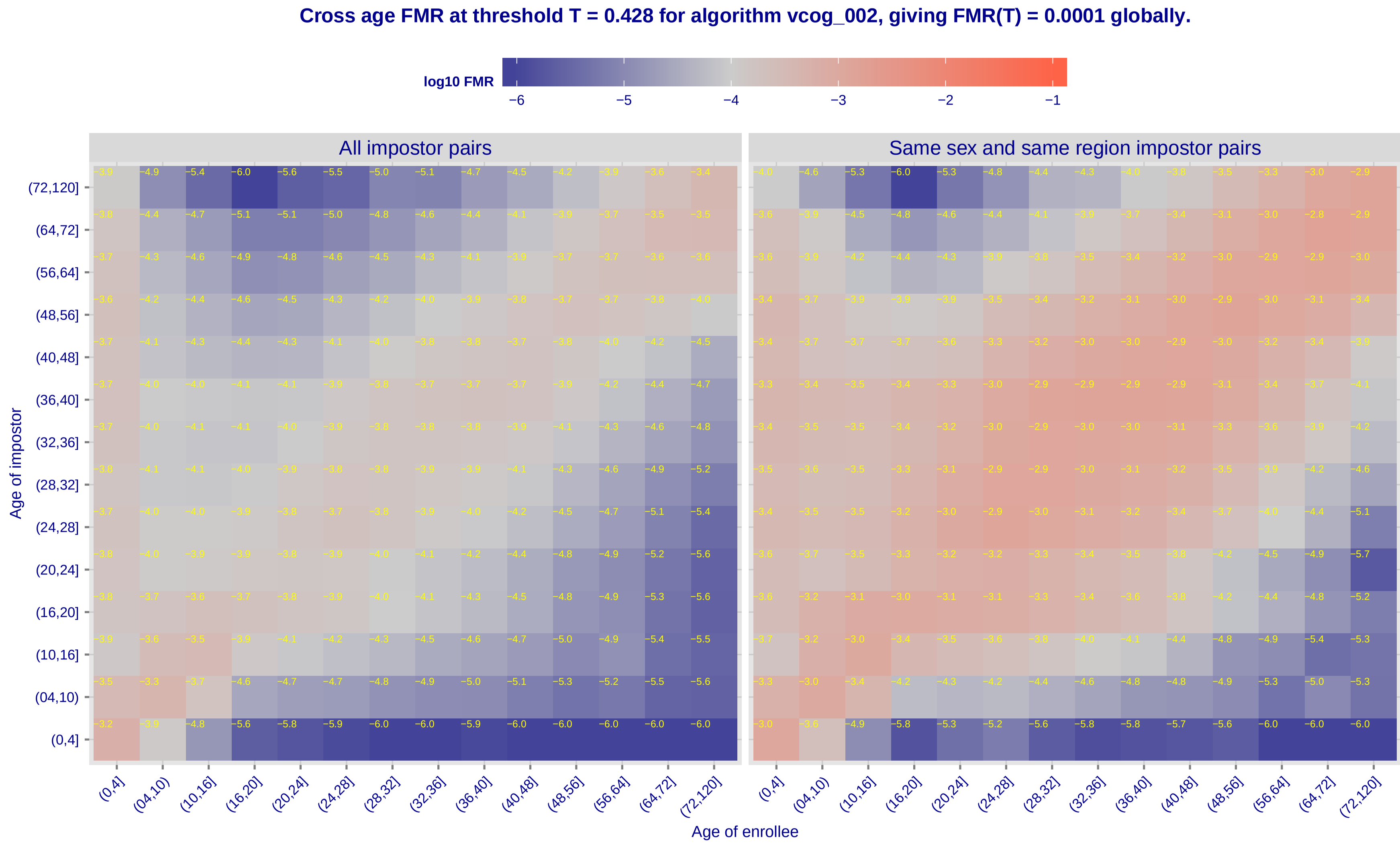


Figure 120: For algorithm vcog-002 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

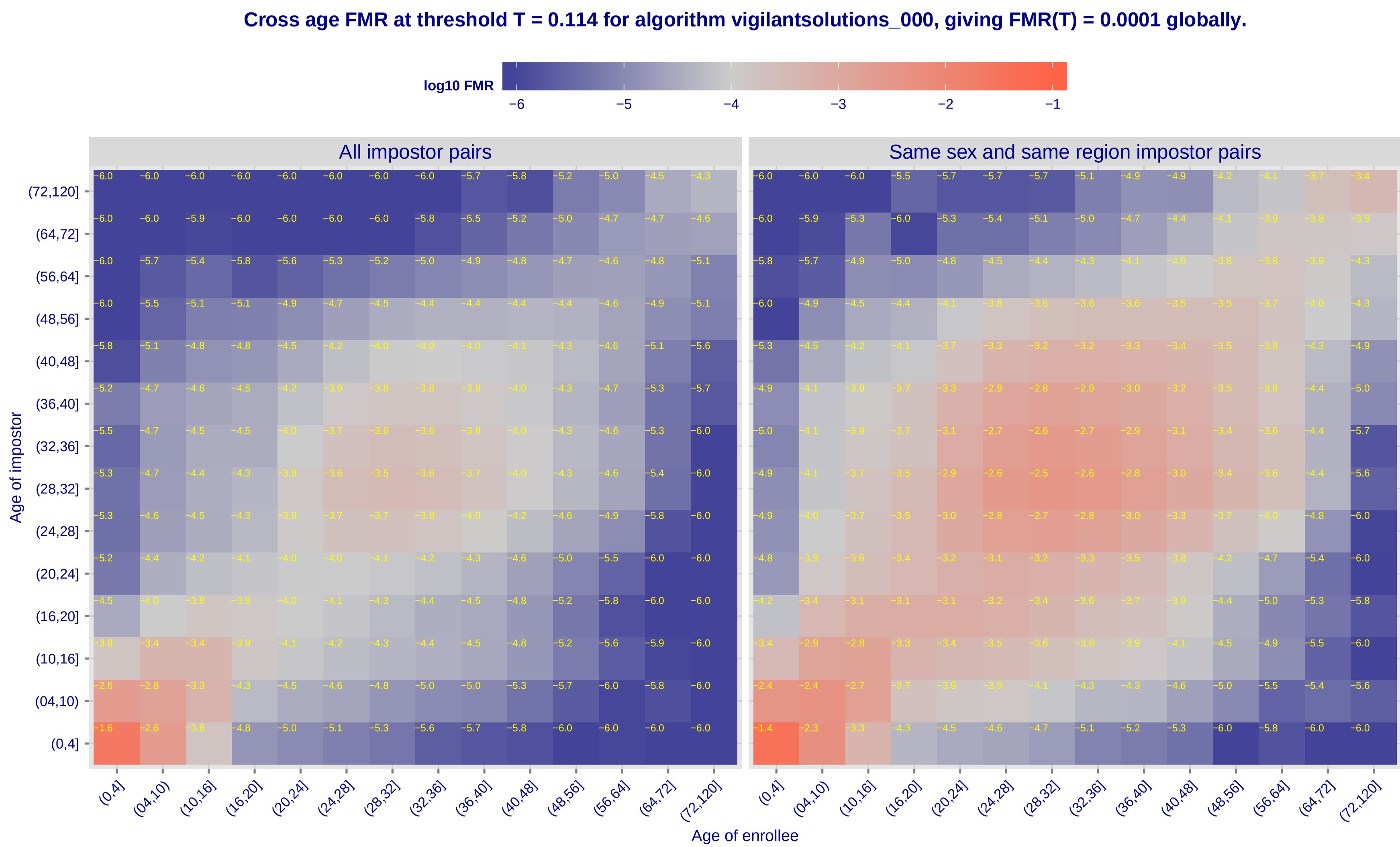


Figure 121: For algorithm vigilantsolutions-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

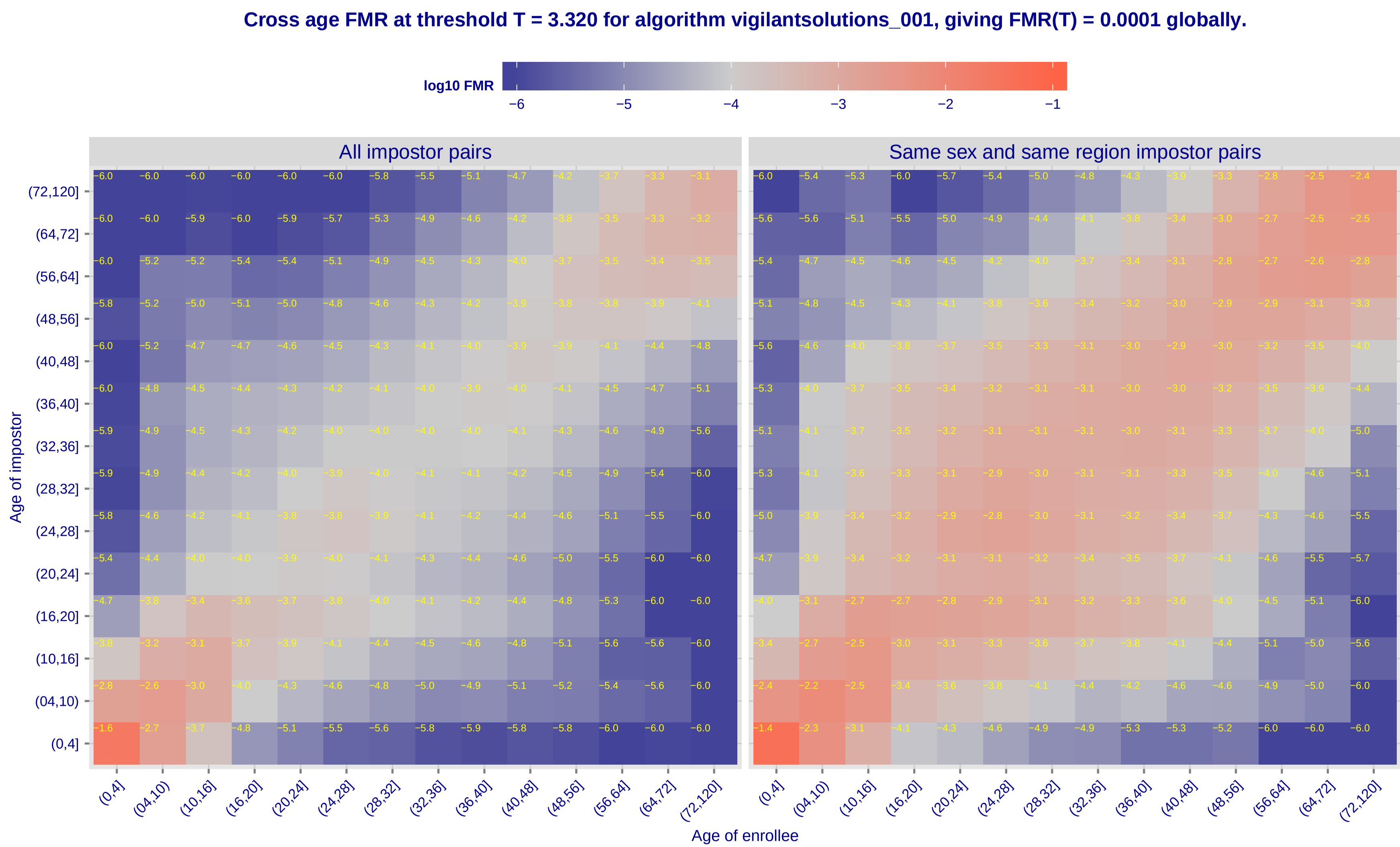


Figure 122: For algorithm vigilantsolutions-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

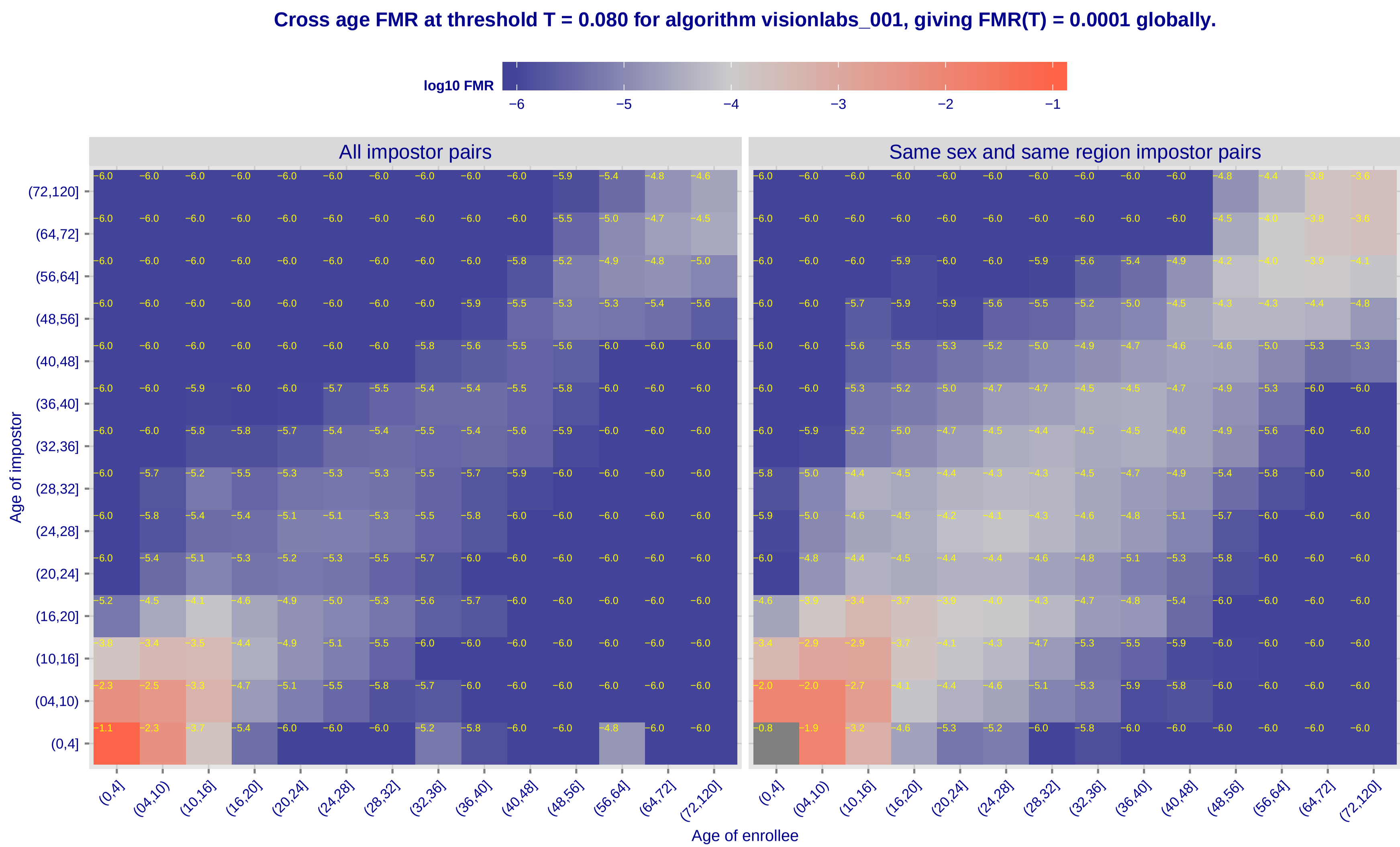


Figure 123: For algorithm visionlabs-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

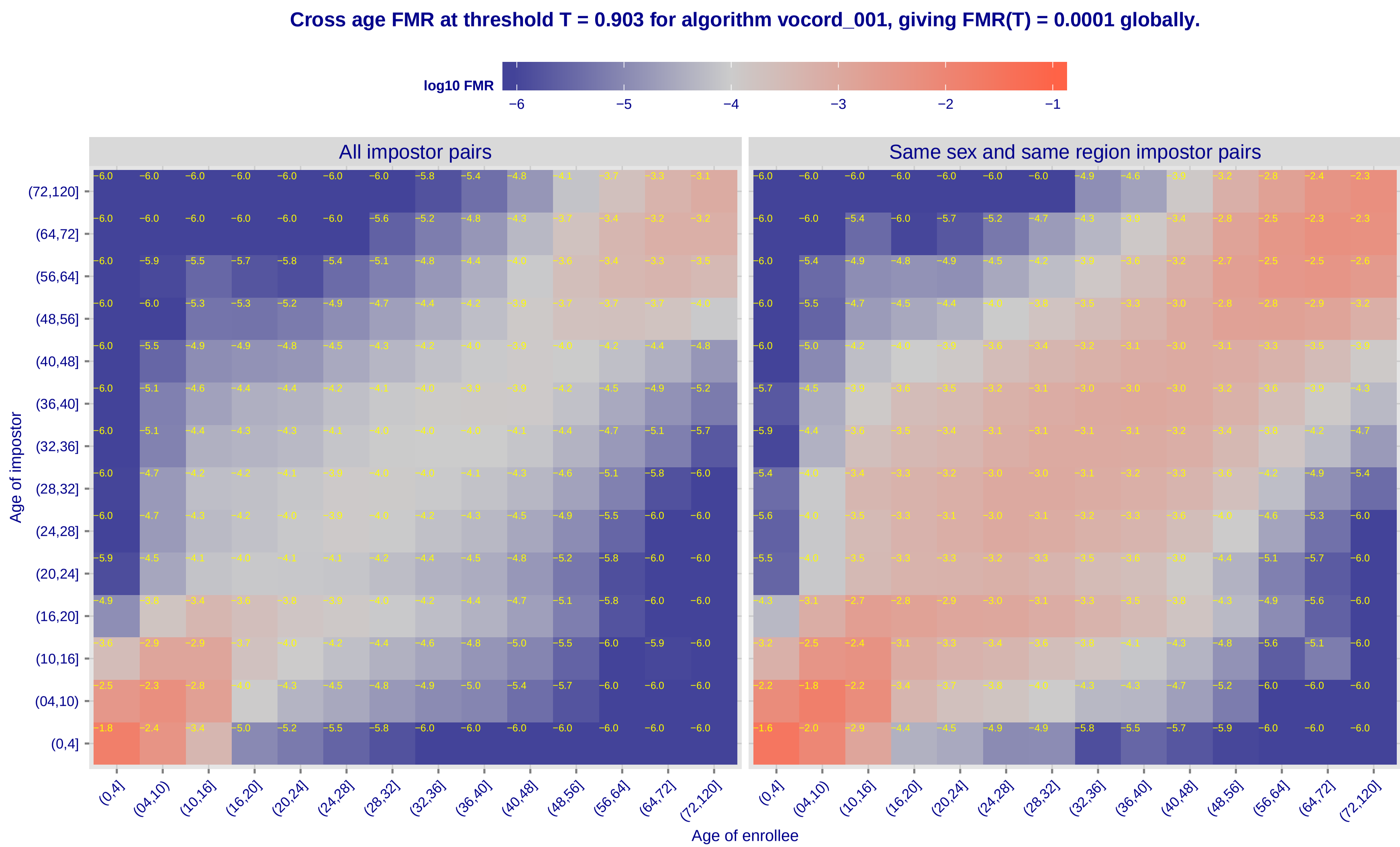


Figure 124: For algorithm vocord-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

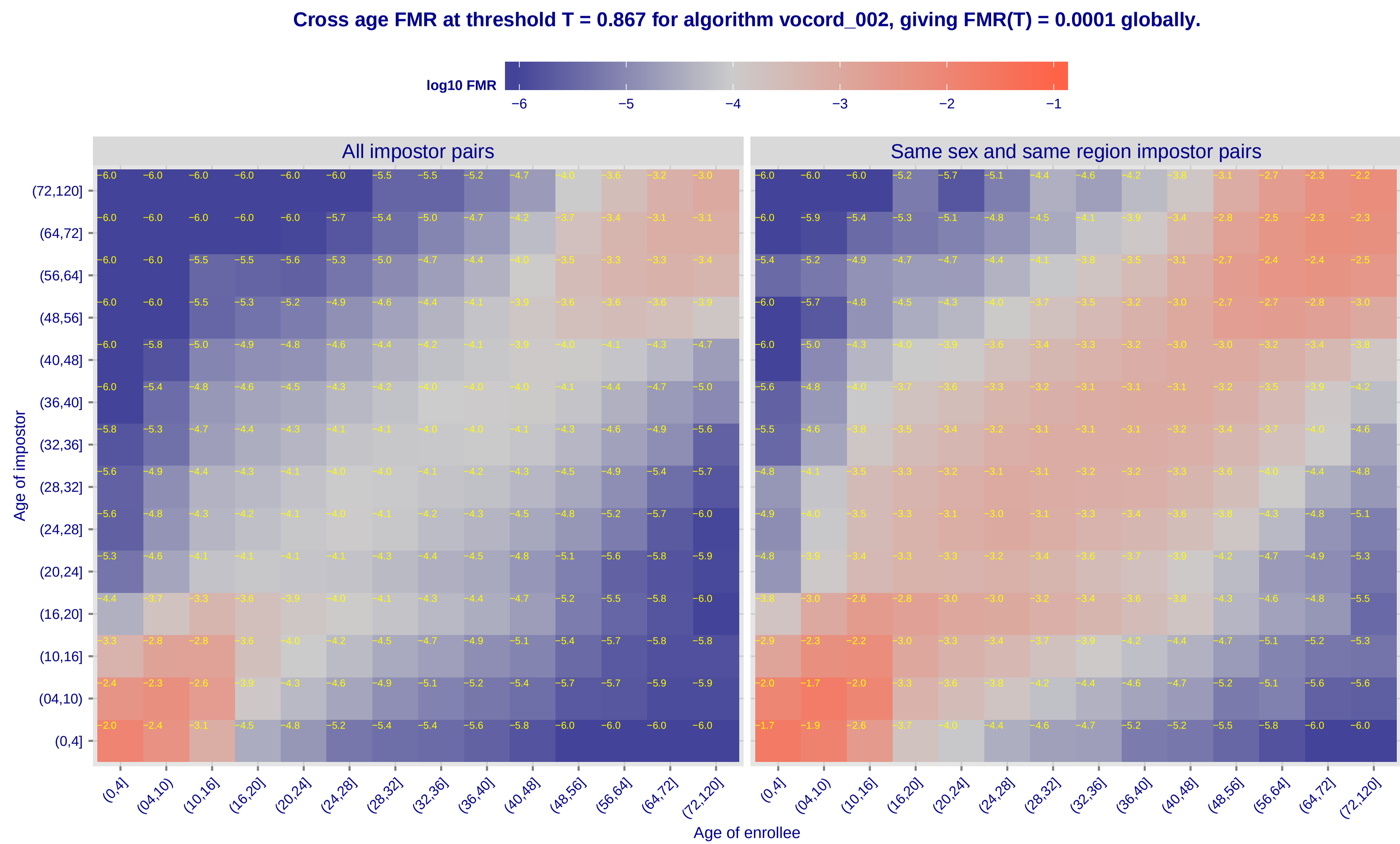


Figure 125: For algorithm vocord-002 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all O(10^{10}) impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

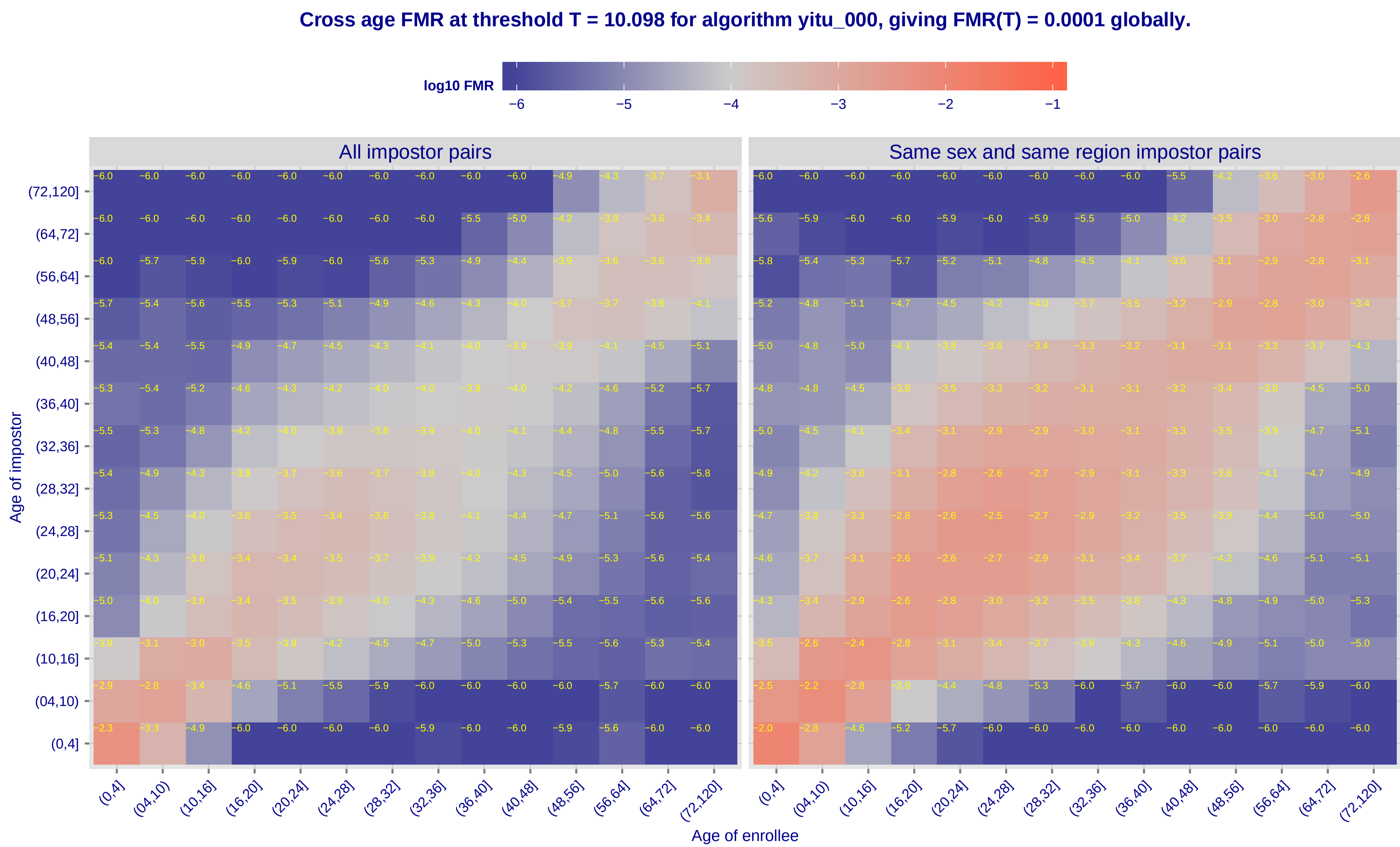


Figure 126: For algorithm yitu-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ impostor comparisons. The text in each box gives the same quantity as that coded by the color. Light colors present a security vulnerability to, for example, a passport gate.

Accuracy Terms + Definitions

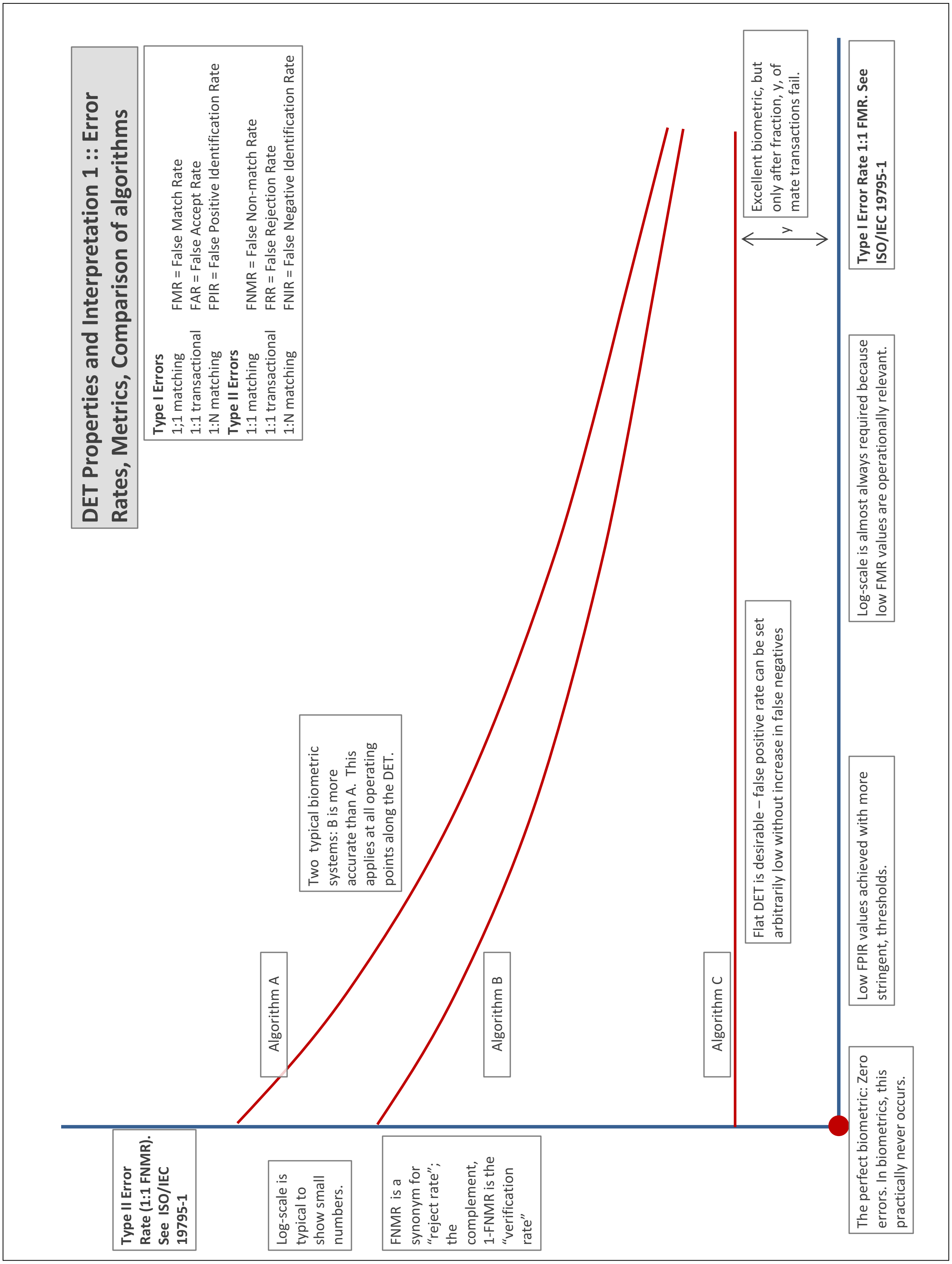
In biometrics, Type II errors occur when two samples of one person do not match – this is called a **false negative**. Correspondingly, Type I errors occur when samples from two persons do match – this is called a **false positive**. Matches are declared by a biometric system when the native comparison score from the recognition algorithm meets some **threshold**. Comparison scores can be either **similarity scores**, in which case higher values indicate that the samples are more likely to come from the same person, or **dissimilarity scores**, in which case higher values indicate different people. Similarity scores are traditionally computed by **fingerprint** and **face** recognition algorithms, while dissimilarities are used in **iris recognition**. In some cases, the dissimilarity score is a distance; this applies only when **metric** properties are obeyed. In any case, scores can be either **mate** scores, coming from a comparison of one person's samples, or **nonmate** scores, coming from comparison of different persons' samples. The words **genuine** or **authentic** are synonyms for mate, and the word **impostor** is used a synonym for nonmate. The words mate and nonmate are traditionally used in identification applications (such as law enforcement search, or background checks) while genuine and impostor are used in verification applications (such as access control).

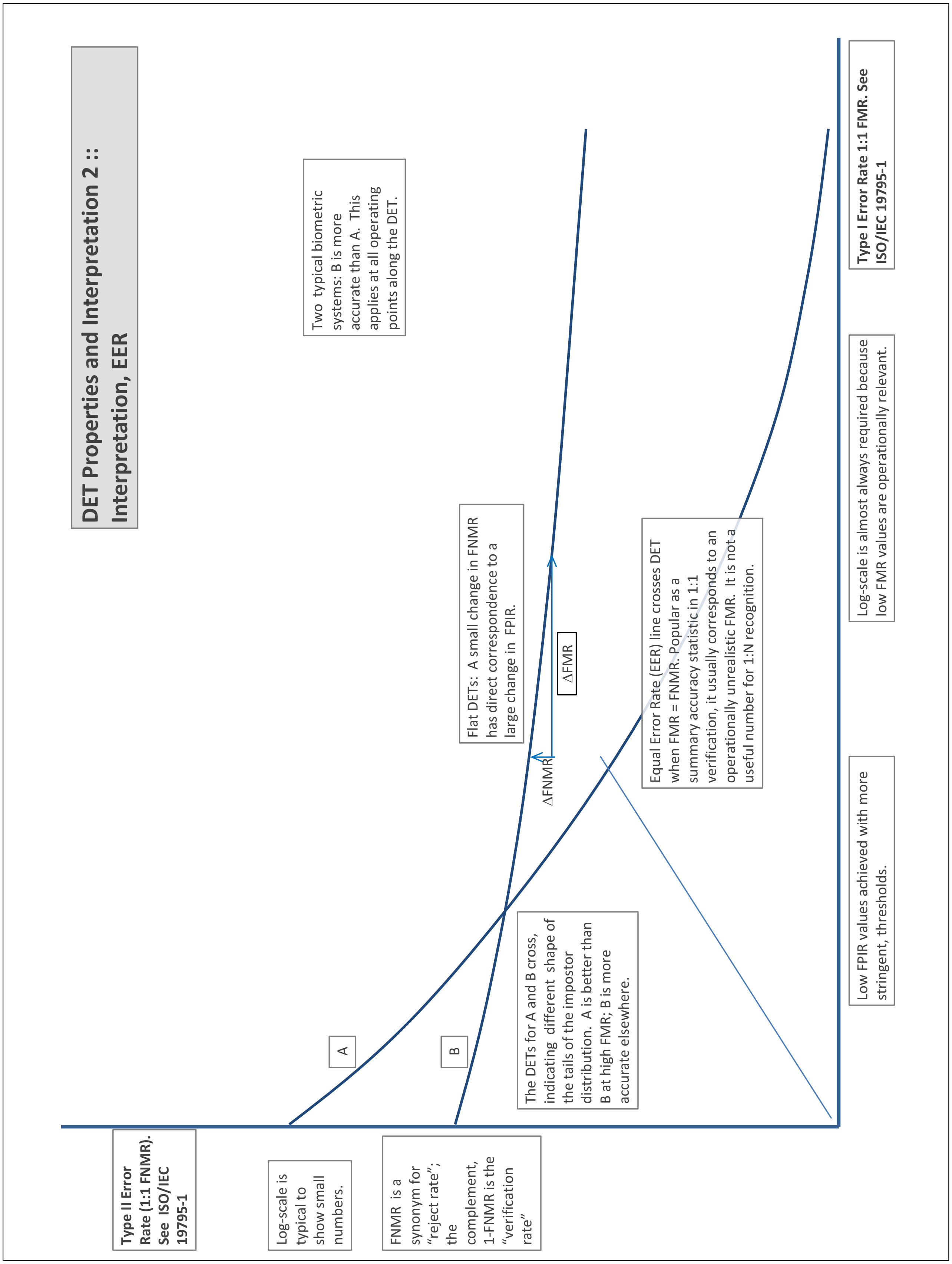
A **error tradeoff** characteristic represents the tradeoff between Type II and Type I classification errors. For verification this plots false non-match rate (FNMR) vs. false match rate (FMR) parametrically with T.

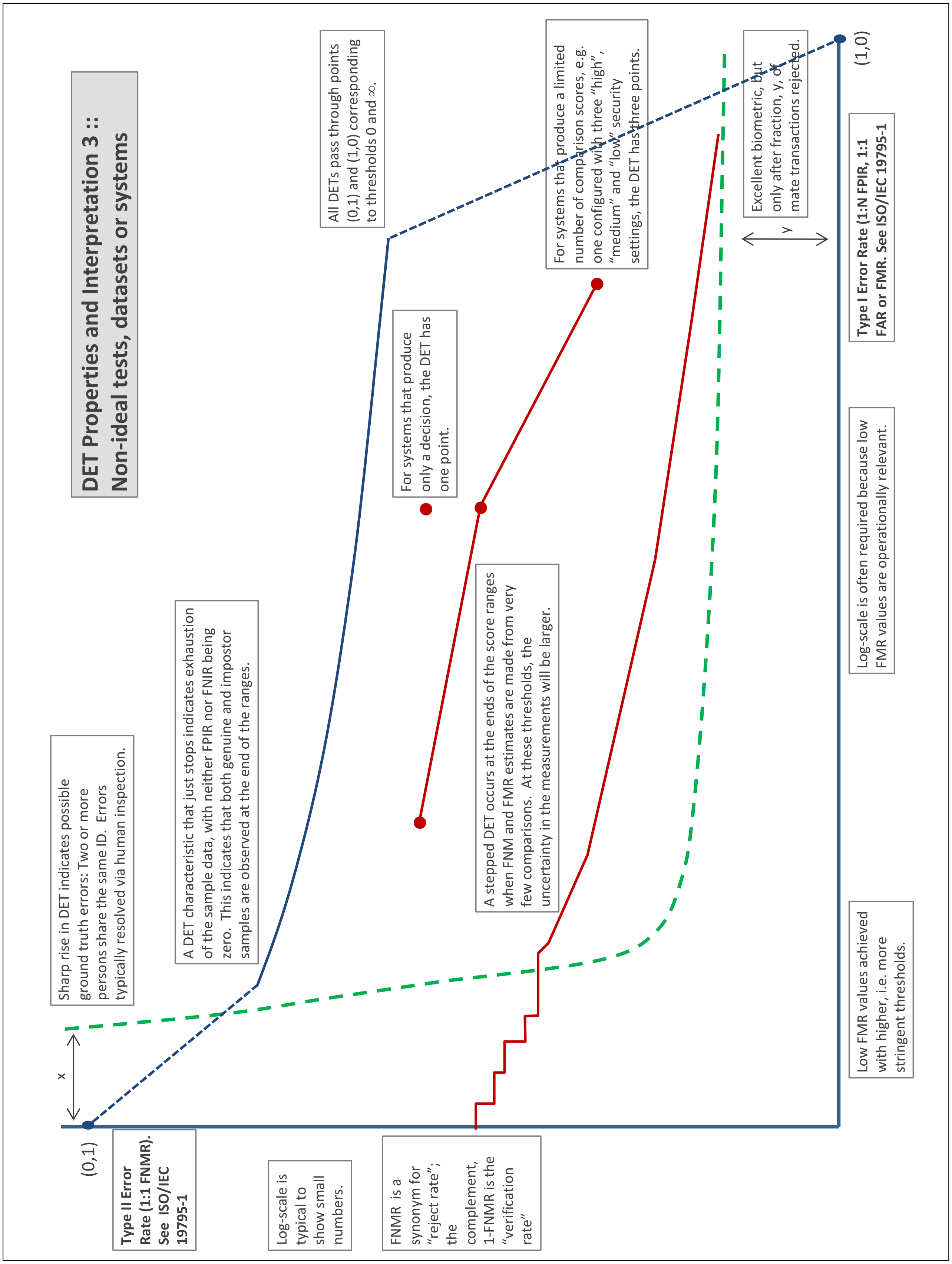
The error tradeoff plots are often called **detection error tradeoff (DET)** characteristics or **receiver operating characteristic** (ROC). These serve the same function but differ, for example, in plotting the complement of an error rate (e.g. TMR = 1 – FNMR) and in transforming the axes most commonly using logarithms, to show multiple decades of FMR. More rarely, the function might be the inverse Gaussian function.

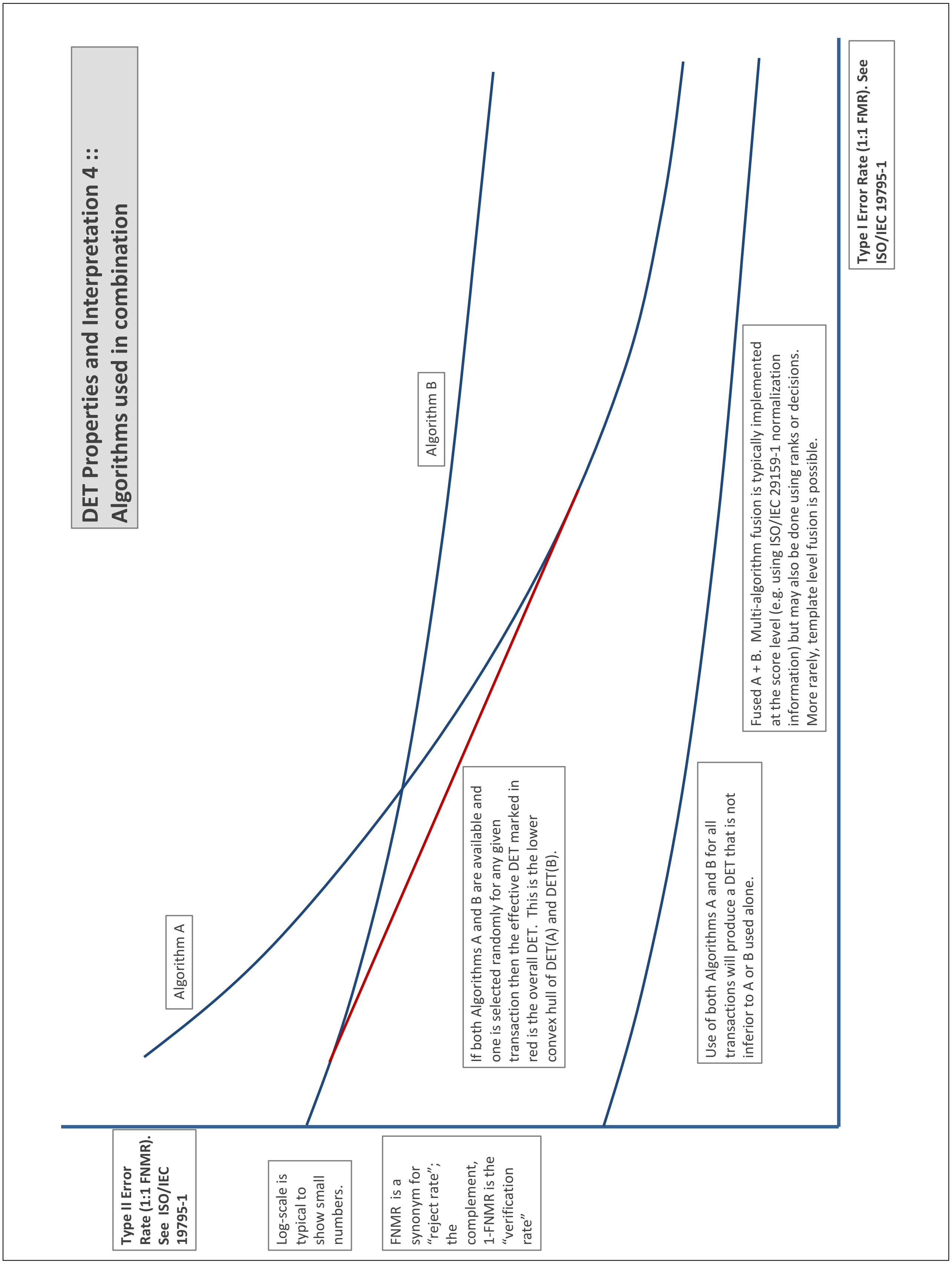
More detail and generality is provided in formal biometrics testing standards, see the various parts of ISO/IEC 19795 Biometrics Testing and Reporting. More terms, including and beyond those to do with accuracy, see ISO/IEC 2382-37 Information technology -- Vocabulary -- Part 37: Harmonized biometric vocabulary

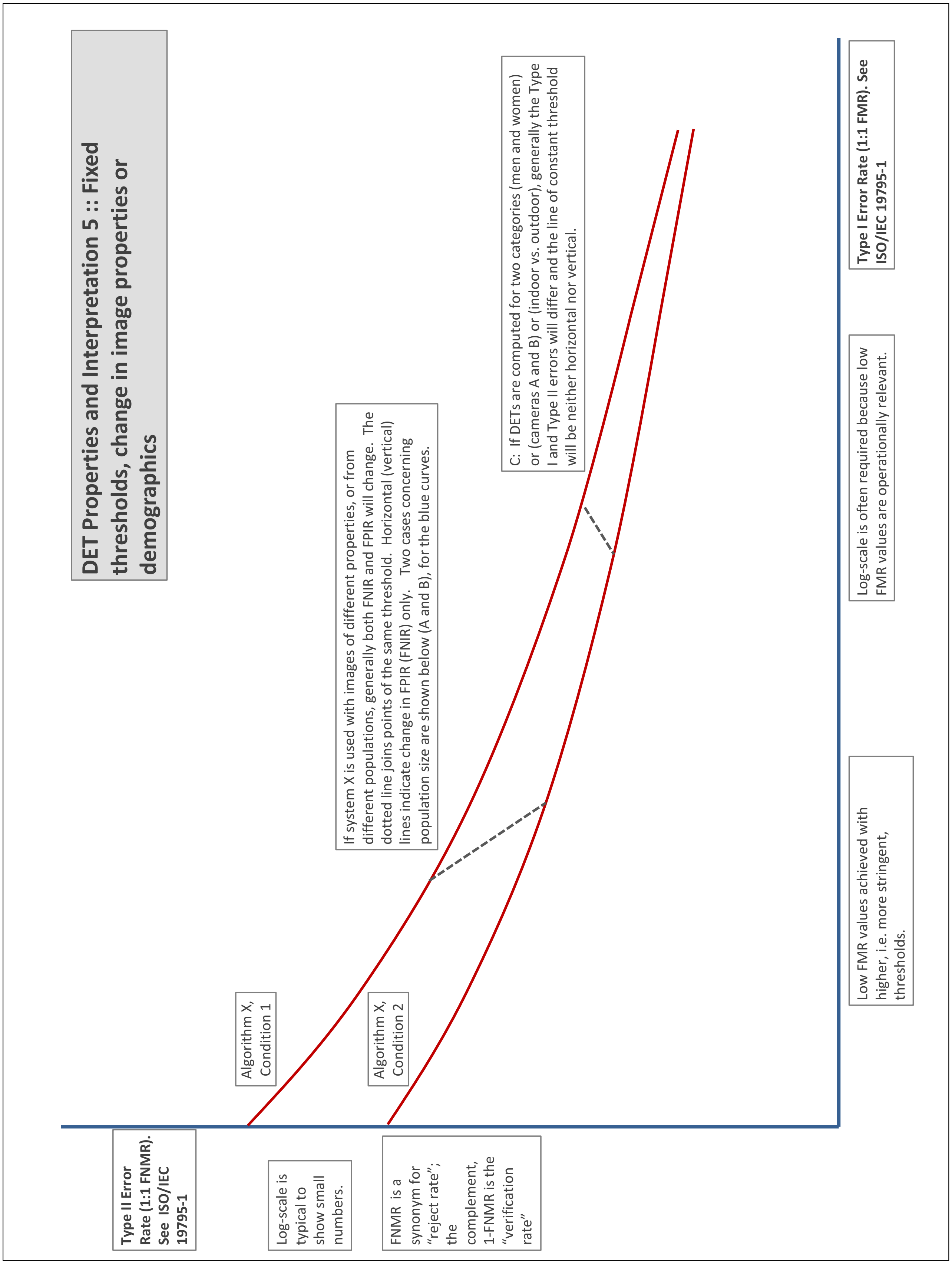
FNMR(T) "False non-match rate"
FMR(T) "False match rate"


DET Properties and Interpretation 1 :: Error Rates, Metrics, Comparison of algorithms
Type I Errors
1;1 matching FMR = False Match Rate
1:1 transactional FAR = False Accept Rate
1:N matching FPIR = False Positive Identification Rate
Type II Errors
1:1 matching FNMR = False Non-match Rate
1:1 transactional FRR = False Rejection Rate
1:N matching FNIR = False Negative Identification Rate
Two typical biometric systems: B is more accurate than A. This applies at all operating points along the DET.
Algorithm A
Algorithm B
Algorithm C
Flat DET is desirable – false positive rate can be set arbitrarily low without increase in false negatives
Excellent biometric, but only after fraction, y, of mate transactions fail.
y
Type I Error Rate 1:1 FMR. See ISO/IEC 19795-1
Log-scale is almost always required because low FMR values are operationally relevant.
Low FPIR values achieved with more stringent, thresholds.
The perfect biometric: Zero errors. In biometrics, this practically never occurs.
Type II Error Rate (1:1 FNMR). See ISO/IEC 19795-1
Log-scale is typical to show small numbers.
FNMR is a synonym for "reject rate", the complement, 1-FNMR is the "verification rate"


DET Properties and Interpretation 2 :: Interpretation, EER
Two typical biometric systems: B is more accurate than A. This applies at all operating points along the DET.
Flat DETs: A small change in FNMR has direct correspondence to a large change in FPIR.
ΔFMR
ΔFNMR
Equal Error Rate (EER) line crosses DET when FMR = FNMR. Popular as a summary accuracy statistic in 1:1 verification, it usually corresponds to an operationally unrealistic FMR. It is not a useful number for 1:N recognition.
A
B
The DETs for A and B cross, indicating different shape of the tails of the impostor distribution. A is better than B at high FMR; B is more accurate elsewhere.
Type I Error Rate 1:1 FMR. See ISO/IEC 19795-1
Log-scale is almost always required because low FMR values are operationally relevant.
Low FPIR values achieved with more stringent, thresholds.
Type II Error Rate (1:1 FNMR). See ISO/IEC 19795-1
Log-scale is typical to show small numbers.
FNMR is a synonym for "reject rate"; the complement, 1-FNMR is the "verification rate"


DET Properties and Interpretation 3 ::
Non-ideal tests, datasets or systems
Sharp rise in DET indicates possible ground truth errors: Two or more persons share the same ID. Errors typically resolved via human inspection.
A DET characteristic that just stops indicates exhaustion of the sample data, with neither FPIR nor FNIR being zero. This indicates that both genuine and impostor samples are observed at the end of the ranges.
All DETs pass through points (0,1) and (1,0) corresponding to thresholds 0 and ∞.
For systems that produce only a decision, the DET has one point.
A stepped DET occurs at the ends of the score ranges when FNM and FMR estimates are made from very few comparisons. At these thresholds, the uncertainty in the measurements will be larger.
For systems that produce a limited number of comparison scores, e.g. one configured with three "high", "medium" and "low" security settings, the DET has three points.
Excellent biometric, but only after fraction, y, of mate transactions rejected.
x
y
(0,1)
(1,0)
Type II Error Rate (1:1 FNMR). See ISO/IEC 19795-1
Log-scale is typical to show small numbers.
FNMR is a synonym for "reject rate"; the complement, 1-FNMR is the "verification rate"
Low FMR values achieved with higher, i.e. more stringent thresholds.
Log-scale is often required because low FMR values are operationally relevant.
Type I Error Rate (1:N FPIR, 1:1 FAR or FMR. See ISO/IEC 19795-1


DET Properties and Interpretation 4 ::
Algorithms used in combination
Algorithm A
Algorithm B
If both Algorithms A and B are available and one is selected randomly for any given transaction then the effective DET marked in red is the overall DET. This is the lower convex hull of DET(A) and DET(B).
Use of both Algorithms A and B for all transactions will produce a DET that is not inferior to A or B used alone.
Fused A + B. Multi-algorithm fusion is typically implemented at the score level (e.g. using ISO/IEC 29159-1 normalization information) but may also be done using ranks or decisions. More rarely, template level fusion is possible.
Type I Error Rate (1:1 FMR). See ISO/IEC 19795-1
Type II Error Rate (1:1 FNMR). See ISO/IEC 19795-1
Log-scale is typical to show small numbers.
FNMR is a synonym for "reject rate", the complement, 1-FNMR is the "verification rate"


DET Properties and Interpretation 5 :: Fixed thresholds, change in image properties or demographics
Type II Error Rate (1:1 FNMR). See ISO/IEC 19795-1
Log-scale is typical to show small numbers.
FNMR is a synonym for "reject rate", the complement, 1-FNMR is the "verification rate"
Algorithm X, Condition 1
Algorithm X, Condition 2
If system X is used with images of different properties, or from different populations, generally both FNIR and FPIR will change. The dotted line joins points of the same threshold. Horizontal (vertical) lines indicate change in FPIR (FNIR) only. Two cases concerning population size are shown below (A and B), for the blue curves.
C: If DETs are computed for two categories (men and women) or (cameras A and B) or (indoor vs. outdoor), generally the Type I and Type II errors will differ and the line of constant threshold will be neither horizontal nor vertical.
Low FMR values achieved with higher, i.e. more stringent, thresholds.
Log-scale is often required because low FMR values are operationally relevant.
Type I Error Rate (1:1 FMR). See ISO/IEC 19795-1