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

Specific hardware and software products identified in this report were used in order to perform the evaluations described in this document. In no case does identification of any commercial product, trade name, or vendor, imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products and equipment identified are necessarily the best available for the purpose.

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.

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

Contents

List of Tables

List of Figures

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

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

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

1 Changelog

1.1 July 29 2017

- Added 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

1.2 June 19 2017

- 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. 1 showing tradeoff between FNMR, template size, template generation time, and match duration.
- Added Fig. 13 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.

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

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
3DiVi	3divi	000	2017-03-16	169360	[5]512 ± 0	[14]285 ± 52	Yes	[1]378 ± 20	[2]375 ± 19
Ayonix	ayonix	000	2017-06-22	58505	[6]1036 ± 0	[1]18 ± 2	No	[2]621 ± 23	[3]620 ± 26
Dermalog	dermalog	001	2017-02-22	0	[7]1043 ± 0	[6]106 ± 1	No	[12]22160 ± 166	[16]22131 ± 131
Dermalog	dermalog	002	2017-02-22	0	[9]1043 ± 0	[3]81 ± 0	No	[13]22169 ± 114	[15]22105 ± 146
Dermalog	dermalog	003	2017-07-10	0	[8]1043 ± 0	[10]121 ± 22	No	[14]22957 ± 93	[17]22808 ± 131
Digital Barriers	barriers	000	2017-05-31	157794	[13]2056 ± 0	[5]104 ± 0	No	[8]13232 ± 166	[11]13226 ± 146
Digital Barriers	barriers	001	2017-07-20	236934	[14]2056 ± 0	[15]294 ± 1	No	[7]12311 ± 164	[9]12347 ± 197
Is It You	isityou	000	2017-06-26	48010	[24]19200 ± 0	[9]113 ± 5	No	[25]237517 ± 1318	[25]237374 ± 1279
ITMO University	itmo	001	2017-06-12	751338	[26]37997 ± 0	[23]870 ± 4	No	[15]29119 ± 1420	[18]27817 ± 340
Morpho	morpho	000	2017-07-11	100806	[1]116 ± 0	[8]109 ± 1	Yes	[4]993 ± 31	[5]1000 ± 34
Neurotechnology	neurotech	000	2017-03-22	62129	[23]7148 ± 0	[20]611 ± 48	No	[23]74288 ± 2194	[23]72879 ± 2640
N-Tech Lab	ntech	000	2017-03-13	191530	[17]2906 ± 1	[13]278 ± 13	No	[17]30787 ± 142	[19]30846 ± 77
N-Tech Lab	ntech	001	2017-05-10	691296	[22]6744 ± 1	[19]587 ± 11	No	[21]67692 ± 833	[22]67486 ± 244
Rank One Computing	rankone	000	2017-03-21	0	[2]144 ± 0	[4]82 ± 9	No	[19]39932 ± 468	[8]8722 ± 171
Rank One Computing	rankone	001	2017-04-12	0	[4]208 ± 0	[2]56 ± 4	No	[22]72754 ± 658	[1]345 ± 29
Samtech InfoNet Limited	samtech	000	2017-05-02	109774	[12]2056 ± 0	[11]262 ± 2	No	[6]4550 ± 26	[7]4541 ± 28
TongYi Transportation Technology	tongyi	001	2017-04-01	625339	[16]2058 ± 0	[16]310 ± 20	No	[10]17769 ± 74	[13]17750 ± 63
TongYi Transportation Technology	tongyi	002	2017-07-15	625336	[15]2058 ± 0	[17]356 ± 35	No	[16]29816 ± 281	[14]17799 ± 127
Tupel	tupel	001	2017-05-05	6347	[20]4142 ± 0	[12]273 ± 0	No	[20]56353 ± 618	[21]59204 ± 563
VCognition	vcog	001	2017-03-28	86103	[18]4126 ± 0	[7]108 ± 17	Yes	[9]16320 ± 197	[12]16426 ± 425
VCognition	vcog	002	2017-06-12	3229434	[27]61504 ± 5	[18]357 ± 25	No	[26]296154 ± 3077	[26]296436 ± 4183
Vigilant Solutions	vigilant	000	2017-03-30	352218	[25]31540 ± 0	[24]884 ± 23	No	[11]18201 ± 94	[10]13030 ± 83
Vigilant Solutions	vigilant	001	2017-06-13	344685	[11]1544 ± 0	[26]921 ± 2	No	[3]644 ± 13	[4]649 ± 16
VisionLabs	visionlabs	001	2017-06-12	343661	[3]204 ± 0	[27]943 ± 8	No	[5]1395 ± 45	[6]1148 ± 53
Vocord	vocord	001	2017-04-21	616989	[21]6194 ± 0	[25]908 ± 16	No	[27]1094730 ± 64282	[27]1107193 ± 66523
Vocord	vocord	002	2017-06-07	918292	[10]1330 ± 0	[22]782 ± 36	Yes	[24]83063 ± 517	[24]83072 ± 714
Shanghai Yitu Technology	yitu	000	2017-05-23	2211068	[19]4130 ± 0	[21]672 ± 2	No	[18]35352 ± 114	[20]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 R Xeon R 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 that 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	FALSE NON-MATCH RATE (FNMR)													
Name	CHILD EXP		MUGSHOT		SELFIE		VISA		VISA		WEBCAM		WILD	
FMR	0.01		0.0001		0.0001		1E-06		0.0001		0.0001		0.0001	
3divi-000/2017-03-16	0.553	3	0.037	6	0.055	8	0.133	7	0.029	6	0.002	3	0.547	4
ayonix-000/2017-06-22	0.843	17	0.309	21	0.360	21	0.487	21	0.230	22	0.172	20	0.807	17
dermalog-001/2017-02-22	0.985	24	0.237	19	0.193	20	0.305	14	0.113	15	0.112	19	0.704	9
dermalog-002/2017-02-22	0.985	23	0.241	20	0.179	19	0.315	15	0.122	16	0.109	18	0.709	11
dermalog-003/2017-07-10	0.845	19	0.202	18	0.115	16	0.280	13	0.112	14	0.041	16	0.693	8
digitalbarriers-000/2017-05-31	0.771	14	0.184	17	0.170	17	0.463	20	0.161	18	0.045	17	0.741	15
digitalbarriers-001/2017-07-20	-	27	0.041	9	0.115	15	0.502	22	0.155	17	0.029	14	0.678	7
isityou-000/2017-06-26	-	25	0.680	24	-	25	0.703	24	0.414	24	0.690	24	1.000	25
itmo-001/2017-06-12	0.797	16	-	27	-	27	0.441	18	0.171	20	-	27	-	27
morpho-000/2017-07-11	0.846	20	0.028	4	0.012	3	0.134	8	0.026	4	0.007	7	0.893	20
neurotechnology-000/2017-03-22	0.845	18	0.062	12	0.052	7	0.237	11	0.052	11	0.005	6	0.943	22
ntechlab-000/2017-03-13	0.533	2	0.044	10	0.014	6	0.086	6	0.027	5	0.003	4	0.367	2
ntechlab-001/2017-05-10	0.472	1	0.030	5	0.014	5	0.083	5	0.025	3	0.003	5	0.319	1
rankone-000/2017-03-21	0.787	15	0.177	16	0.092	14	0.276	12	0.071	12	0.021	11	0.723	12
rankone-001/2017-04-12	0.858	21	0.091	14	0.176	18	0.420	17	0.171	21	0.024	13	0.842	18
samtech-000/2017-05-02	0.765	13	0.044	11	0.063	11	0.443	19	0.161	19	0.021	12	0.878	19
tongyitrans-001/2017-04-01	0.743	9	0.041	8	0.063	10	0.072	4	0.038	10	0.009	8	0.704	10
tongyitrans-002/2017-07-15	0.746	10	0.039	7	0.063	12	0.066	3	0.030	7	0.010	9	0.725	13
tupel-001/2017-05-05	-	26	0.641	23	-	26	1.000	27	0.680	27	-	26	-	26
vcog-001/2017-03-28	0.686	6	-	26	0.427	22	0.892	25	0.409	23	0.302	21	-	24
vcog-002/2017-06-12	0.752	11	0.692	25	0.666	24	0.903	26	0.504	26	0.559	23	0.778	16
vigilantsolutions-000/2017-03-30	0.894	22	0.595	22	0.643	23	0.688	23	0.415	25	0.401	22	0.915	21
vigilantsolutions-001/2017-06-13	0.730	8	0.101	15	0.061	9	0.348	16	0.105	13	0.016	10	0.729	14
visionlabs-001/2017-06-12	0.561	4	0.024	3	0.014	4	0.180	10	0.030	8	0.001	2	0.591	5
vocord-001/2017-04-21	0.695	7	0.063	13	0.069	13	0.141	9	0.035	9	0.036	15	0.654	6
vocord-002/2017-06-07	0.762	12	0.019	2	0.012	1	0.034	2	0.013	1	-	25	0.948	23
yitu-000/2017-05-23	0.586	5	0.017	1	0.012	2	0.033	1	0.021	2	0.000	1	0.431	3

Table 2: FNMR is the proportion of mated comparisons below a threshold.

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

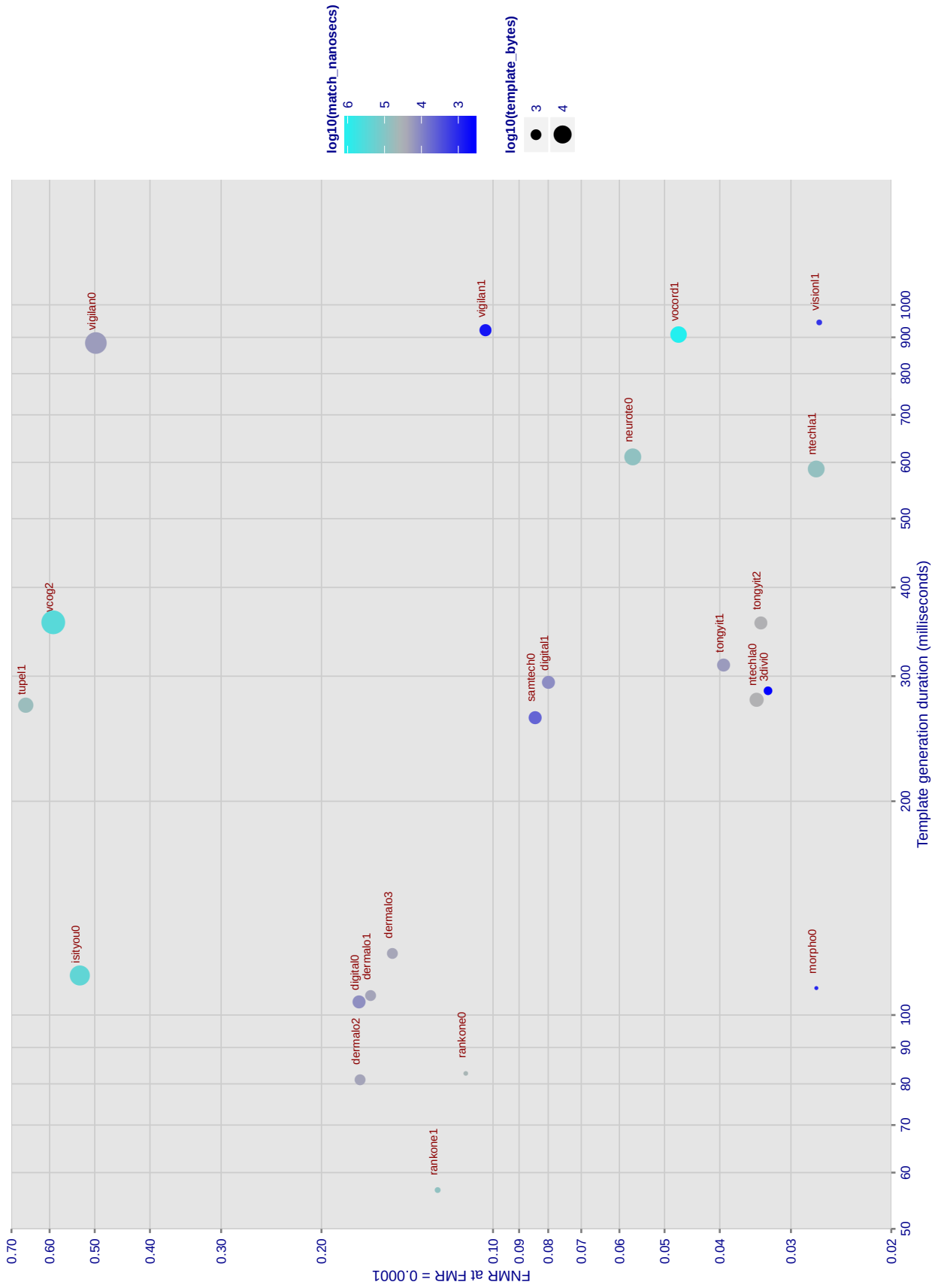


Figure 1: The points show false non-match rates (FNMR) versus the duration of the template generation operation. FNMR is the geometric mean of FNMR values for visa and mugshot images (from Figs. 2 and 3) at a false match rate (FMR) of 0.0001. Template generation time is a median estimated over 640 x 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 template comparison duration - which span three orders of magnitude.

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

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.

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

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, with significant imbalance due to visa issuance patterns.
- ▷ The images are of subjects 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).

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

- ▷ 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 ”wild” - see Table 4 of the FRVT API.

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 arbitrarily unconstrained. Pose varies considerably around all yaw and pitch axes. 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	14	0.0019	15	0.0202	15	0.0008	9	0.0020	13	0.2070	17
ayonix-000	0.0000	1	0.0109	27	0.0751	22	0.0137	30	0.0109	17	0.0000	2
dermalog-001	0.9109	29	0.0045	21	0.0954	25	0.0013	12	0.0471	25	0.3979	21
dermalog-002	0.9109	30	0.0045	22	0.0954	26	0.0013	13	0.0471	26	0.3979	22
dermalog-003	0.0434	8	0.0007	6	0.0000	1	0.0025	22	0.0007	7	0.0701	10
digitalbarriers-000	0.5469	25	0.0043	19	0.0925	23	0.0019	20	0.0184	21	0.5170	27
isityou-000	0.4714	23	0.0022	17	0.0665	21	0.0010	10	0.0116	19	0.4586	25
itmo-001	0.5751	28	0.0103	26	-	26	0.0047	28	-	26	0.7752	30
morpho-000	0.0000	2	0.0000	1	0.0000	4	0.0000	1	0.0000	4	0.0000	4
neurotechnology-000	0.0000	4	0.0000	2	0.0000	5	0.0000	2	0.0000	5	0.0163	7
ntechlab-000	0.2496	17	0.0015	13	0.0058	12	0.0016	16	0.0007	9	0.1099	12
ntechlab-001	0.0926	12	0.0009	7	0.0029	7	0.0005	7	0.0007	6	0.0584	9
rankone-000	0.0187	7	0.0005	4	0.0000	6	0.0003	5	0.0007	10	0.2349	18
rankone-001	0.0012	6	0.0001	3	0.0000	3	0.0000	3	0.0000	3	0.0858	11
samtech-000	0.5474	26	0.0052	23	0.0491	19	0.0042	27	0.0252	23	0.7023	29
tongyitrans-001	0.0000	5	0.0068	24	0.0462	18	0.0040	26	0.0055	16	0.0000	5
tongyitrans-002	0.3609	21	0.0078	25	0.0462	16	0.0040	25	0.0055	15	0.0000	3
tupel-001	-	26	0.0030	18	0.0000	2	0.0018	19	0.0000	2	-	26
vcog-001	0.1579	13	-	26	0.0058	11	0.0018	17	0.0000	1	-	26
vcog-002	0.2209	15	0.0021	16	0.0087	13	0.0019	21	0.0007	8	0.1672	14
vigilantsolutions-000	0.5580	27	0.0018	14	0.0462	17	0.0007	8	0.0109	18	0.5927	28
vigilantsolutions-001	0.3585	20	0.0010	9	0.0116	14	0.0004	6	0.0048	14	0.3262	20
visionlabs-001	0.2699	18	0.0014	10	0.0058	10	0.0014	15	0.0020	12	0.1803	15
vocord-001	0.4732	24	0.0158	28	0.0520	20	0.0038	24	0.0348	24	0.4494	24
vocord-002	0.3782	22	0.0015	11	0.0029	9	0.0037	23	0.0171	20	0.1992	16
yitu-000	0.3475	19	0.0015	12	0.0029	8	0.0013	14	0.0014	11	0.1591	13

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.
The effects of FTE are included in the accuracy results later in this report by regarding any template comparison that involves an failed template is taken to produce a low similarity score.

4.4 Recognition accuracy

Core algorithm accuracy is stated via:

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

- The summary table of Figure 2;
- The visa image DETs of Figure 2;
- The mugshot DETs of Figure 3 ;
- The selfie-portrait DETs of Figure 4;
- The webcam-portrait DETs of Figure 5;
- The child-exploitation DET of Figure 7;
- The child-exploitation CMC of Figure 8.

Figure 12 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 14 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.

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

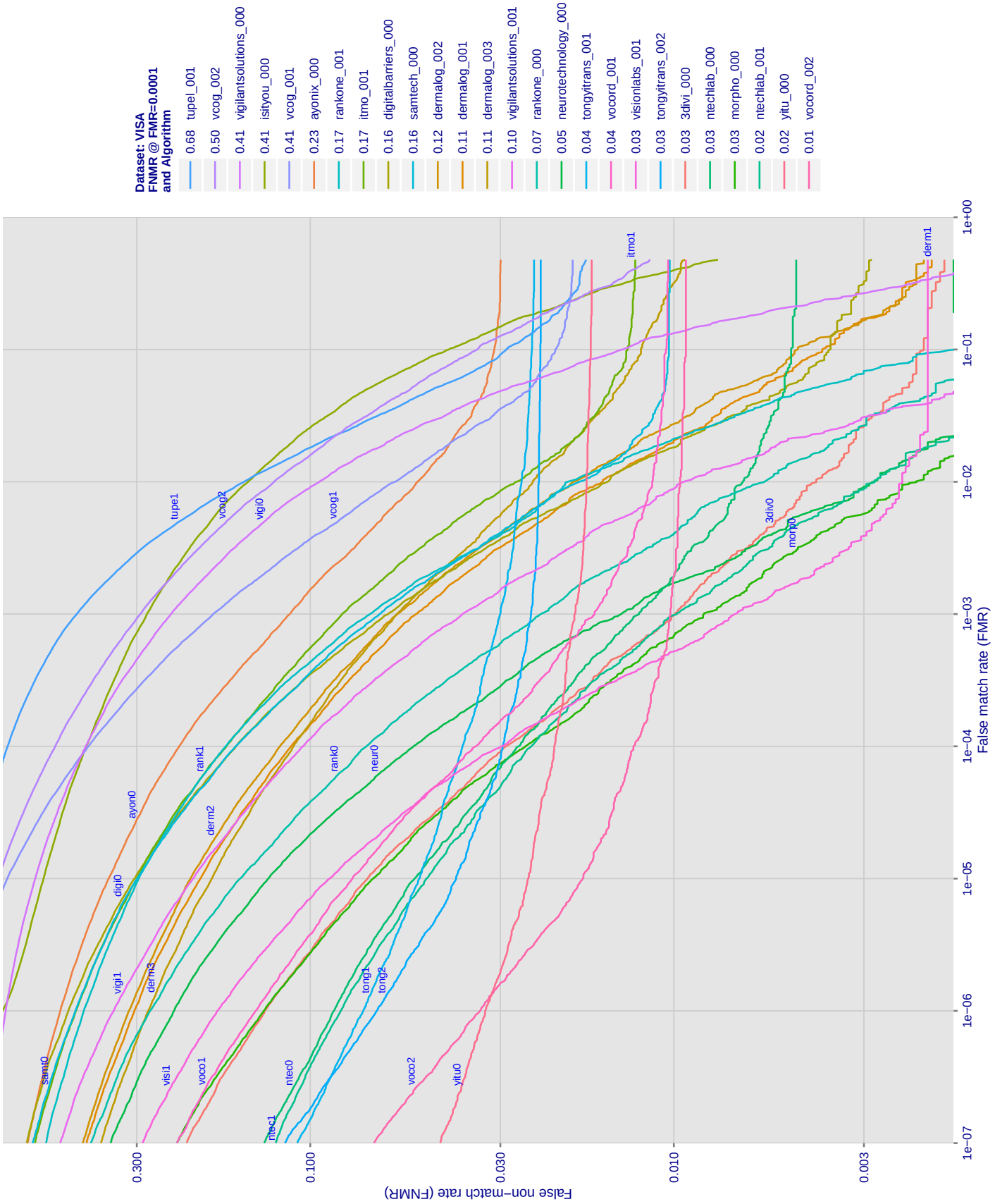


Figure 2: For the visa images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of FMR.

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

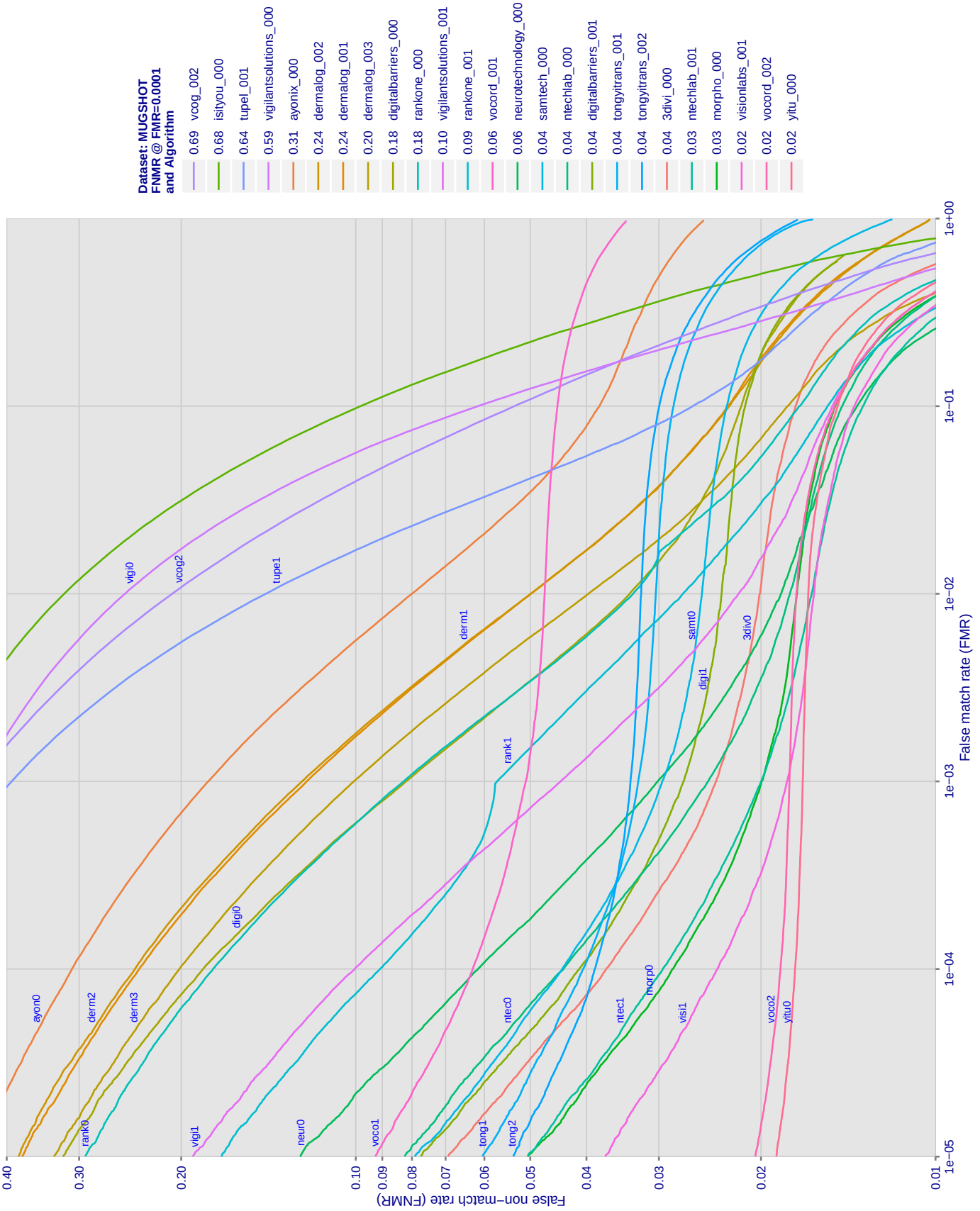


Figure 3: For the mugshot images, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show decades of FMR.

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

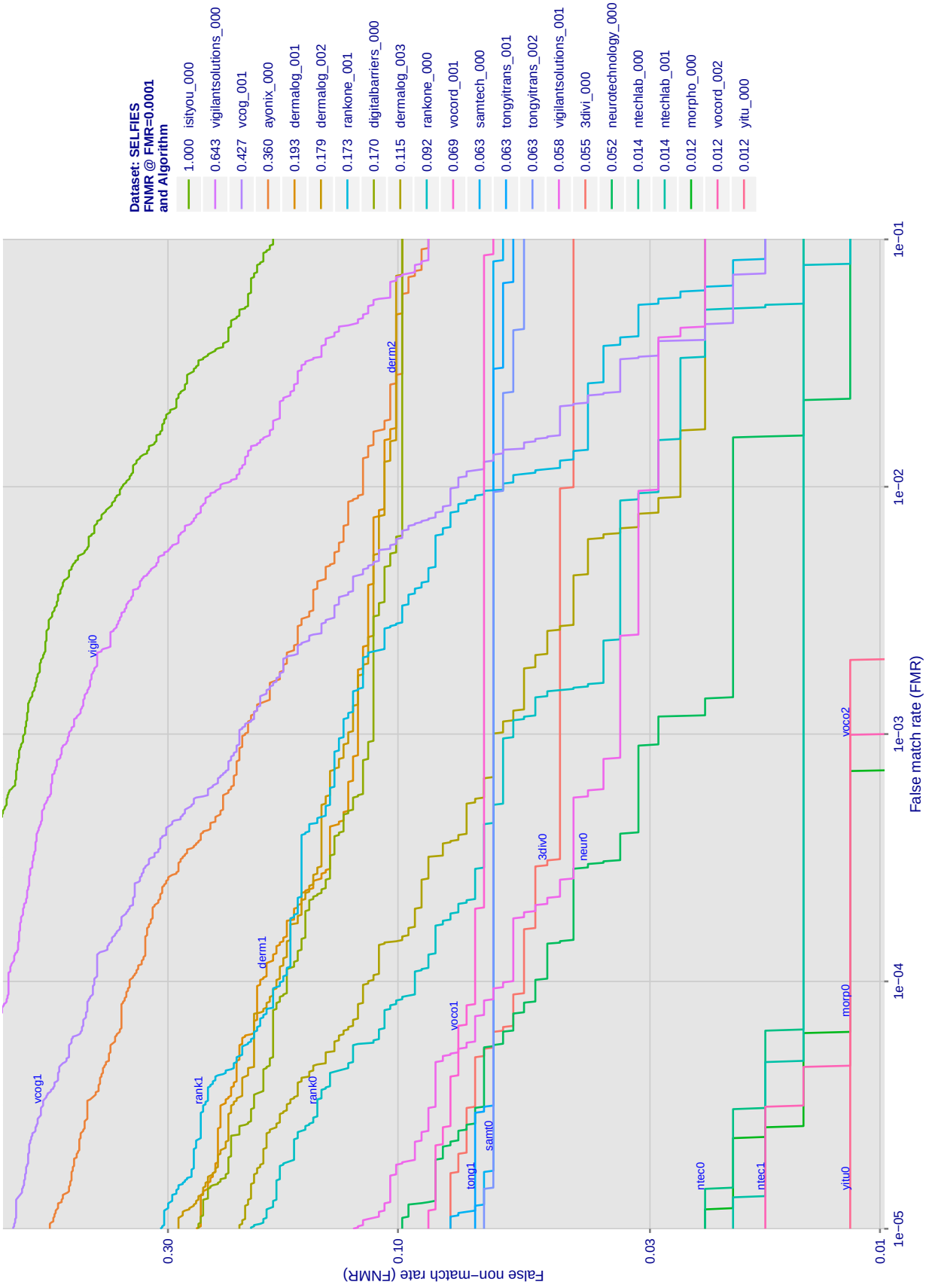


Figure 4: For the selfie-to-portrait comparisons, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show several decades of FMR. Caution: The FNMR values here are optimistic statements 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.

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

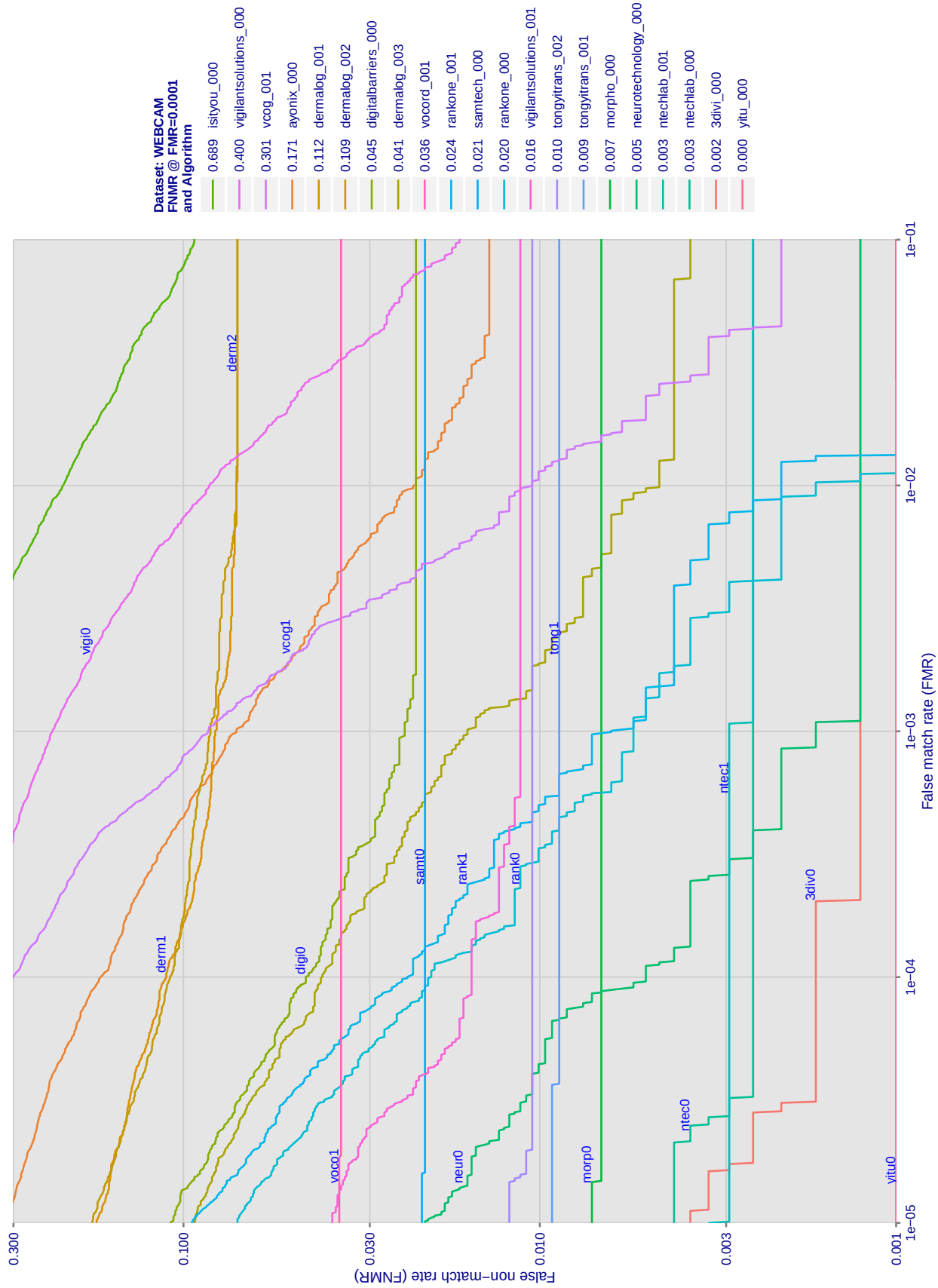


Figure 5: For the webcam-to-portrait comparisons, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show several decades of FMR. Caution: The FNMR values here are optimistic statements 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.

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

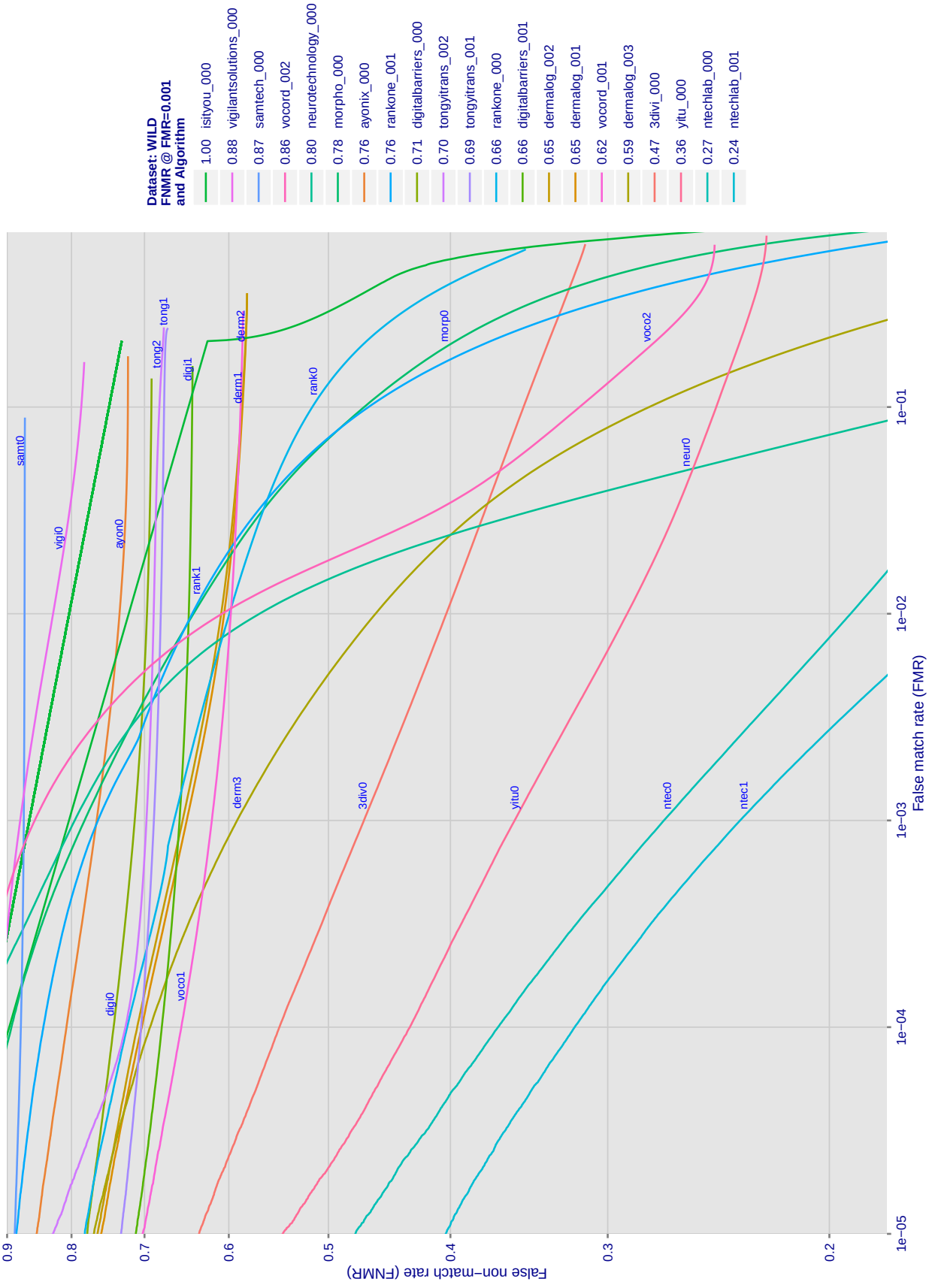


Figure 6: For the wild image comparisons, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show several decades of FMR.

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

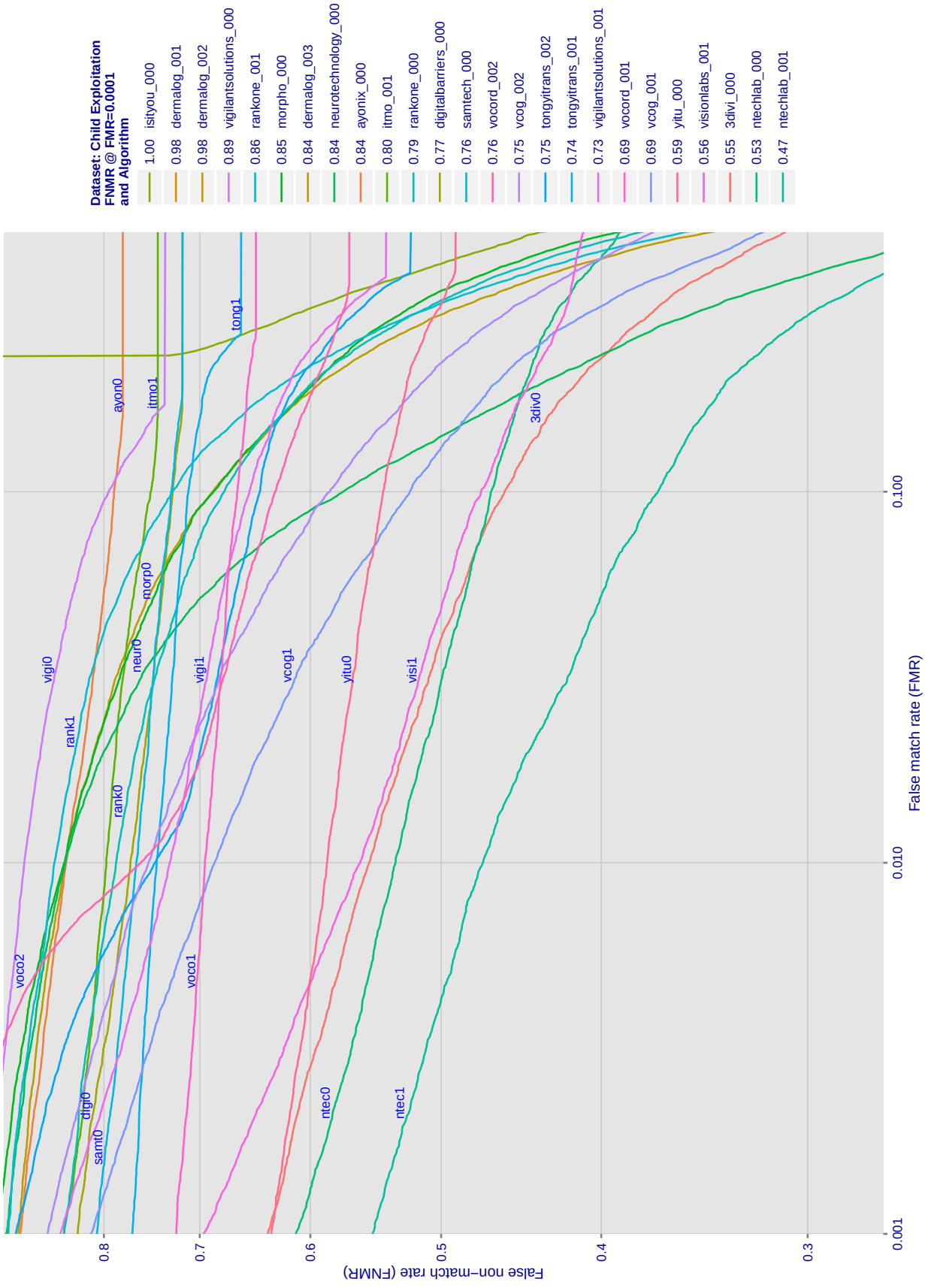


Figure 7: For child exploitation images, detection error tradeoff (DET) *characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. The scales are logarithmic in order to show many decades of* FMR. *Accuracy is poor because many images have adverse quality characteristics, and because detection and enrollment fails.*

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

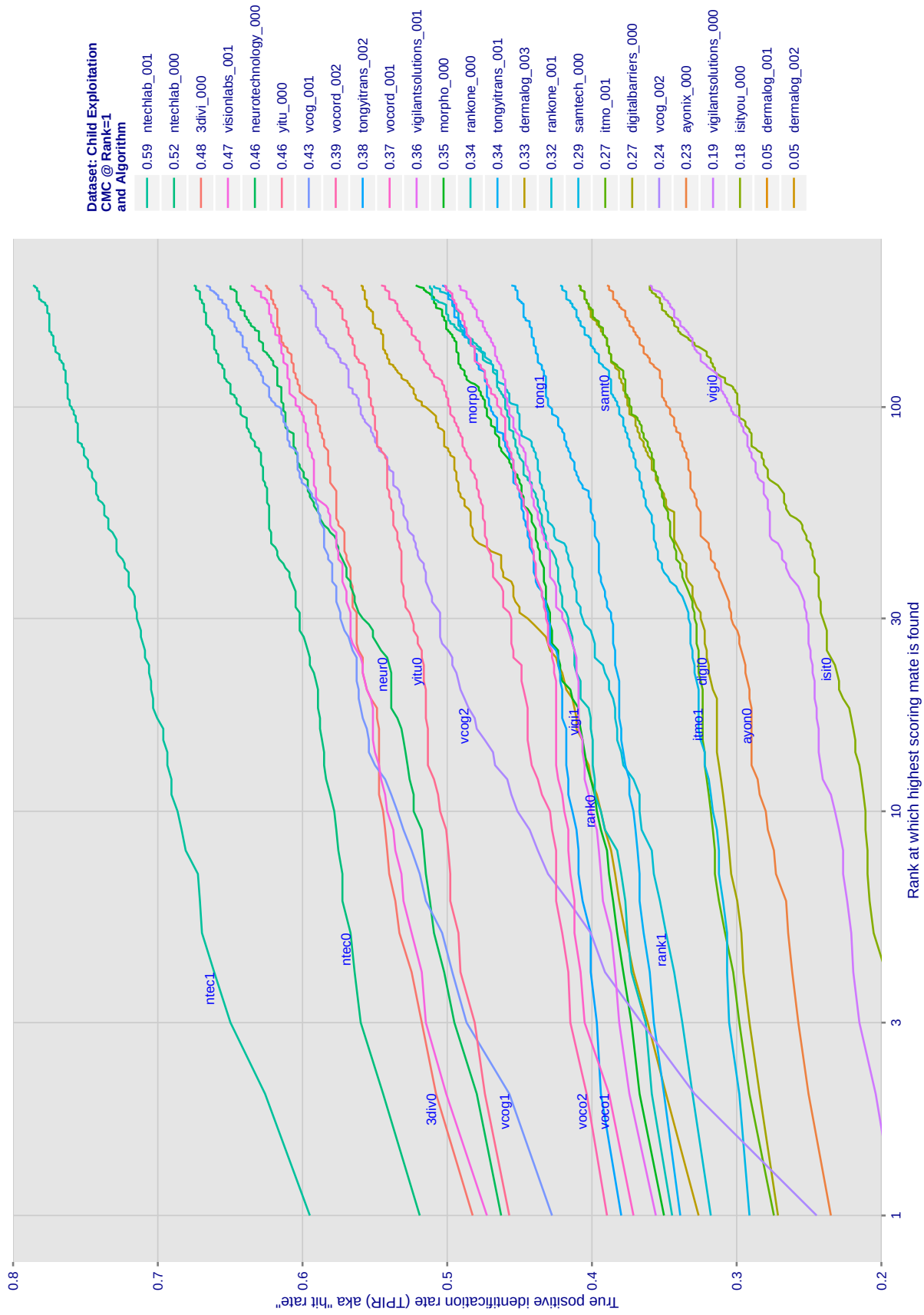


Figure 8: For child exploitation images, cumulative match characteristics (CMC) showing false negative identification rate vs. rank. This is simulation of a one-to-many search experiment - see discussion in section 4.2. The scales are logarithmic in order to show the effect of long candidate lists. Accuracy is poor but much improved relative to the 1:1 DETs of Fig. 7 because a search can succeed if any of a subject's several enrolled images matches the search image with a high score.

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

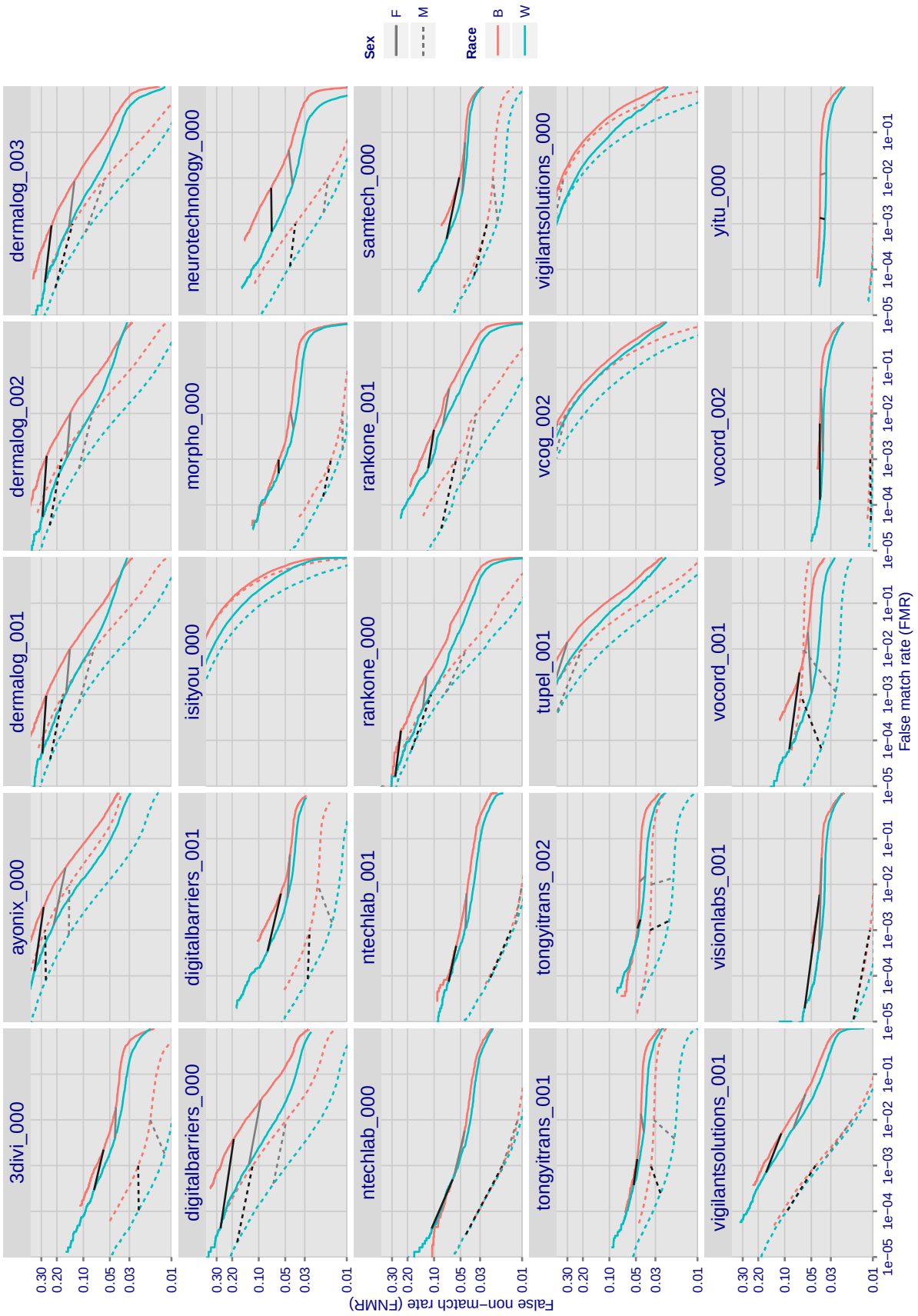


Figure 9: For the mugshot images, error tradeoff characteristics for white females, black females, black males and white males. The grey lines correspond to fixed thresholds, showing how both FNMR and FMR vary at one operating threshold. Important: Many of the plots will naively be read as saying whites 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 14. As access control systems almost always operate at a fixed threshold, the naive interpretation is incorrect.

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

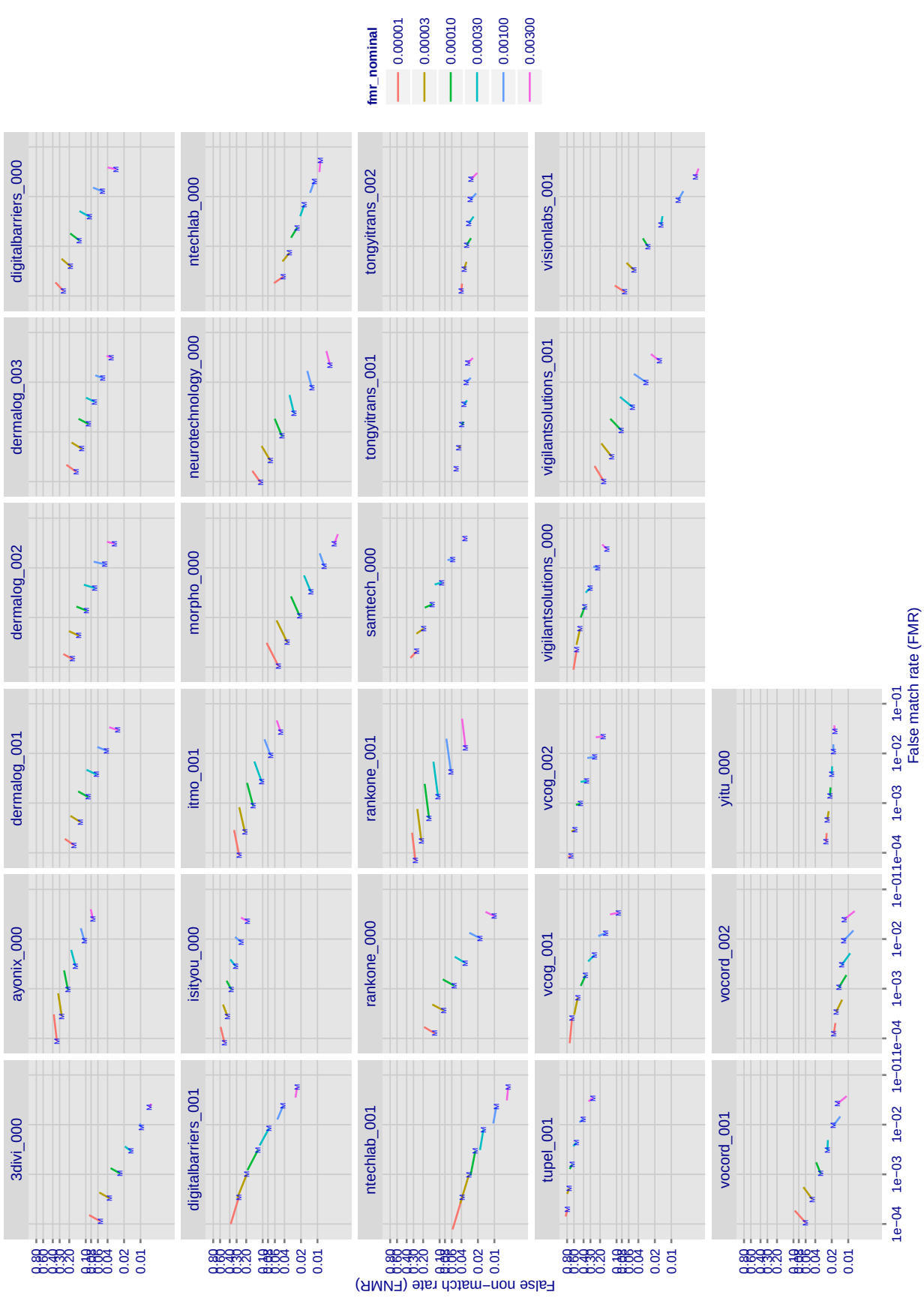


Figure 10: For the visa images, FNMR and FMR at six operating points along the DET characteristic. At each point a line is drawn between (FMR,FNMR)$_{\text{MALE}}$ and (FMR,FNMR)$_{\text{FEMALE}}$ showing how which sex has lower FMR and/or FNMR. The "M" label denotes male, the other end of the line corresponds 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).

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

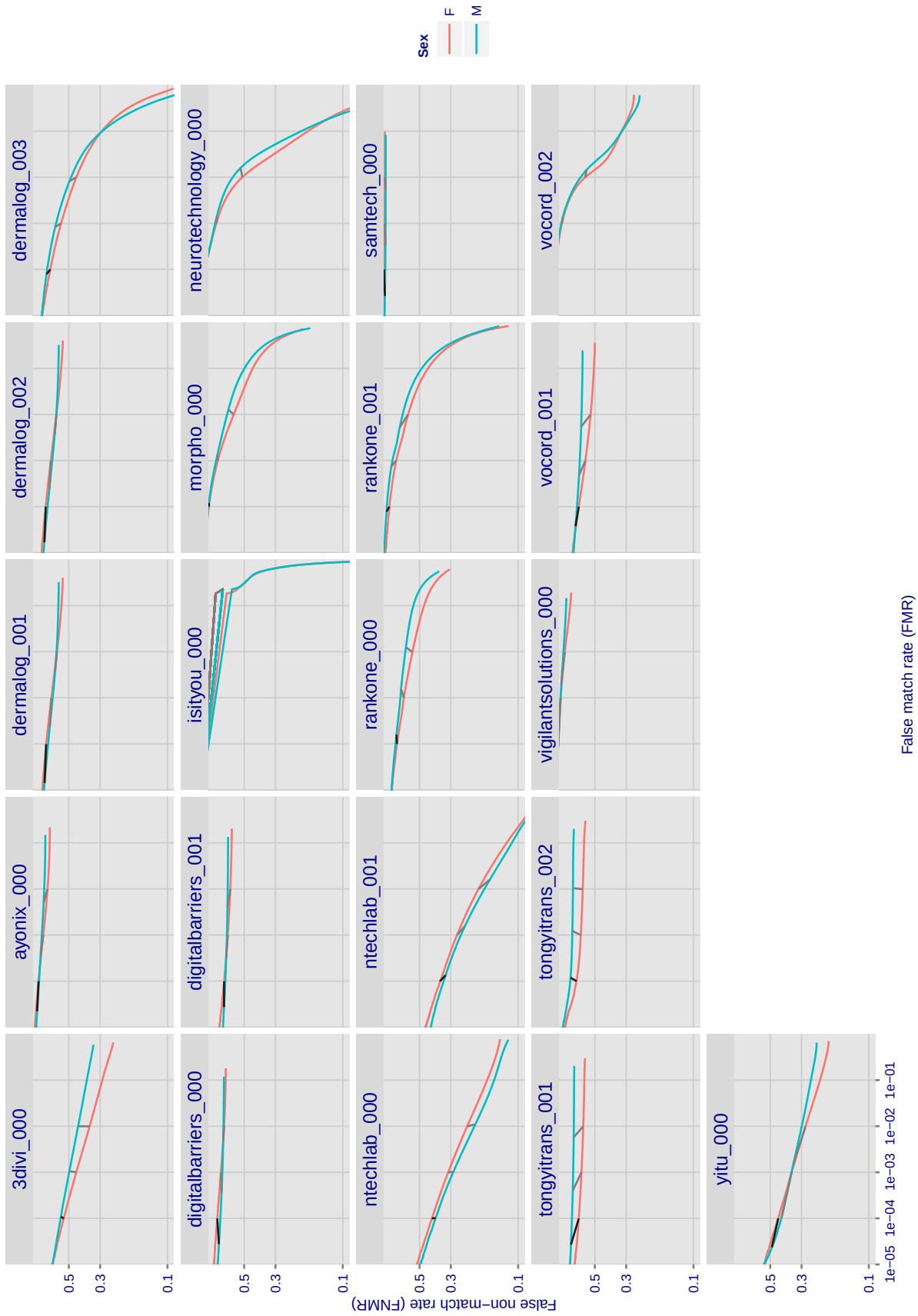


Figure 11: For the wild image comparisons, detection error tradeoff (DET) characteristics showing false non-match rate vs. false match rate plotted parametrically on threshold, T. Error rates are higher here than in the generic wild DET (Fig 6) becuase the impostor pairs here are same-sex only. The scales are logarithmic in order to show several decades of FMR.

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

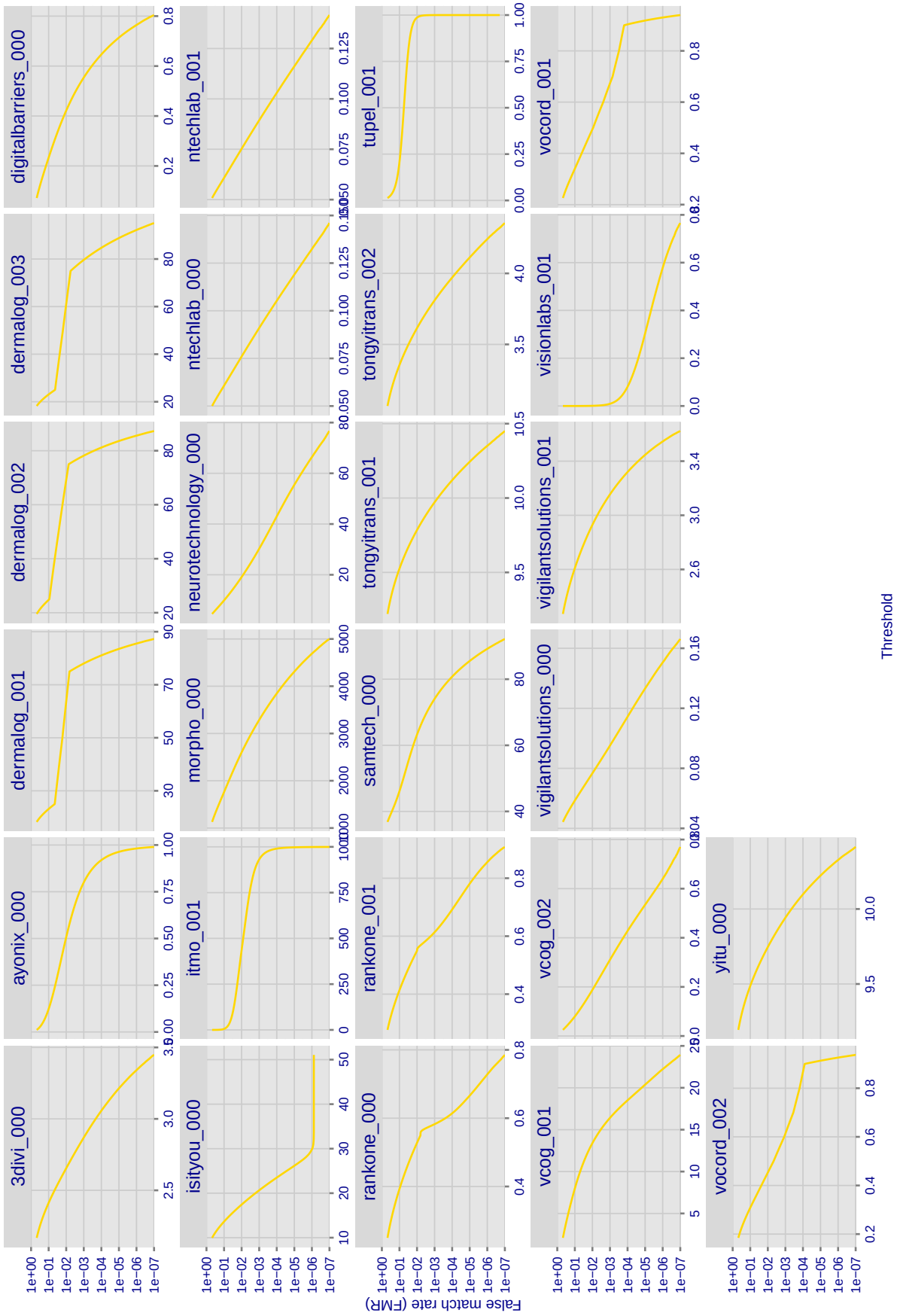


Figure 12: For the visa images, the false match calibration curves show false match rate vs. threshold. These curves apply to zero-effort impostors. As shown later (sec. 4.6), FMR is higher for demographic-matched impostors.

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

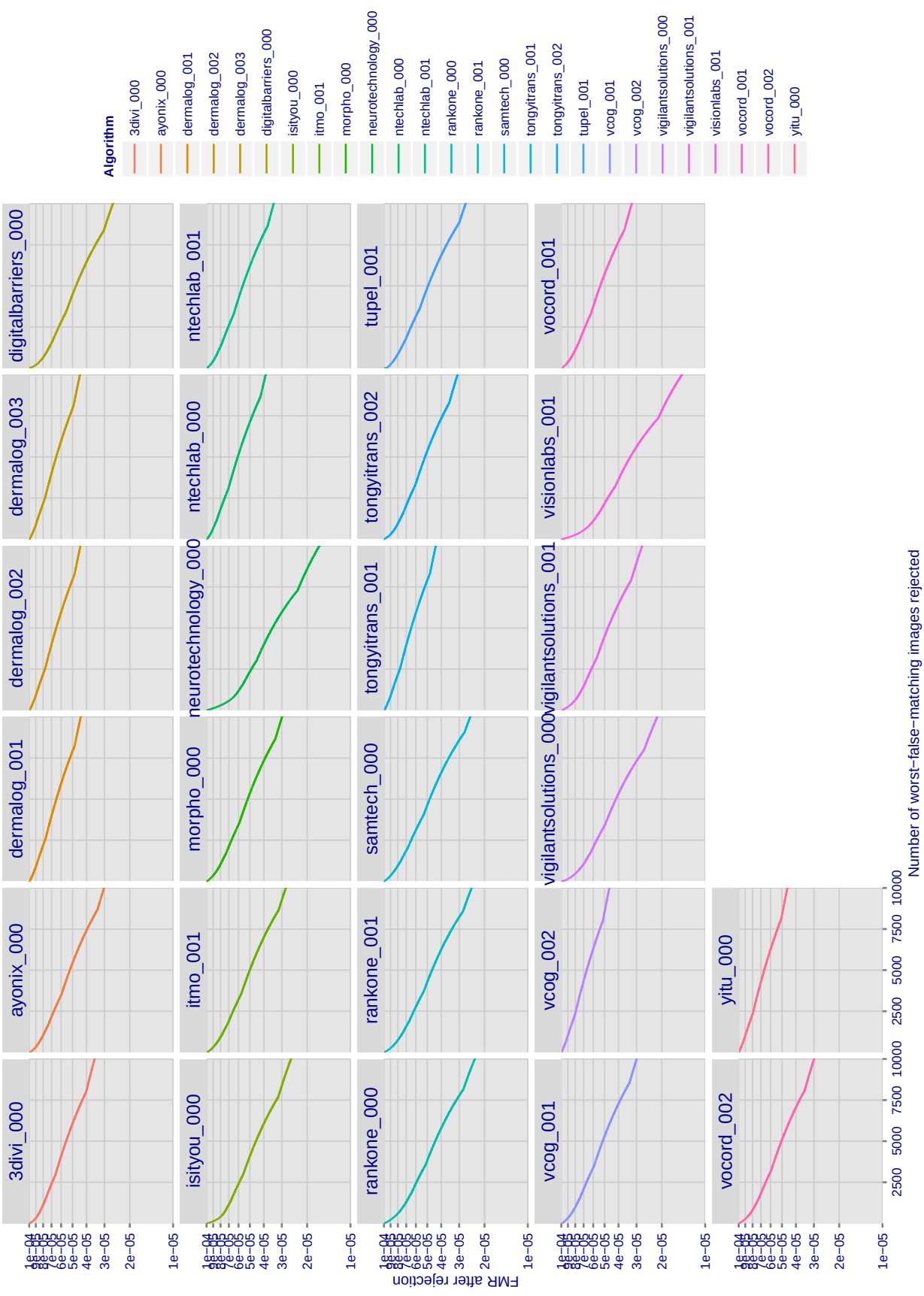


Figure 13: For the visa images, the curves show how false matches are concentrated in certain images. Specifically each line plots FMR(k) with k the number of images rejected in decreasing order of how many false matches that image was involved in. $FMR(0) = 10^{-4}$. In terms of the biometric zoo, the 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.

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

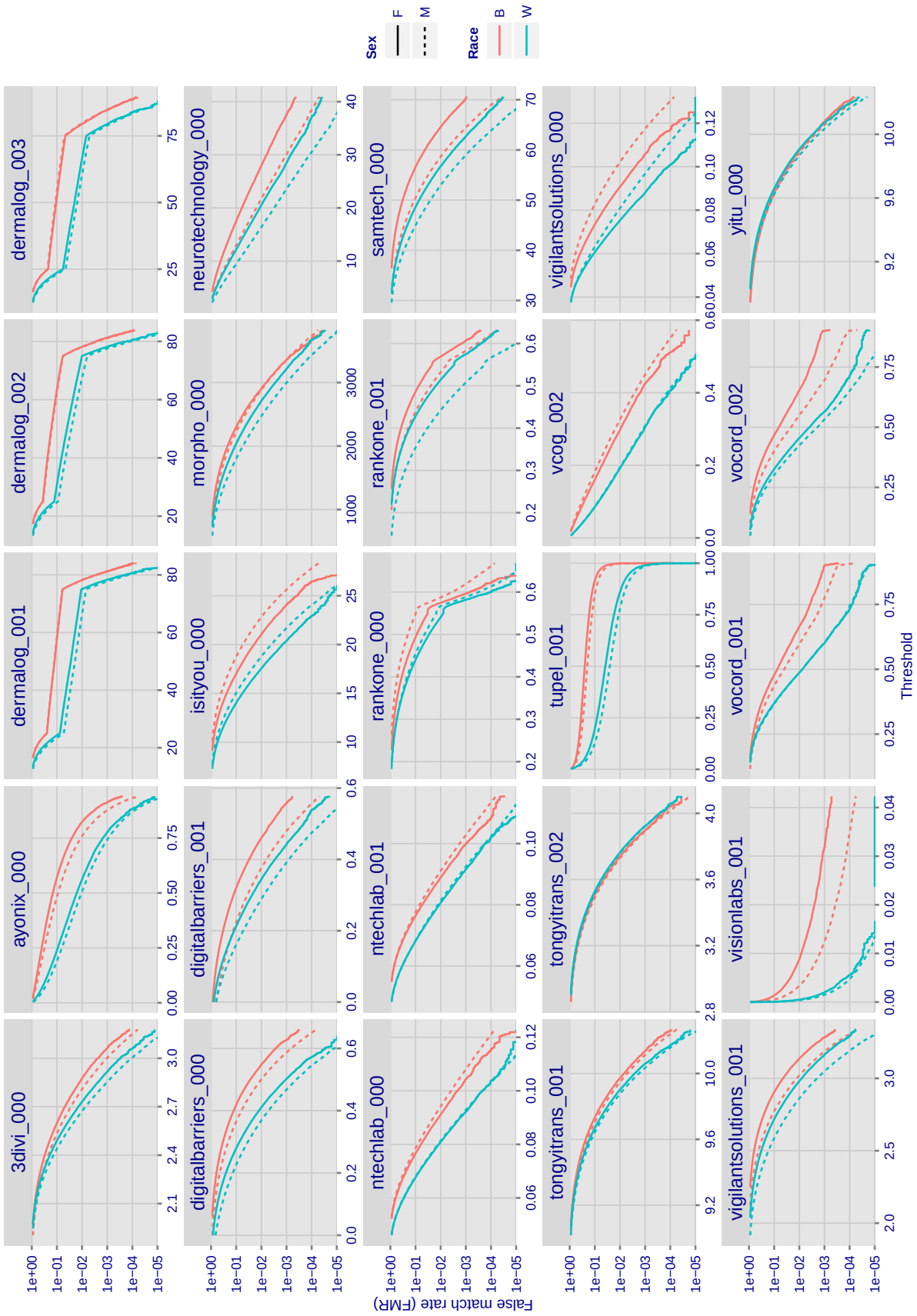


Figure 14: For the mugshot images, the false match calibration curves show false match rate vs. threshold. Separate curves appear for white females, black females, black males and white males.

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

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 15 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.

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

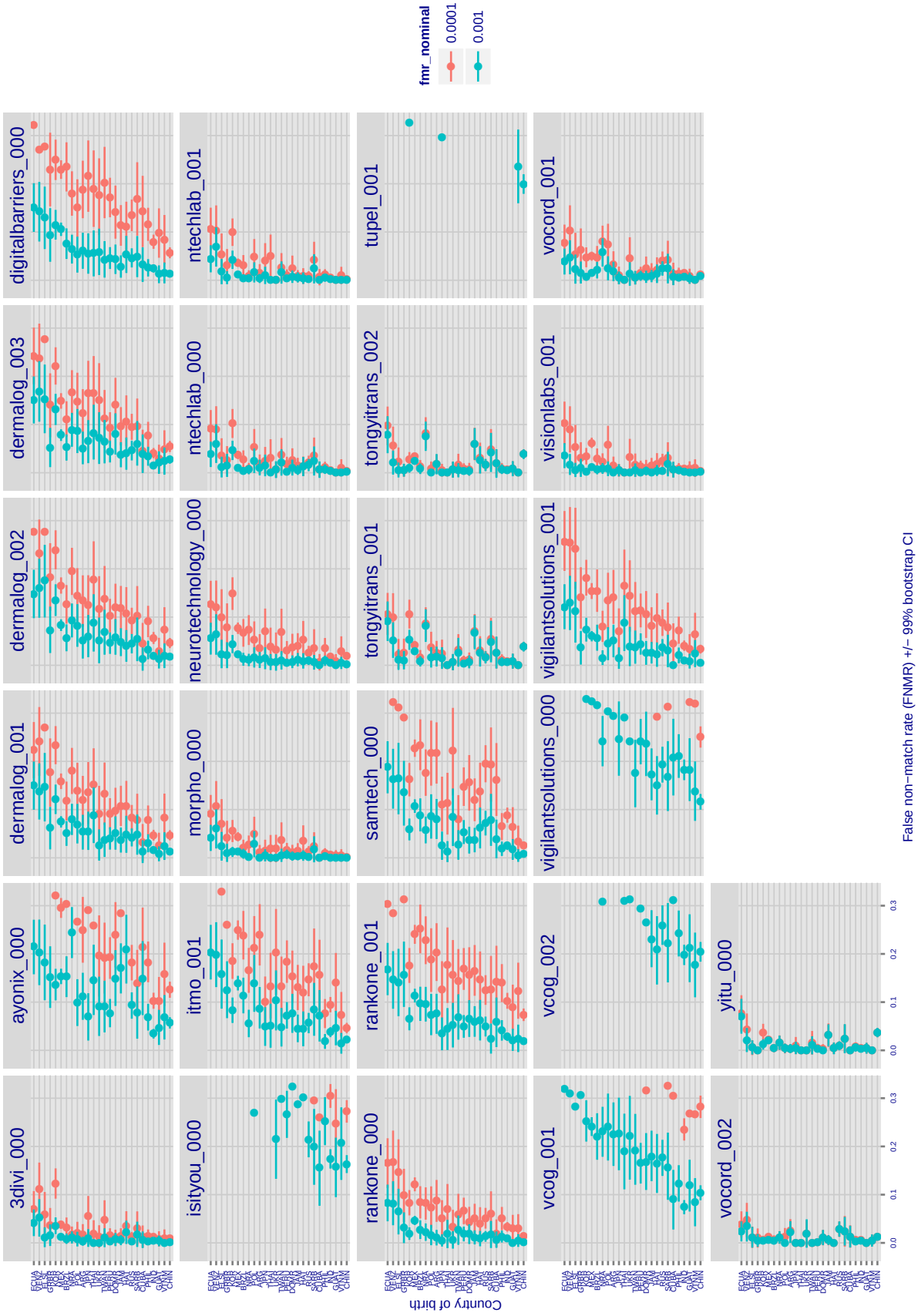


Figure 15: For the visa images, the dots show FNMR by country of birth for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all $O(10^{10})$ impostor scores. The figures shows an order of magnitude variation in FNMR across country of birth, these effects are due to quality variations. The least accurate countries vary by algorithm.

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

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. Only those countries with at least 30 individuals are included in the analysis.

Results: For the visa images, Figure 16 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 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.

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

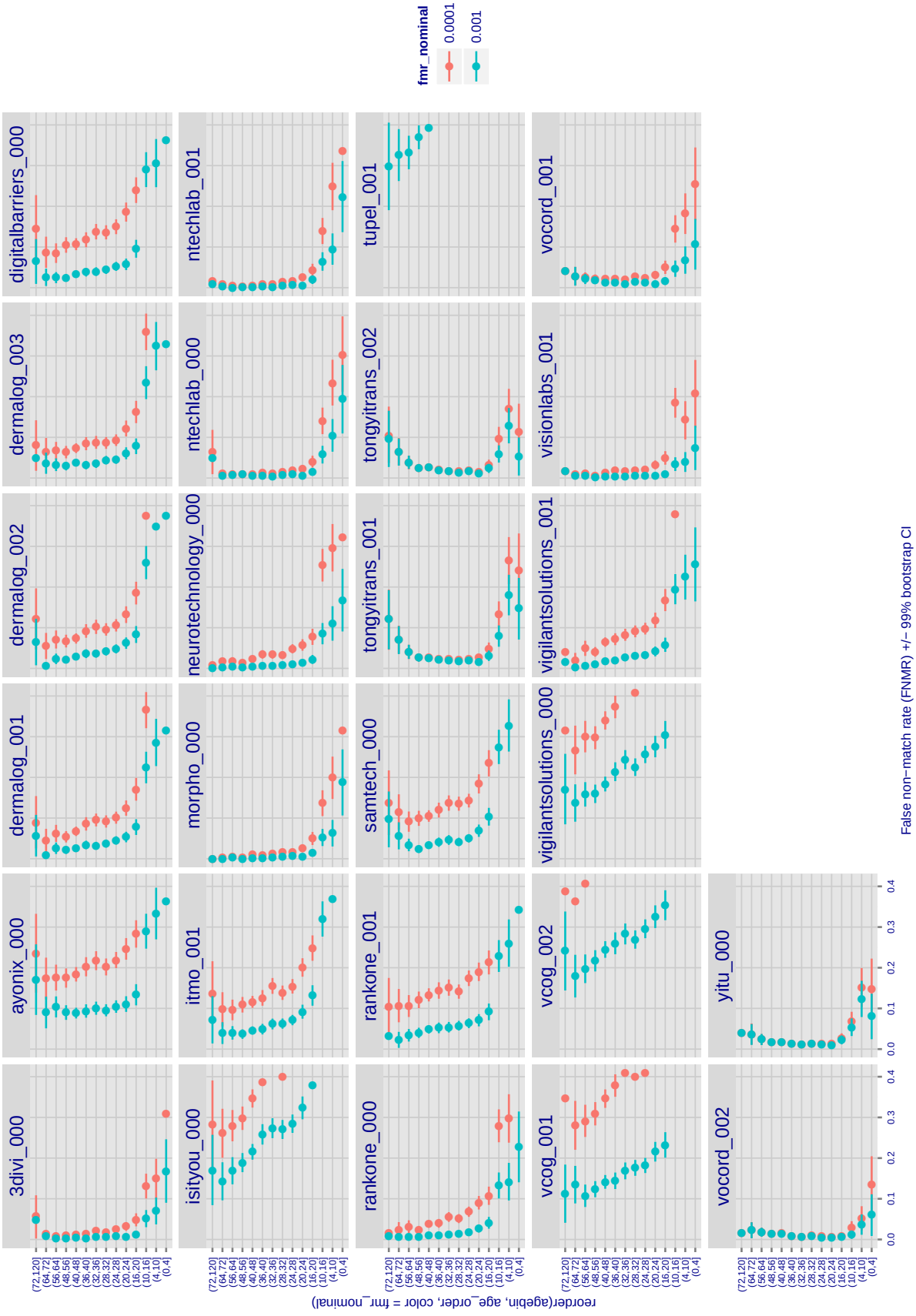


Figure 16: For the visa images, the dots show FNMR by age group for two operating thresholds corresponding to FMR = $\{0.001, 0.0001\}$ computed over all $O(10^{10})$ impostor scores. Given a pair of face images taken at different times, we assign a false non-match to the bin that is the arithmetic average 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.

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

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.

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

  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 71.

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.

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

▷ 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.

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

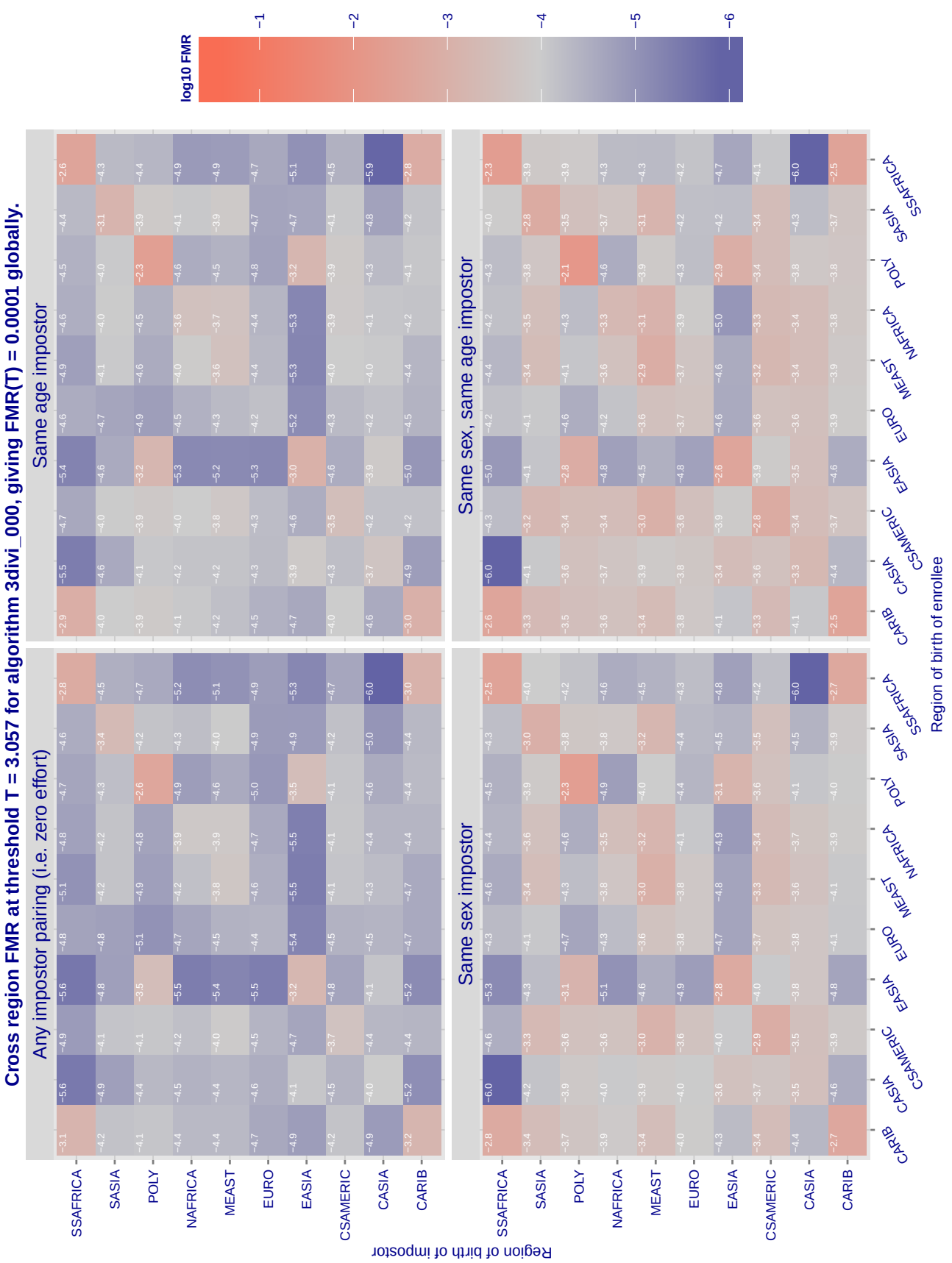


Figure 17: For algorithm 3divi-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR 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.

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

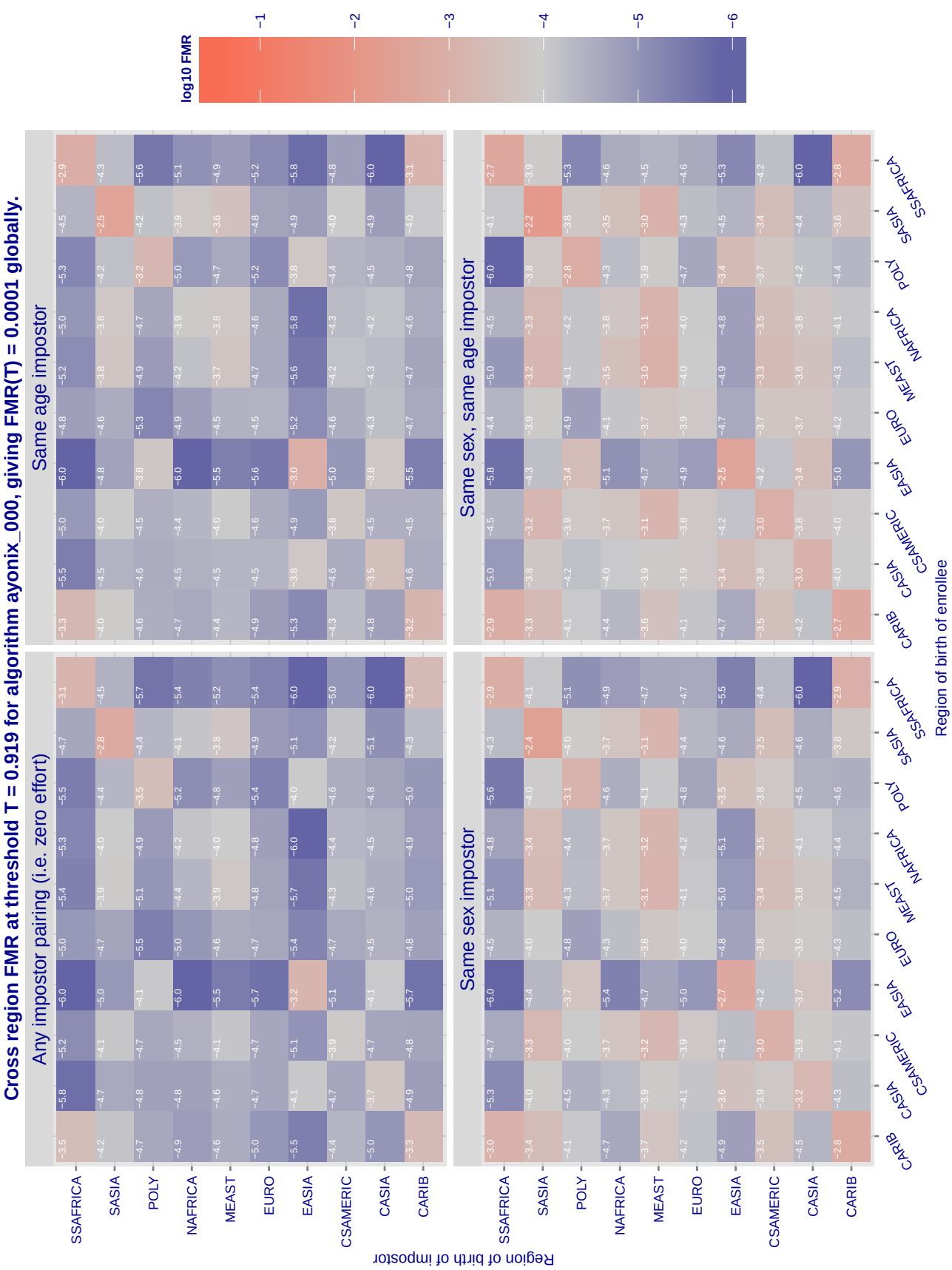


Figure 18: For algorithm ayonix-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

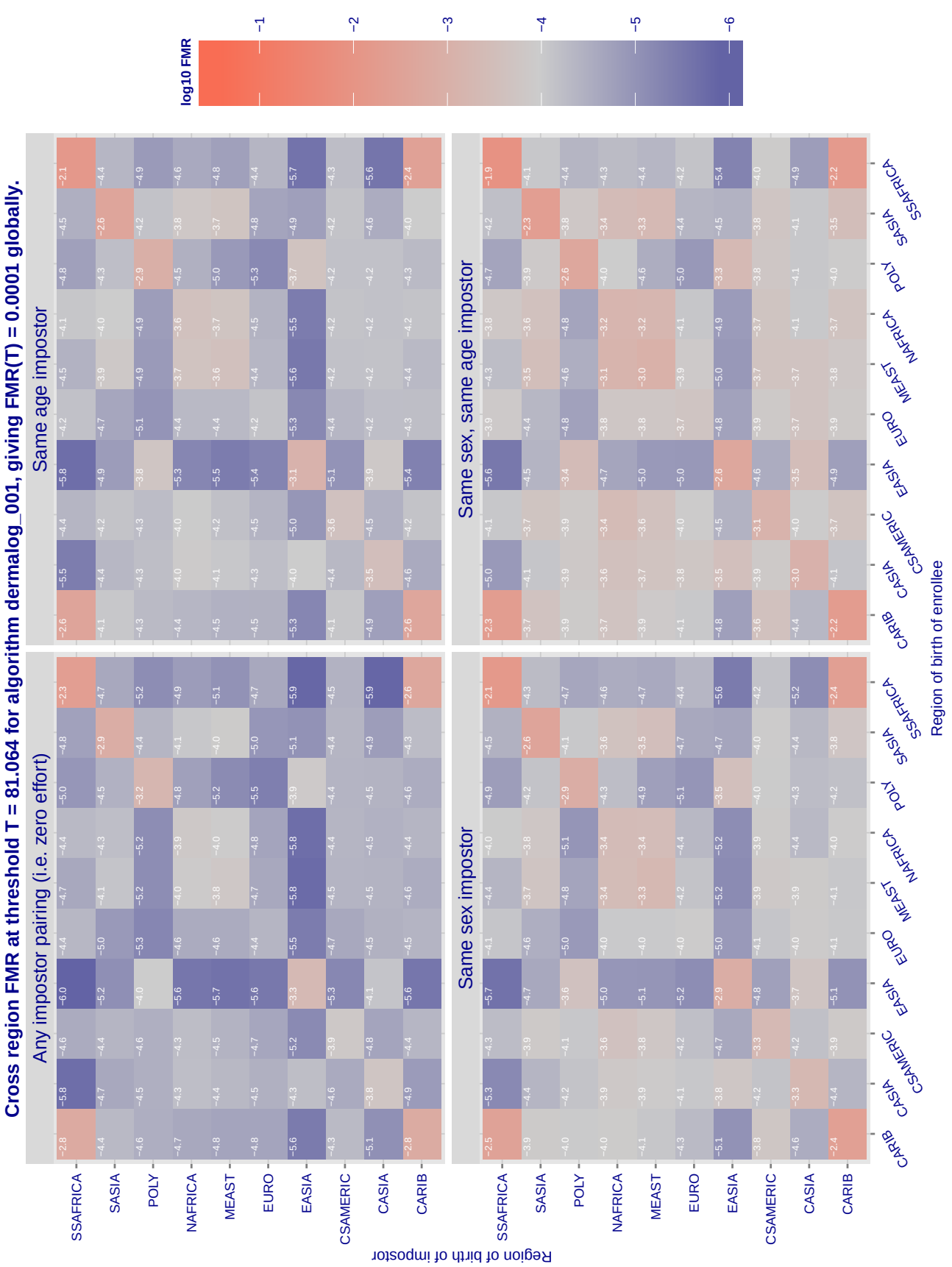


Figure 19: For algorithm dermalog-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

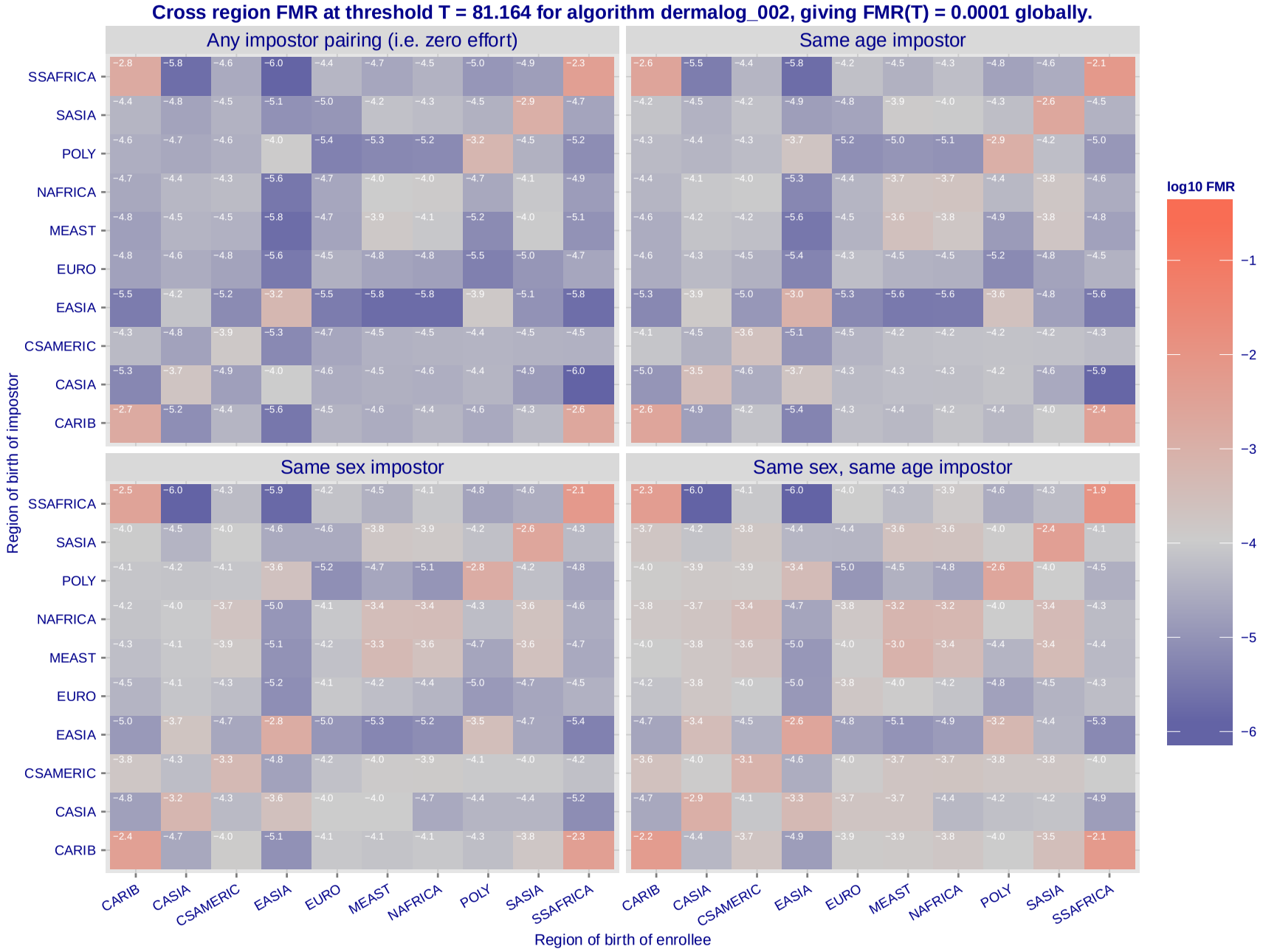


Figure 20: For algorithm dermalog-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

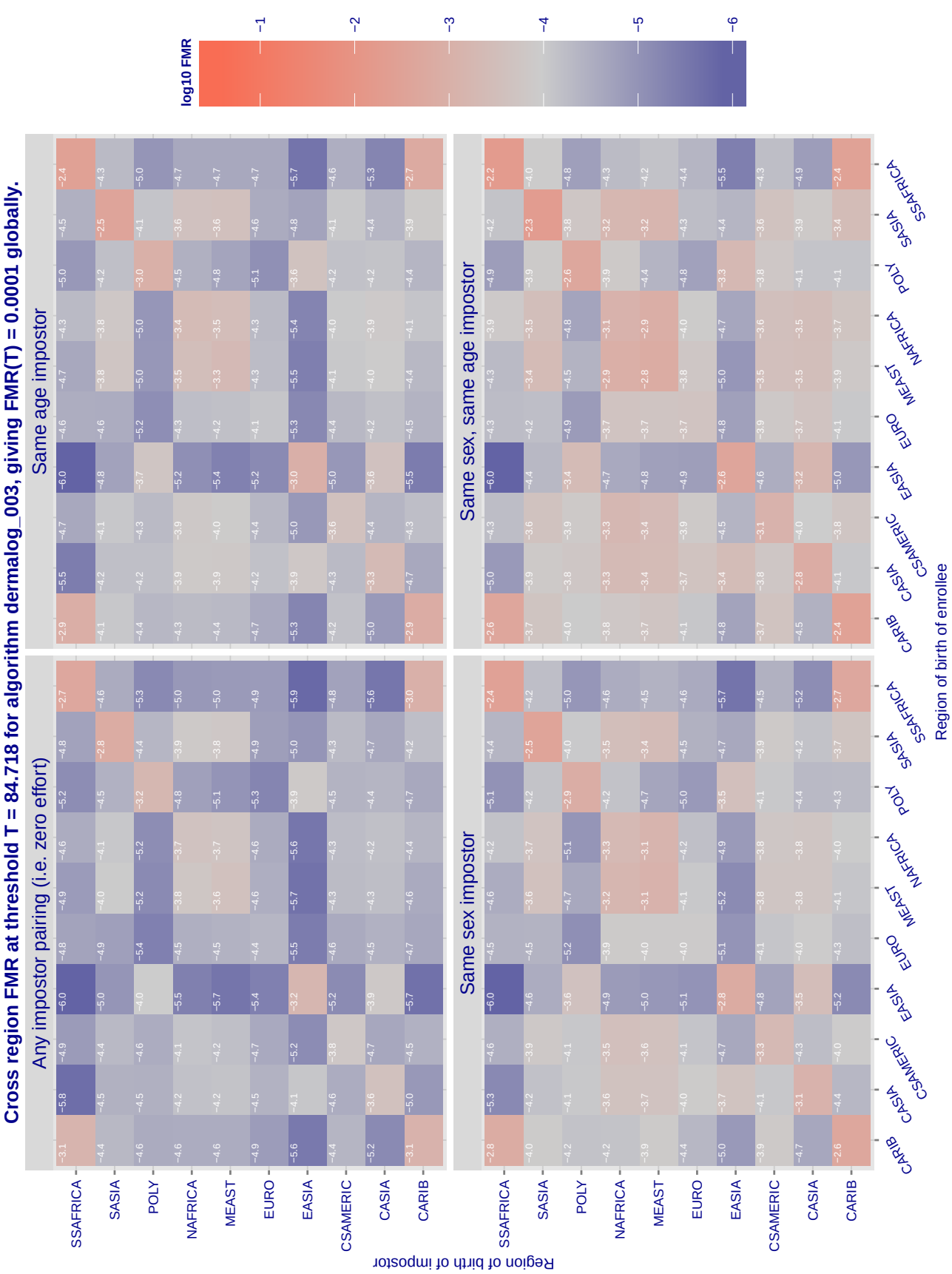


Figure 21: For algorithm dermalog-003 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

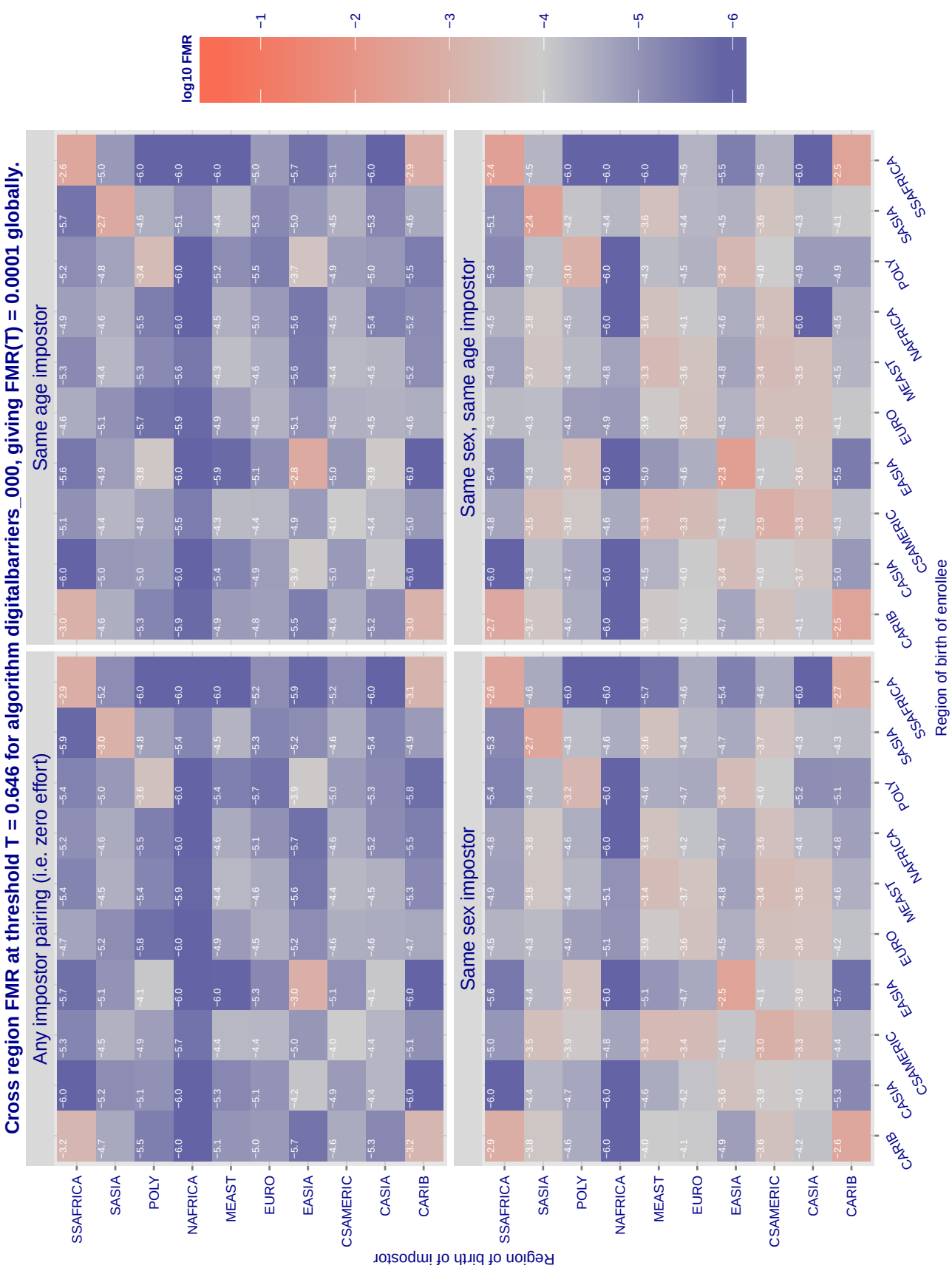


Figure 22: For algorithm digitalbarriers-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the 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.

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

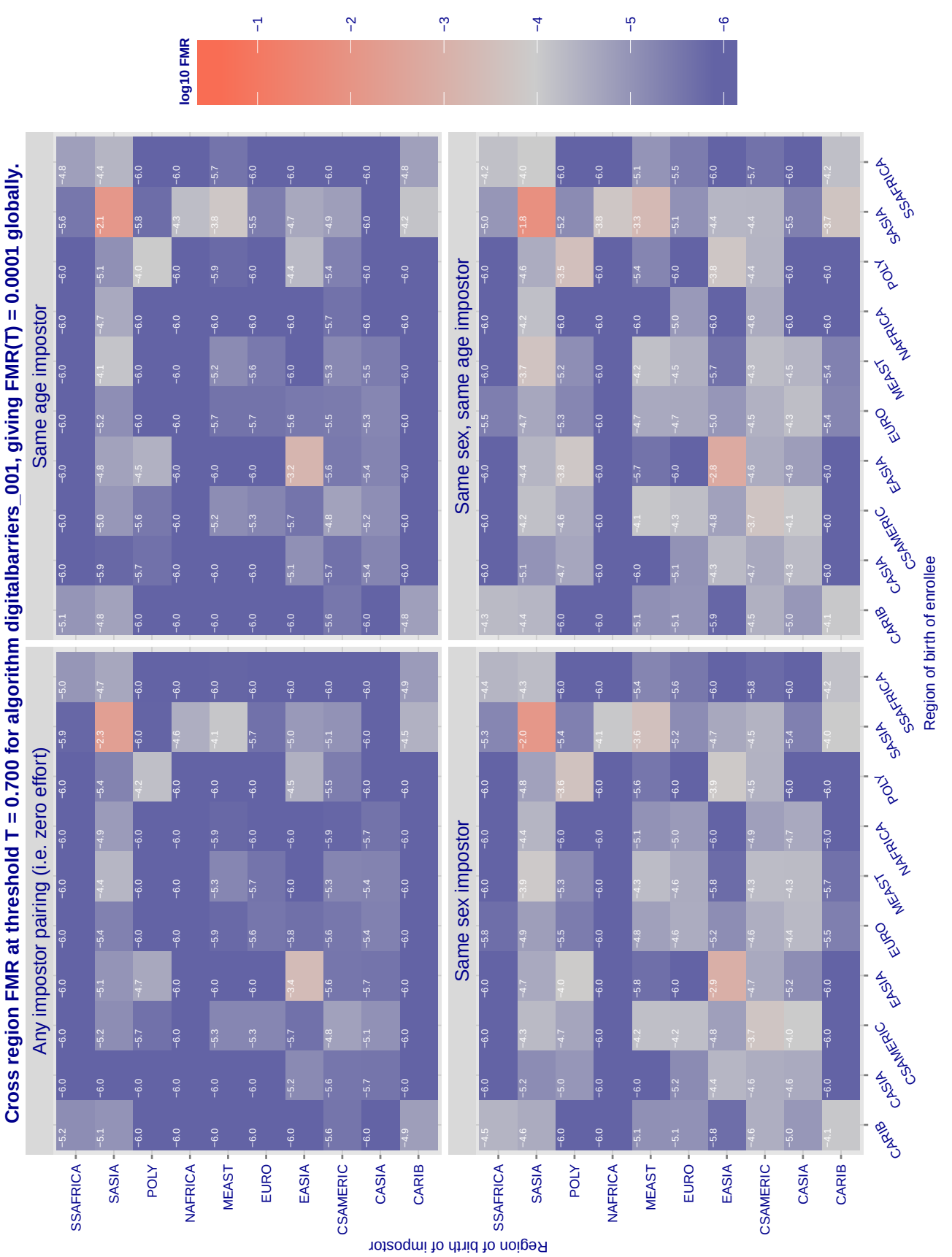


Figure 23: For algorithm digitalbarriers-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the 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.

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

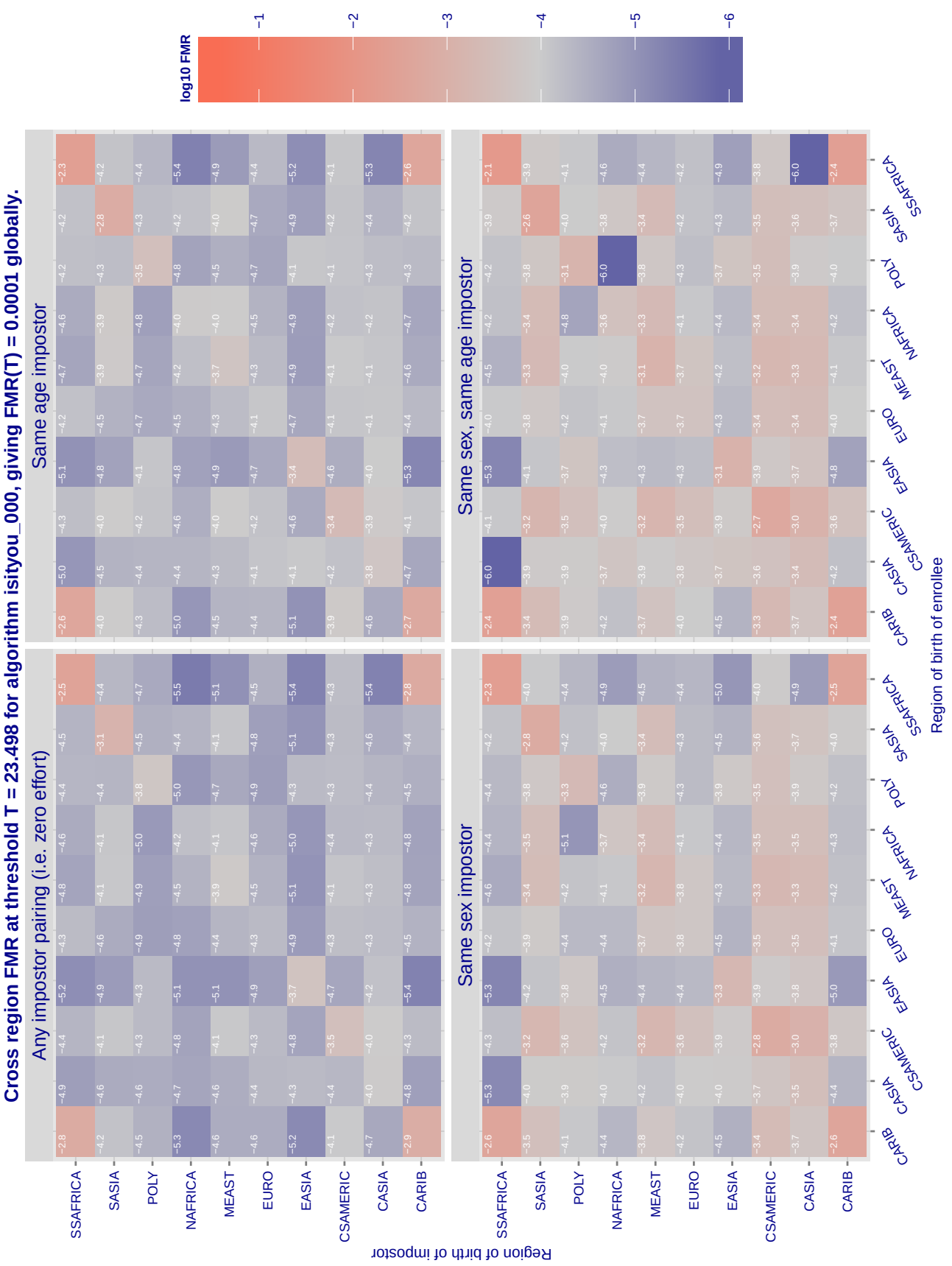


Figure 24: For algorithm isityou-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

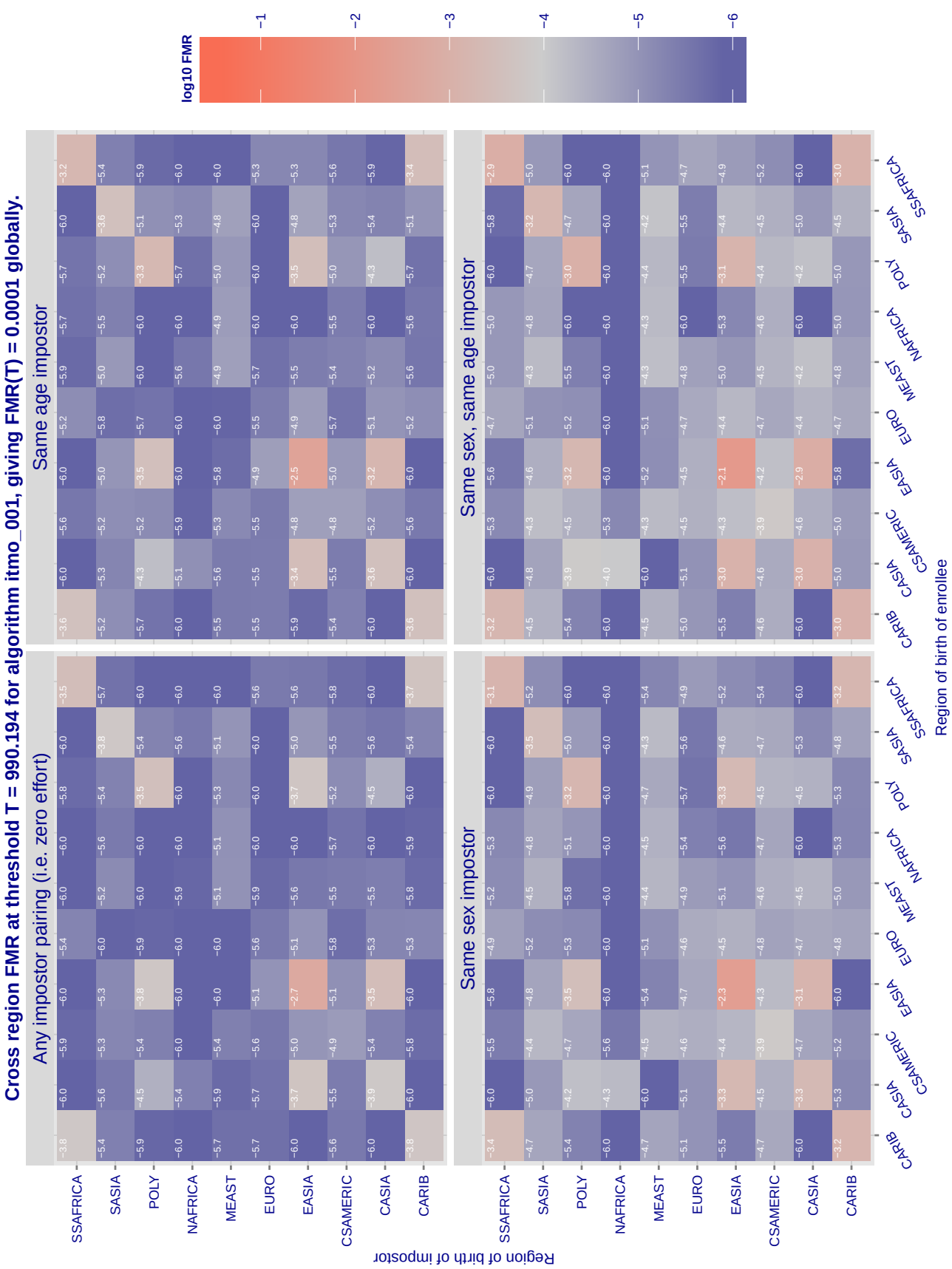


Figure 25: For algorithm itmo-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR 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.

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

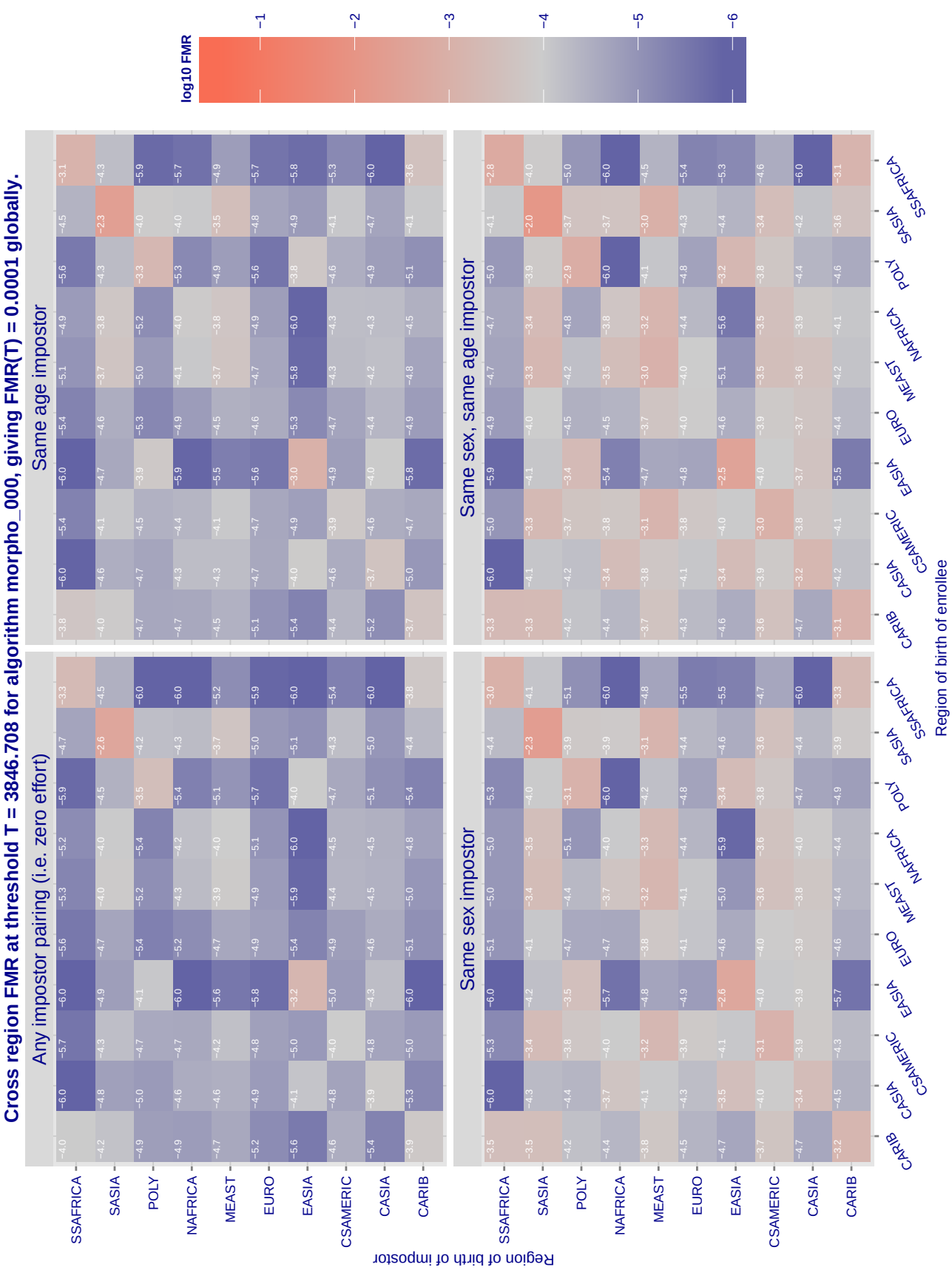


Figure 26: For algorithm morpho-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

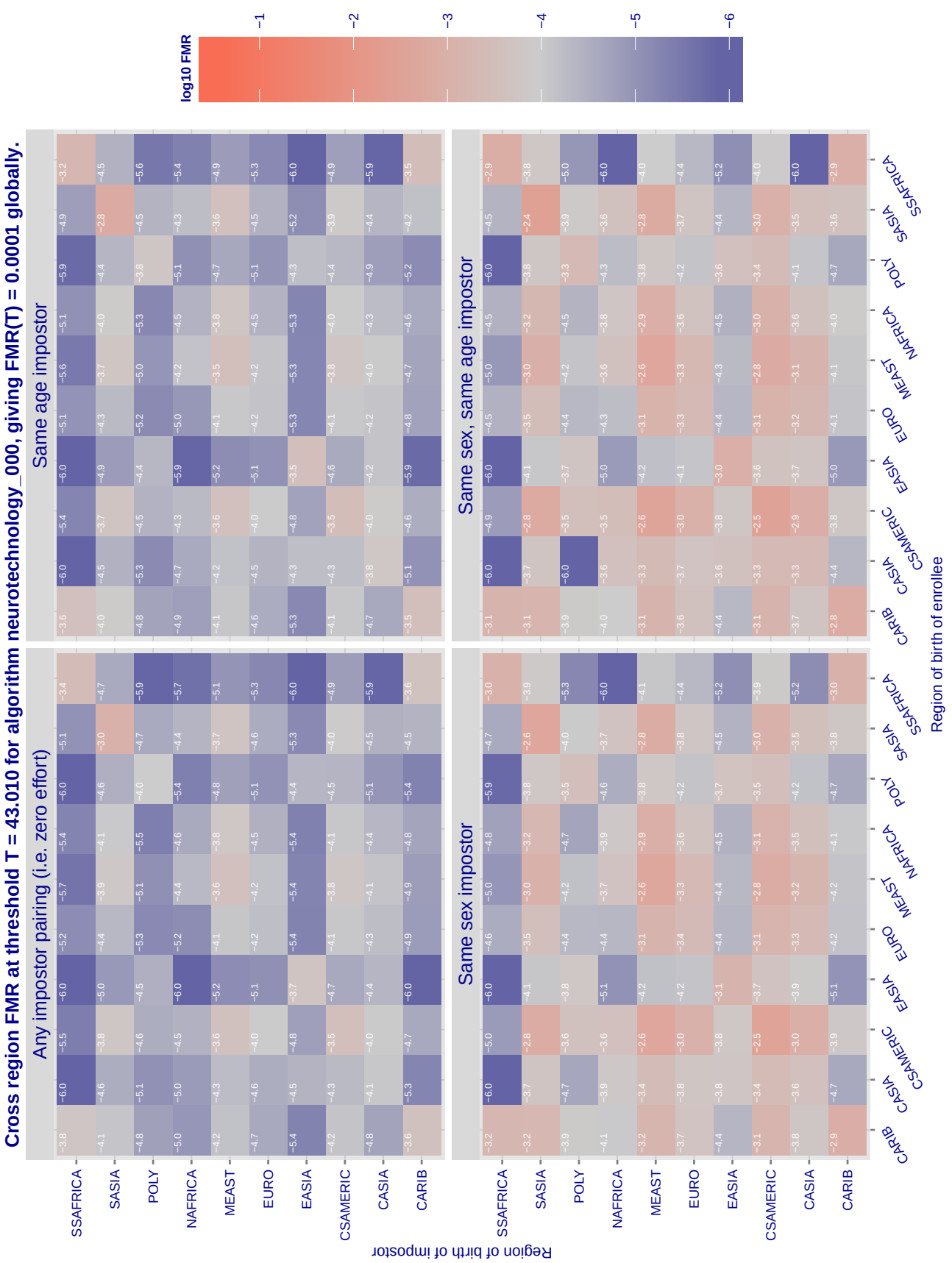


Figure 27: For algorithm neurotechnology-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the 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.

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

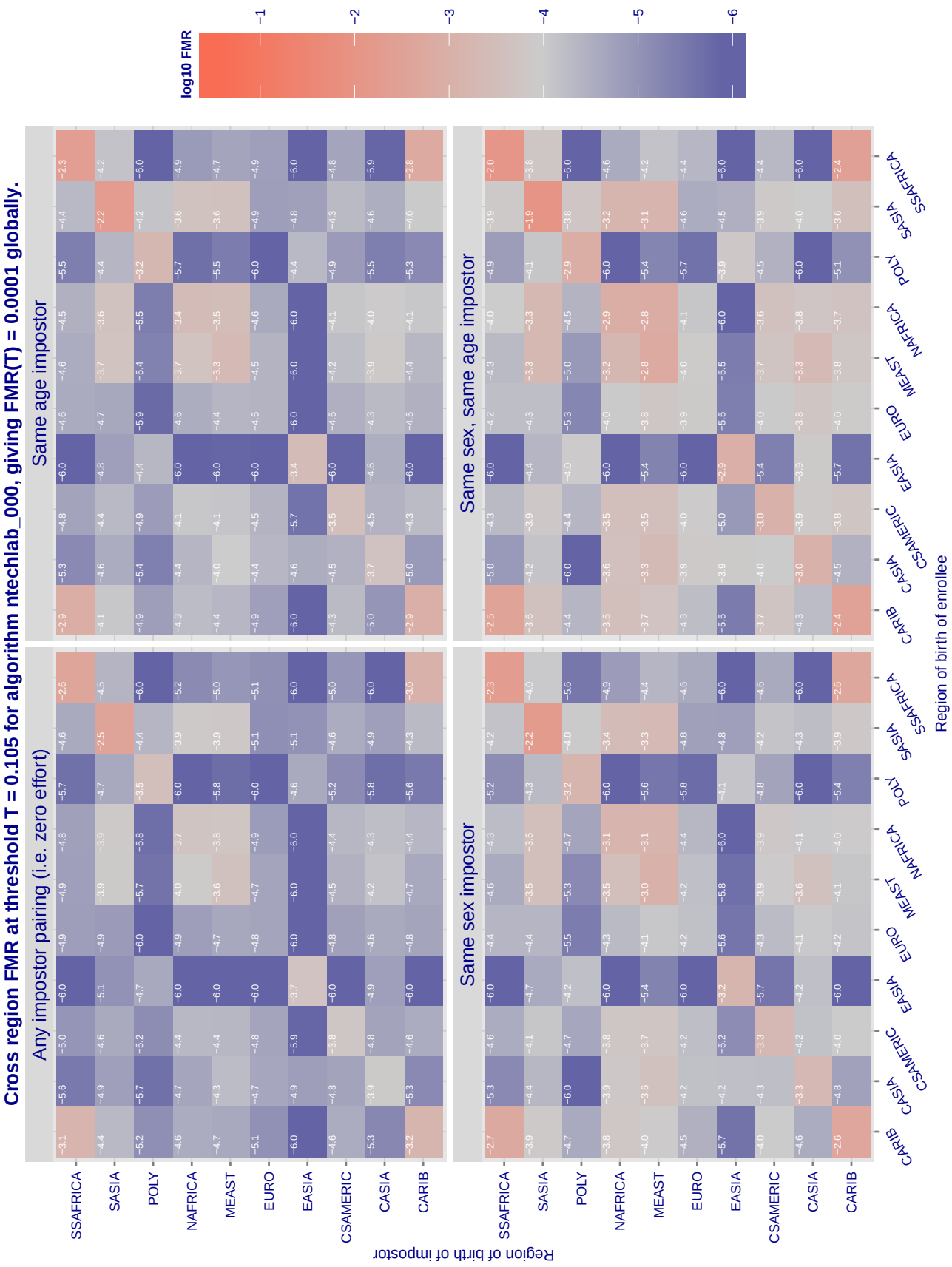


Figure 28: For algorithm ntechlab-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

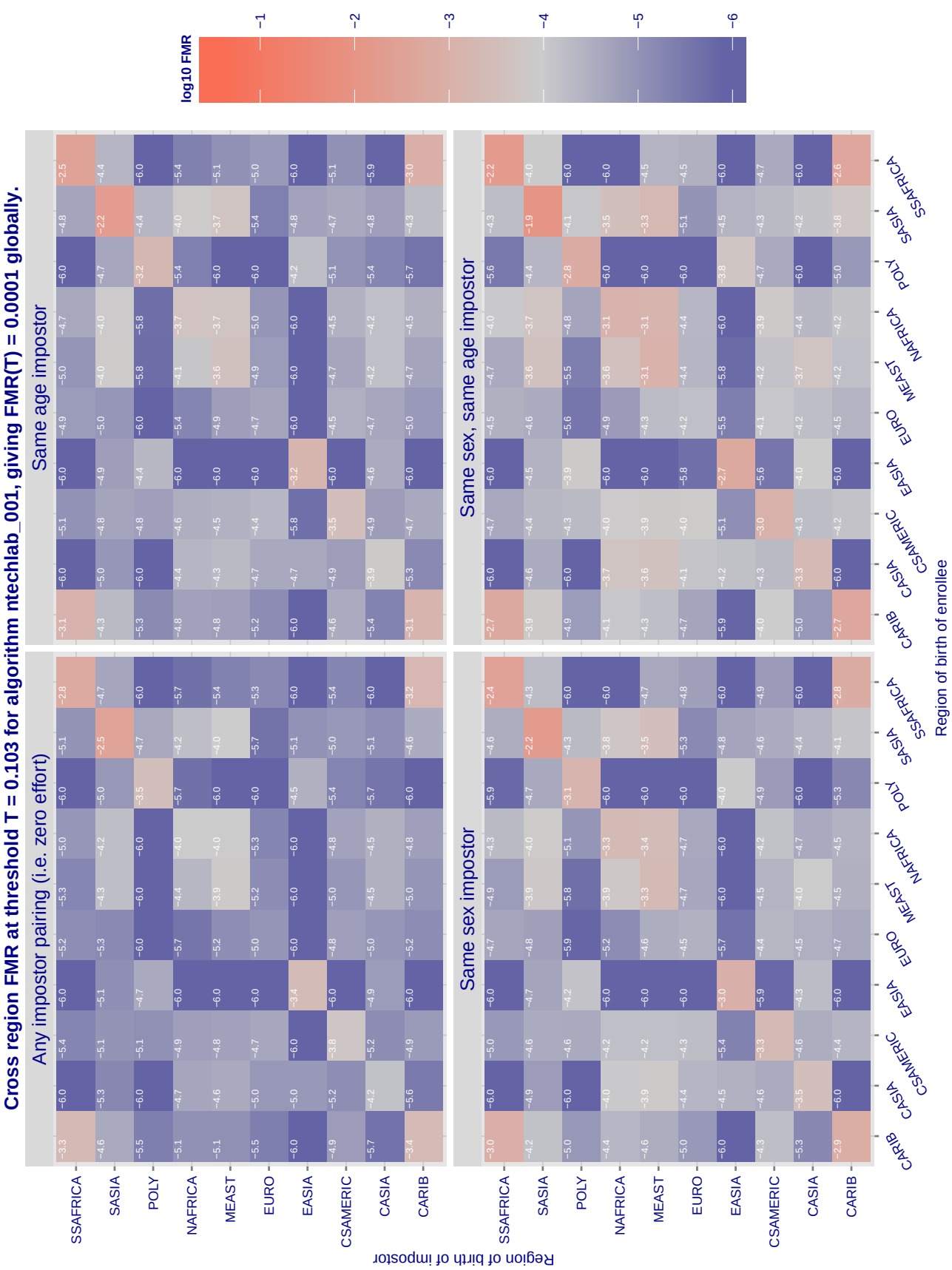


Figure 29: For algorithm ntechlab-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

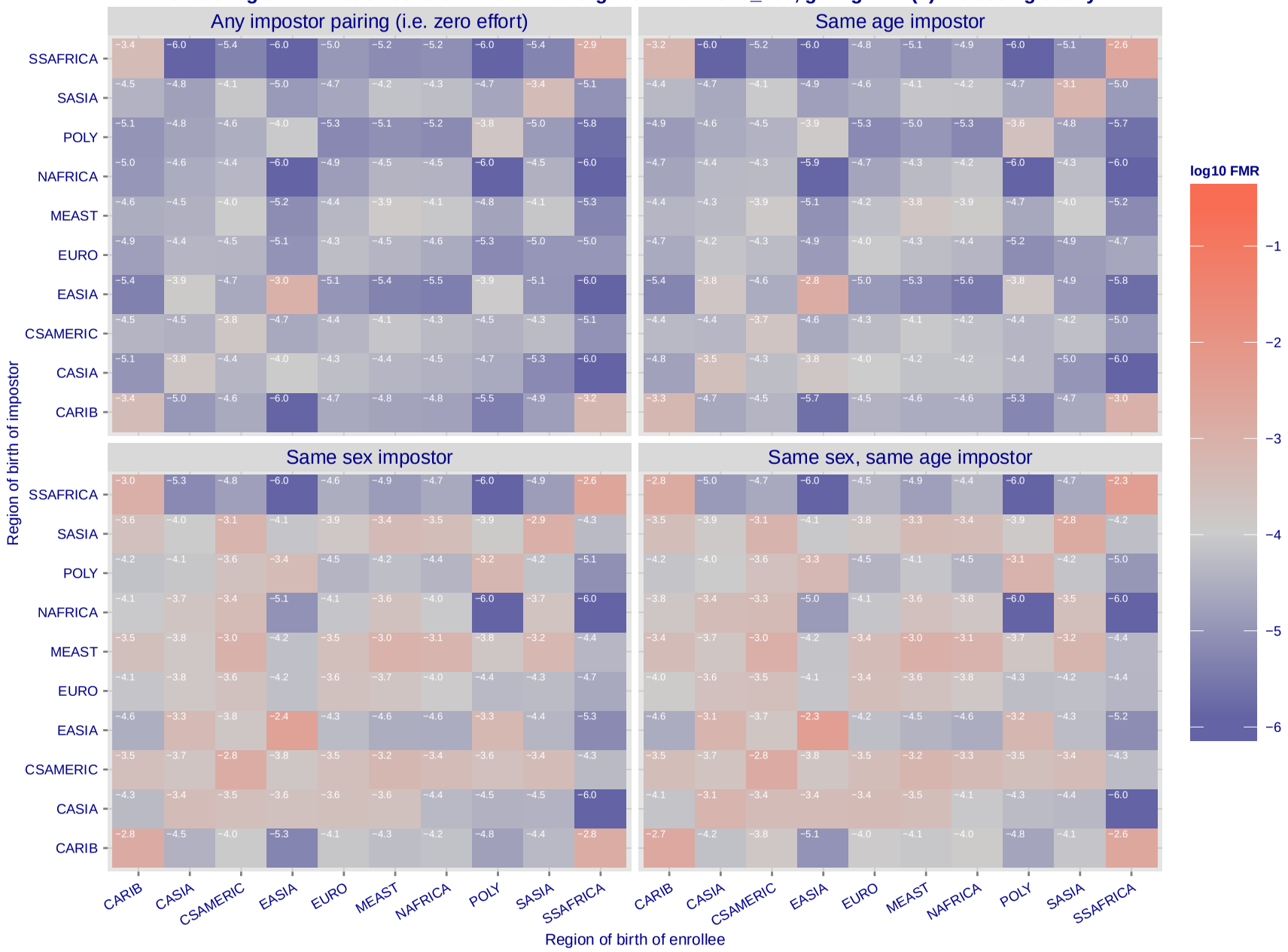


Figure 30: For algorithm rankone-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

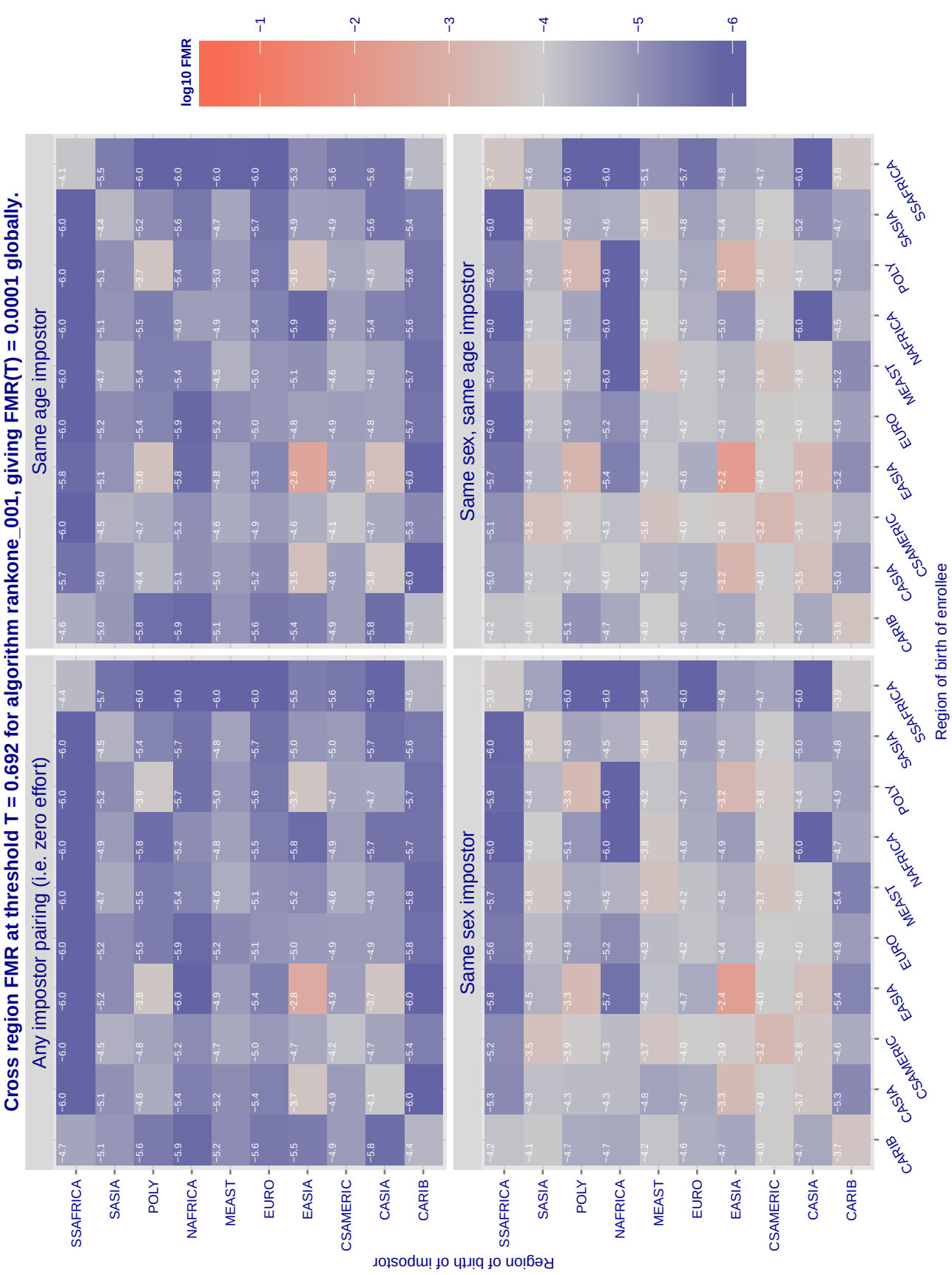


Figure 31: For algorithm rankone-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

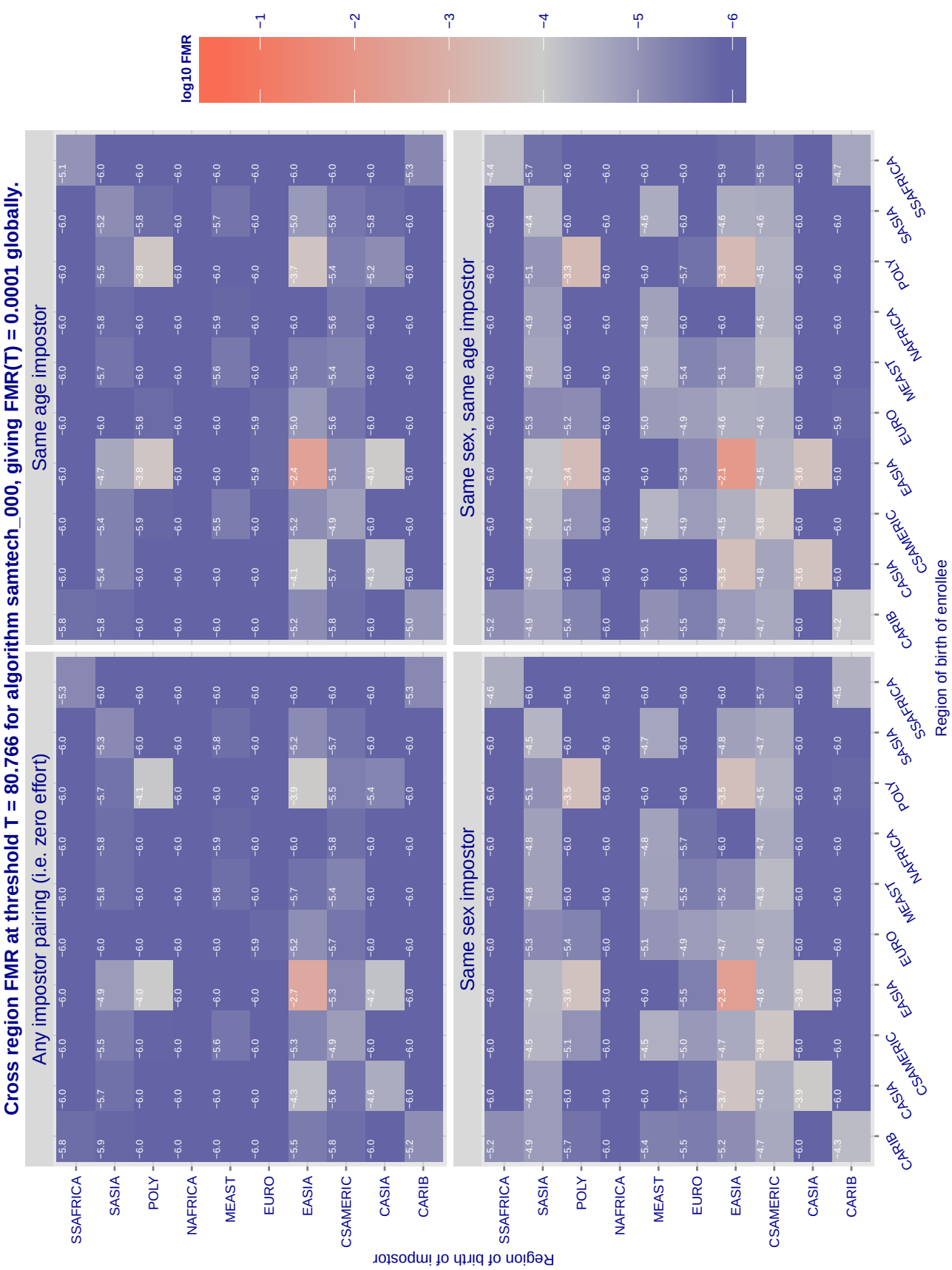


Figure 32: For algorithm samtech-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

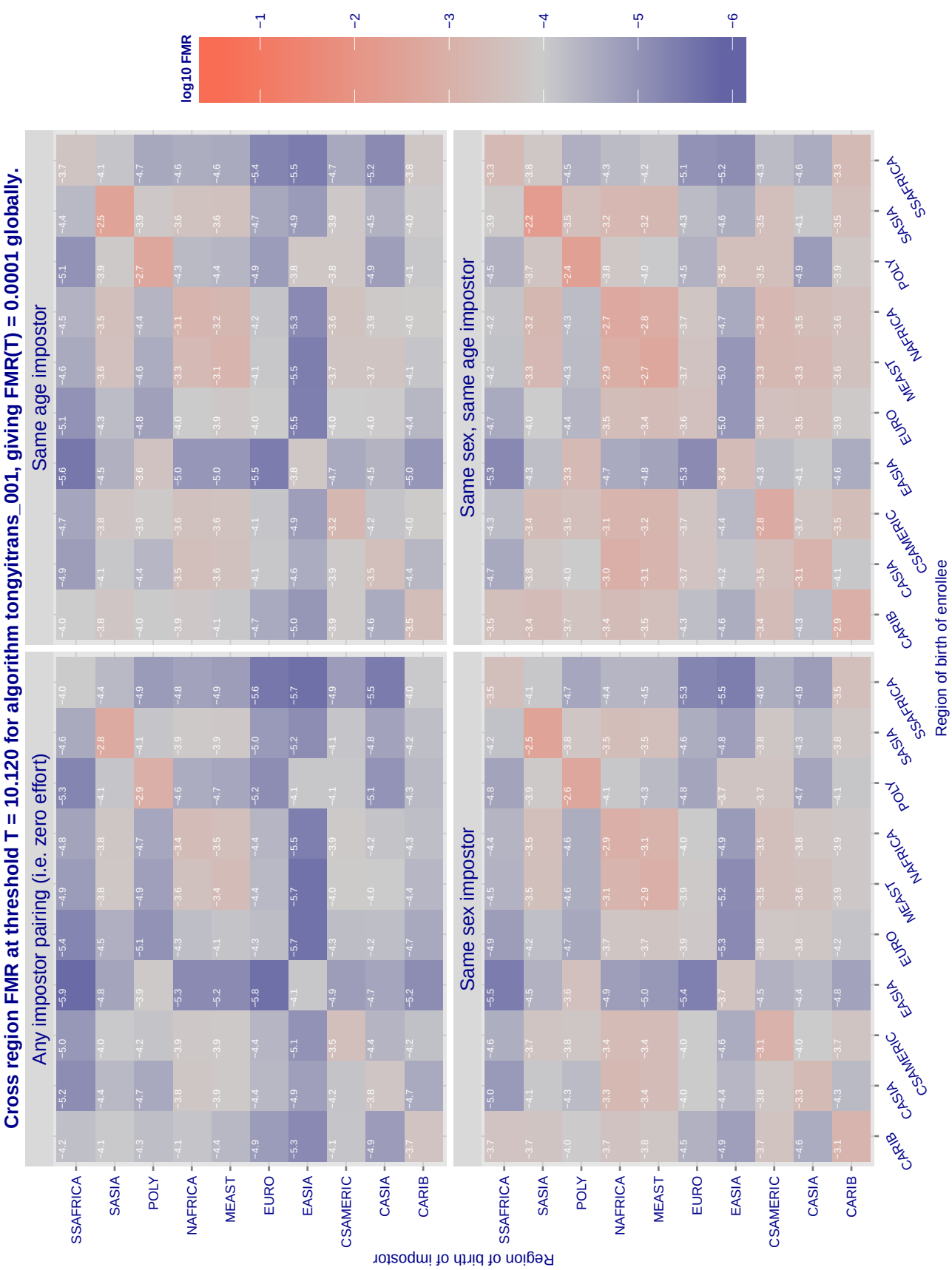


Figure 33: For algorithm tongyitrans-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

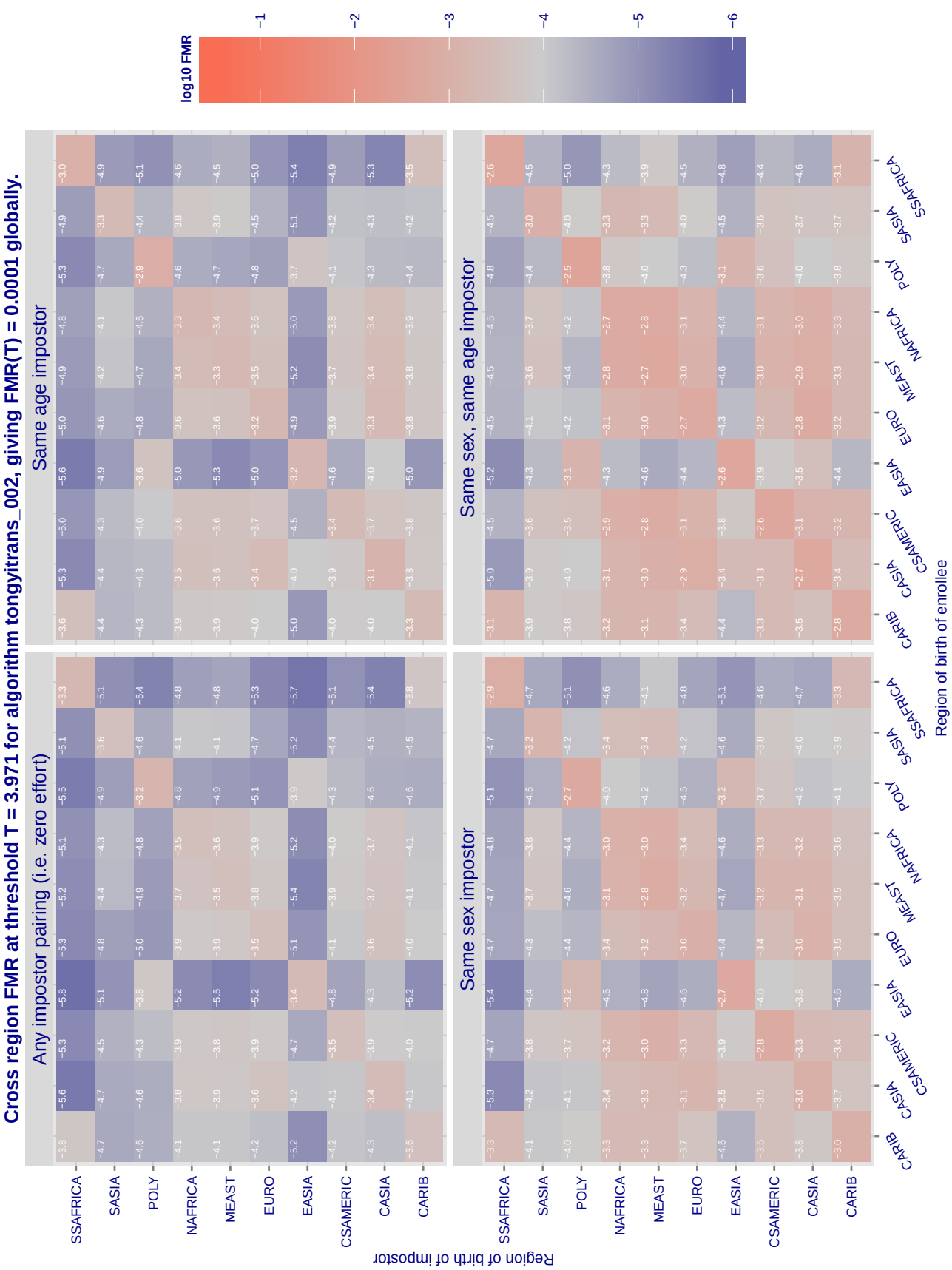


Figure 34: For algorithm tongyitrans-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

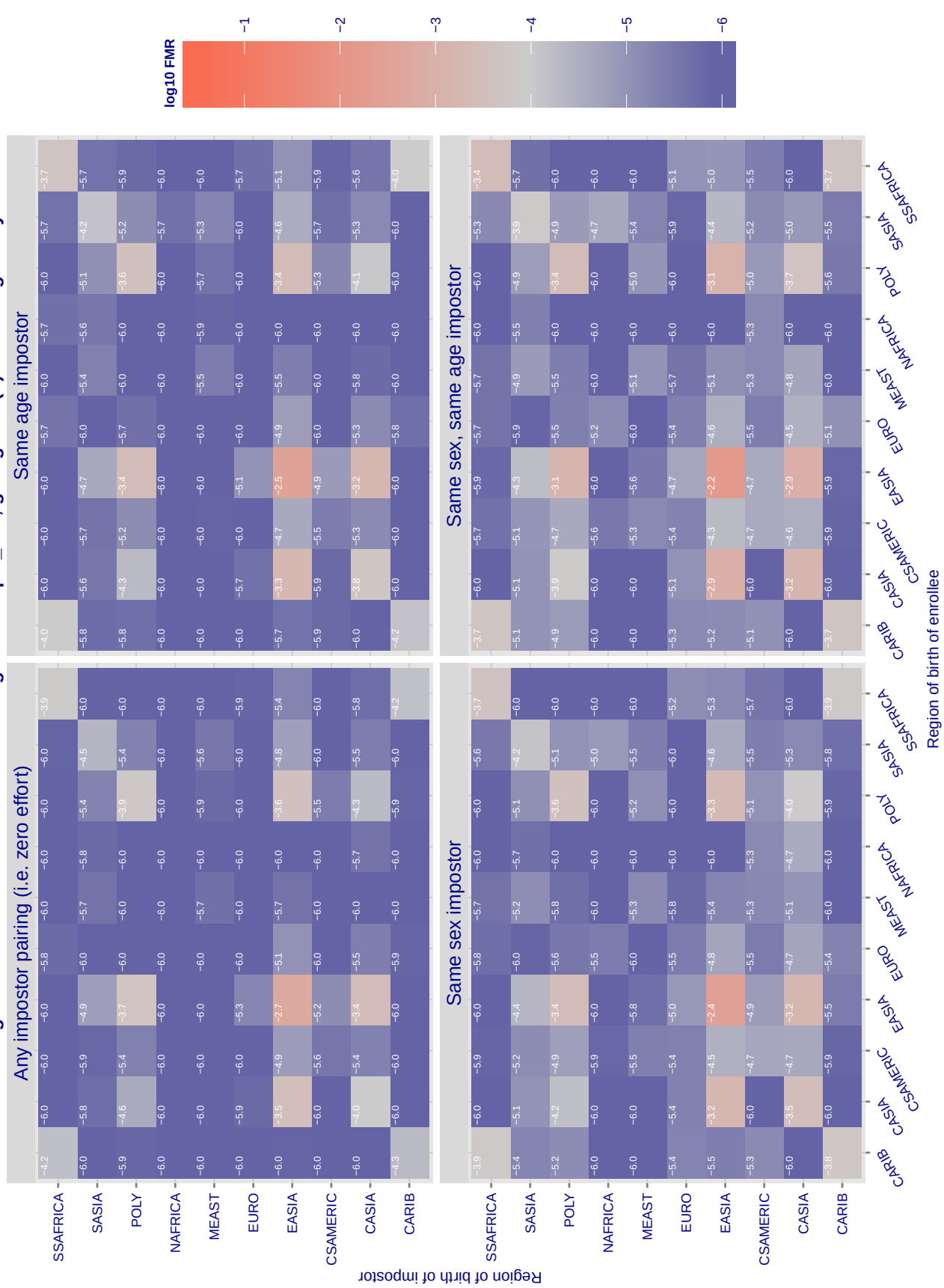


Figure 35: For algorithm tupel-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR 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.

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

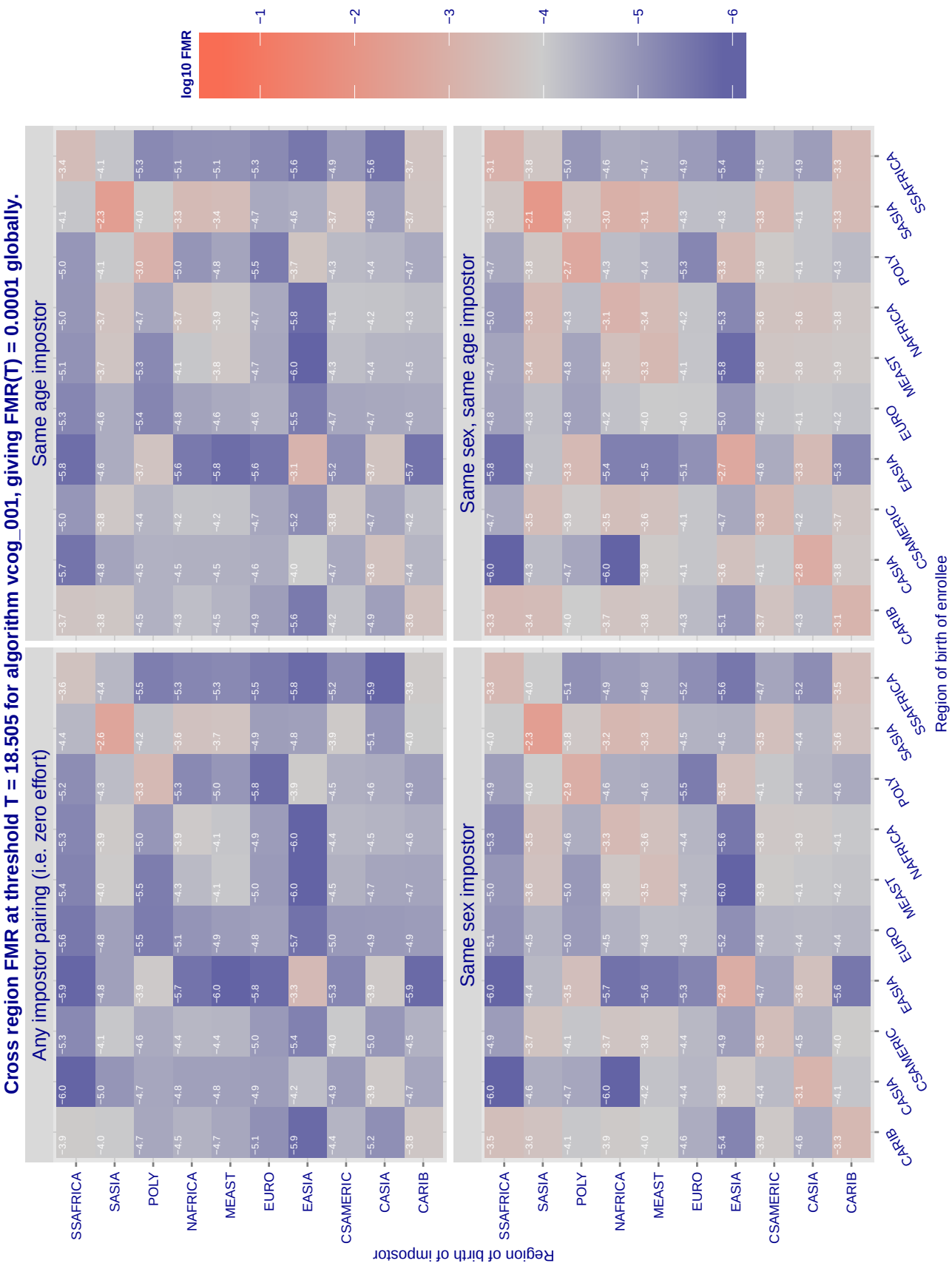


Figure 36: For algorithm vcog-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR 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.

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

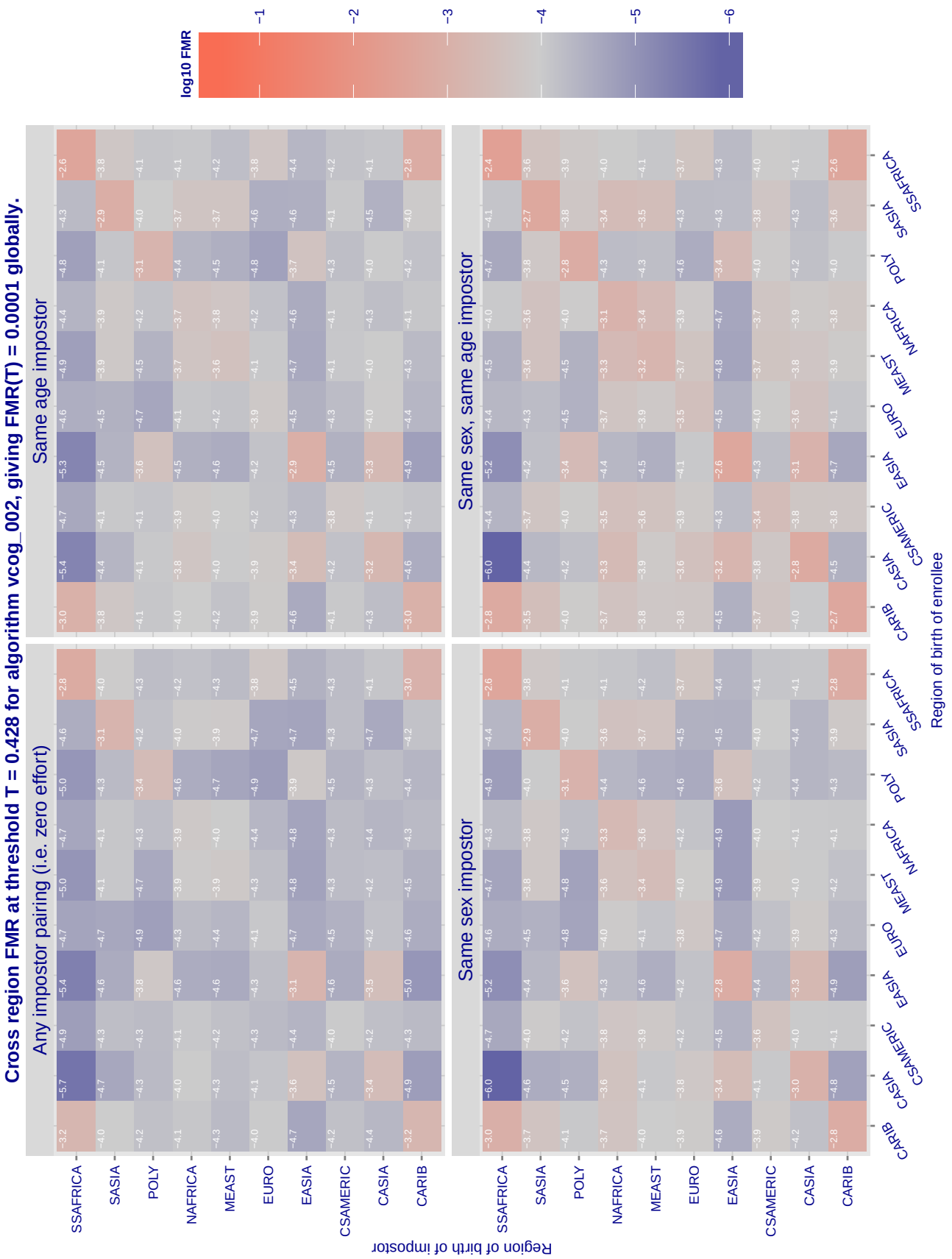


Figure 37: For algorithm vcog-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR 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.

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

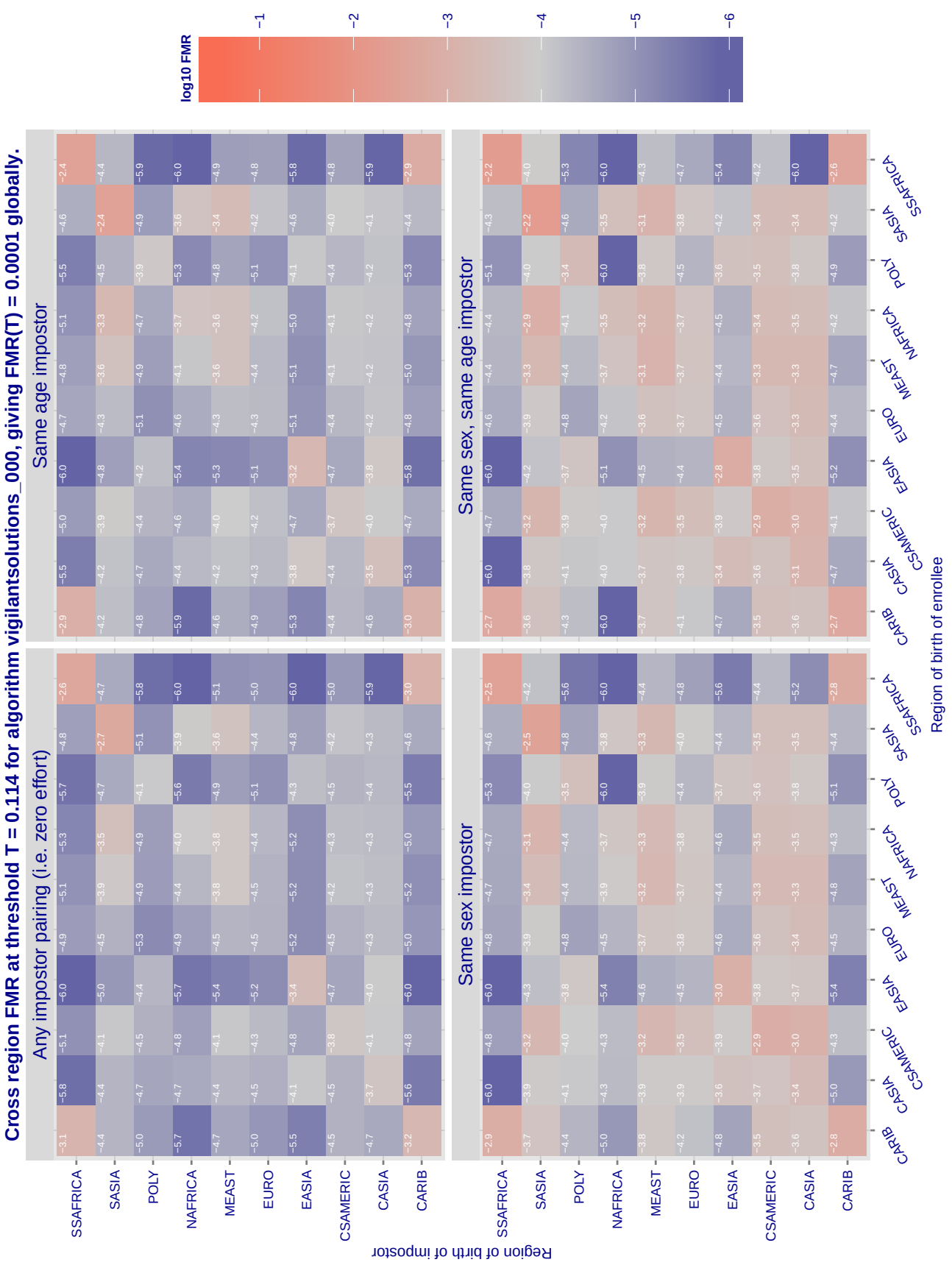


Figure 38: For algorithm vigilantsolutions-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give 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.

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

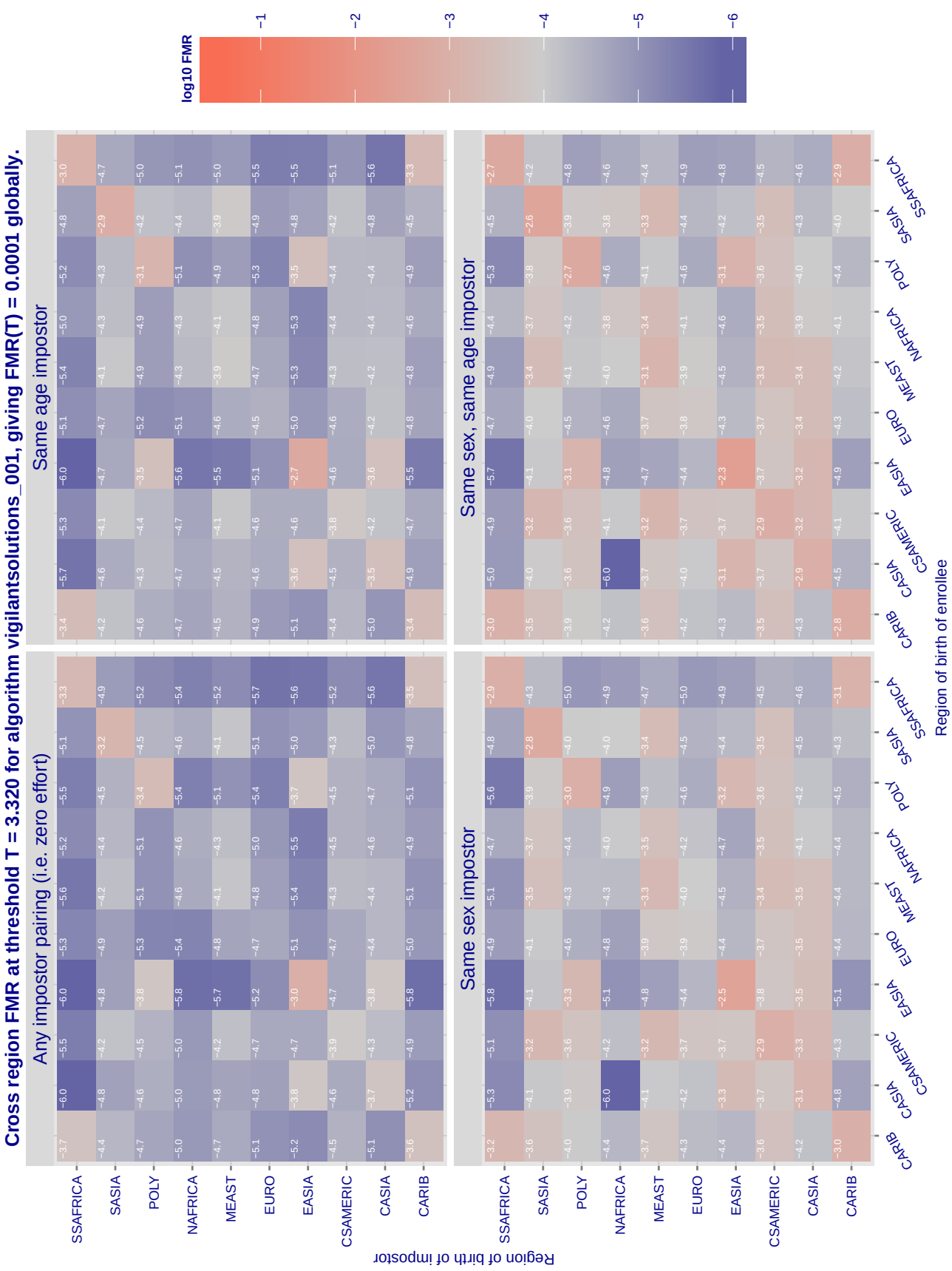


Figure 39: For algorithm vigilantsolutions-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give 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.

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

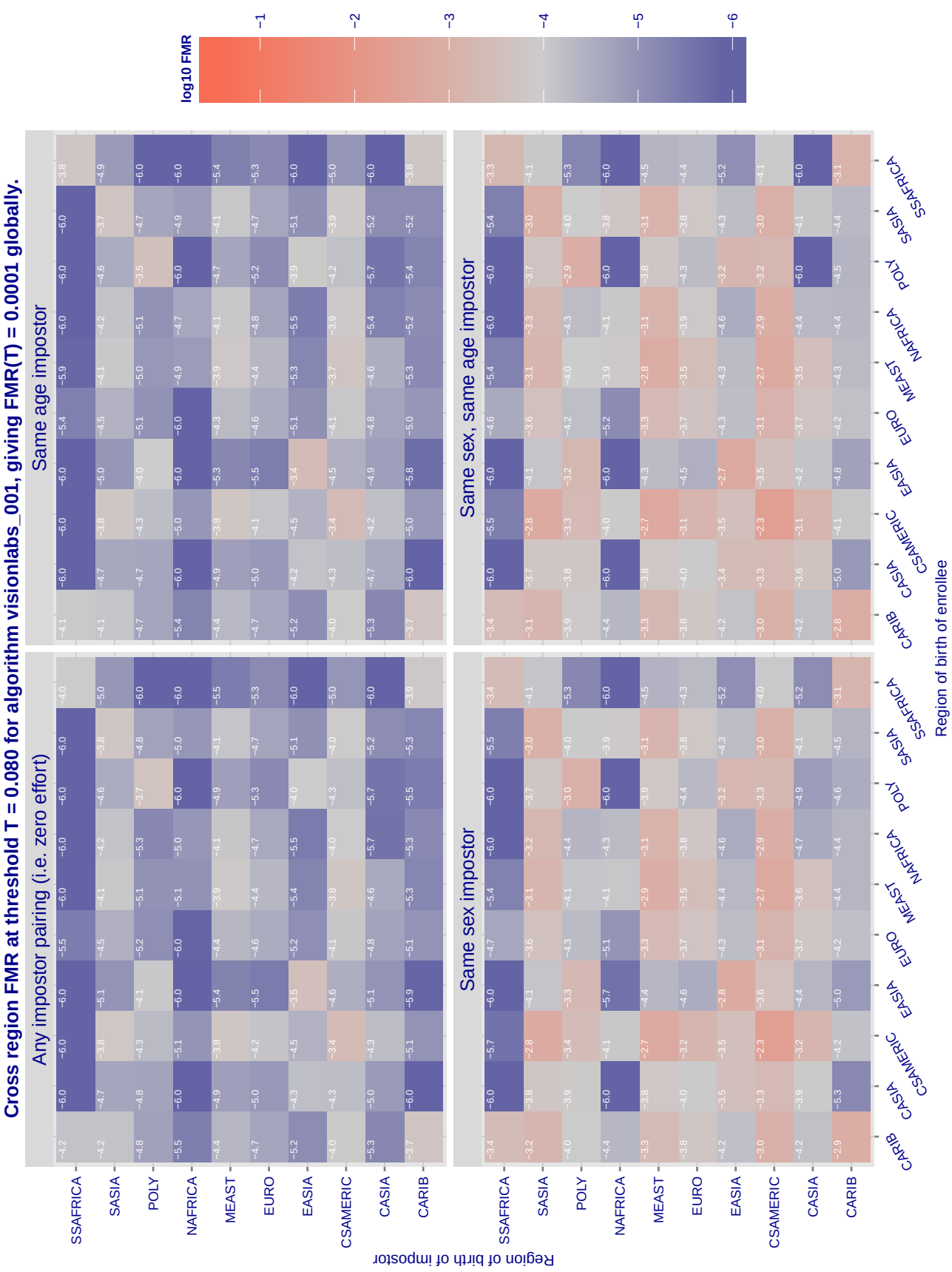


Figure 40: For algorithm visionlabs-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

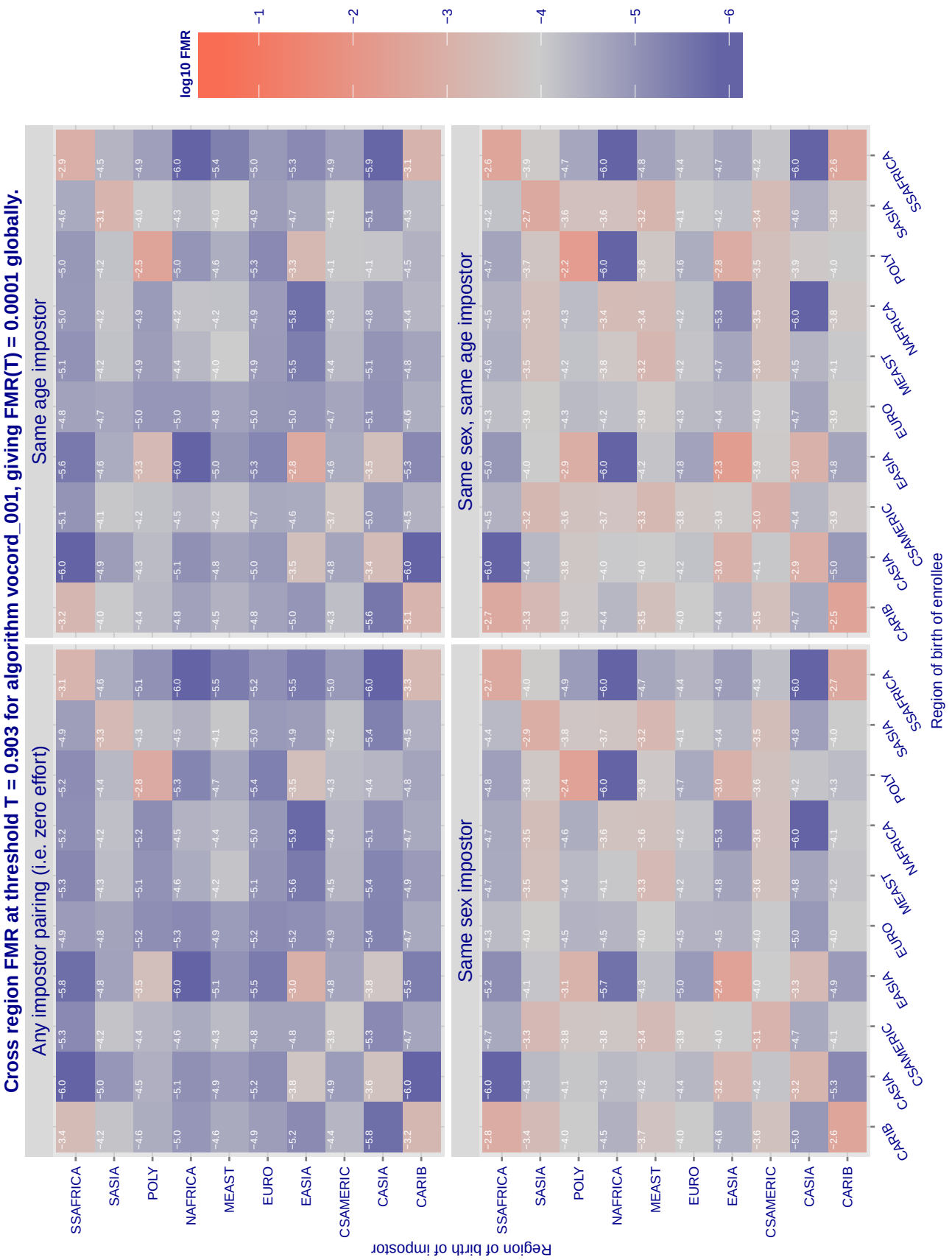


Figure 41: For algorithm vocord-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

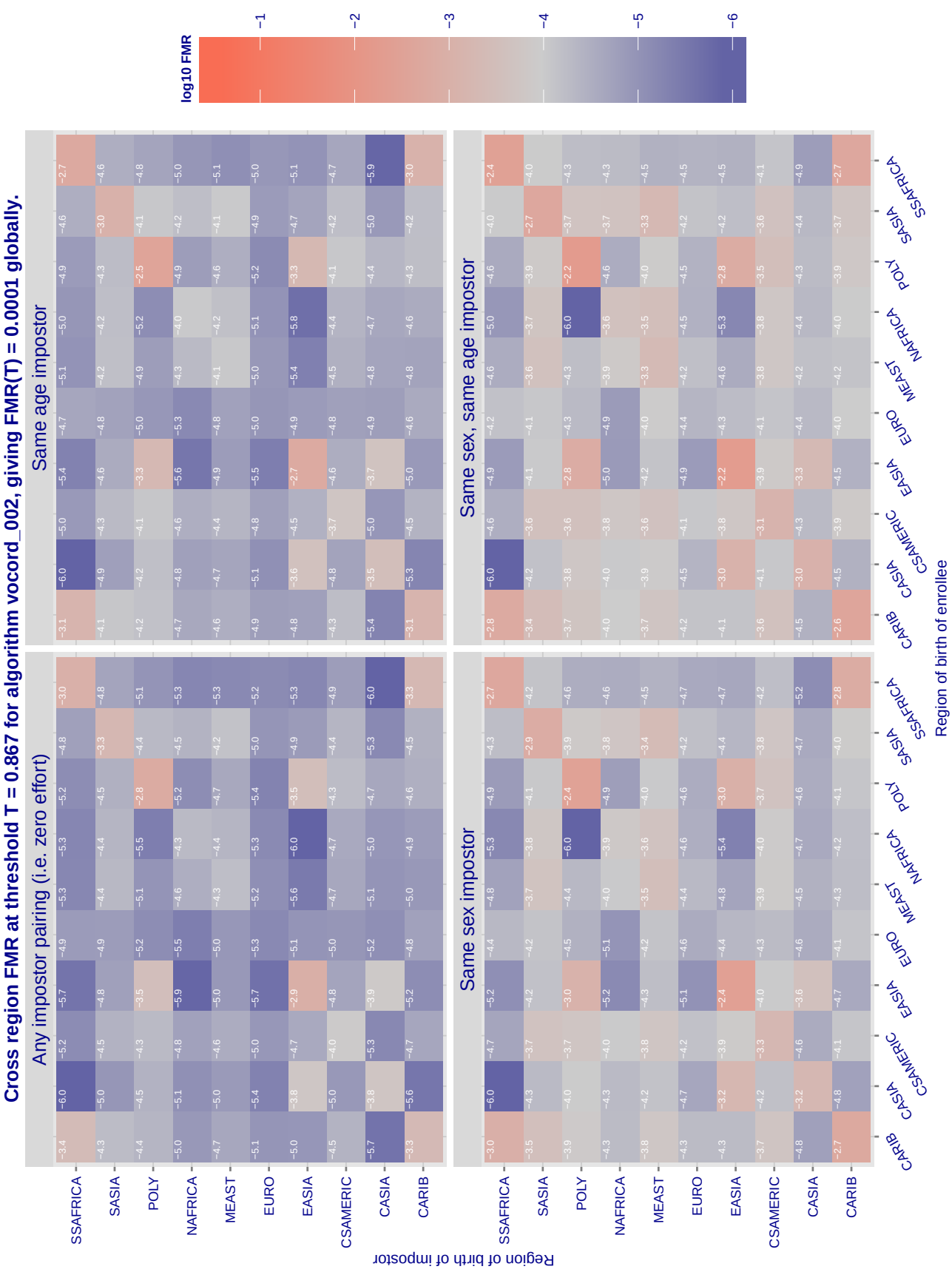


Figure 42: For algorithm vocord-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

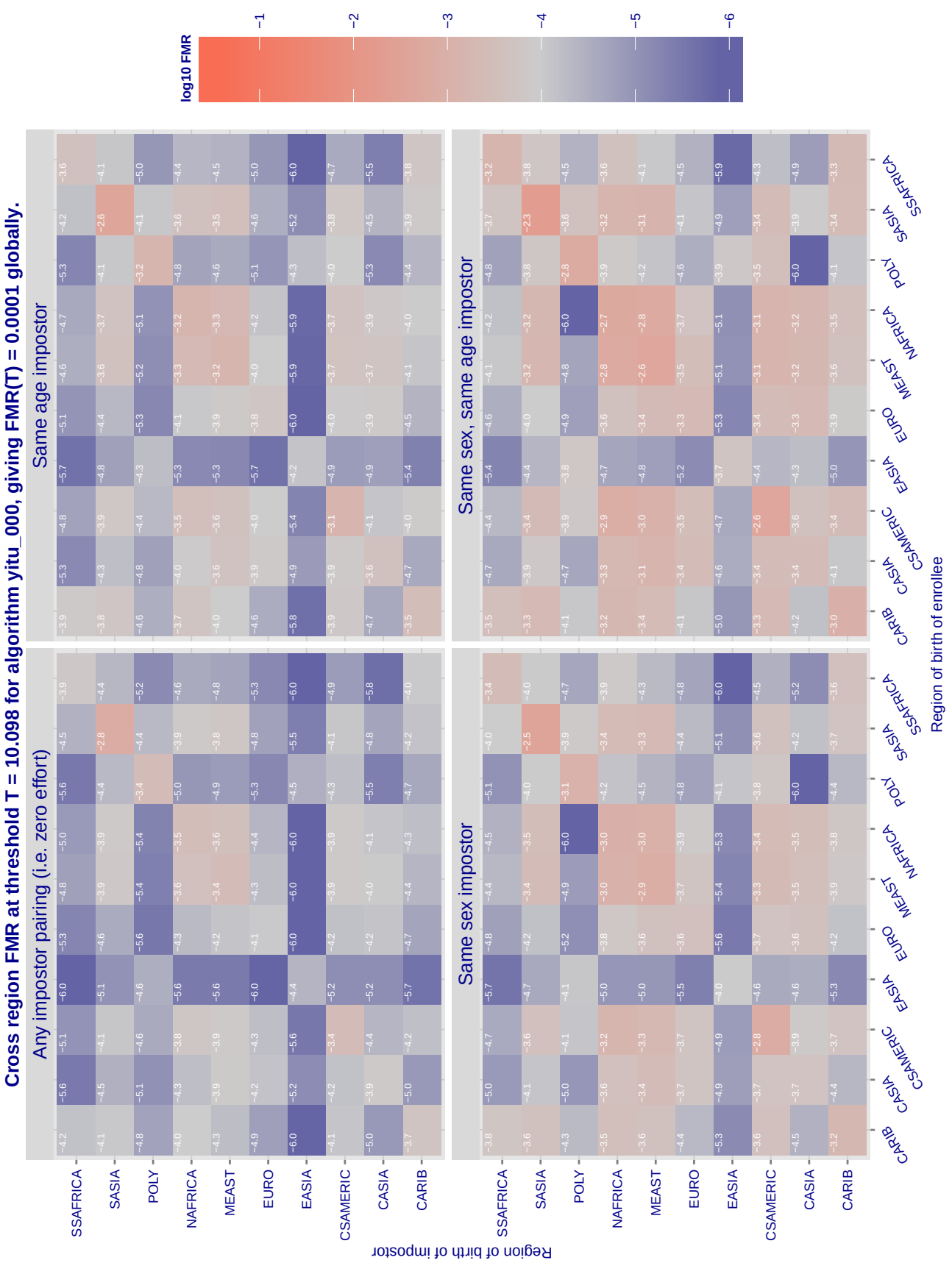


Figure 43: For algorithm yitu-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given region pair. False matches are counted against a recognition threshold fixed globally to give the target FMR 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.

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

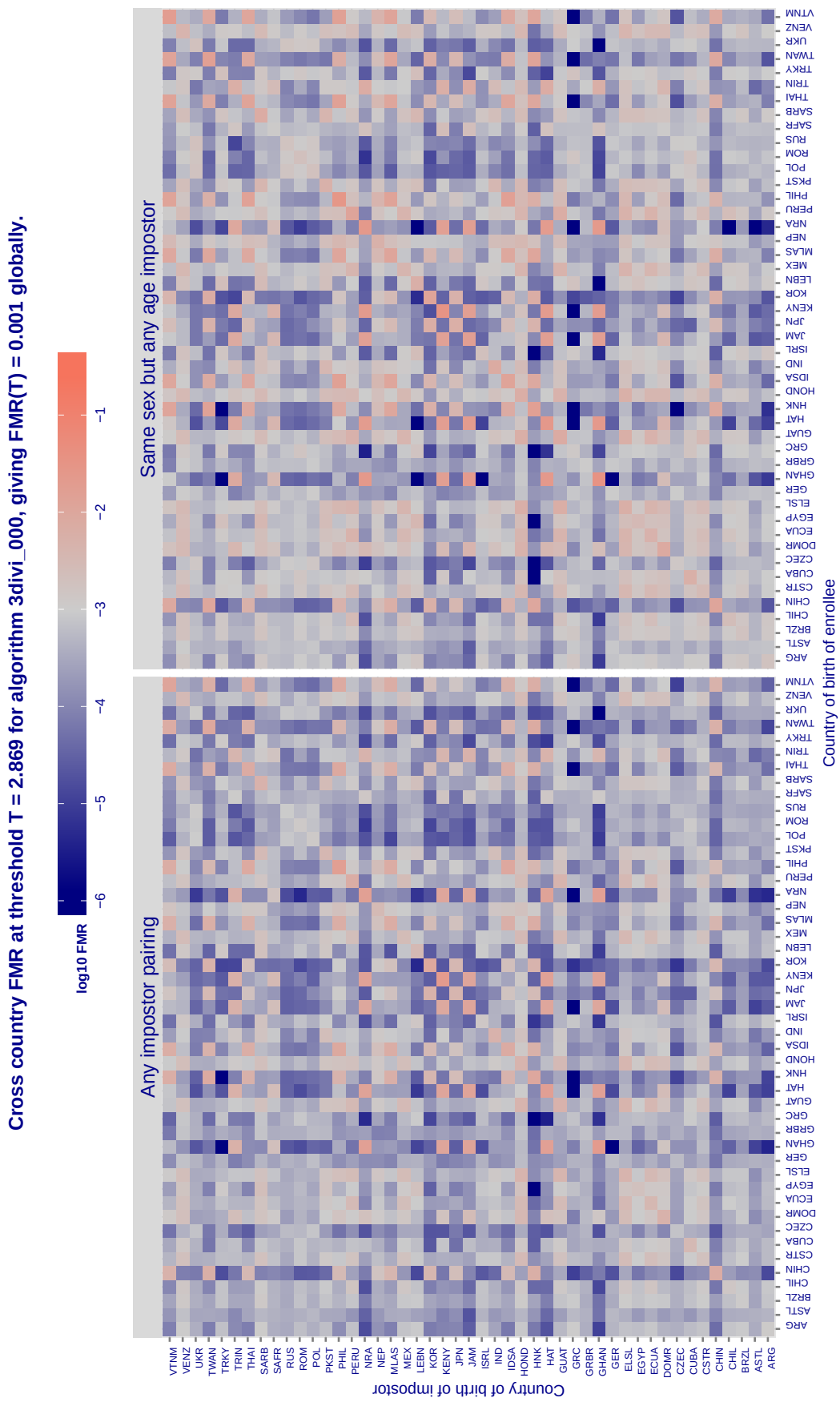


Figure 44: For algorithm 3divi-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

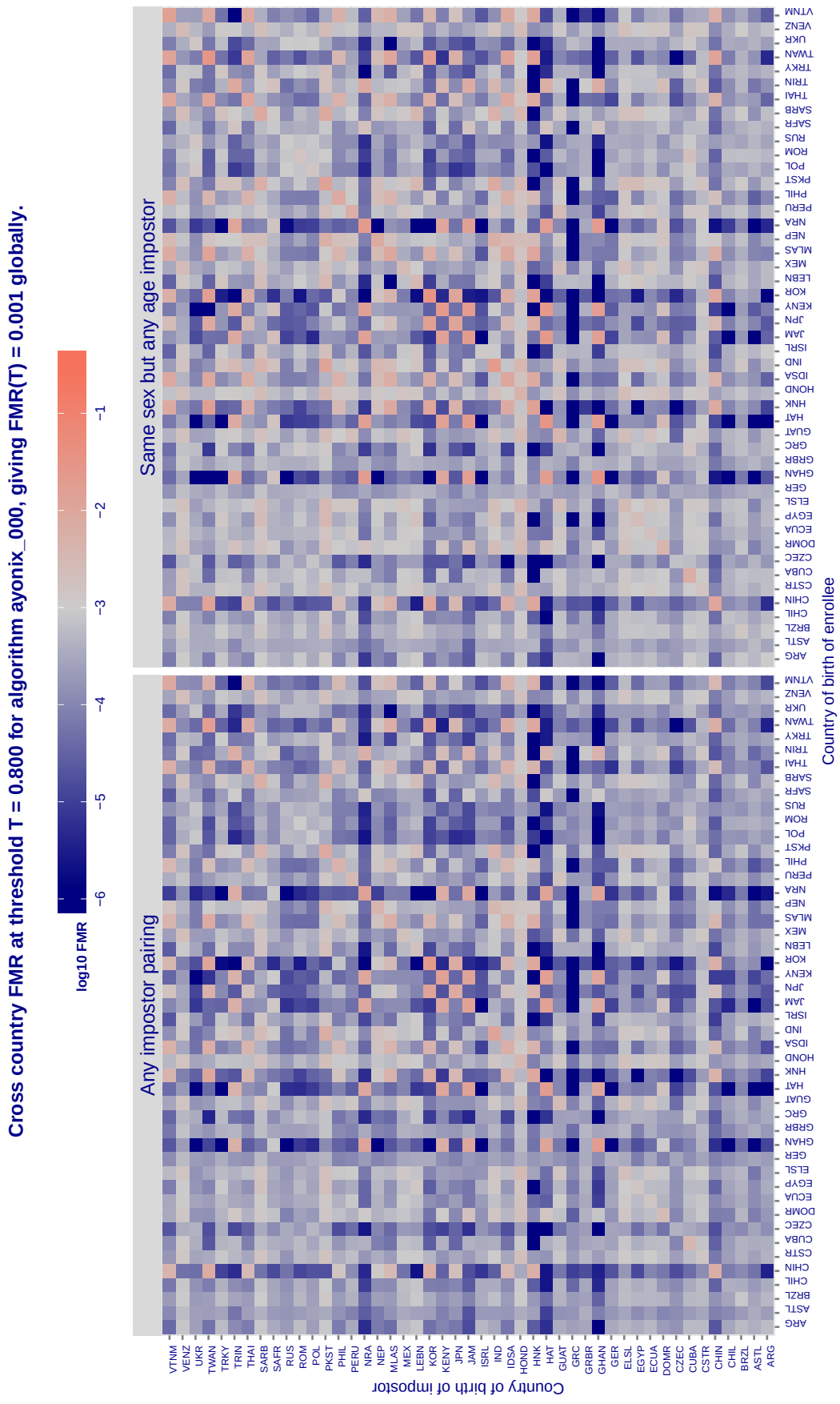


Figure 45: For algorithm ayonix-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

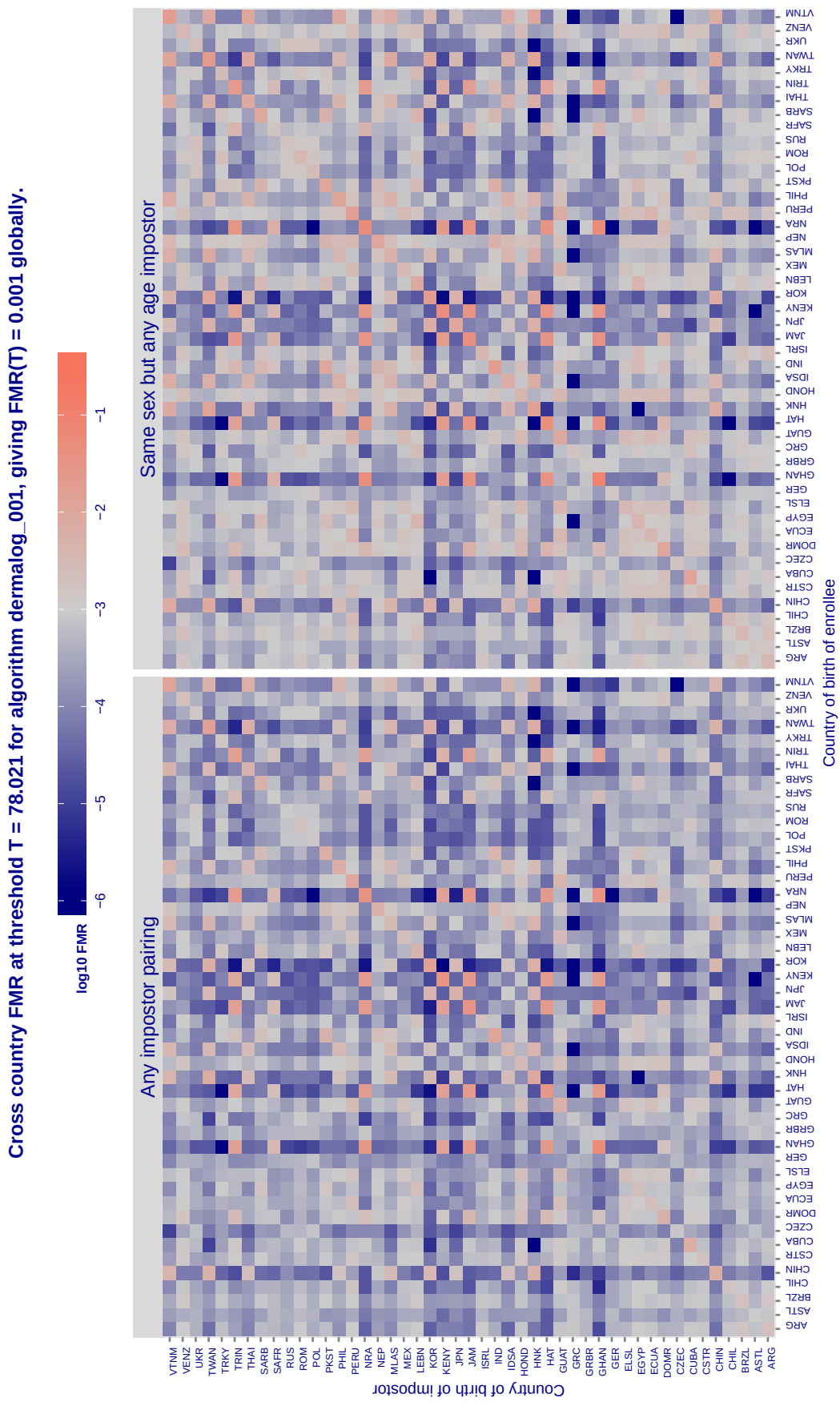


Figure 46: For algorithm dermalog-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

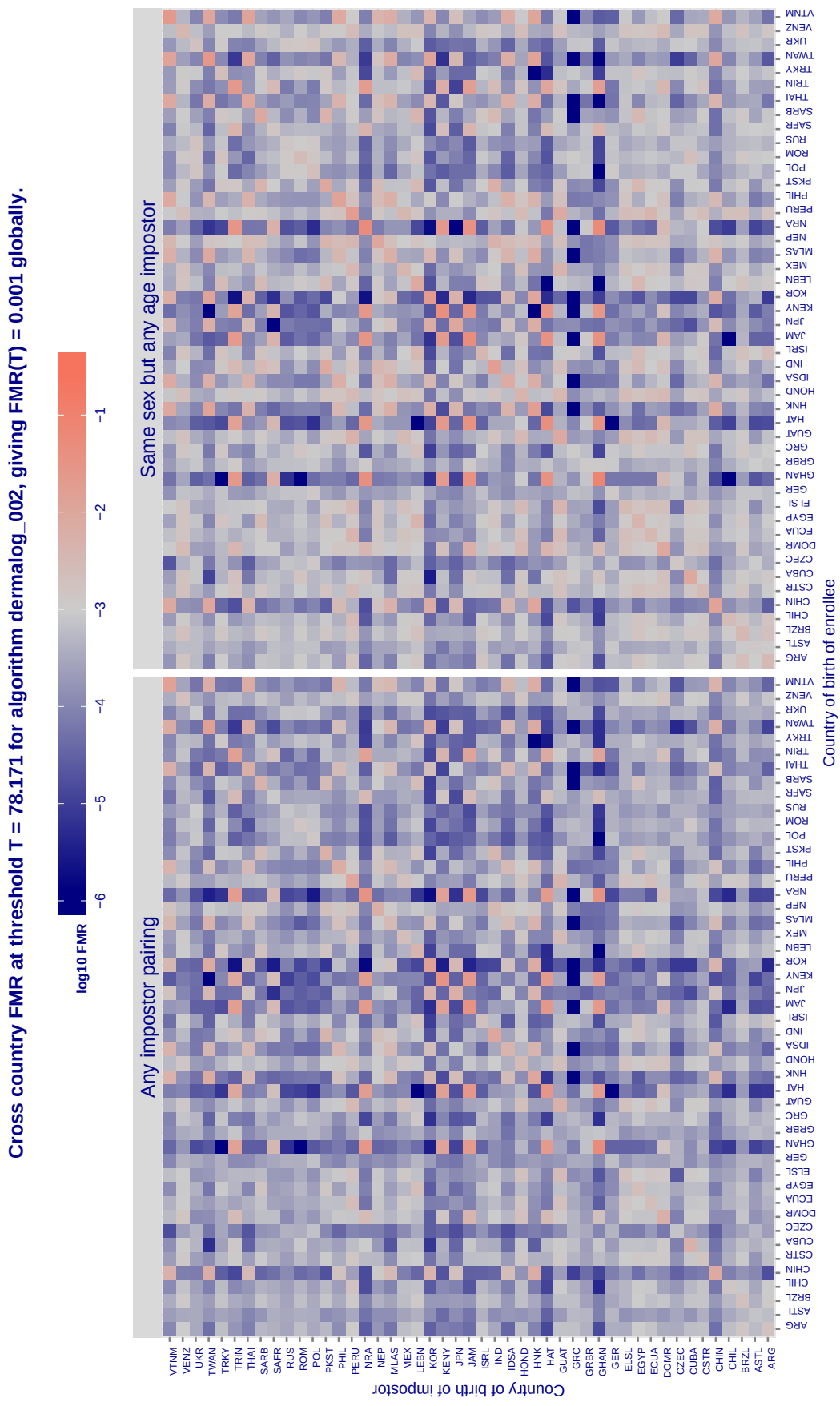


Figure 47: For algorithm dermalog-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

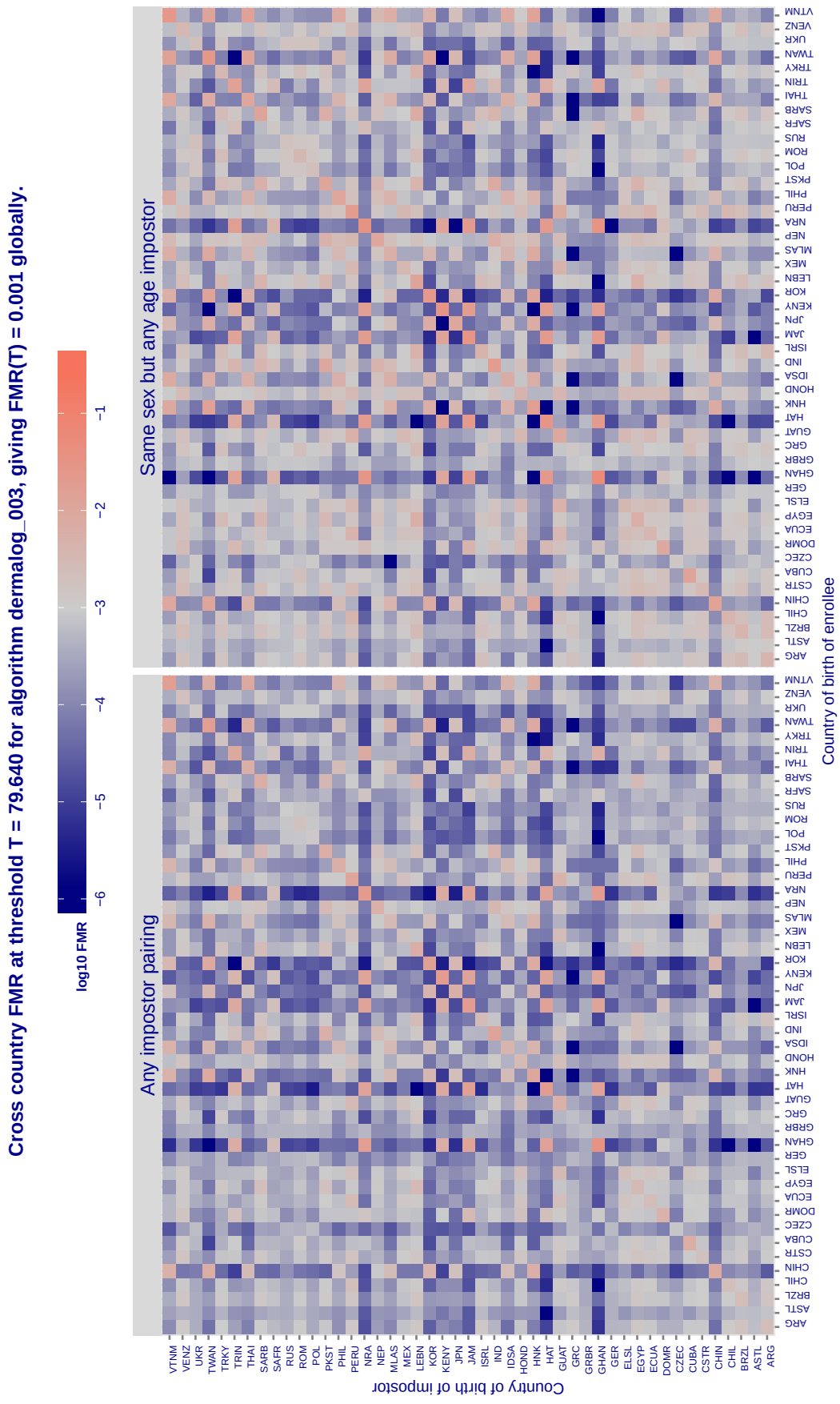


Figure 48: For algorithm dermalog-003 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

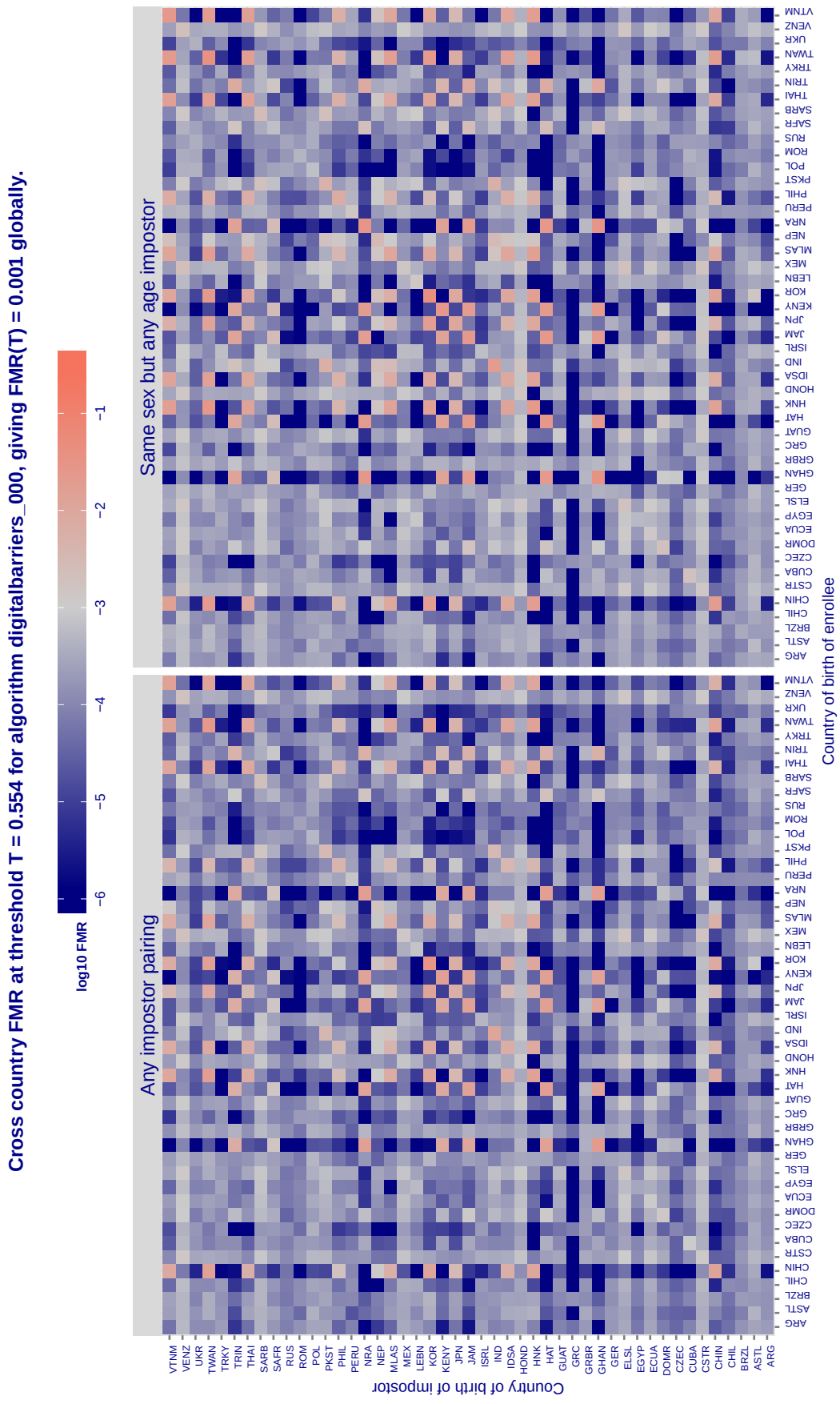


Figure 49: For algorithm digitalbarriers-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the 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.

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

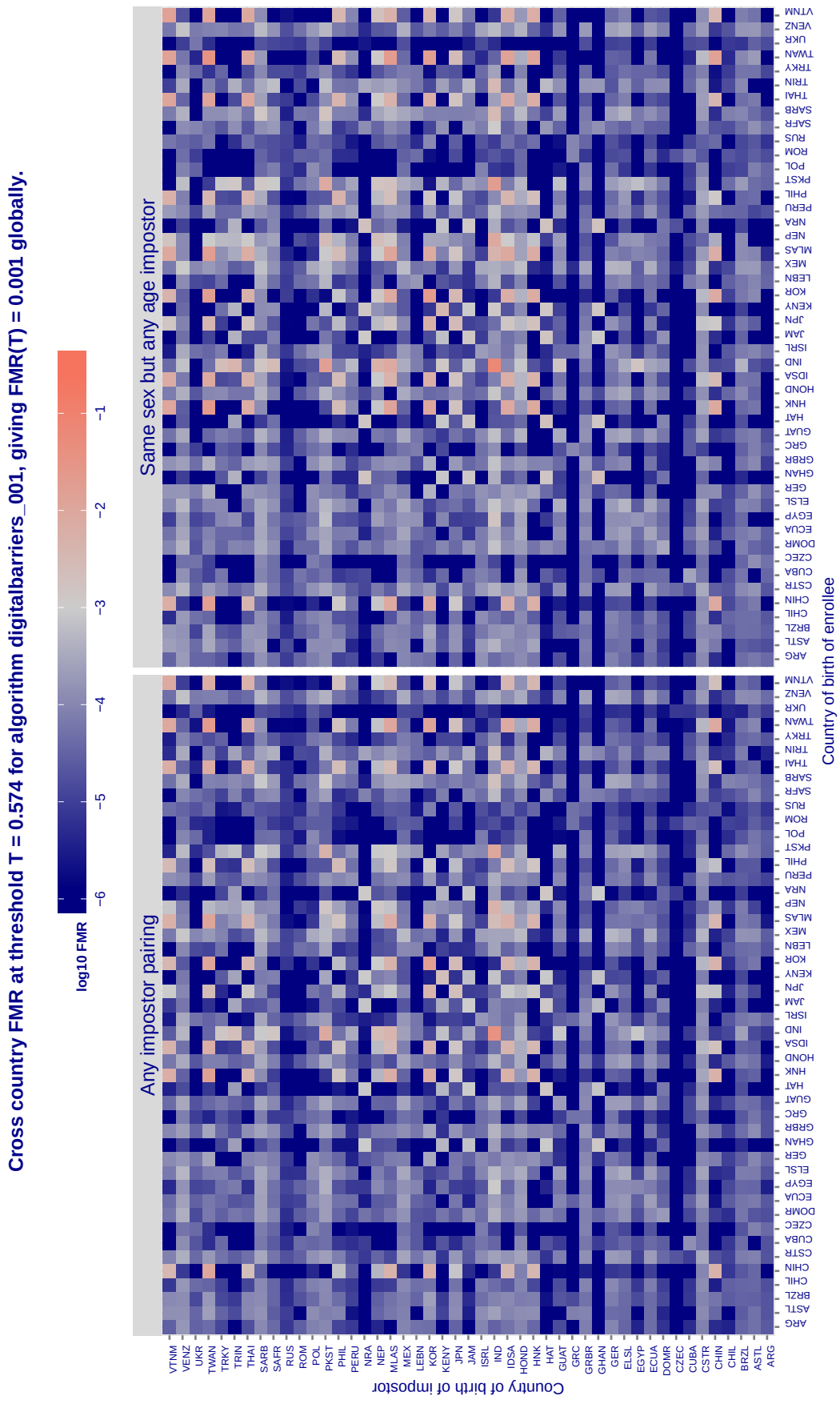


Figure 50: For algorithm digitalbarriers-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the 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.

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

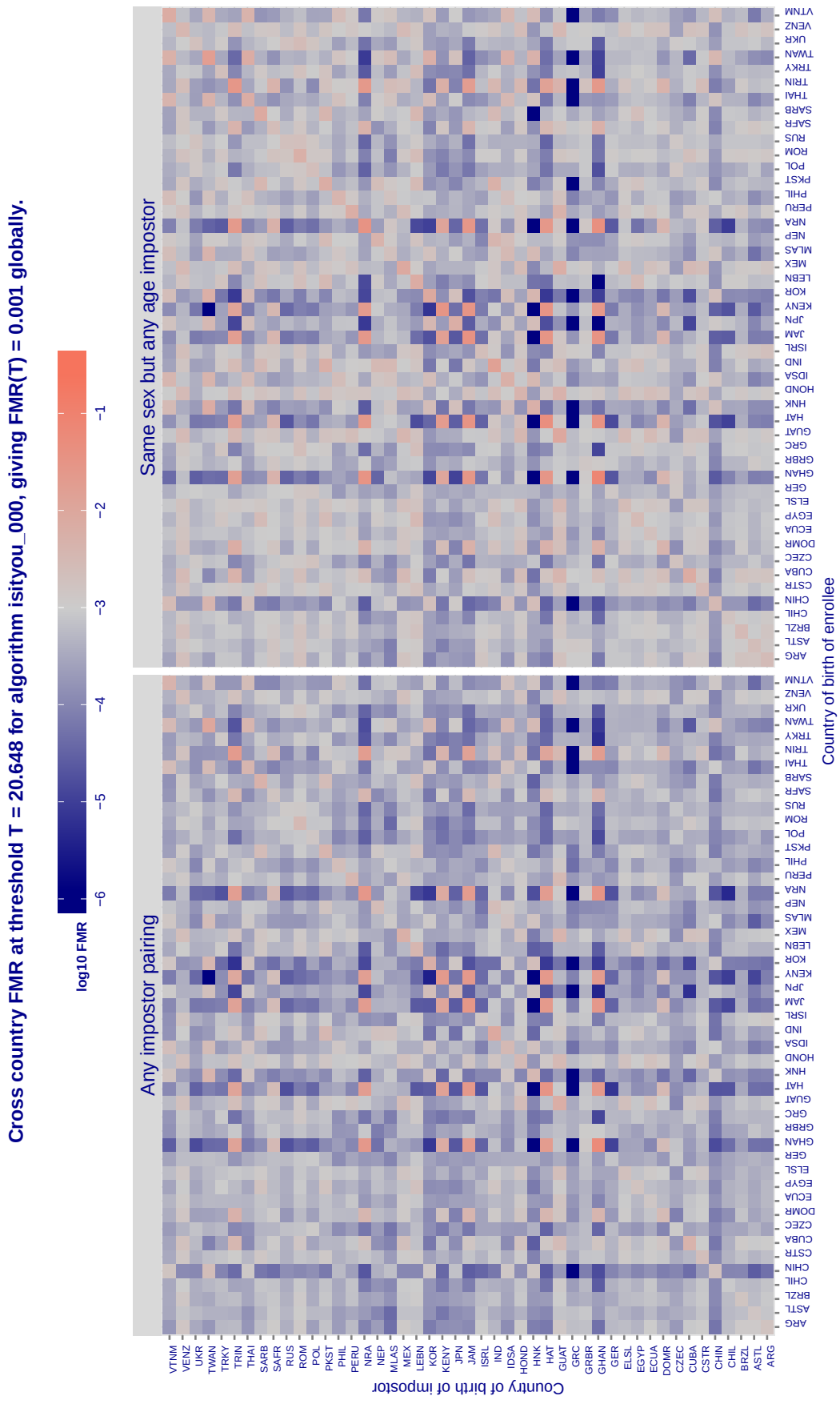


Figure 51: For algorithm isityou-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

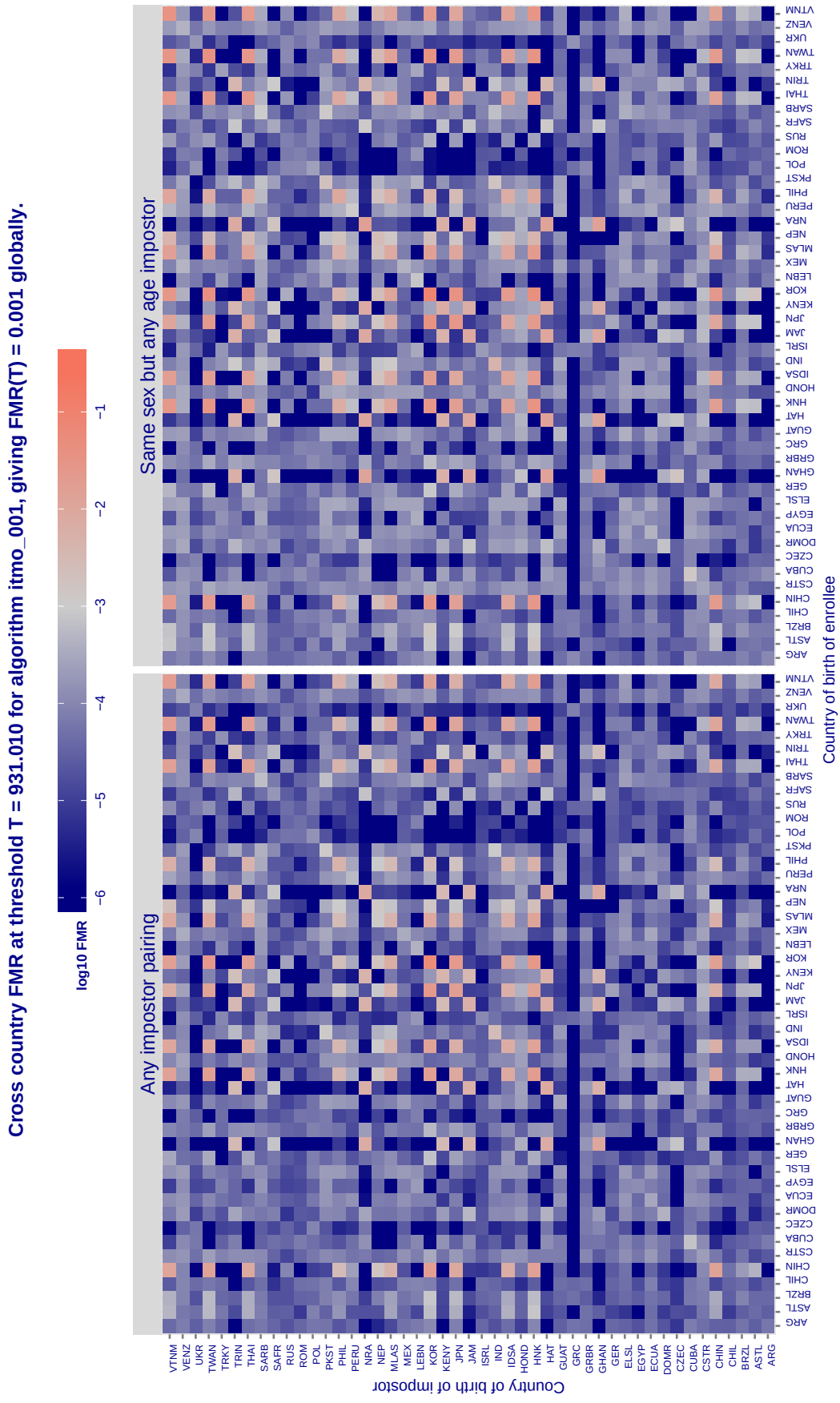


Figure 52: For algorithm itmo-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR 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.

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

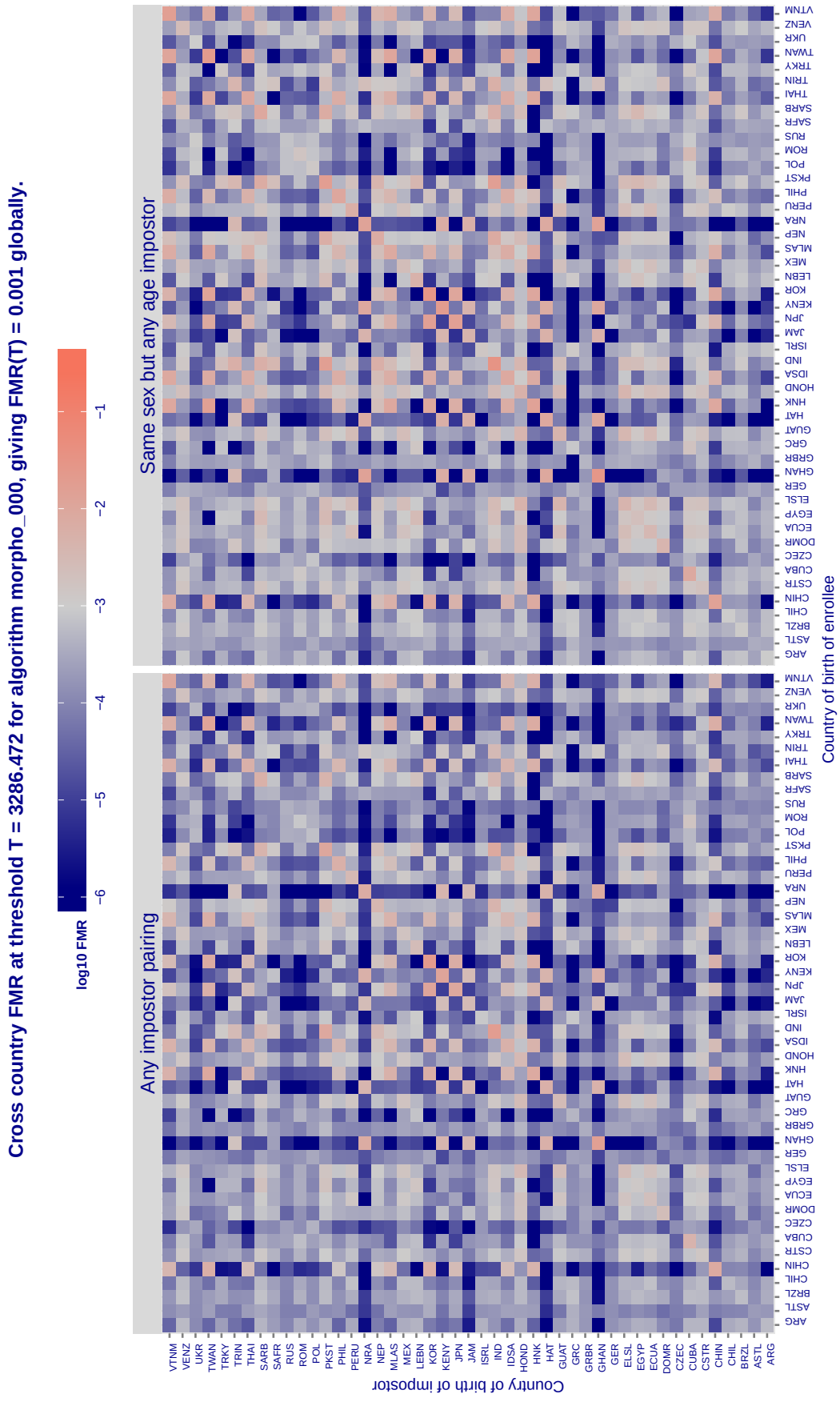


Figure 53: For algorithm morpho-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

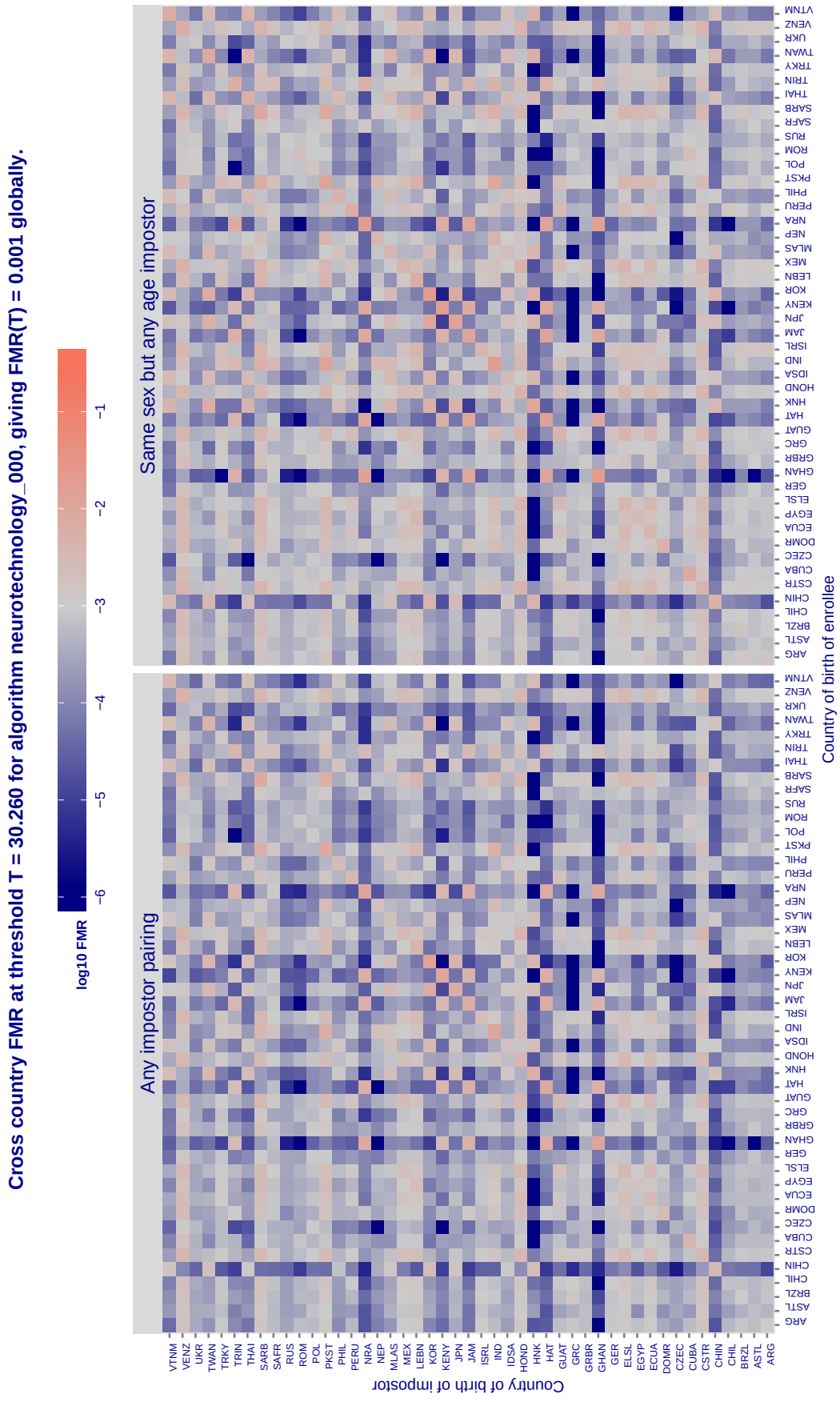


Figure 54: For algorithm neurotechnology-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the 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.

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

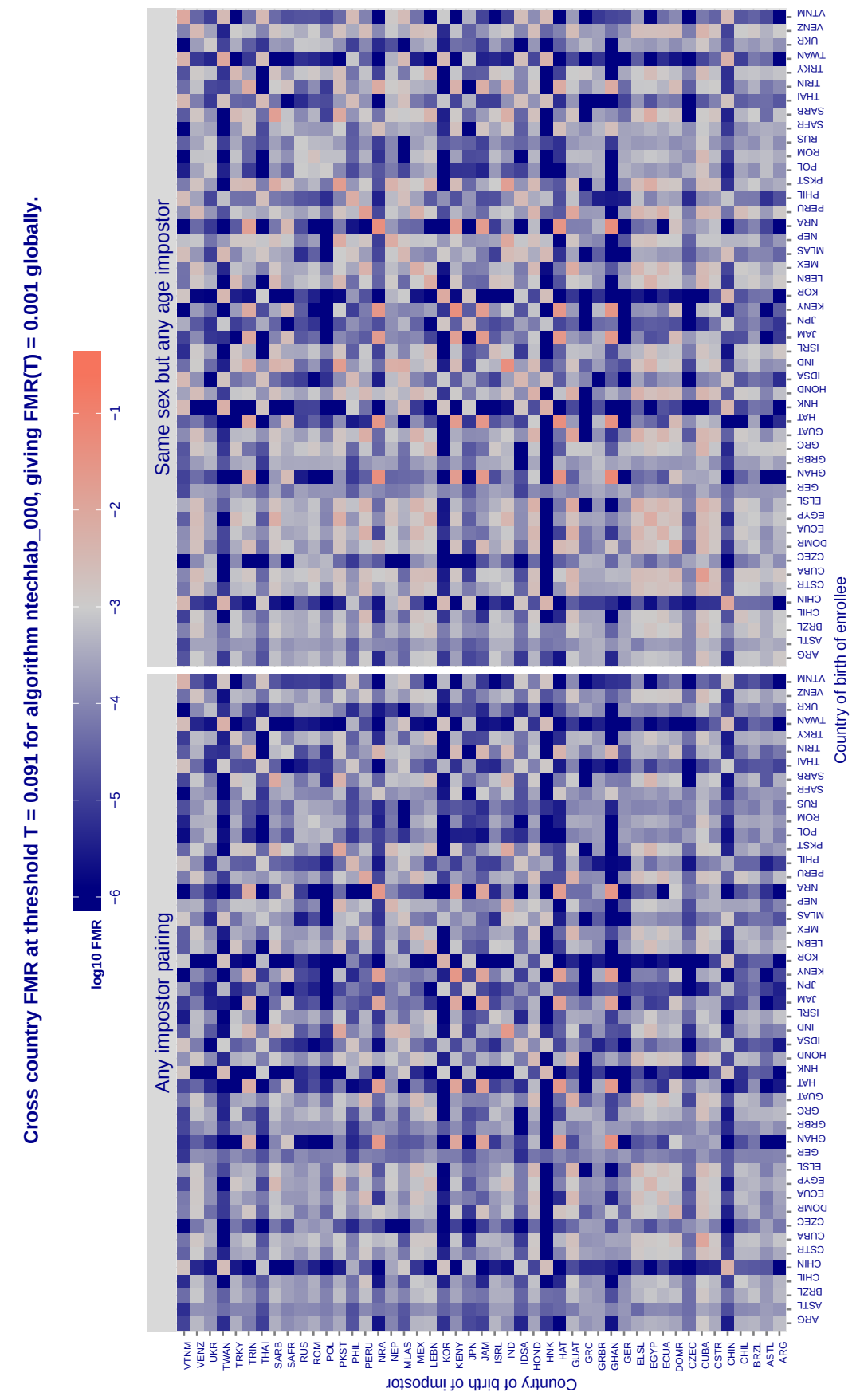


Figure 55: For algorithm ntechlab-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

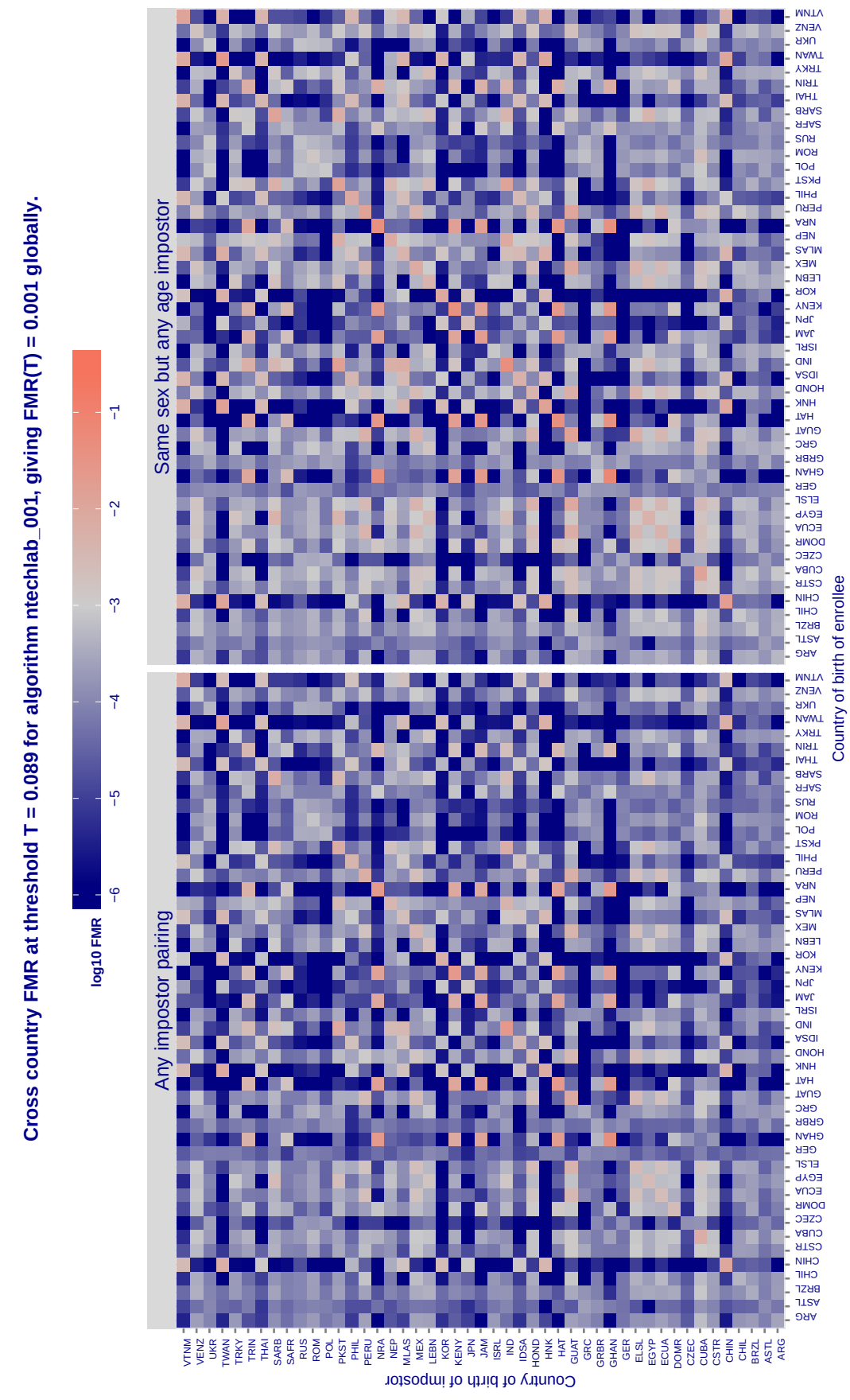


Figure 56: For algorithm ntechlab-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

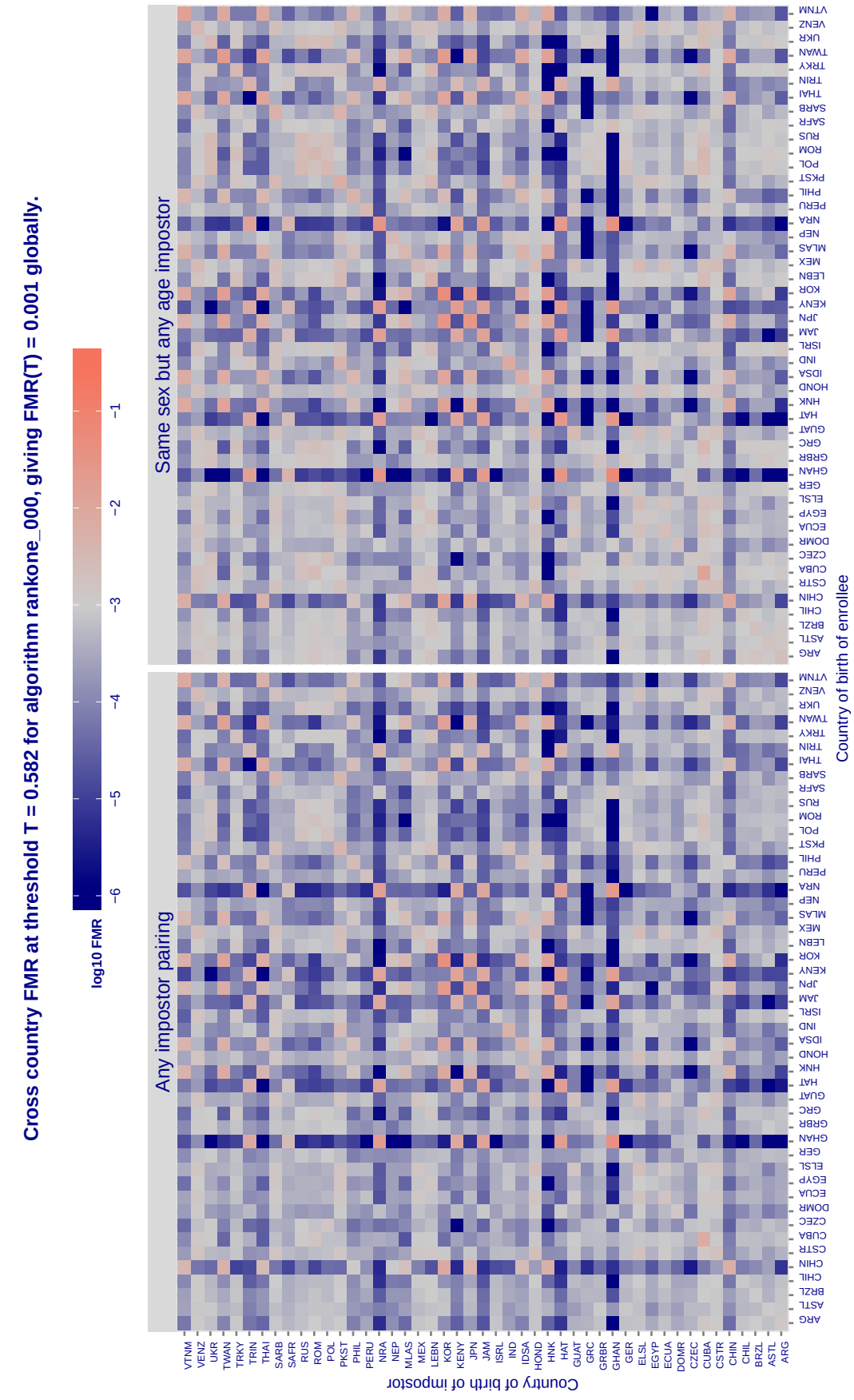


Figure 57: For algorithm rankone-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

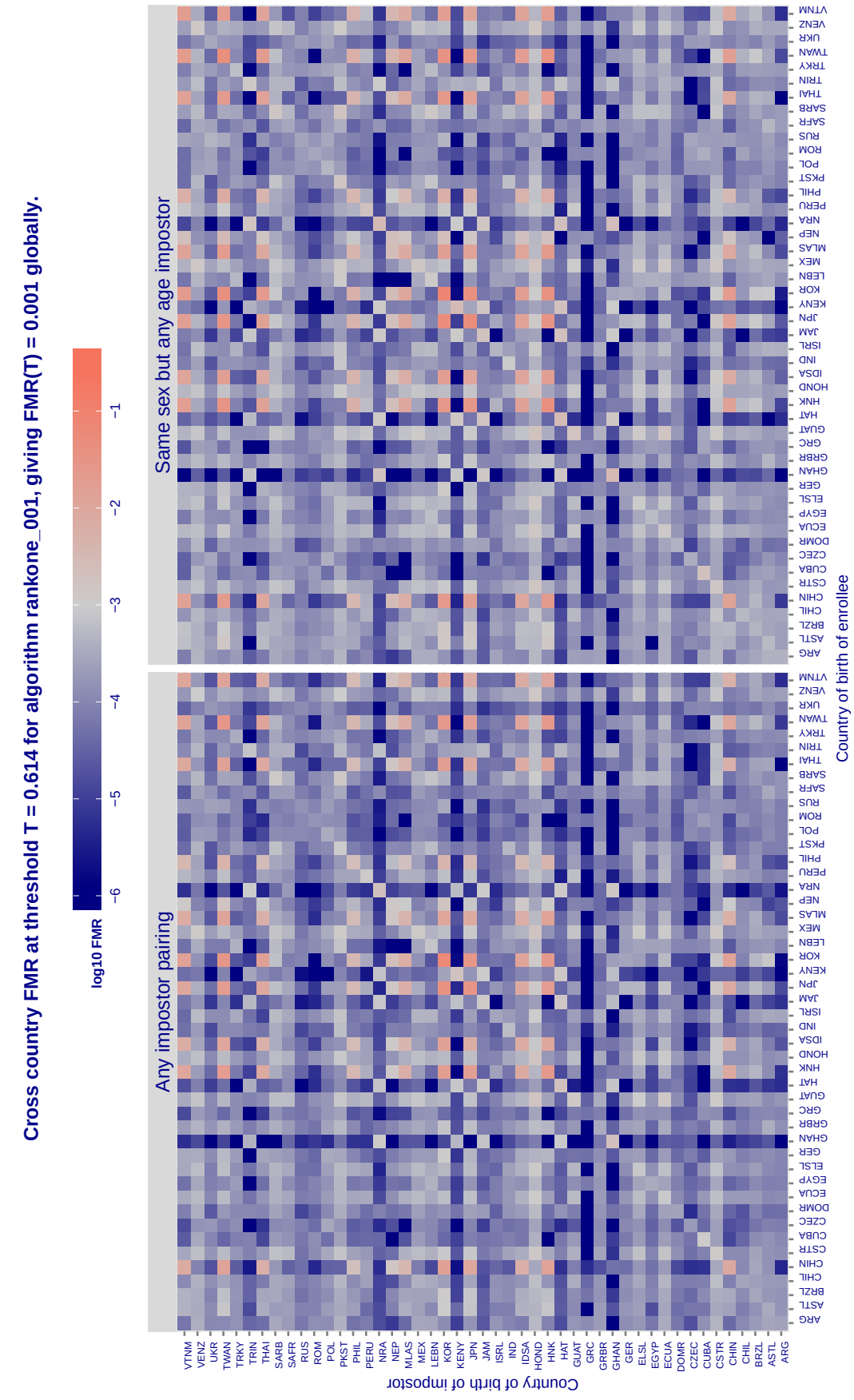


Figure 58: For algorithm rankone-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

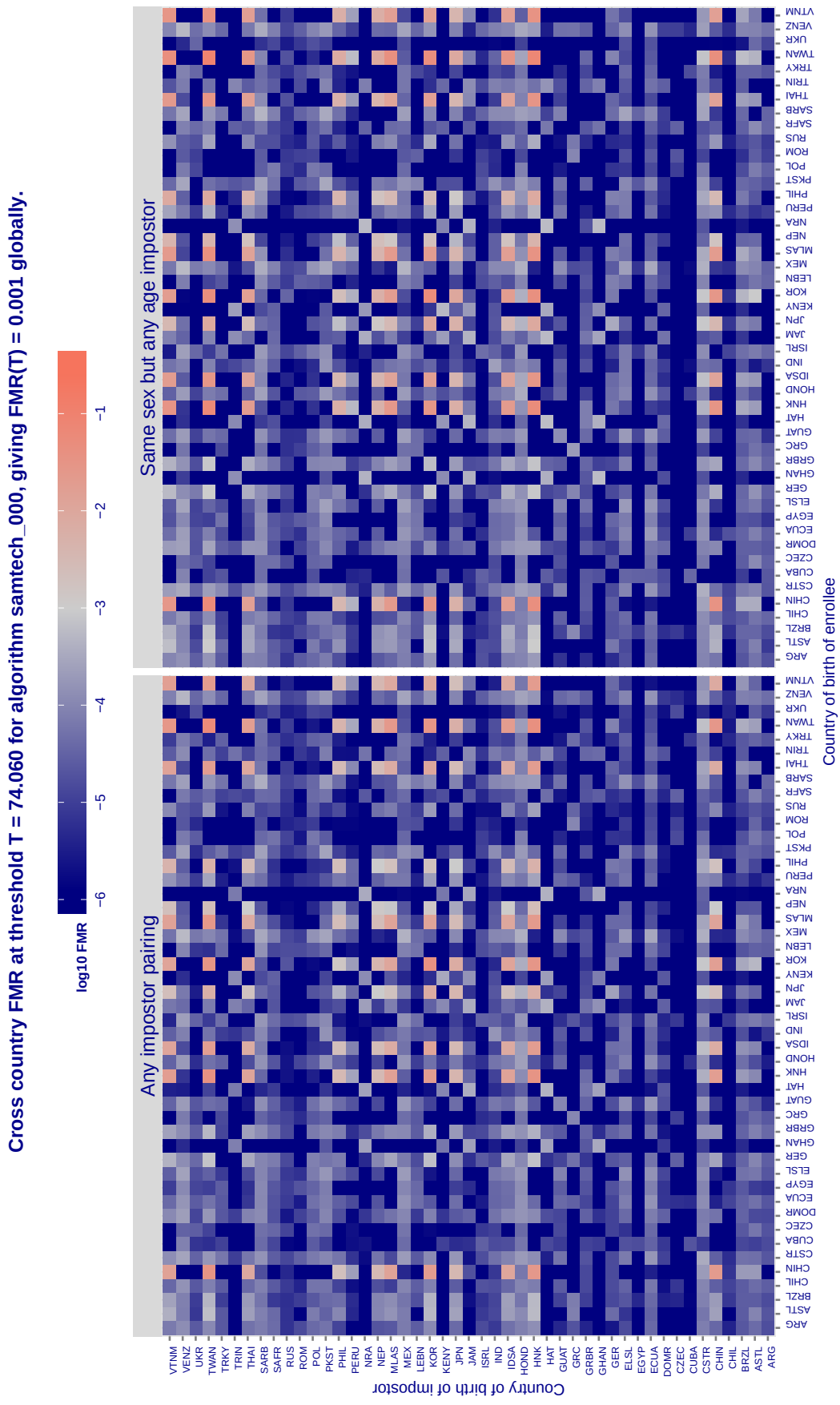


Figure 59: For algorithm samtech-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

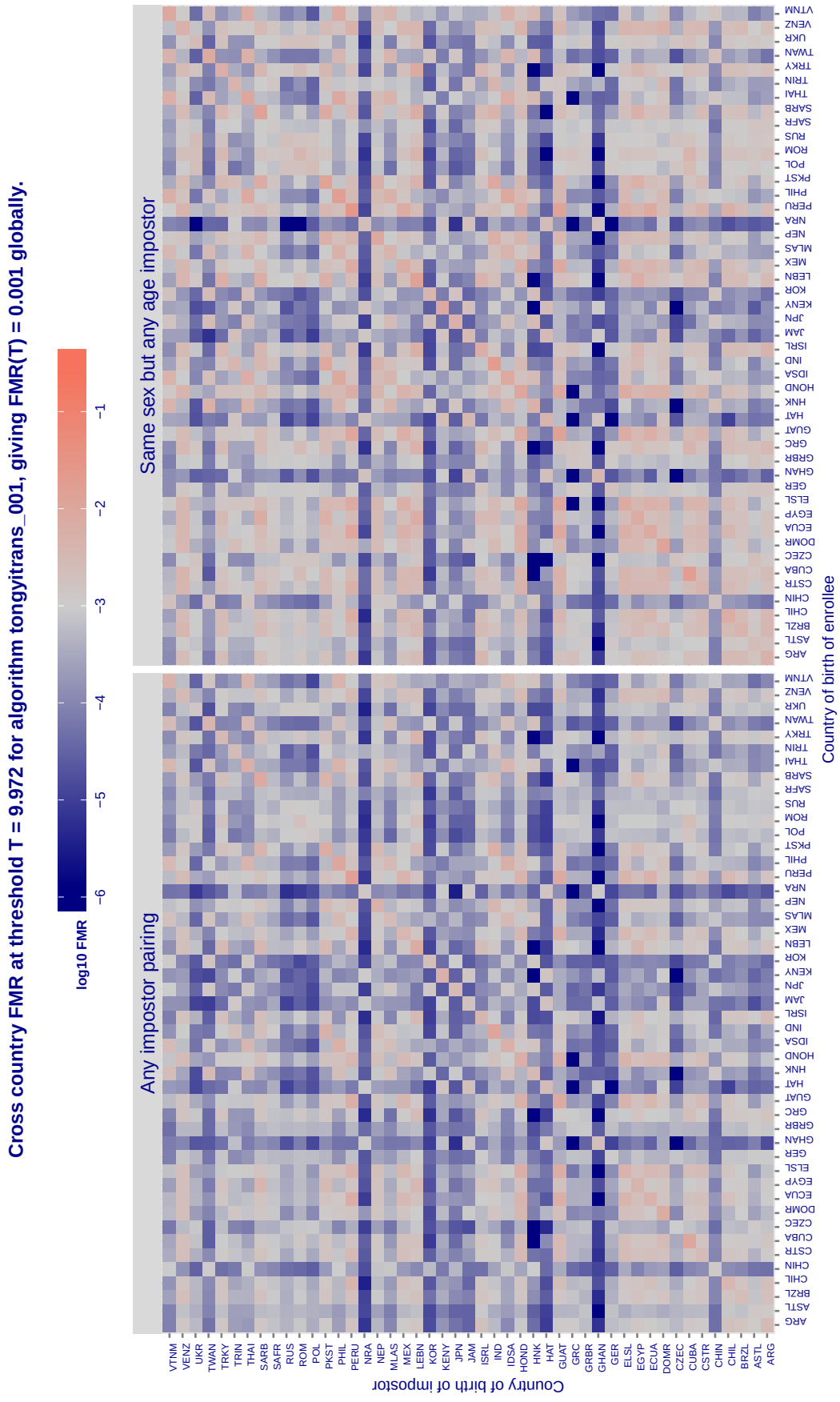


Figure 60: For algorithm tongyitrans-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

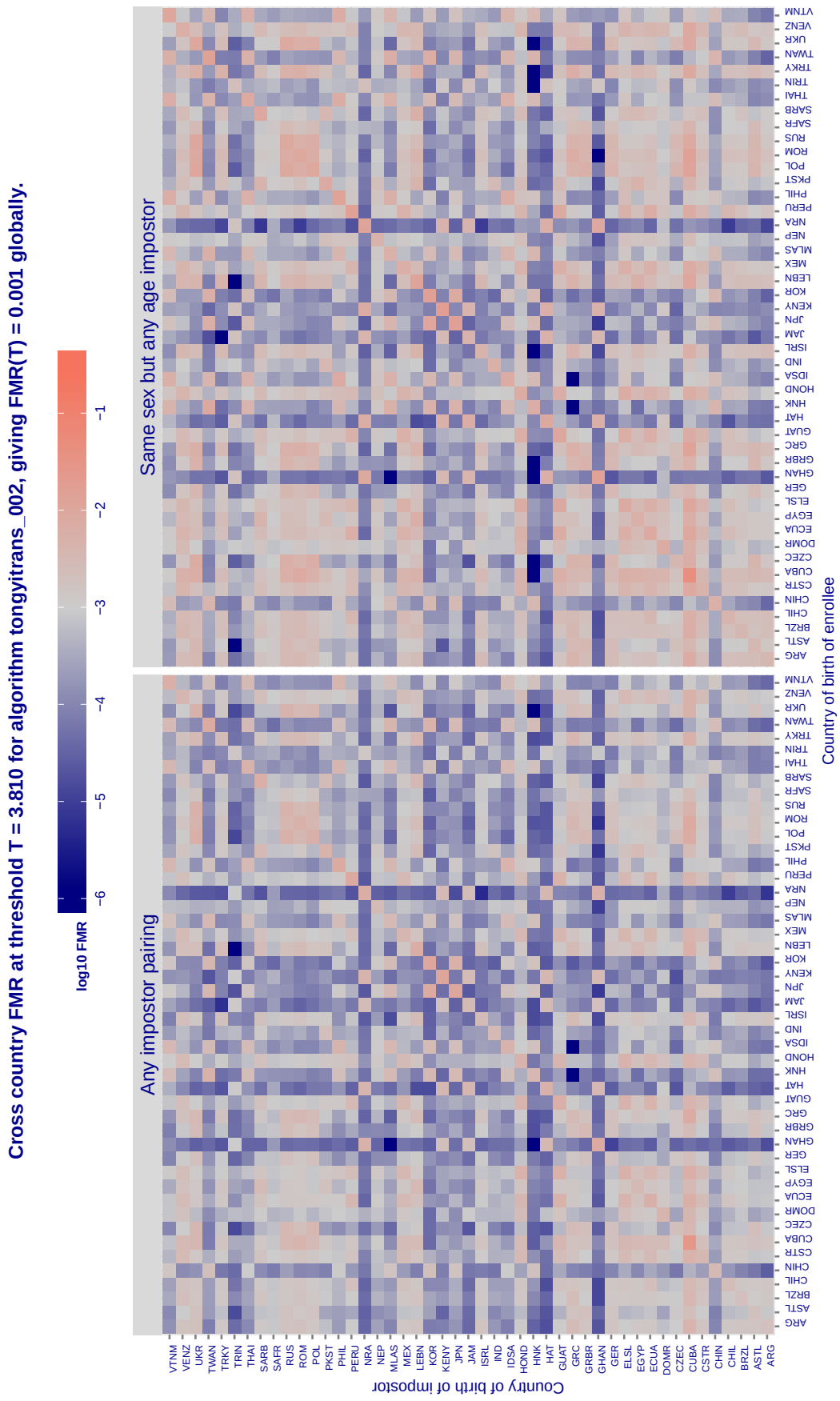


Figure 61: For algorithm tongyitrans-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

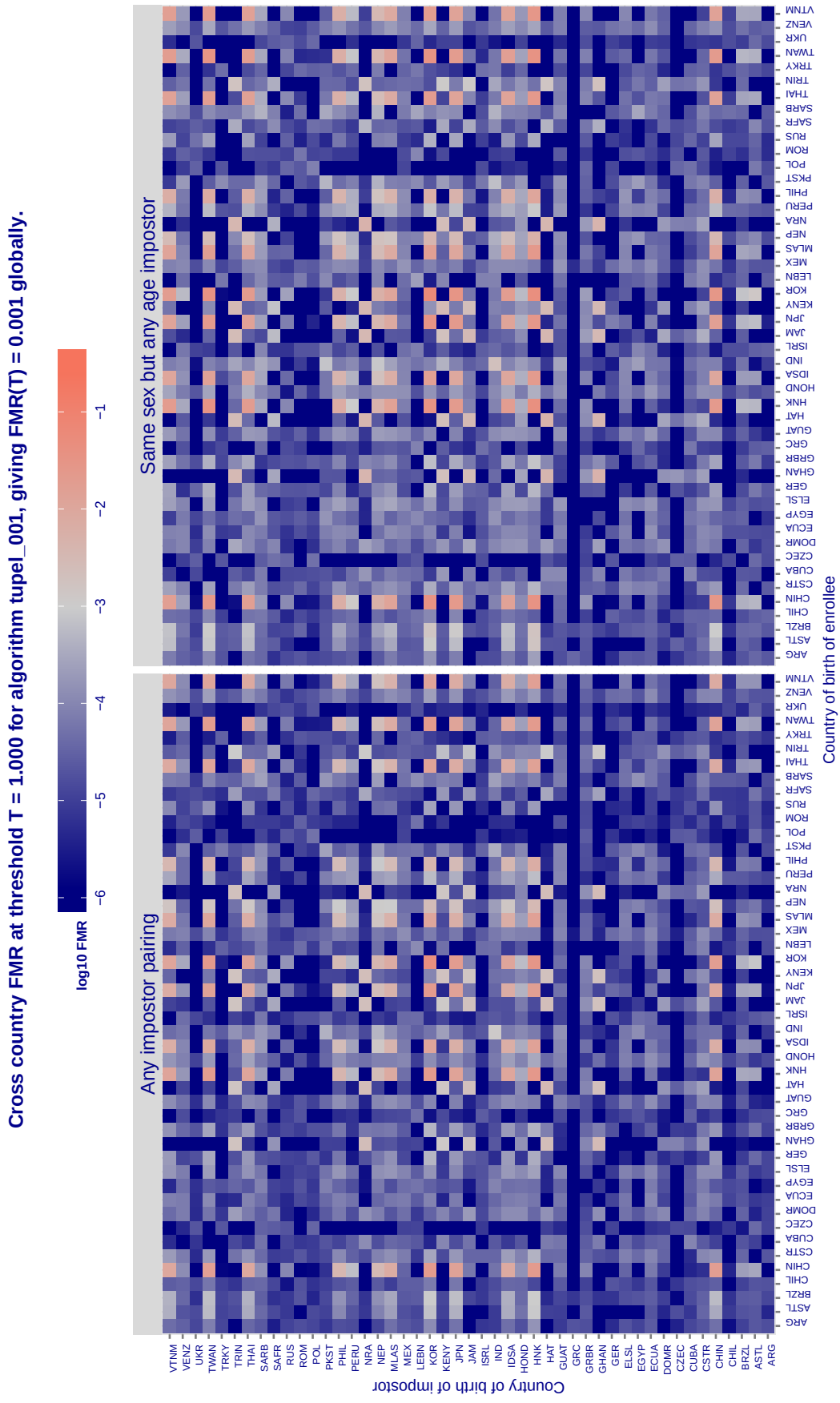


Figure 62: For algorithm tupel-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

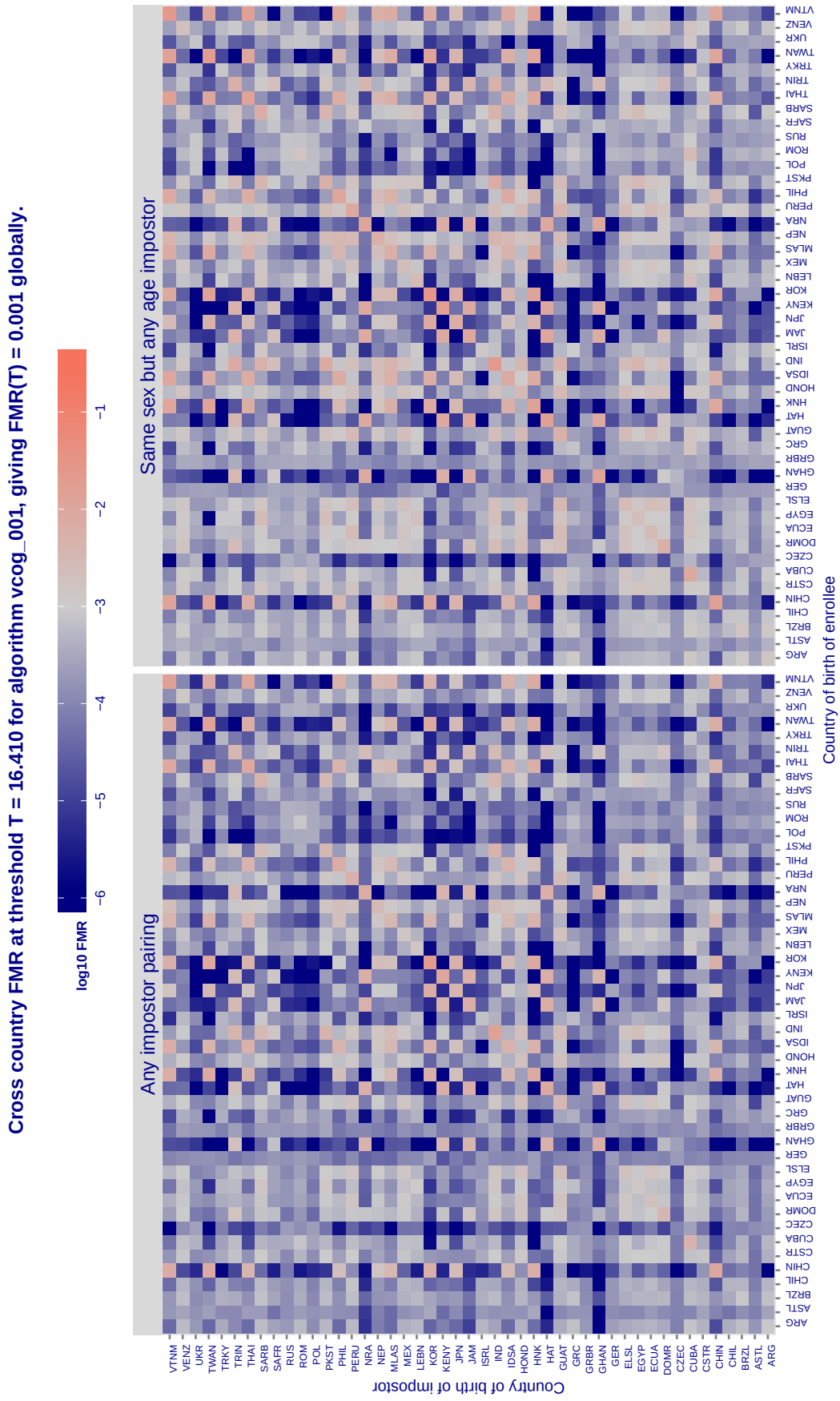


Figure 63: For algorithm vcog-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR 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.

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

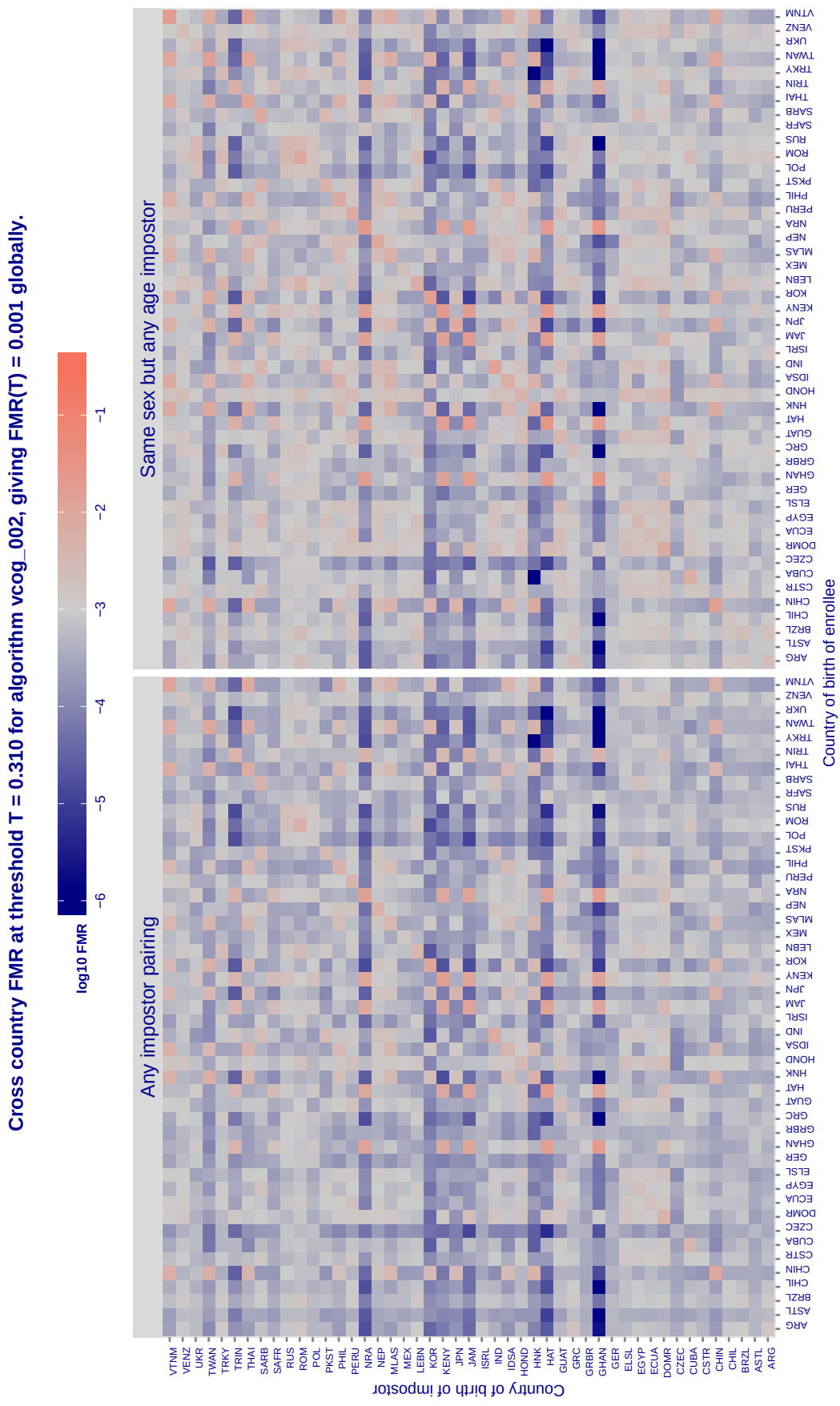


Figure 64: For algorithm vcog-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR 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.

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

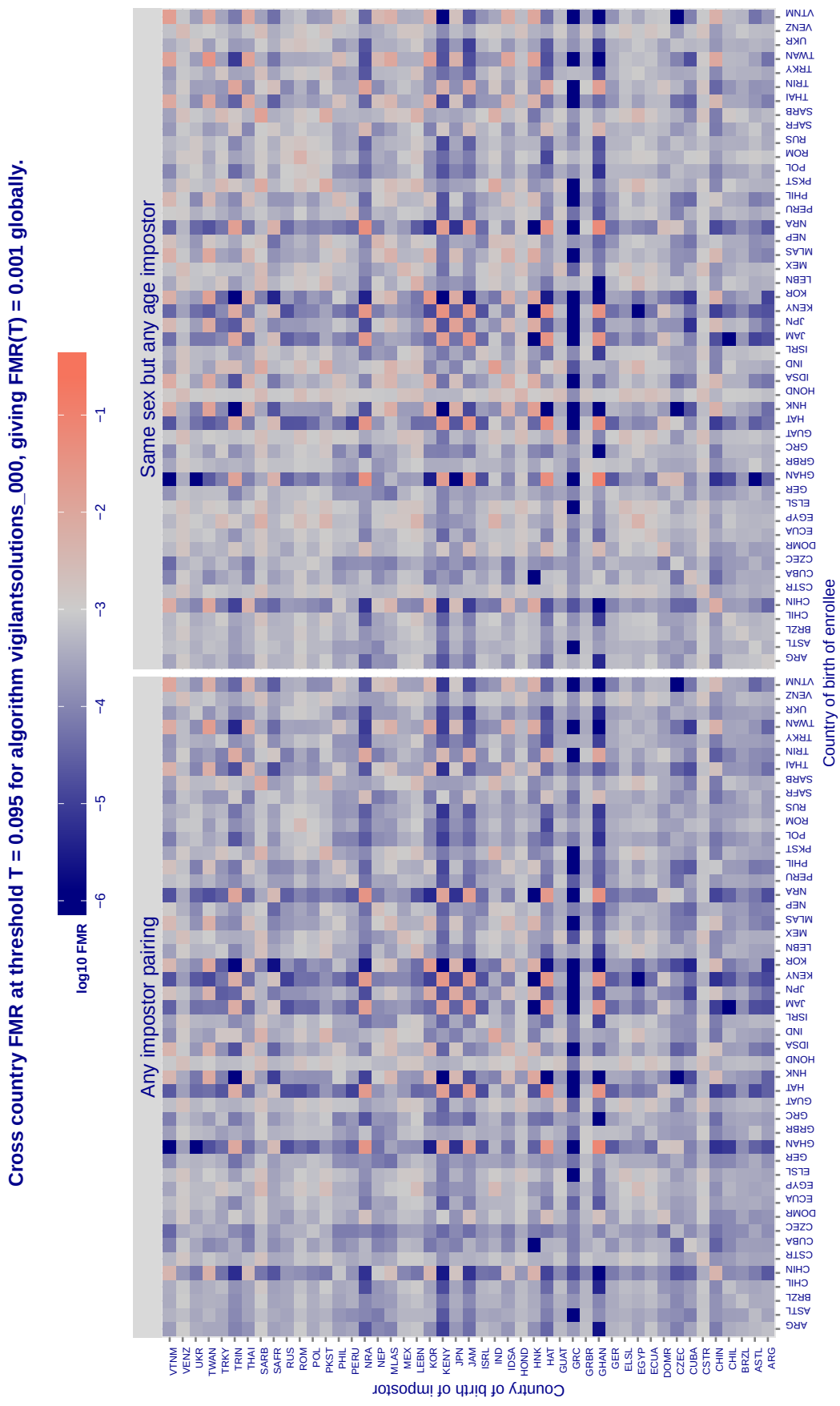


Figure 65: For algorithm vigilantsolutions-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give 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.

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

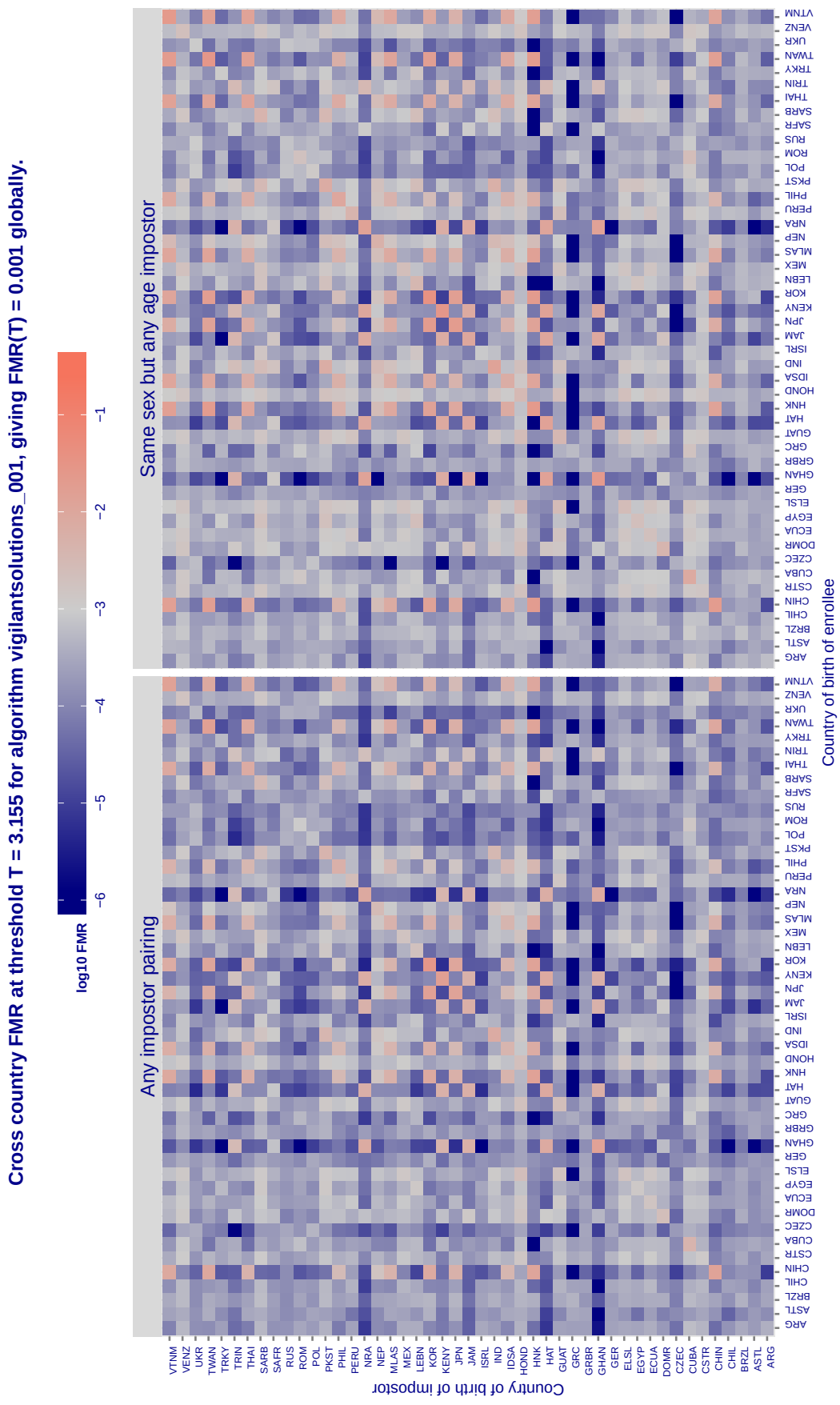


Figure 66: For algorithm vigilantsolutions-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give 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.

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

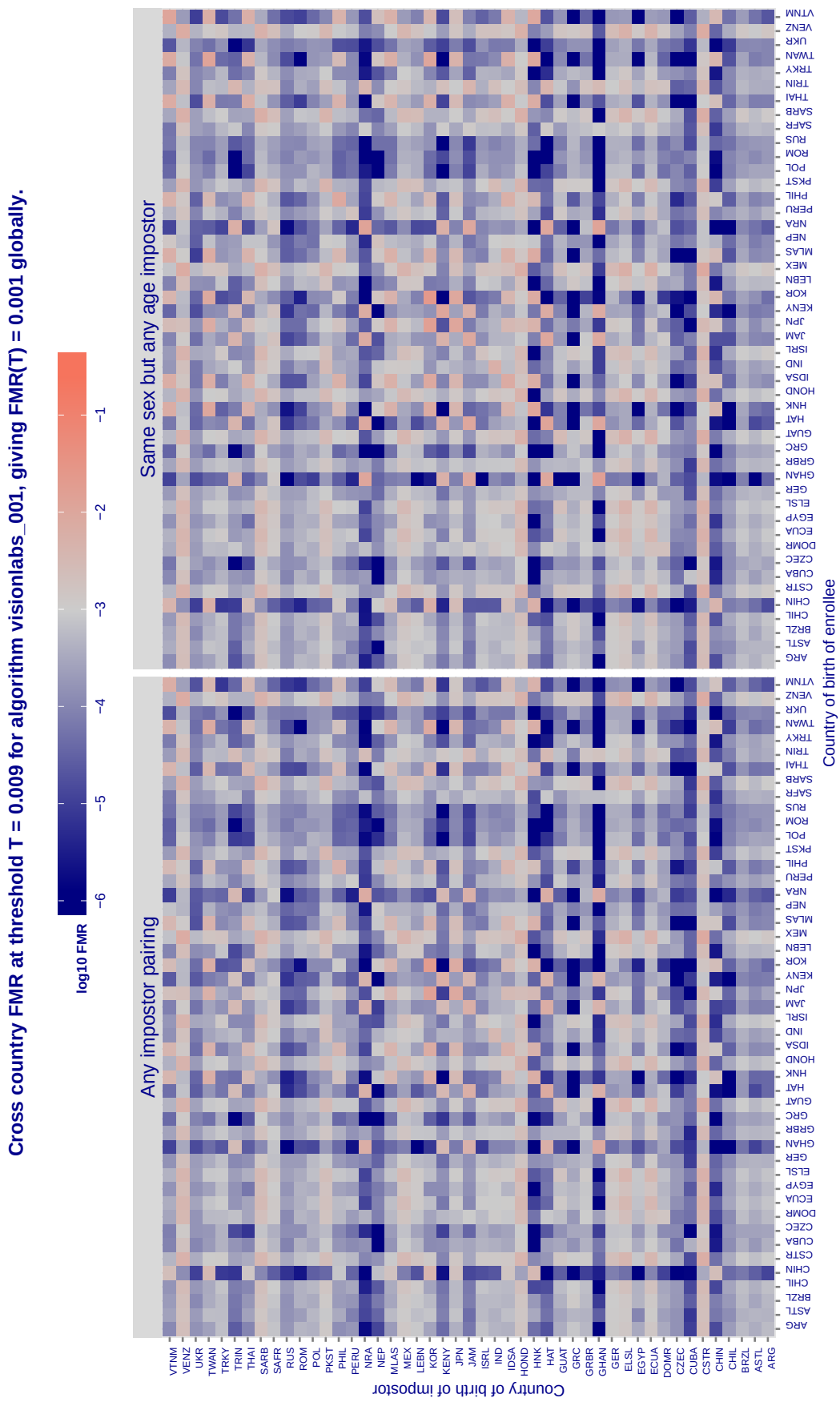


Figure 67: For algorithm visionlabs-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

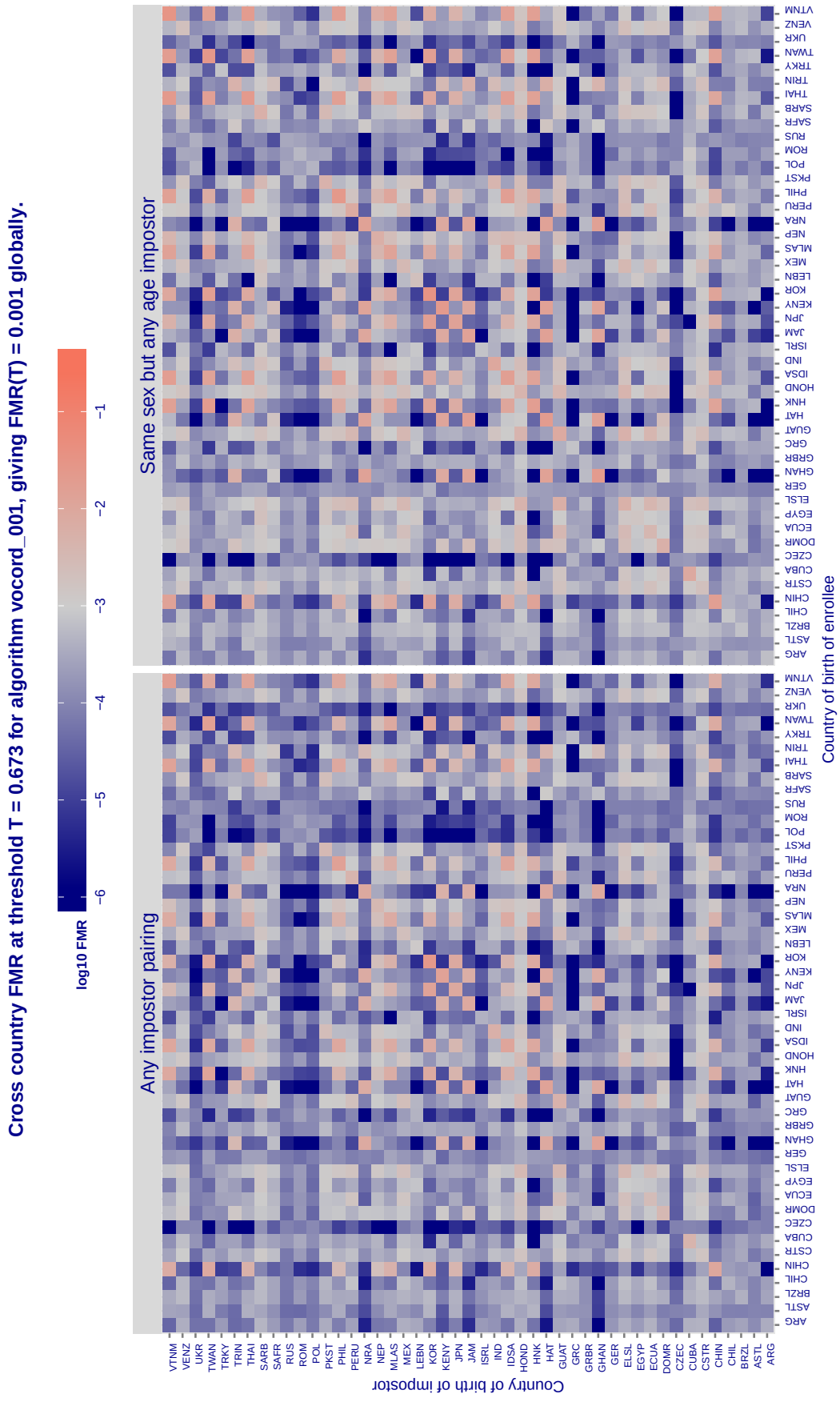


Figure 68: For algorithm vocord-001 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

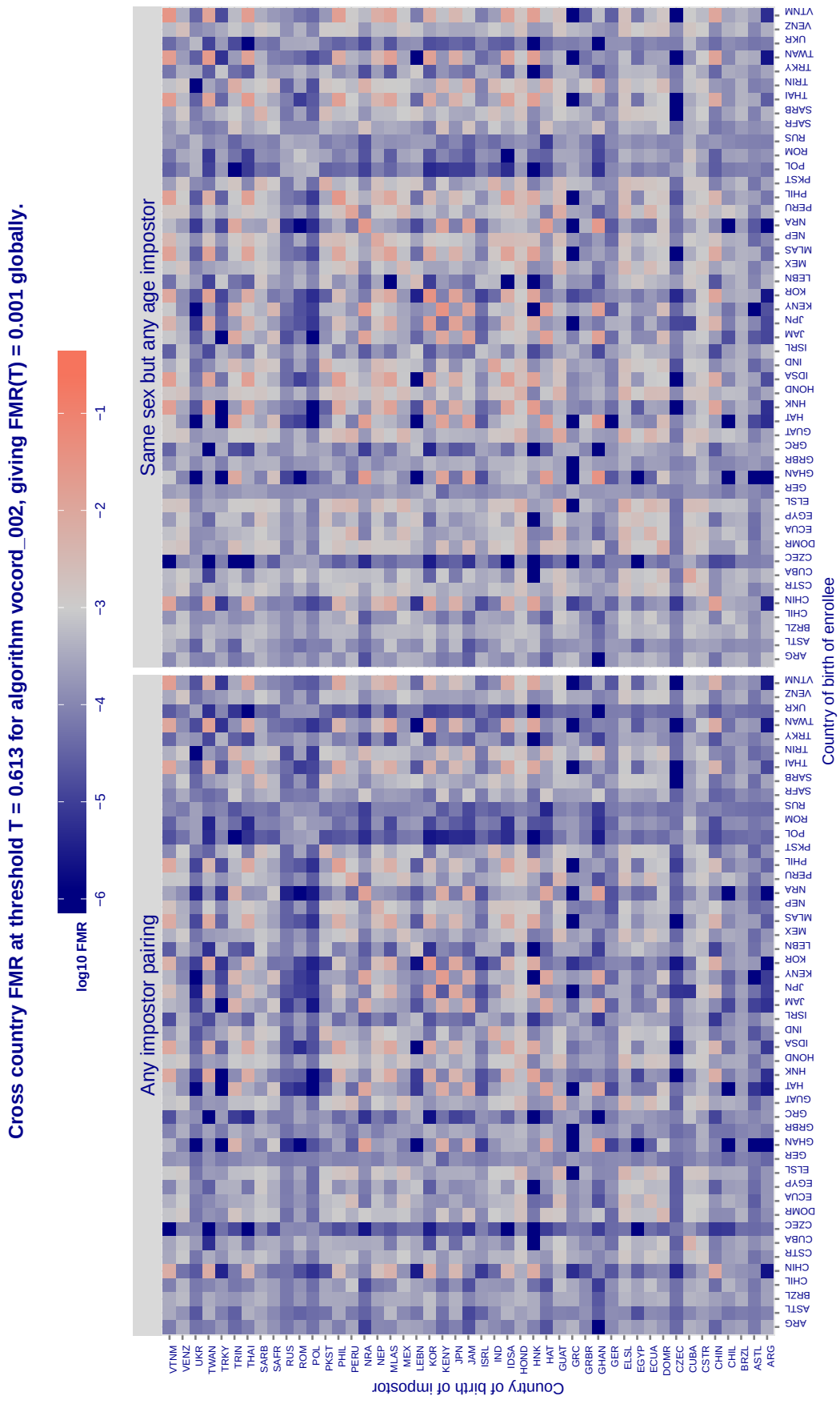


Figure 69: For algorithm vocord-002 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target 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.

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

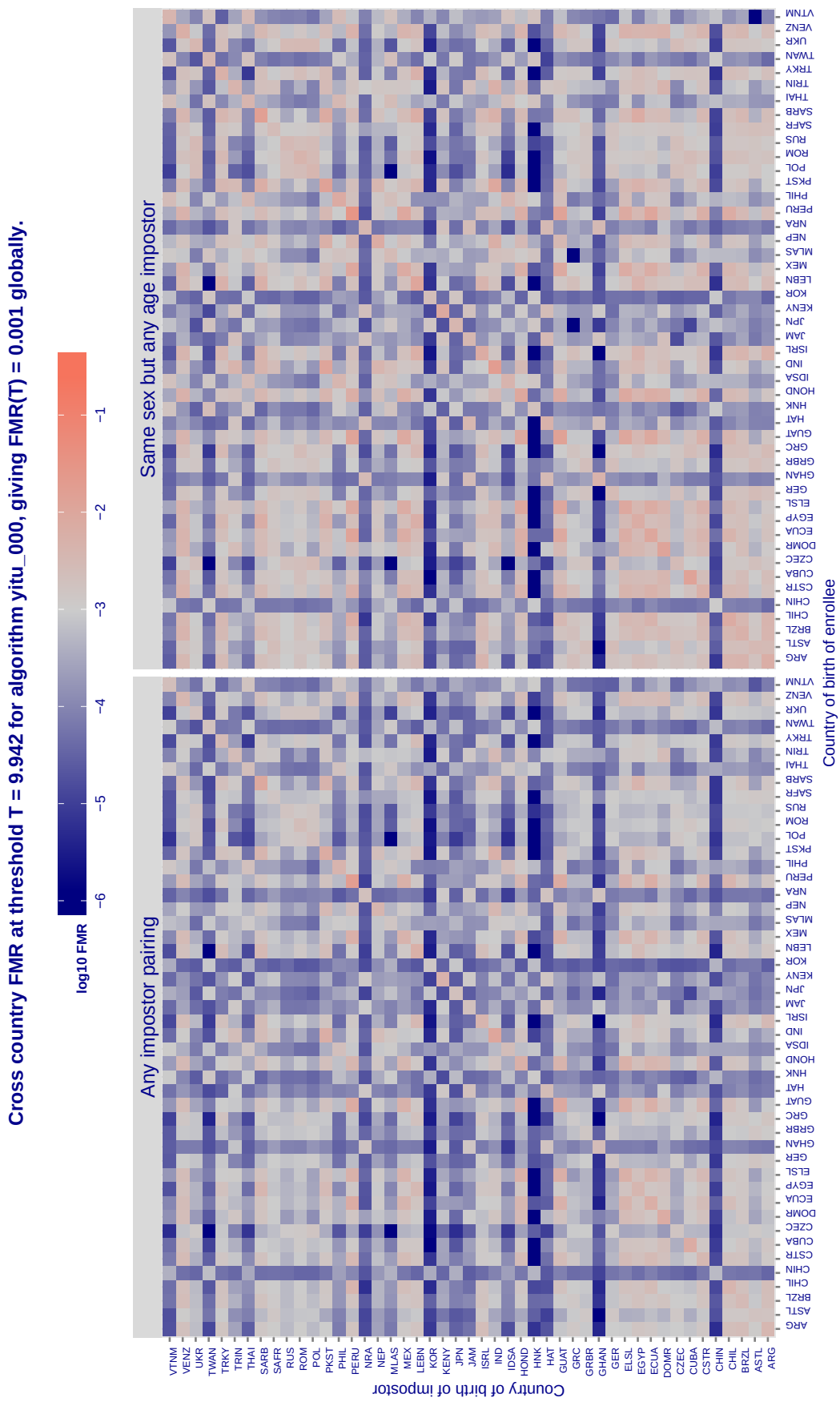


Figure 70: For algorithm yitu-000 operating on visa images, the heatmap shows false match rates observed over impostor comparisons of faces from different individuals who were born in the given country pair. False matches are counted against a recognition threshold fixed globally to give the target FMR 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.

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

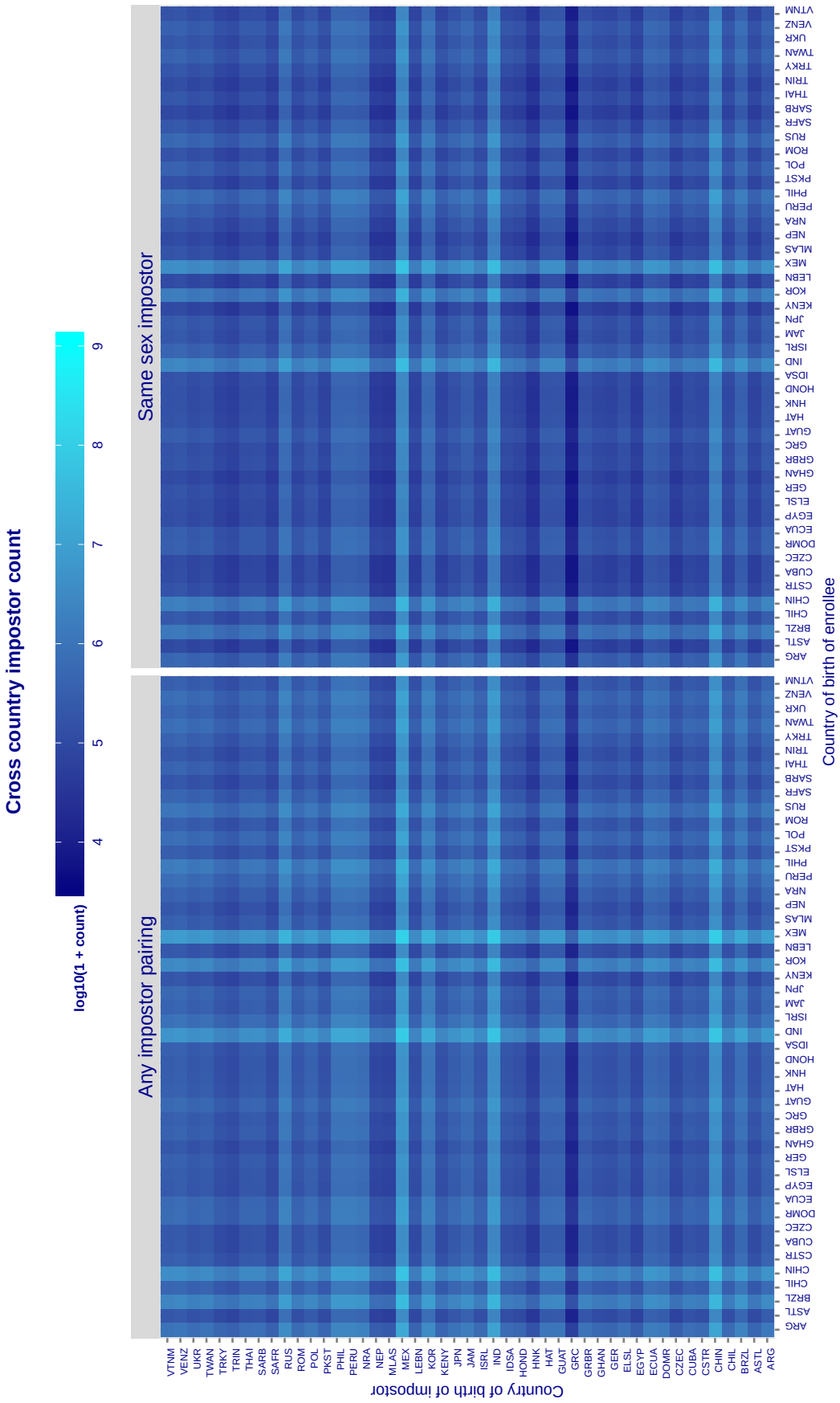


Figure 71: For visa images, the heatmap shows The count of impostor comparisons of faces from different individuals who were born in the given country pair.

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

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 $\Delta \log_{10}$ 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 $\Delta \log_{10}$ 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.

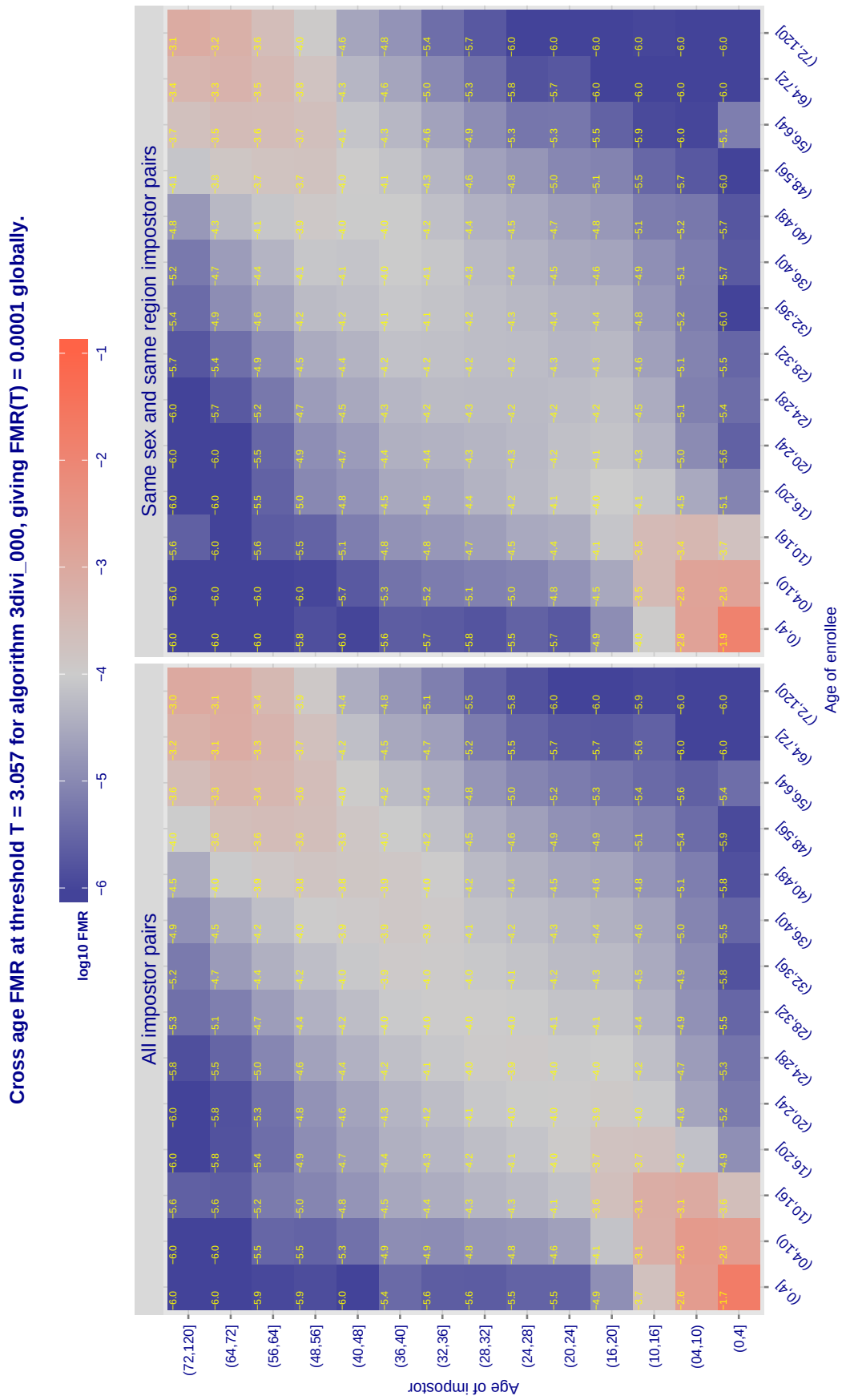


Figure 72: For algorithm 3divi-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

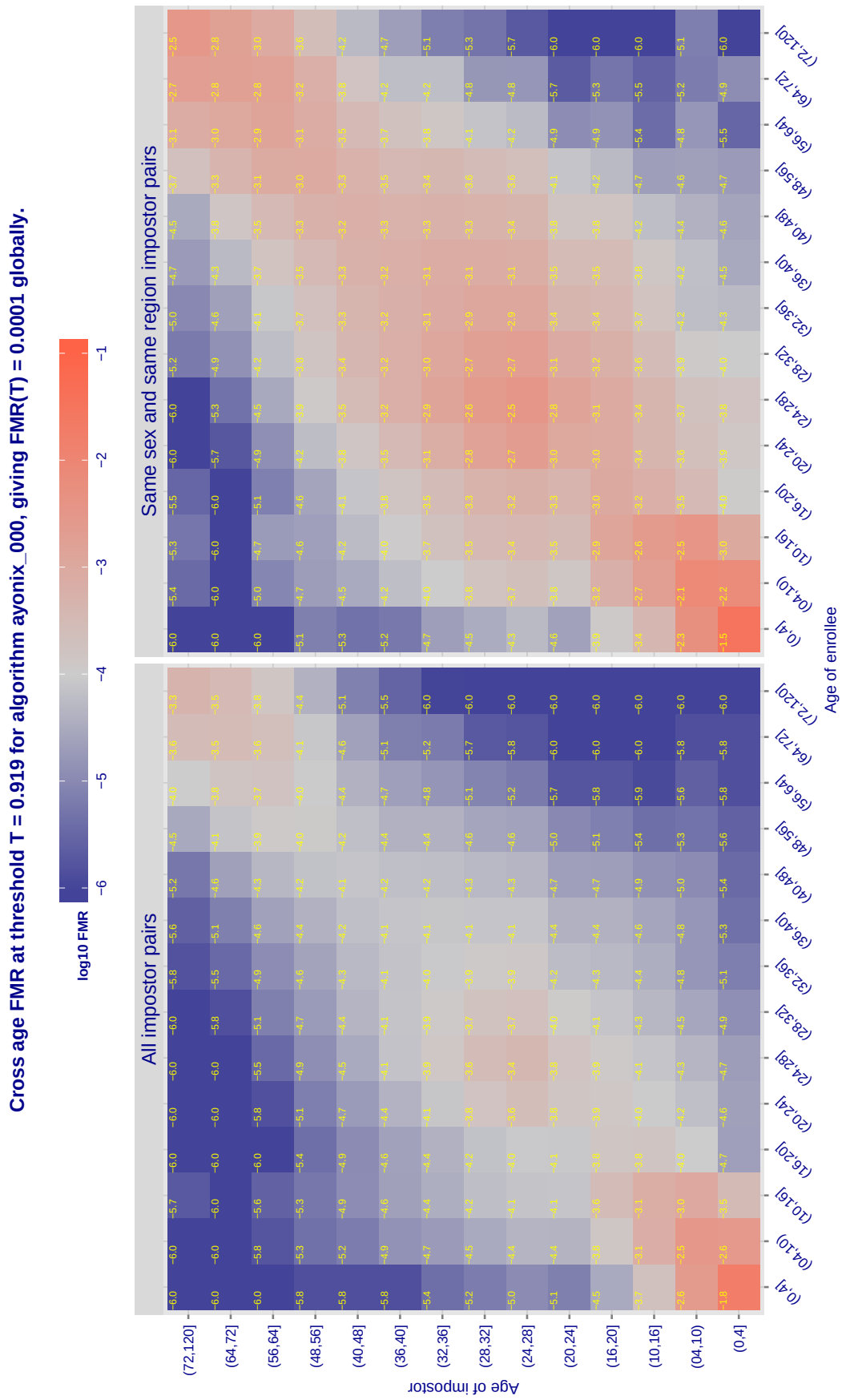


Figure 73: For algorithm ayonix-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

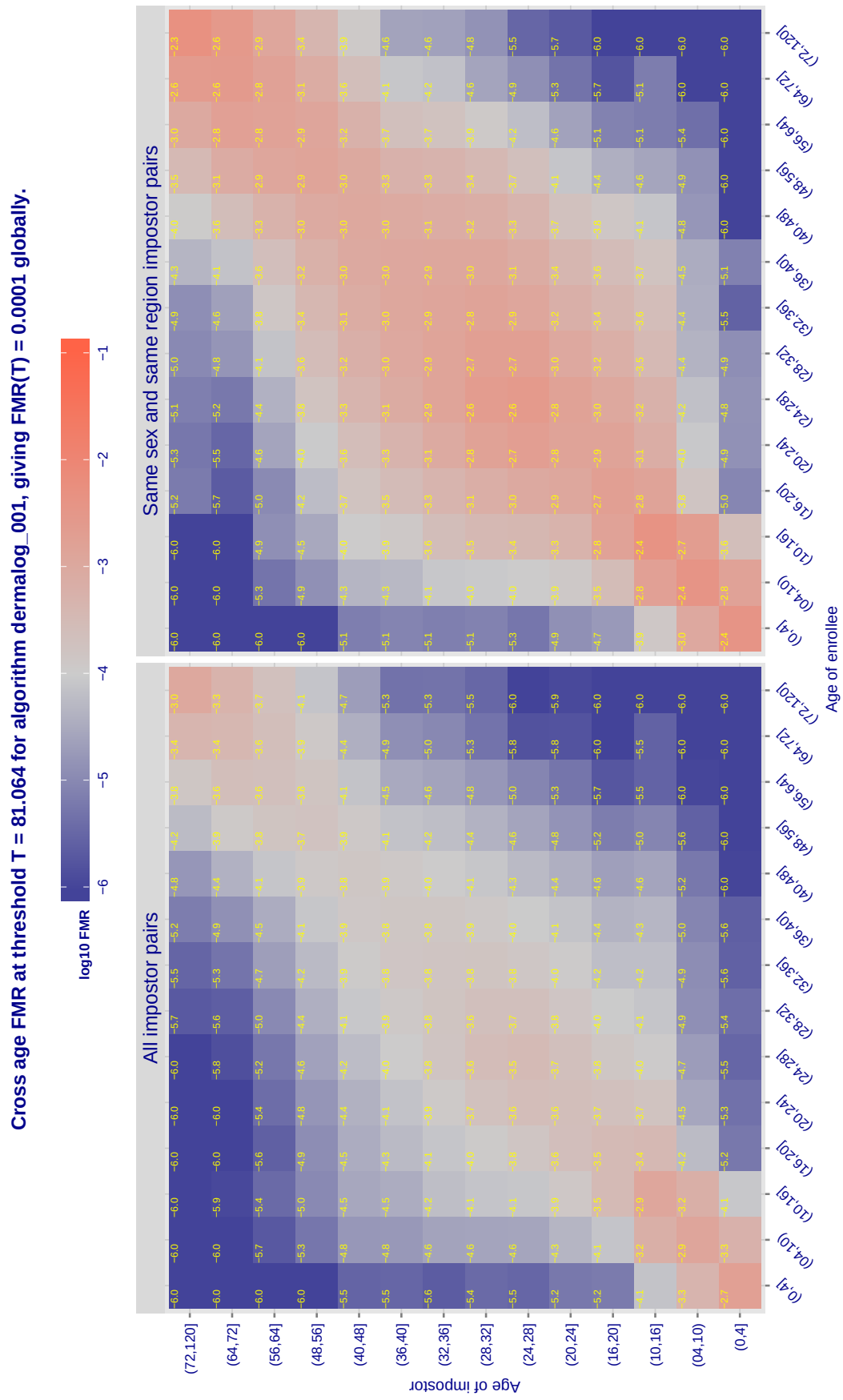


Figure 74: For algorithm dermalog-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

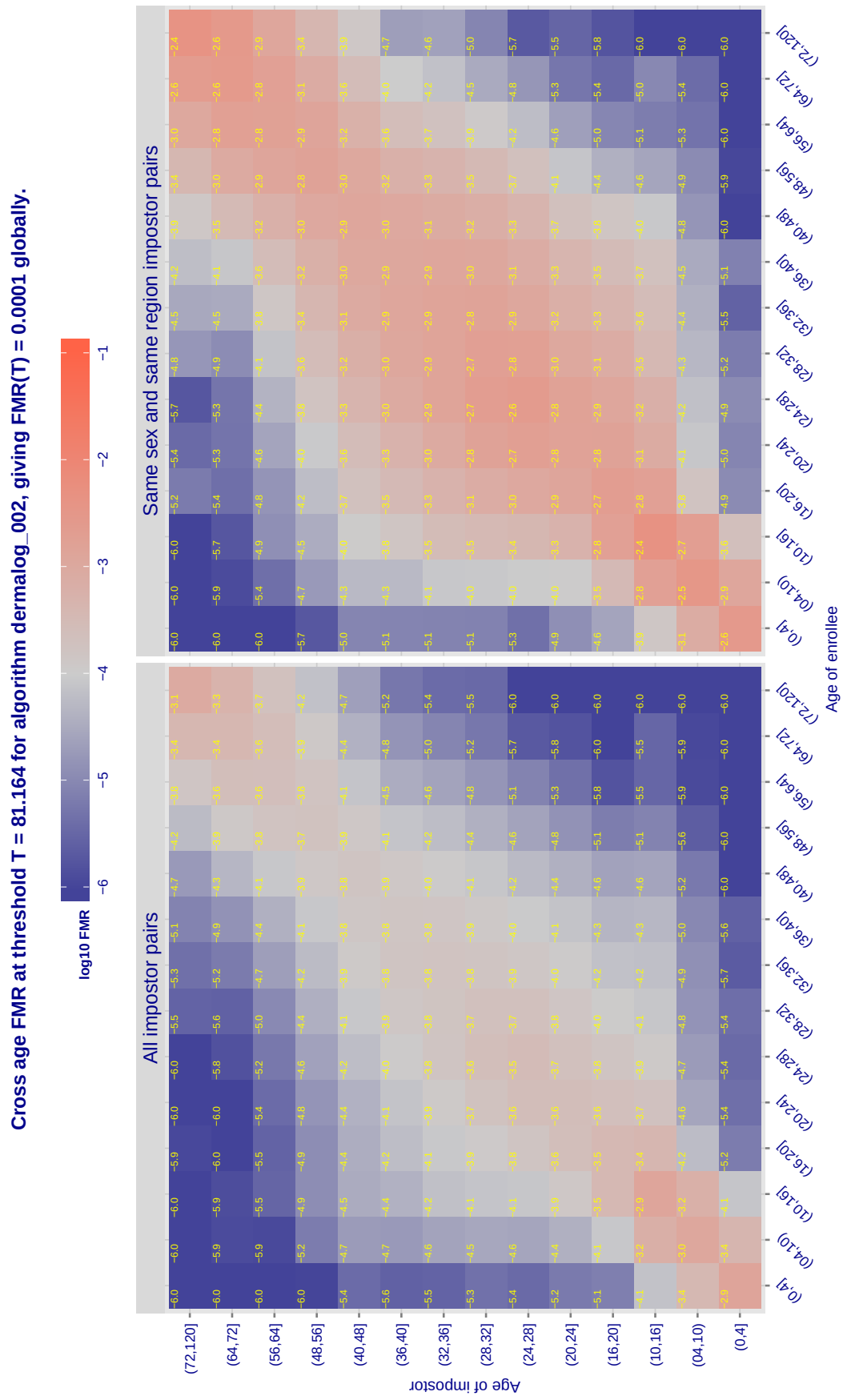


Figure 75: For algorithm dermalog-002 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

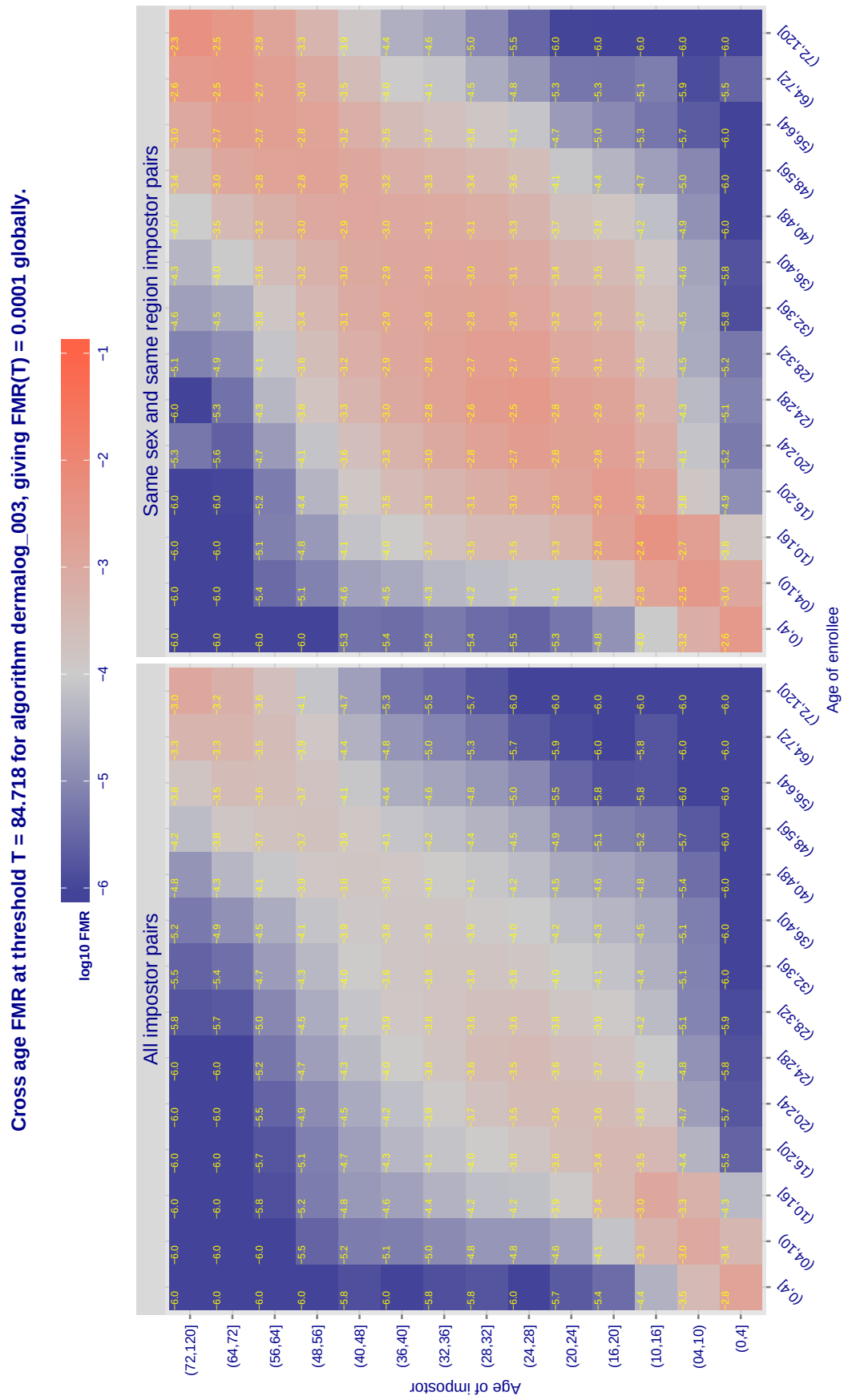


Figure 76: For algorithm dermalog-003 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

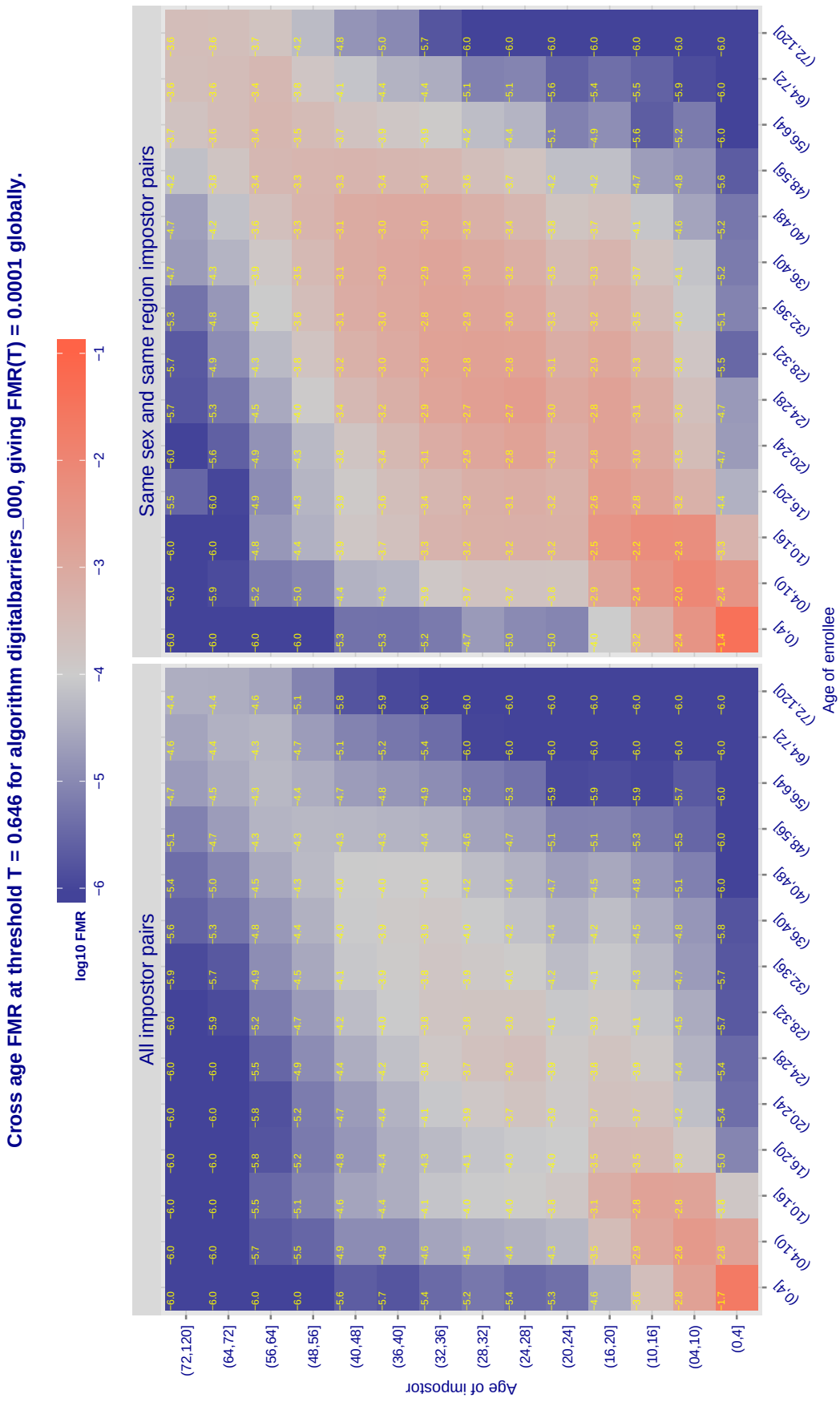


Figure 77: For algorithm digitalbarriers-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

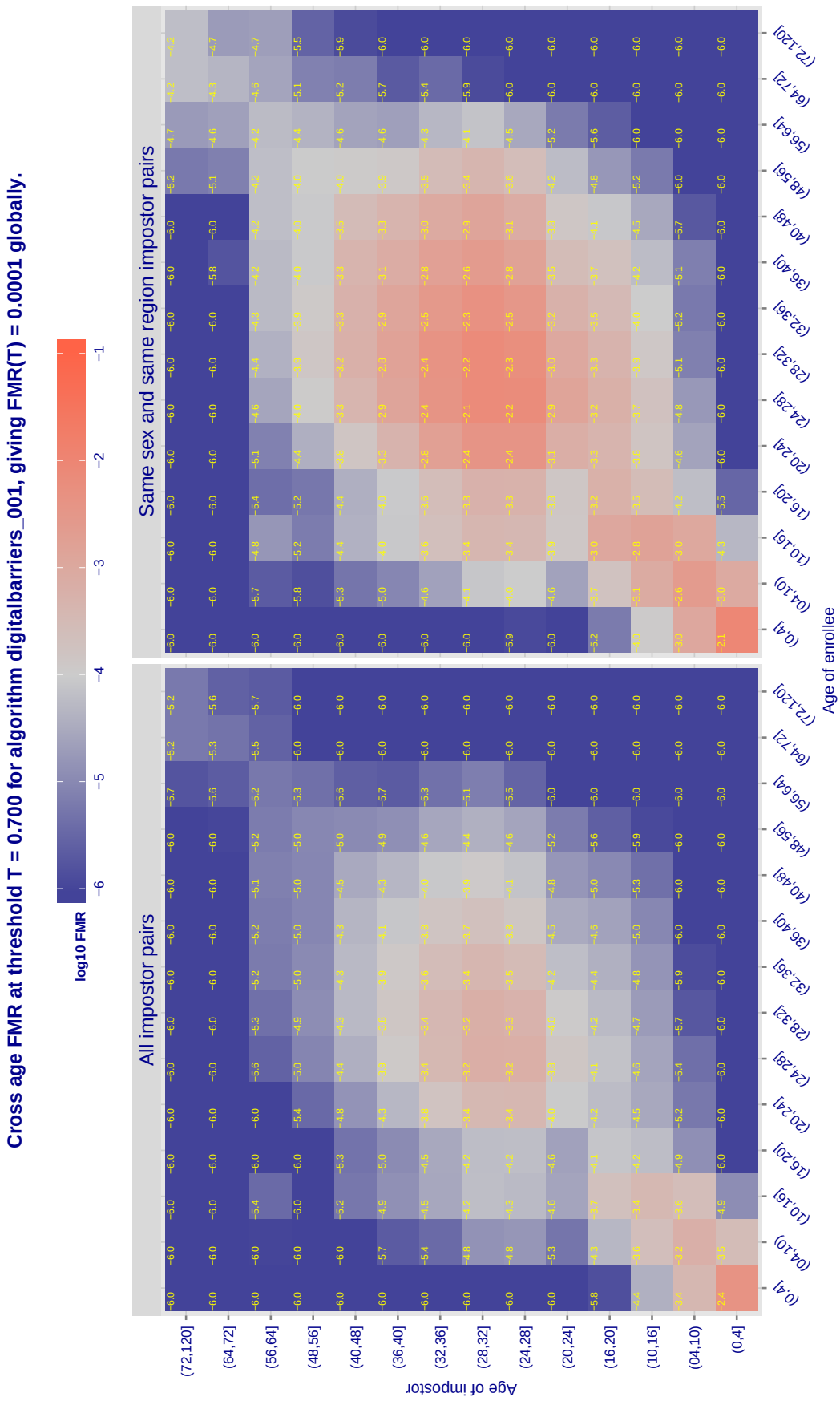


Figure 78: For algorithm digitalbarriers-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

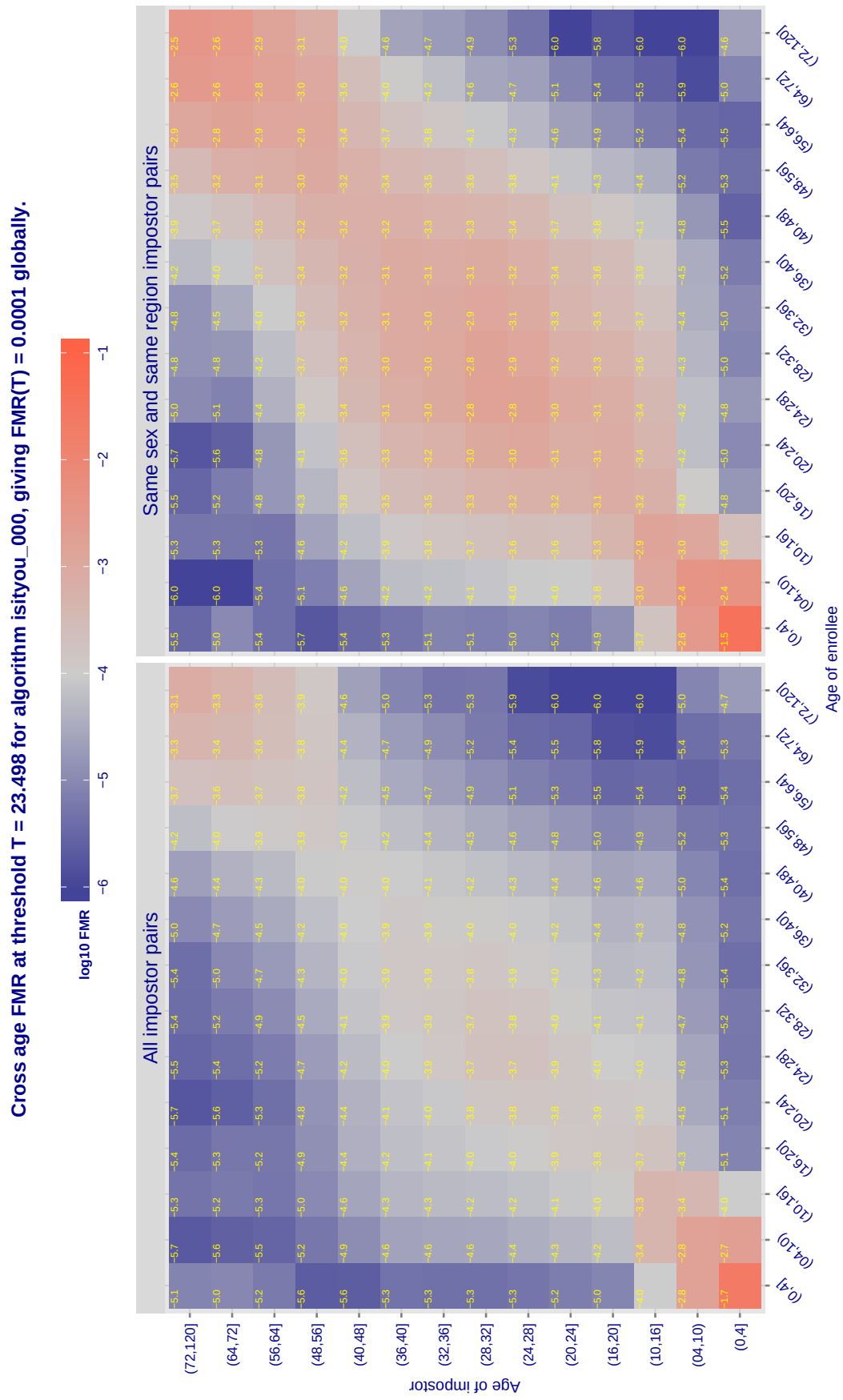


Figure 79: For algorithm isityou-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

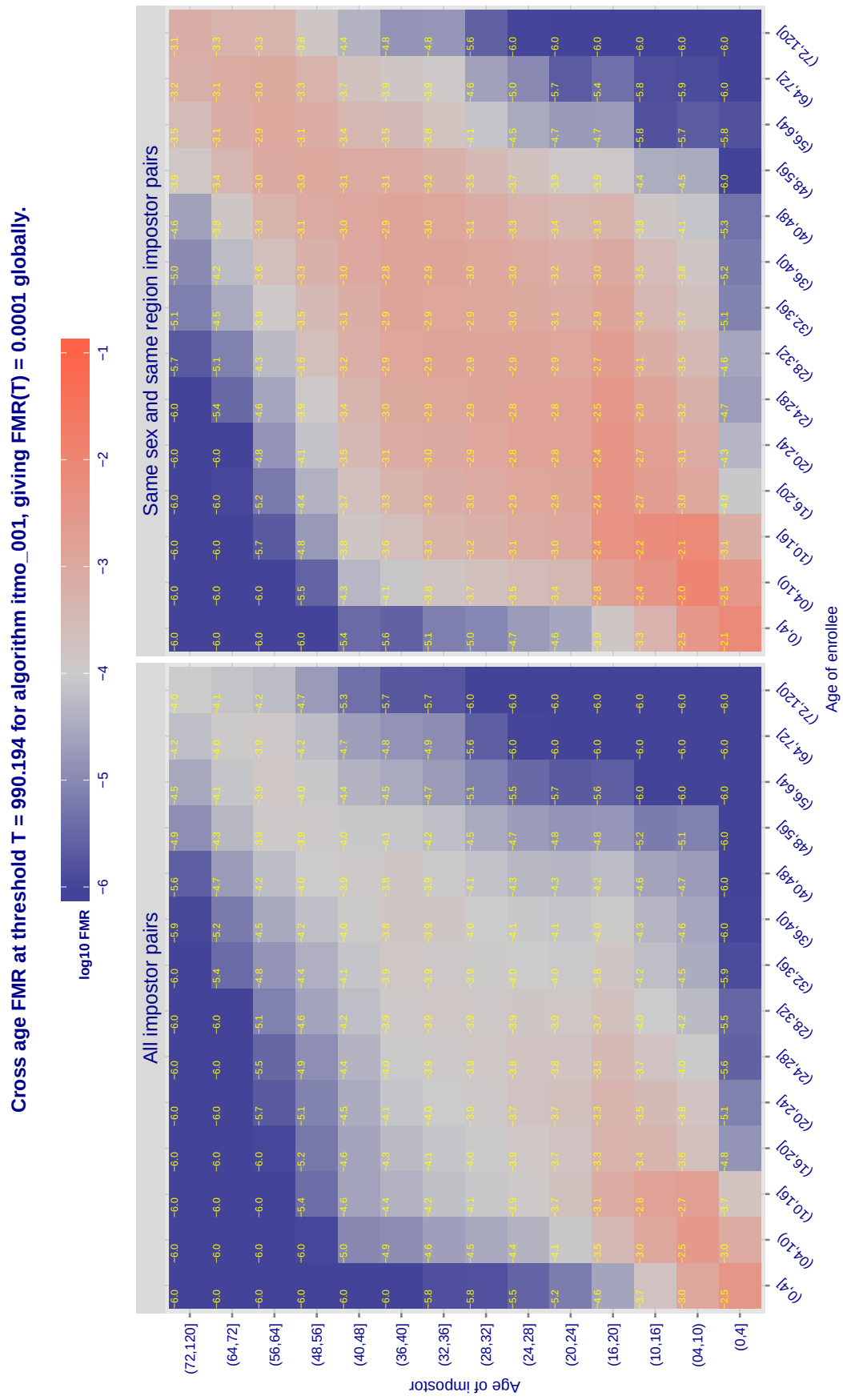


Figure 80: For algorithm itmo-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

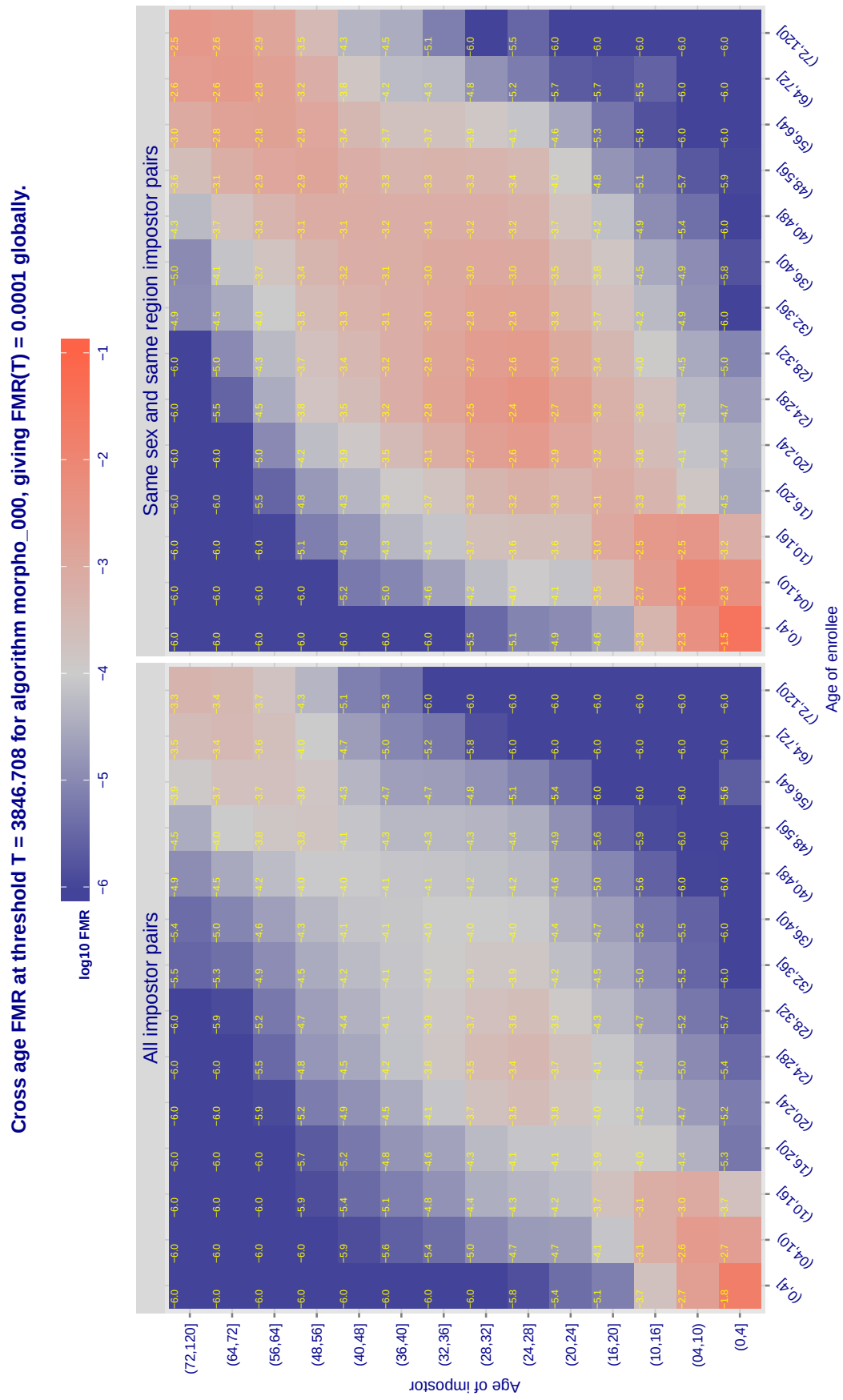


Figure 81: For algorithm morpho-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

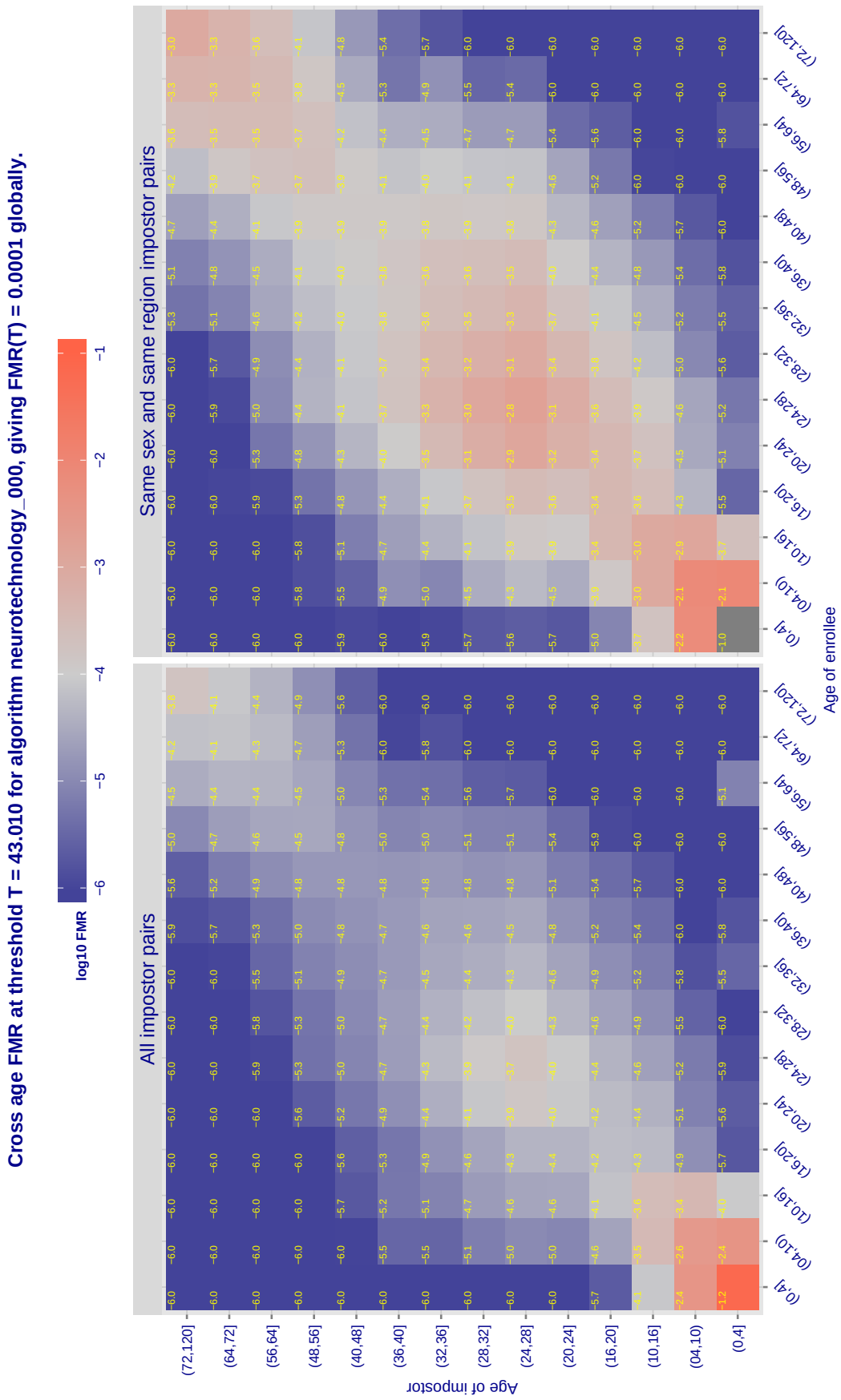


Figure 82: For algorithm neurotechnology-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

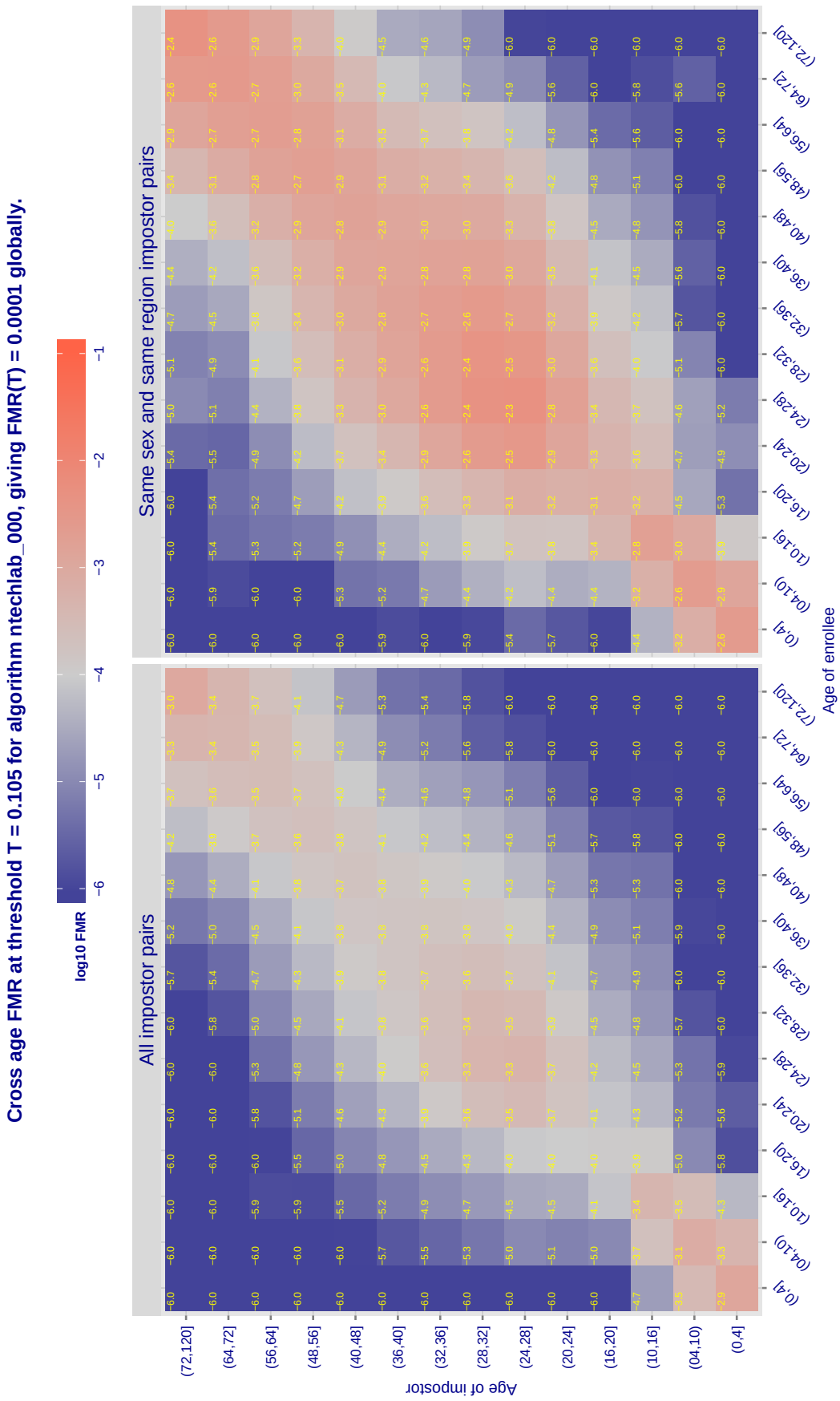


Figure 83: For algorithm ntechlab-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

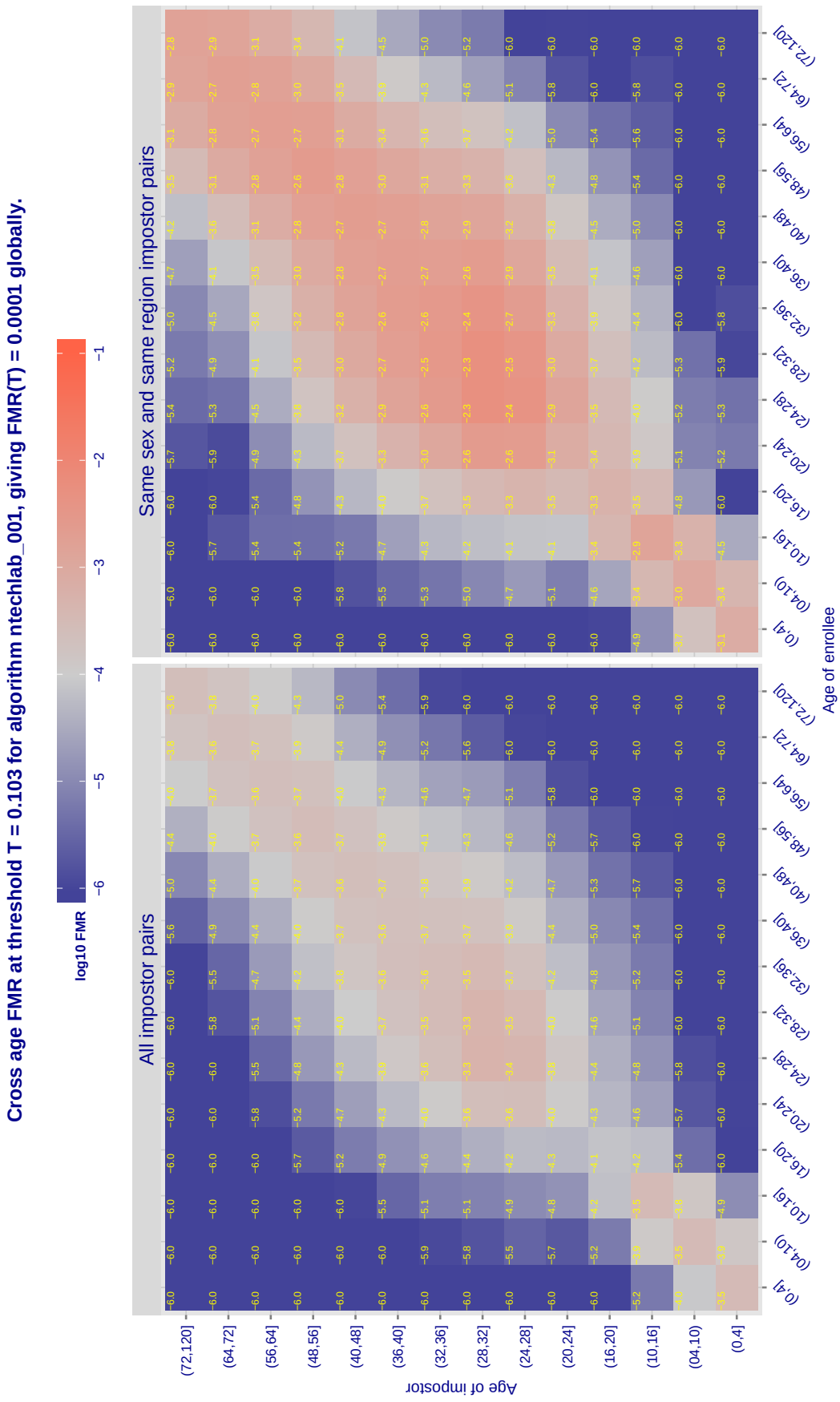


Figure 84: For algorithm ntechlab-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

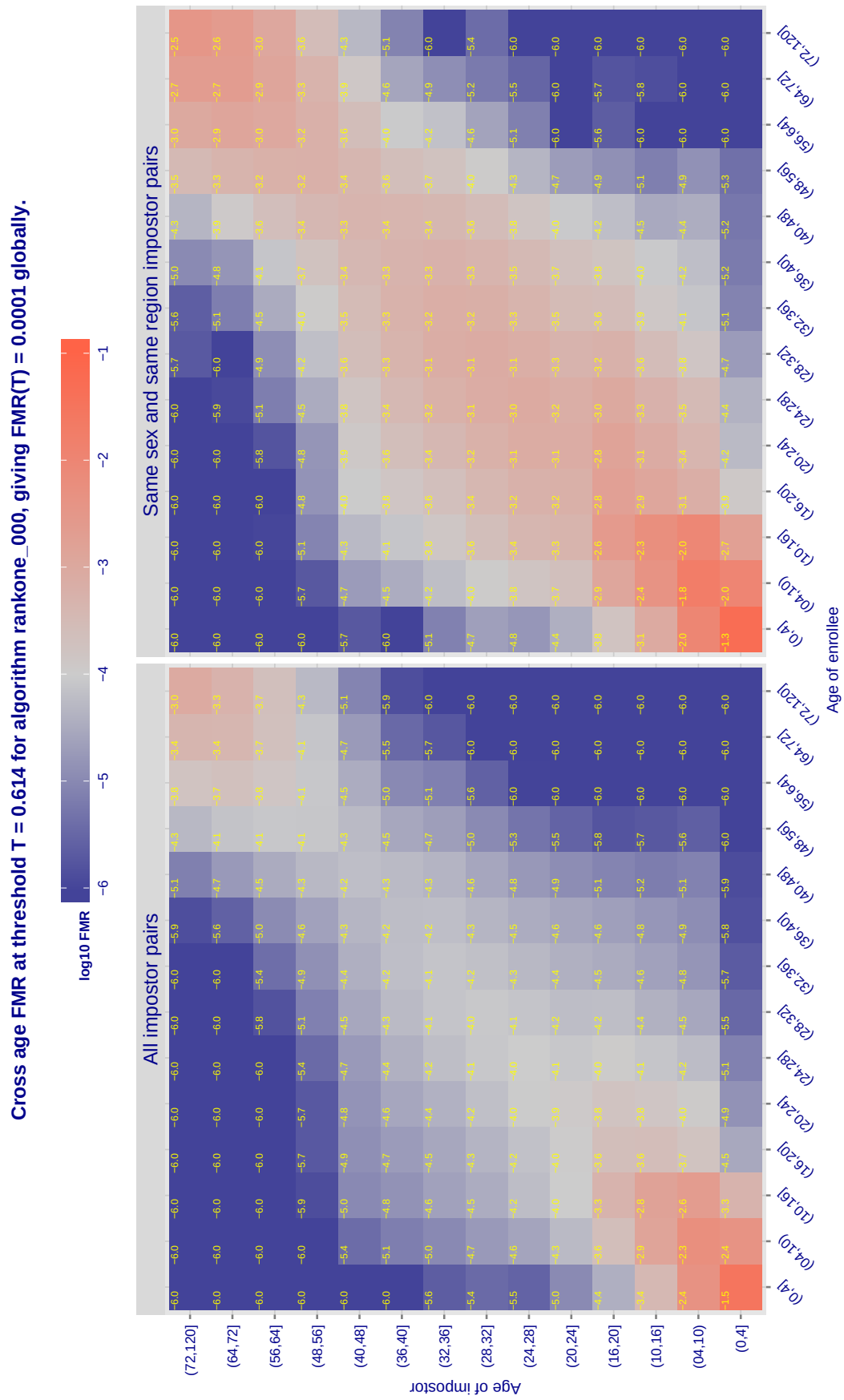


Figure 85: For algorithm rankone-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

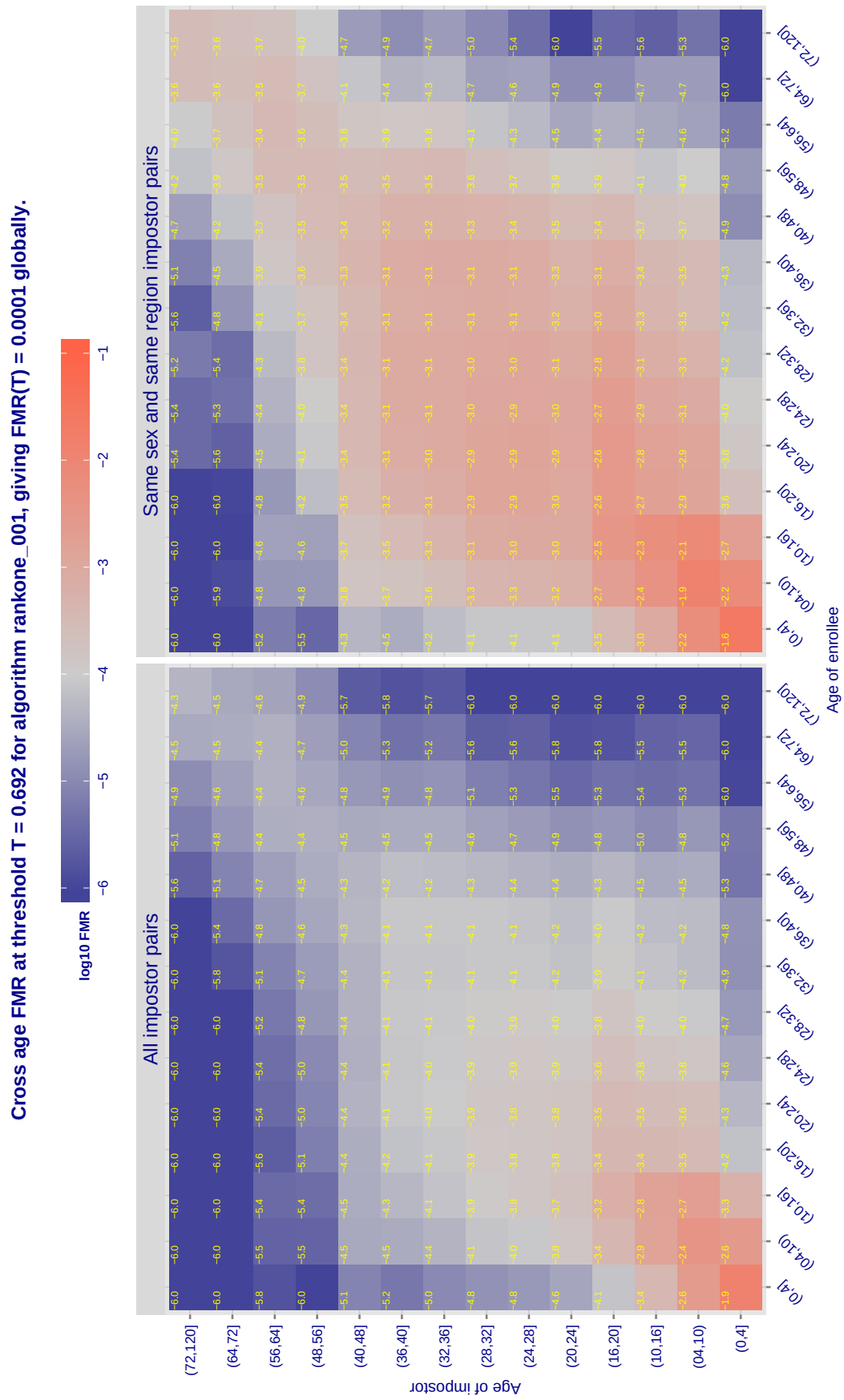


Figure 86: For algorithm rankone-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

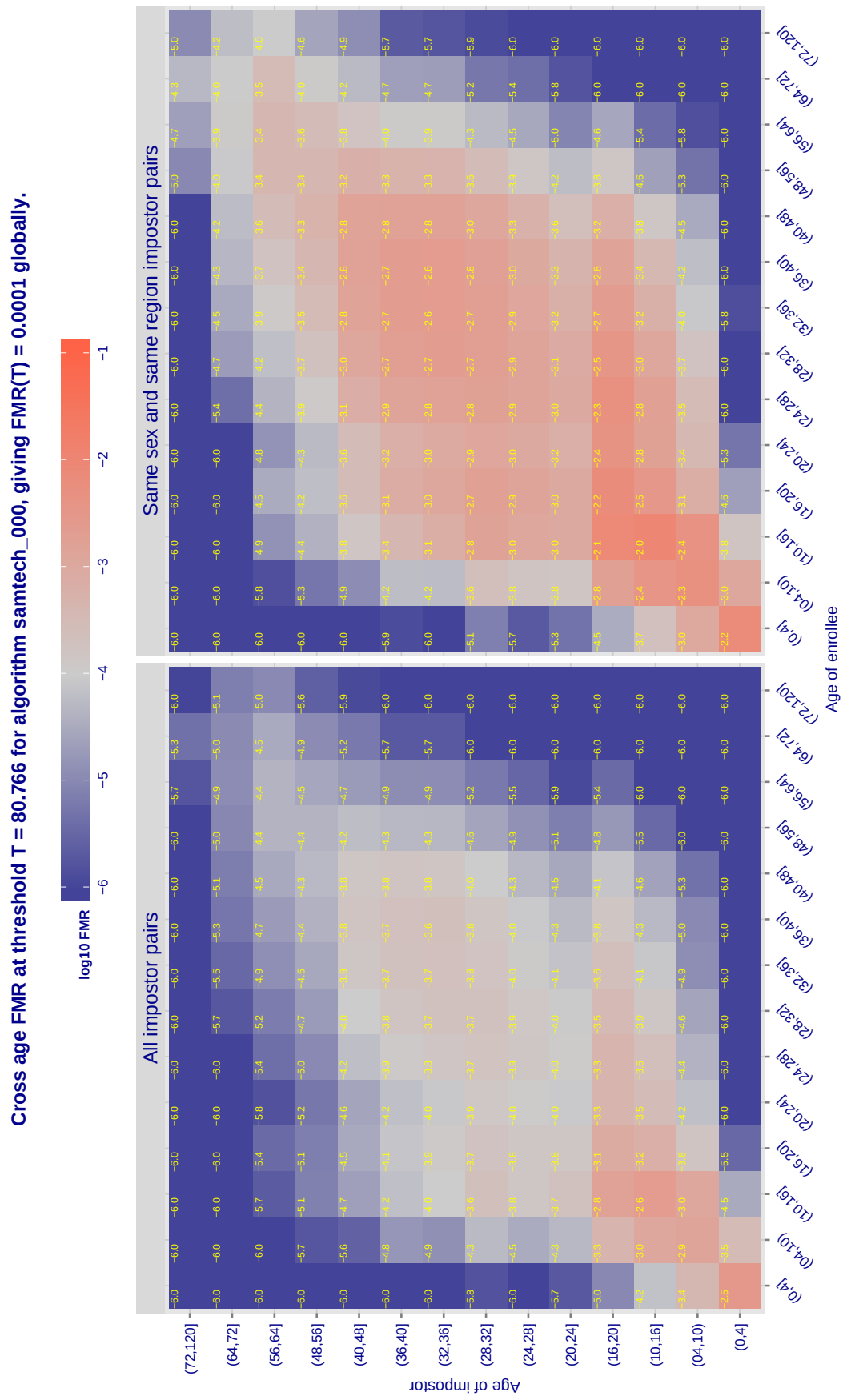


Figure 87: For algorithm samtech-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

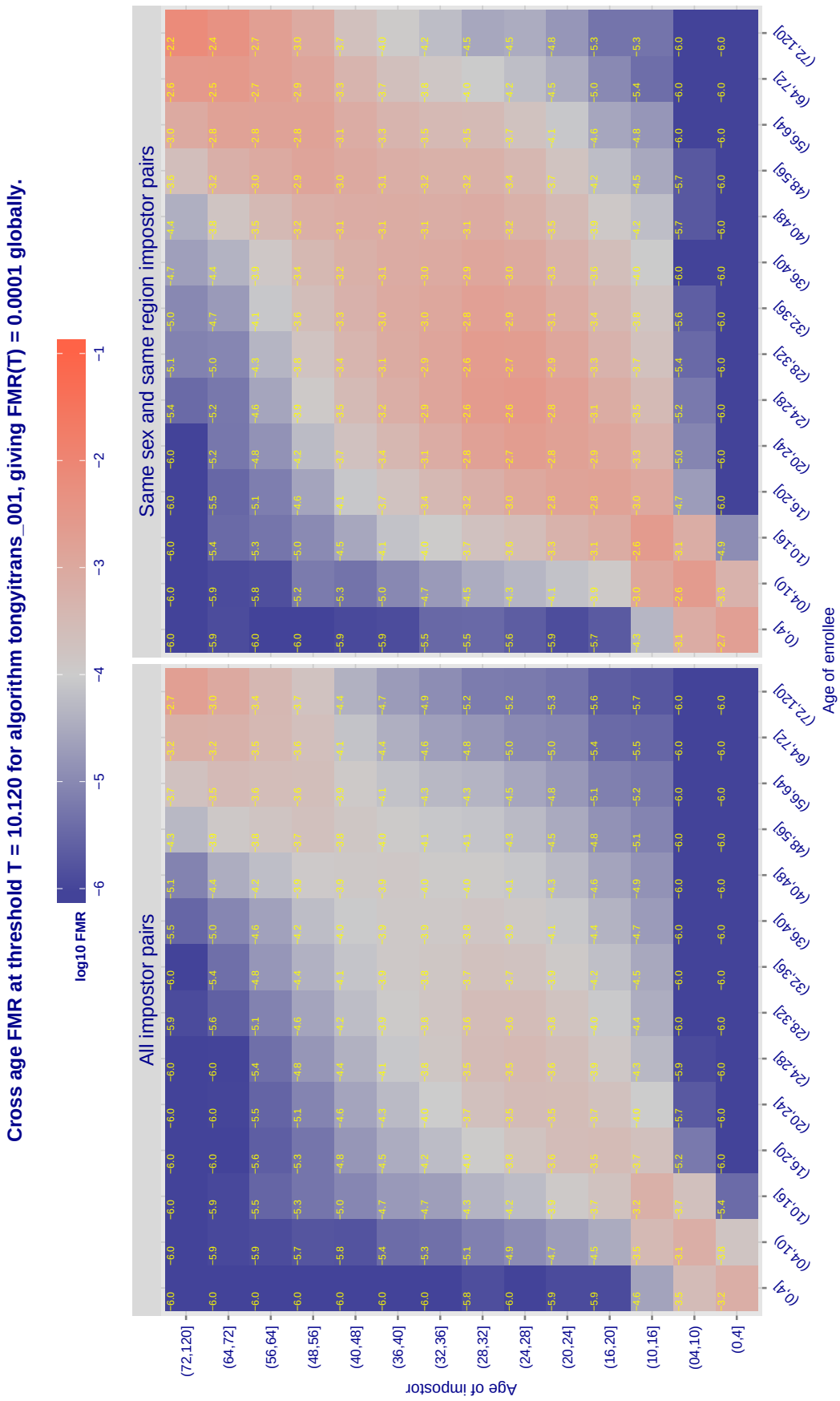


Figure 88: For algorithm tongyitrans-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

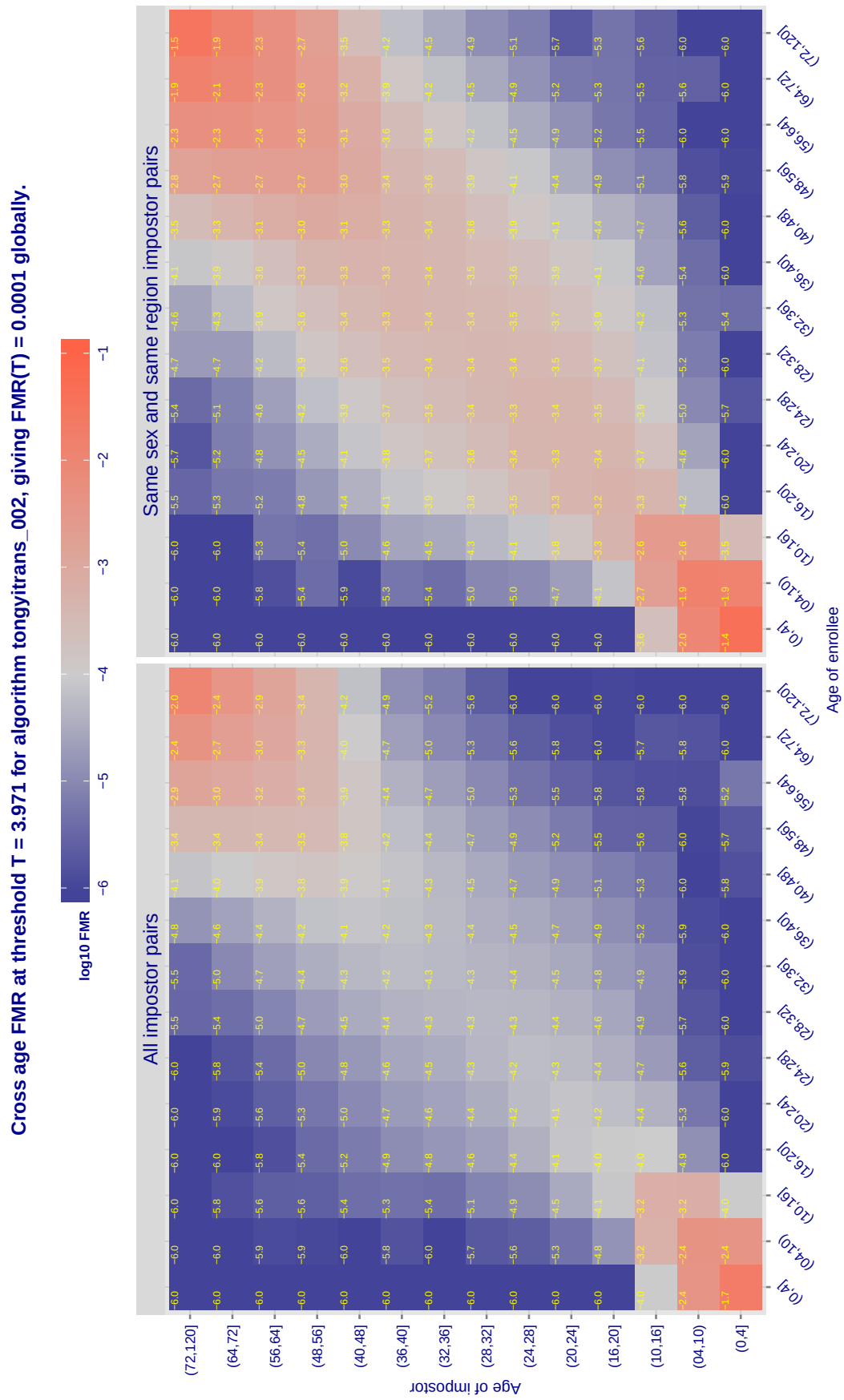


Figure 89: For algorithm tongyitrans-002 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

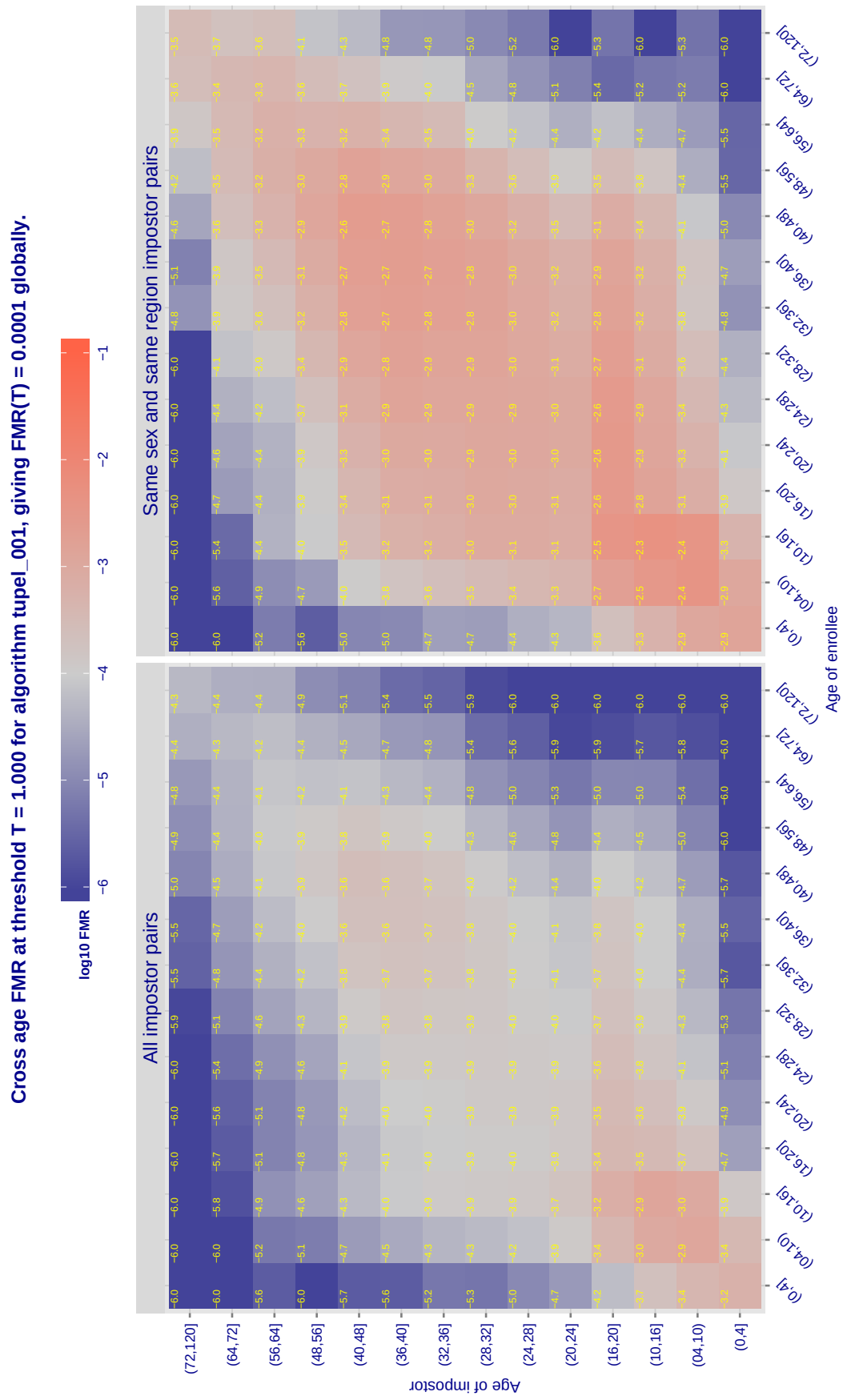


Figure 90: For algorithm tupel-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

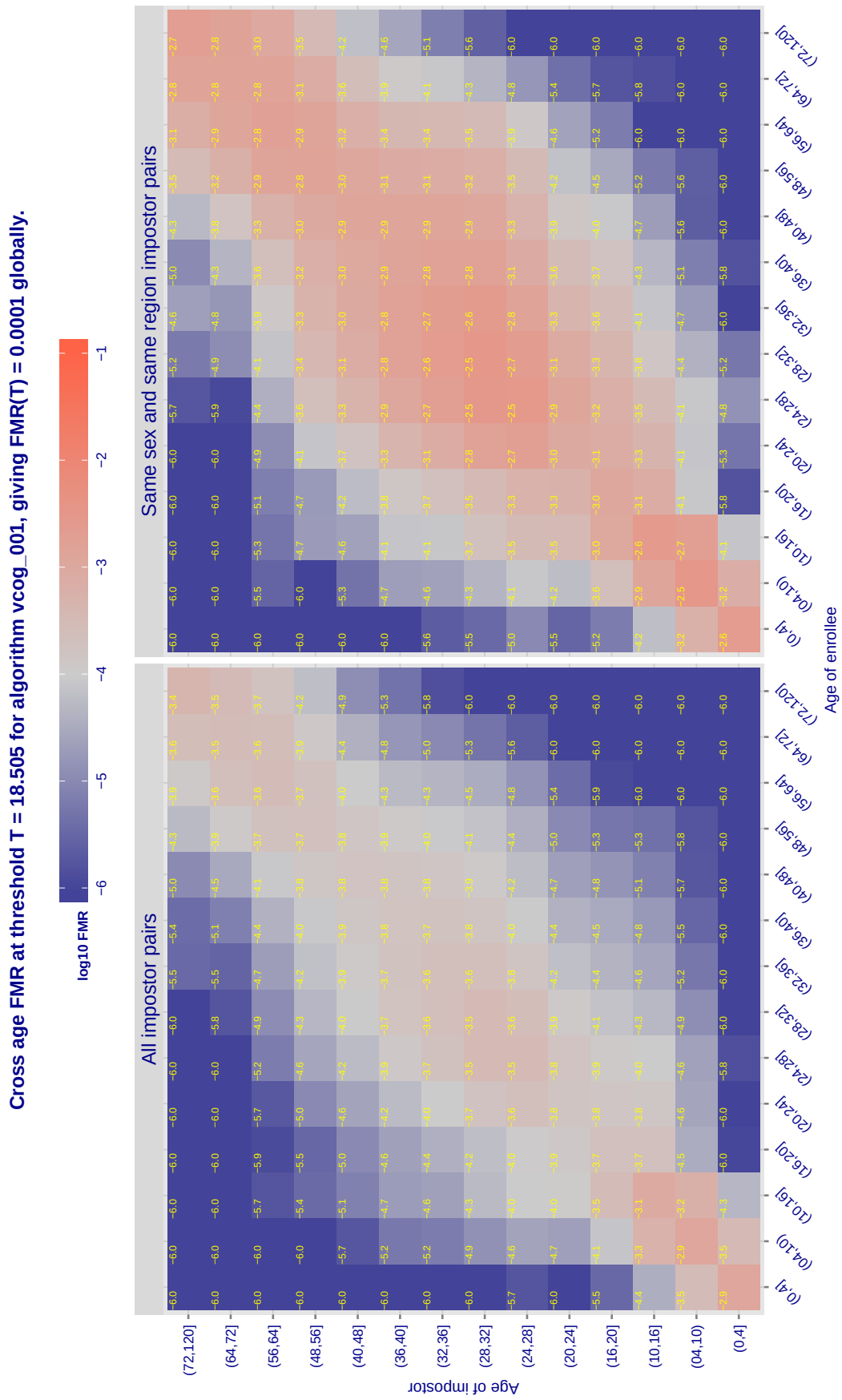


Figure 91: For algorithm vcog-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

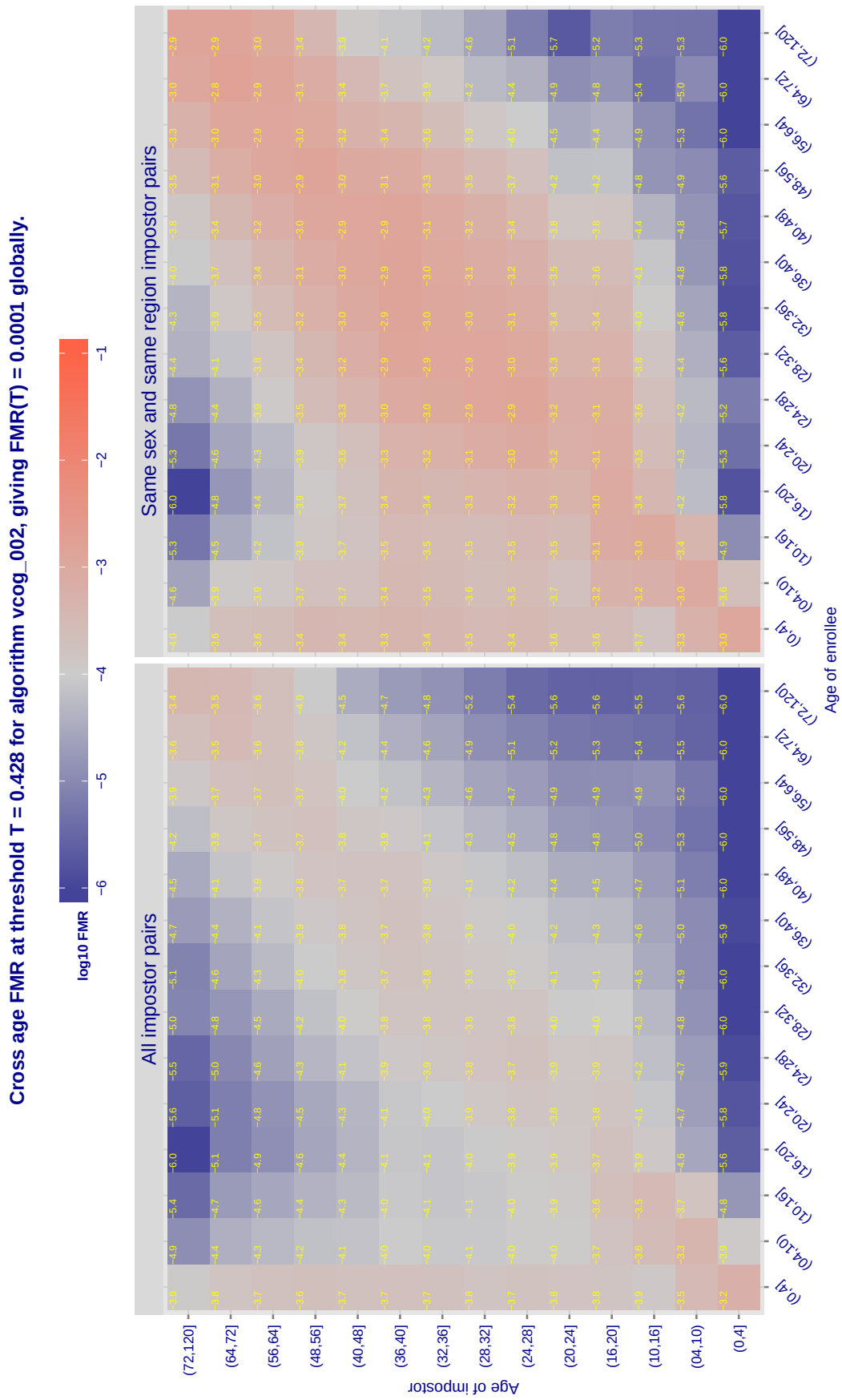


Figure 92: For algorithm vcog-002 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

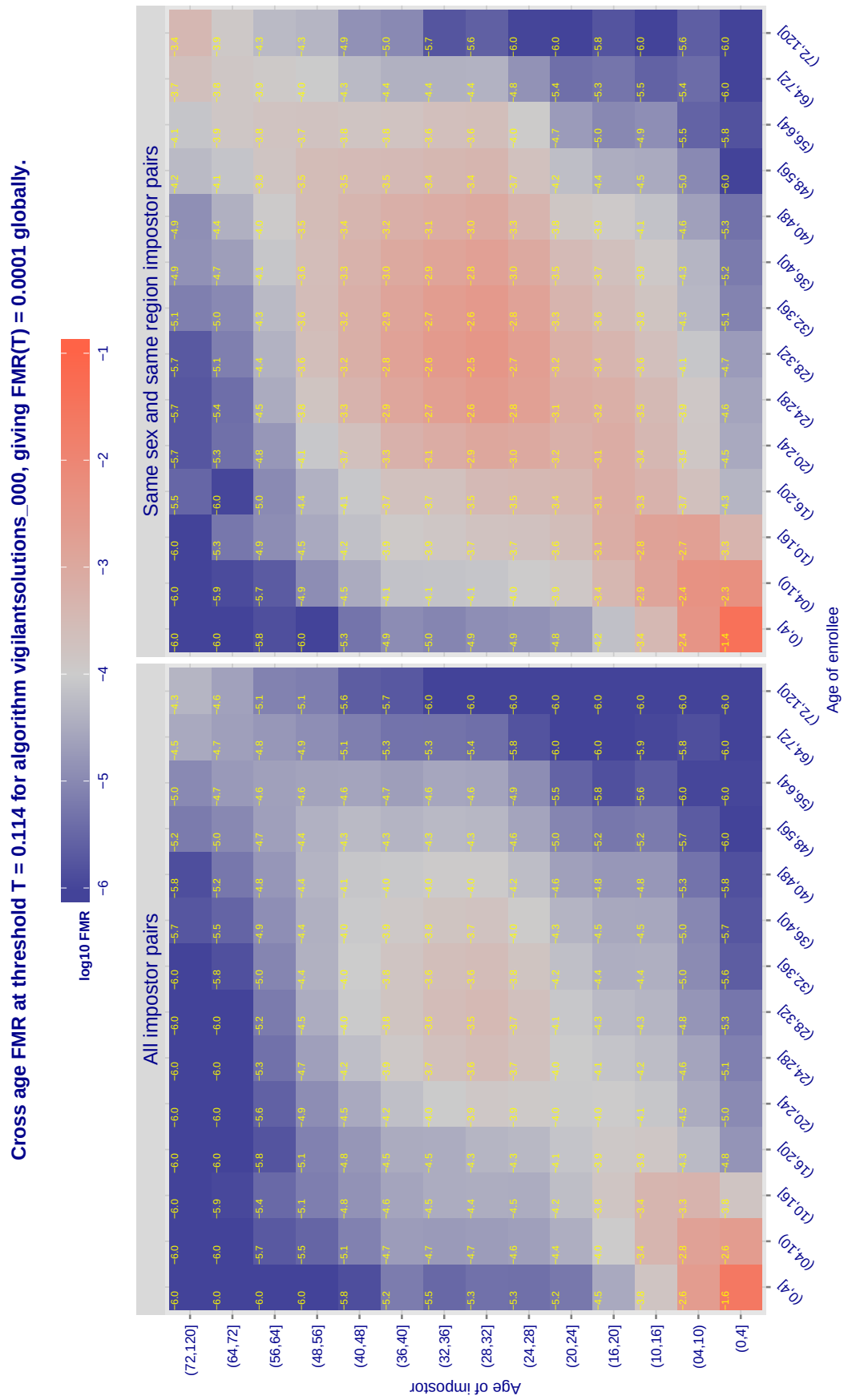


Figure 93: For algorithm vigilantsolutions-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over 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.

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

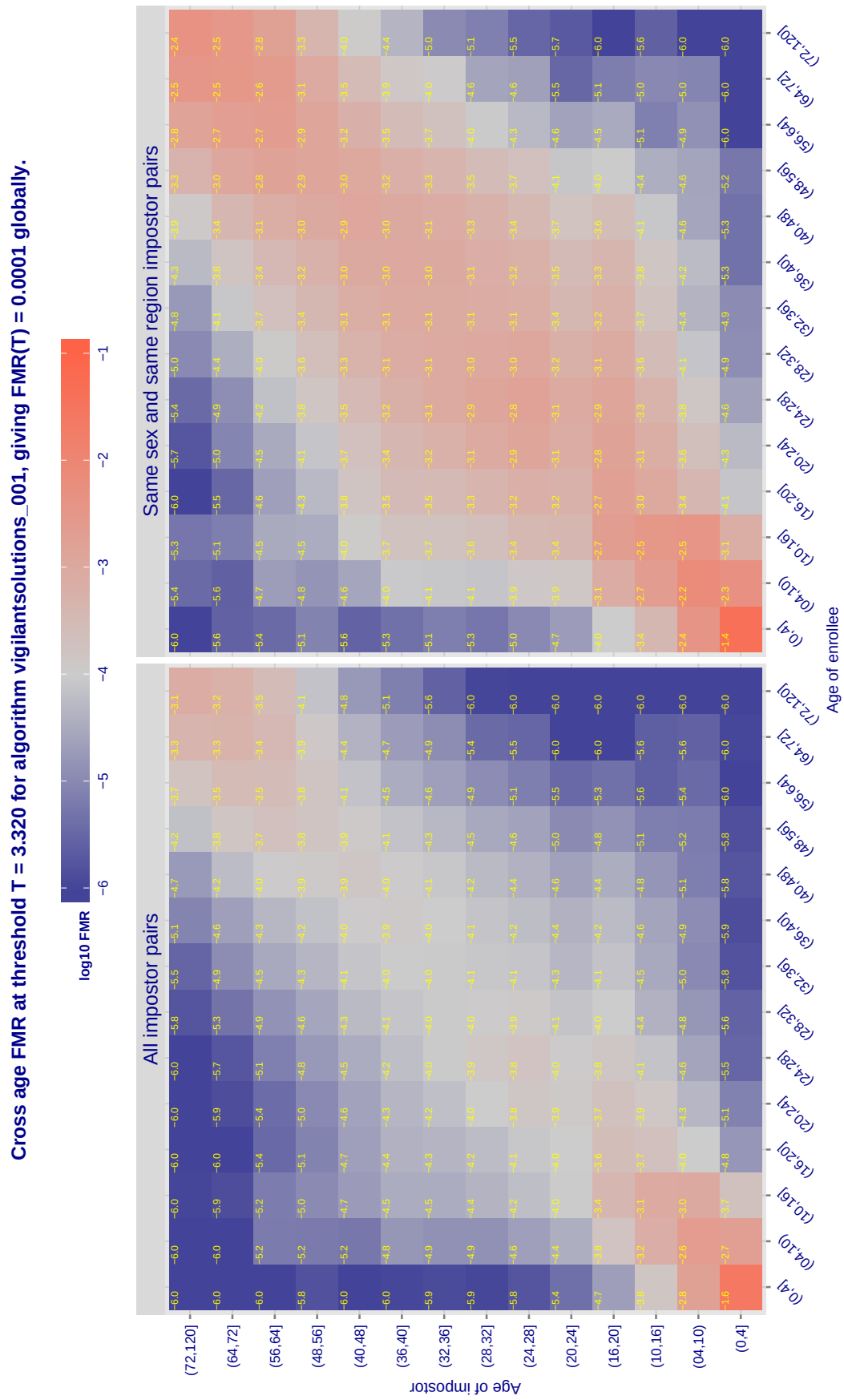


Figure 94: For algorithm vigilantsolutions-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over 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.

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

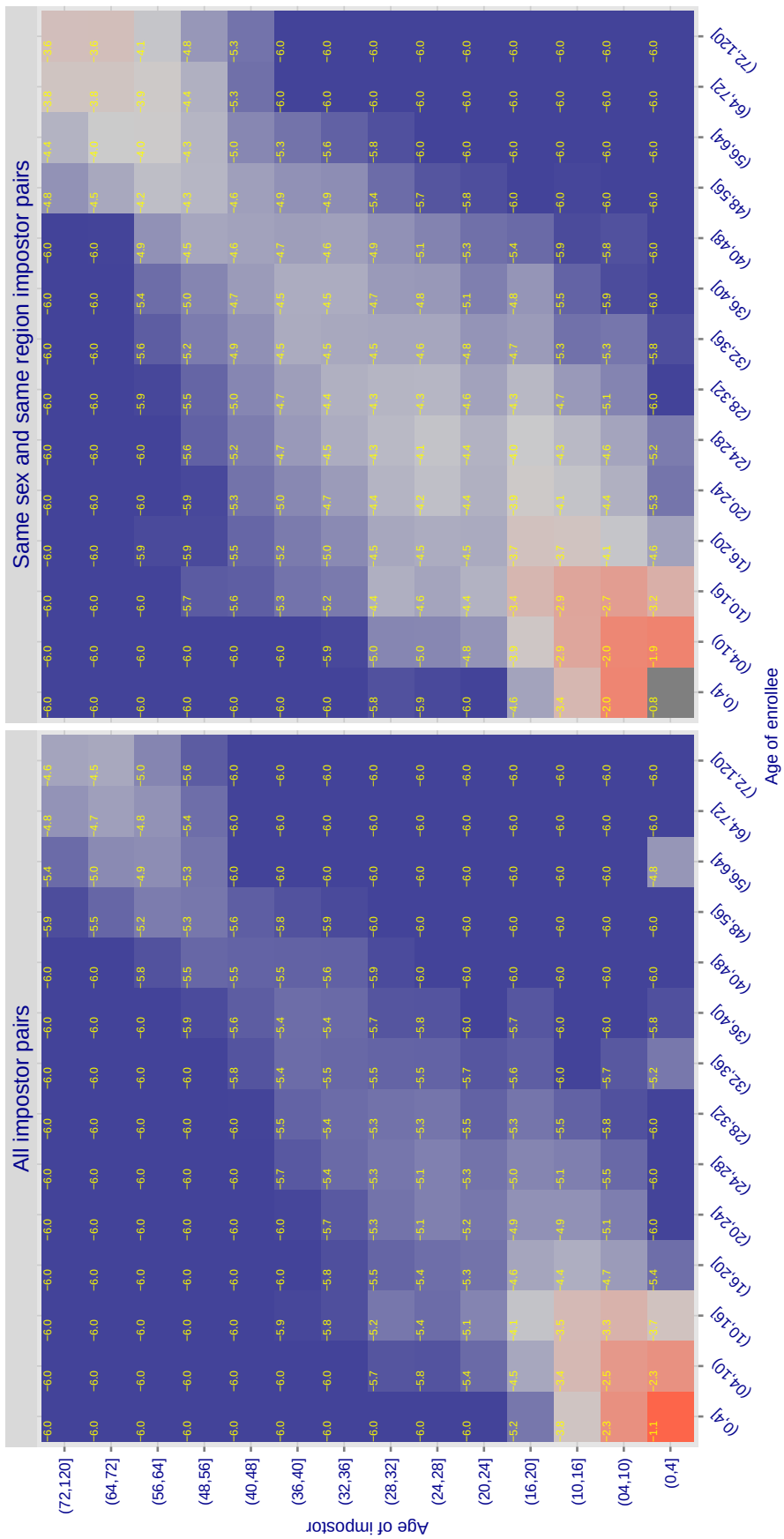


Figure 95: For algorithm visionlabs-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

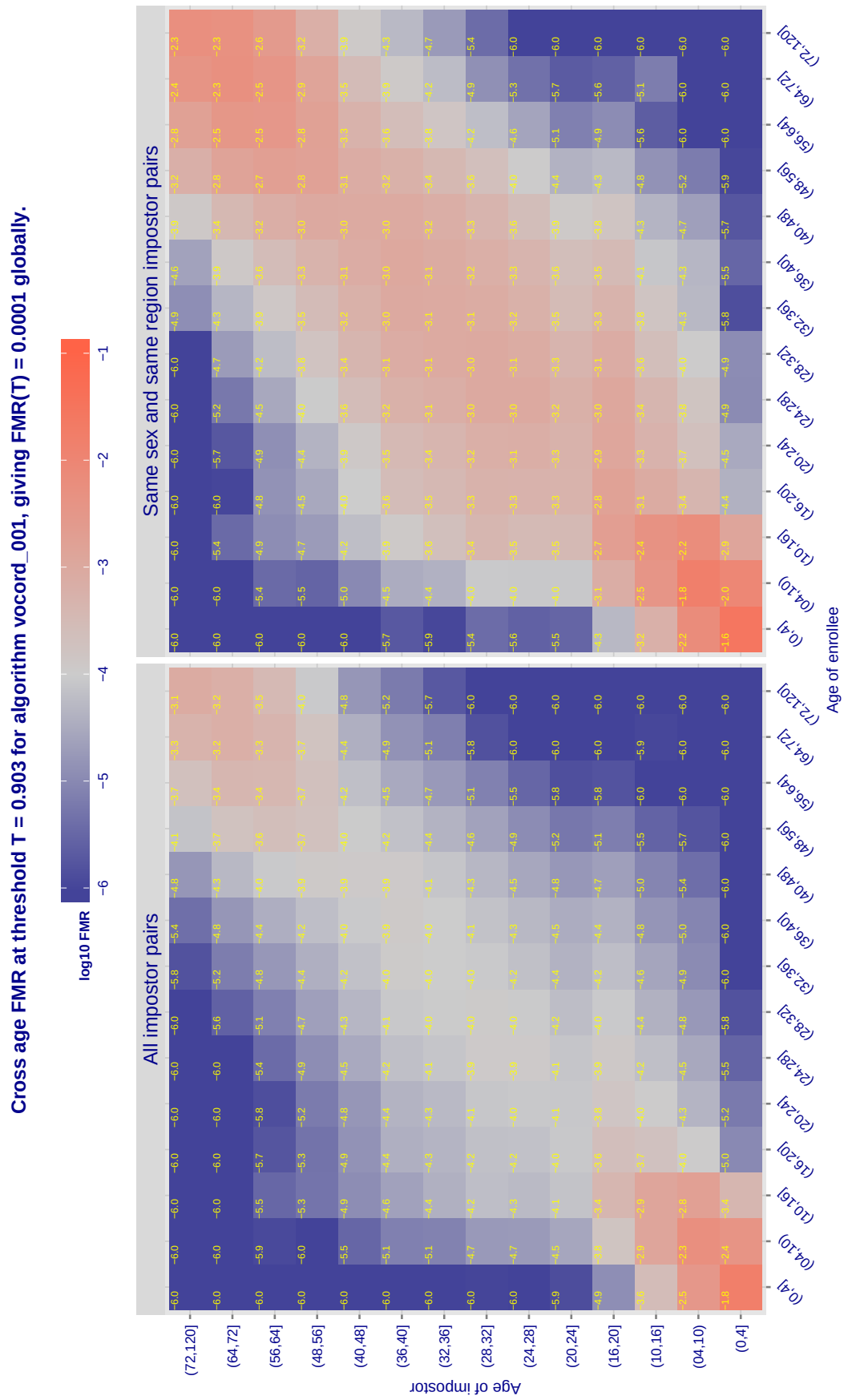


Figure 96: For algorithm vocord-001 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

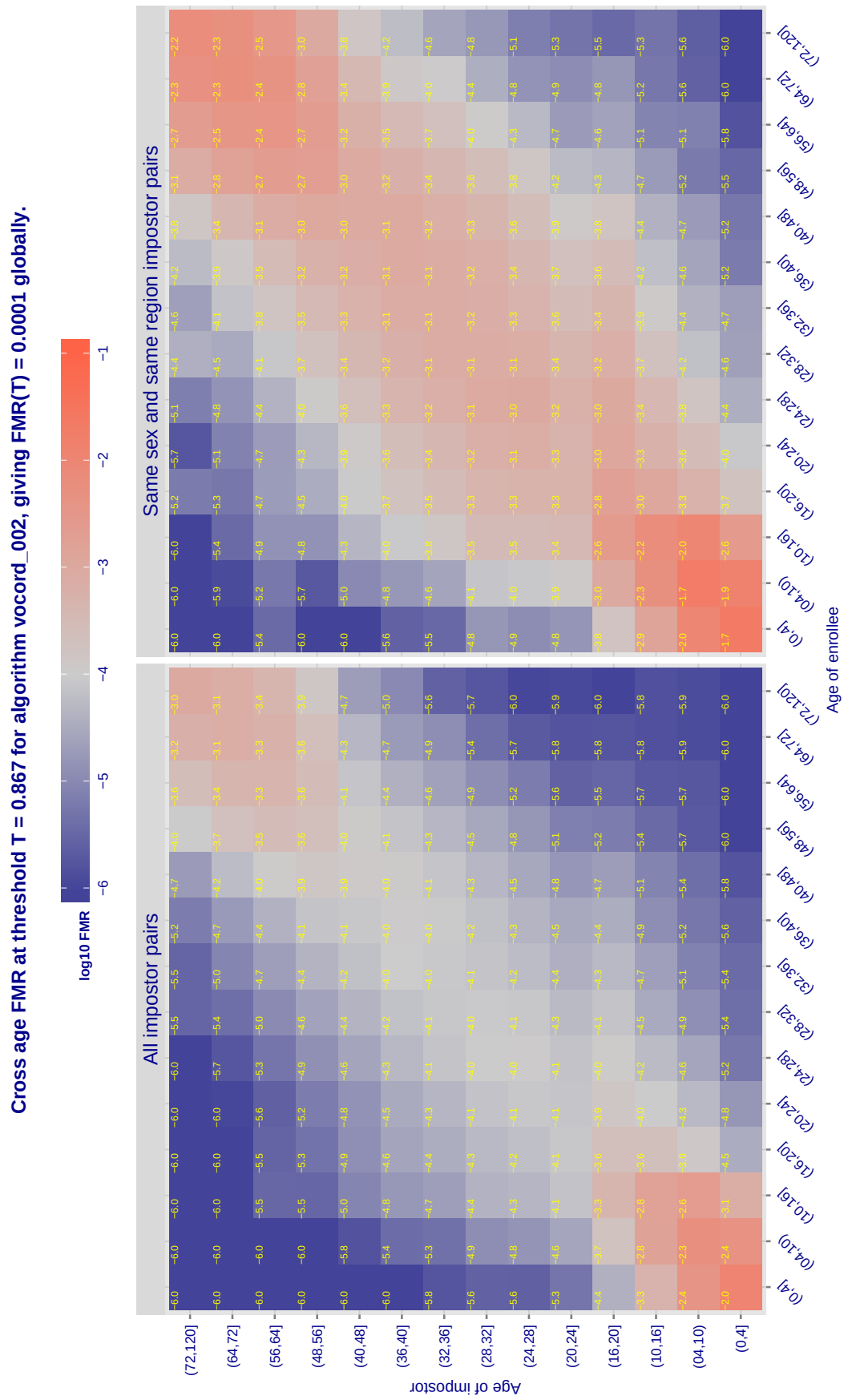


Figure 97: For algorithm vocord-002 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

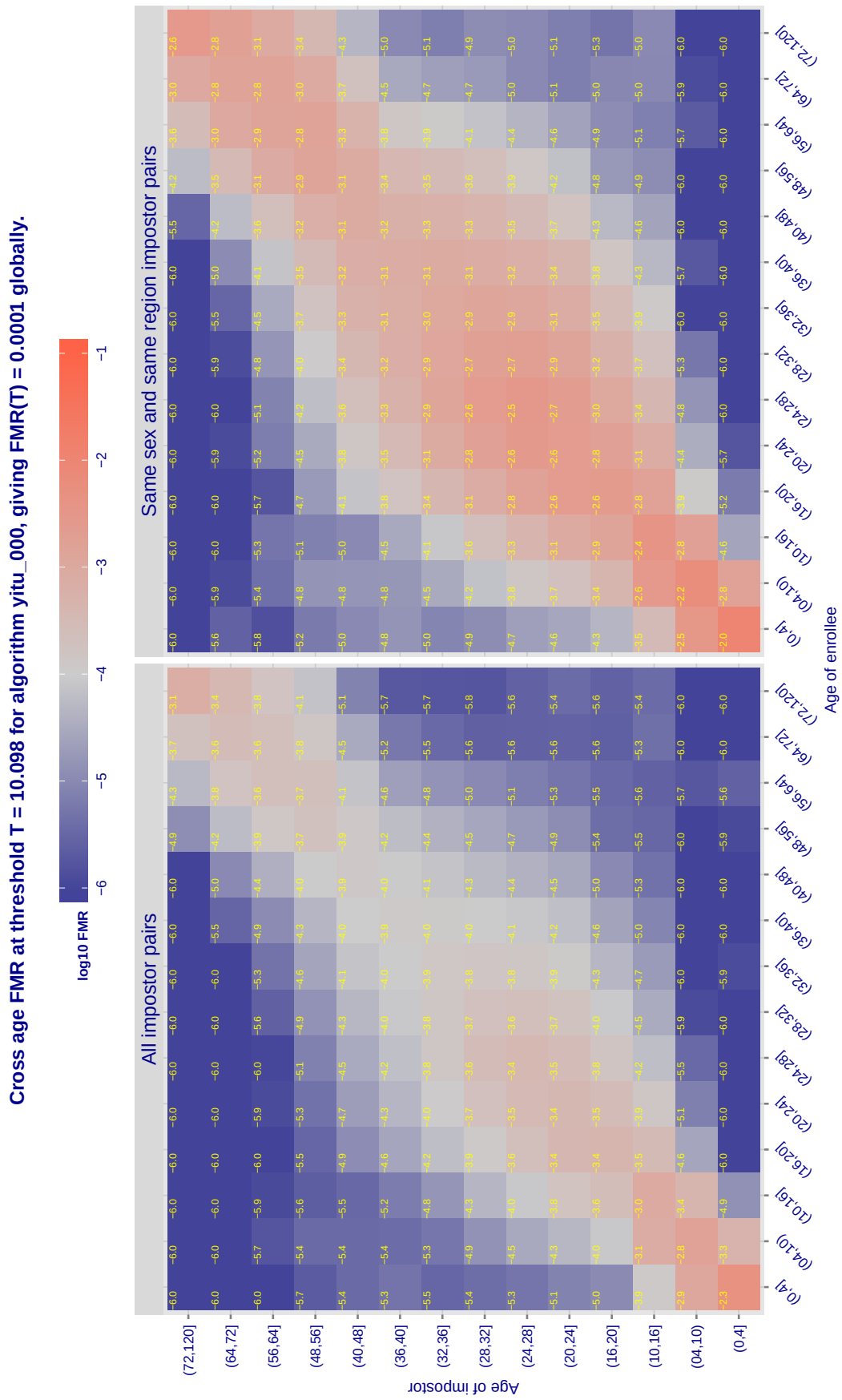


Figure 98: For algorithm yitu-000 operating on visa images, the heatmap shows false match observed over impostor comparisons of faces from different individuals who have the given age pair. False matches are counted against a recognition threshold fixed globally to give FMR = 0.001 over all $O(10^{10})$ 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.

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

Accuracy Terms + Definitions

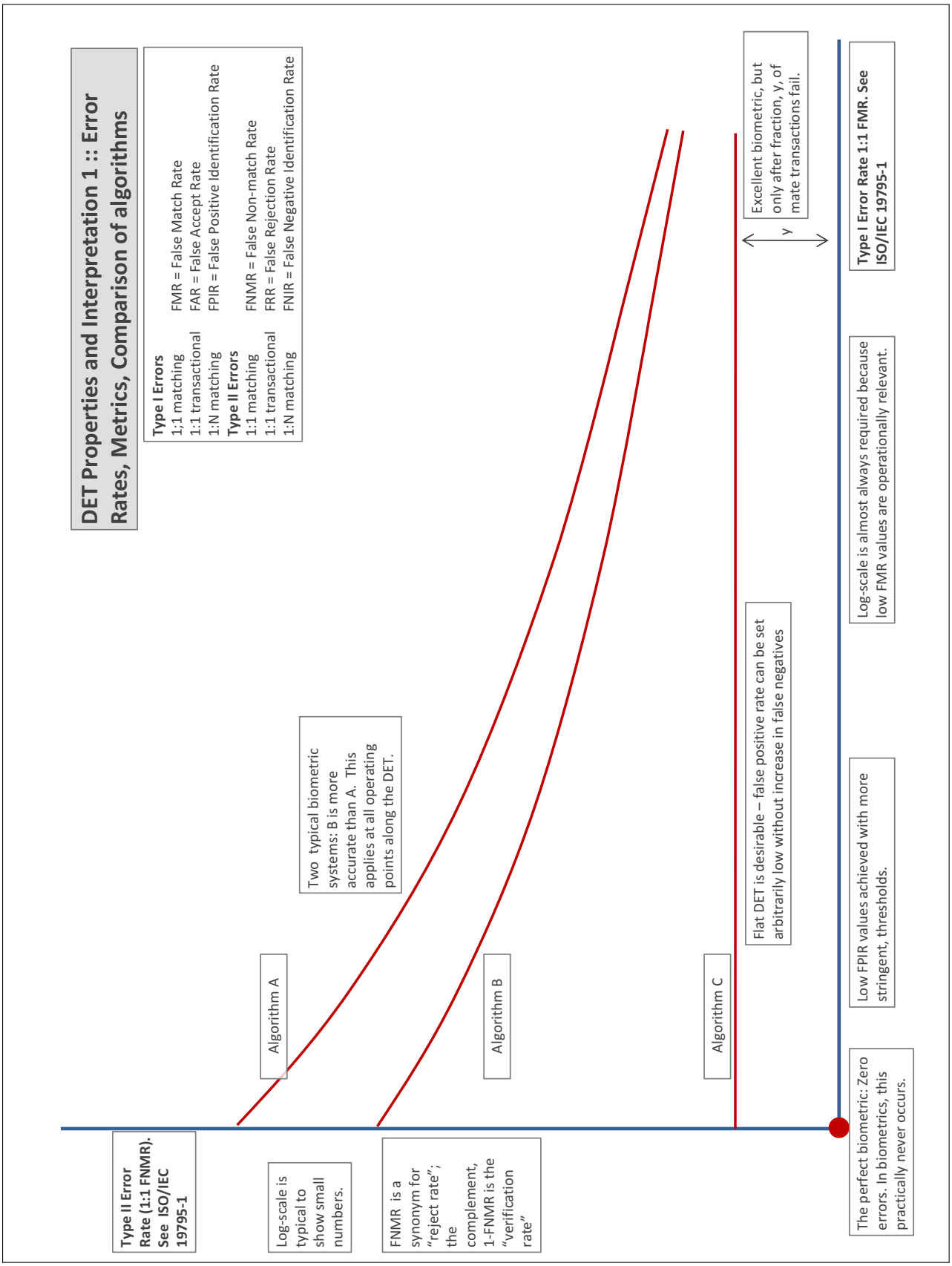
In biometrics, Type II errors occur when two samples of one person do not match – this is called a **false negative**. Correspondingly, Type I errors occur when samples from two persons do match – this is called a **false positive**. Matches are declared by a biometric system when the native comparison 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.

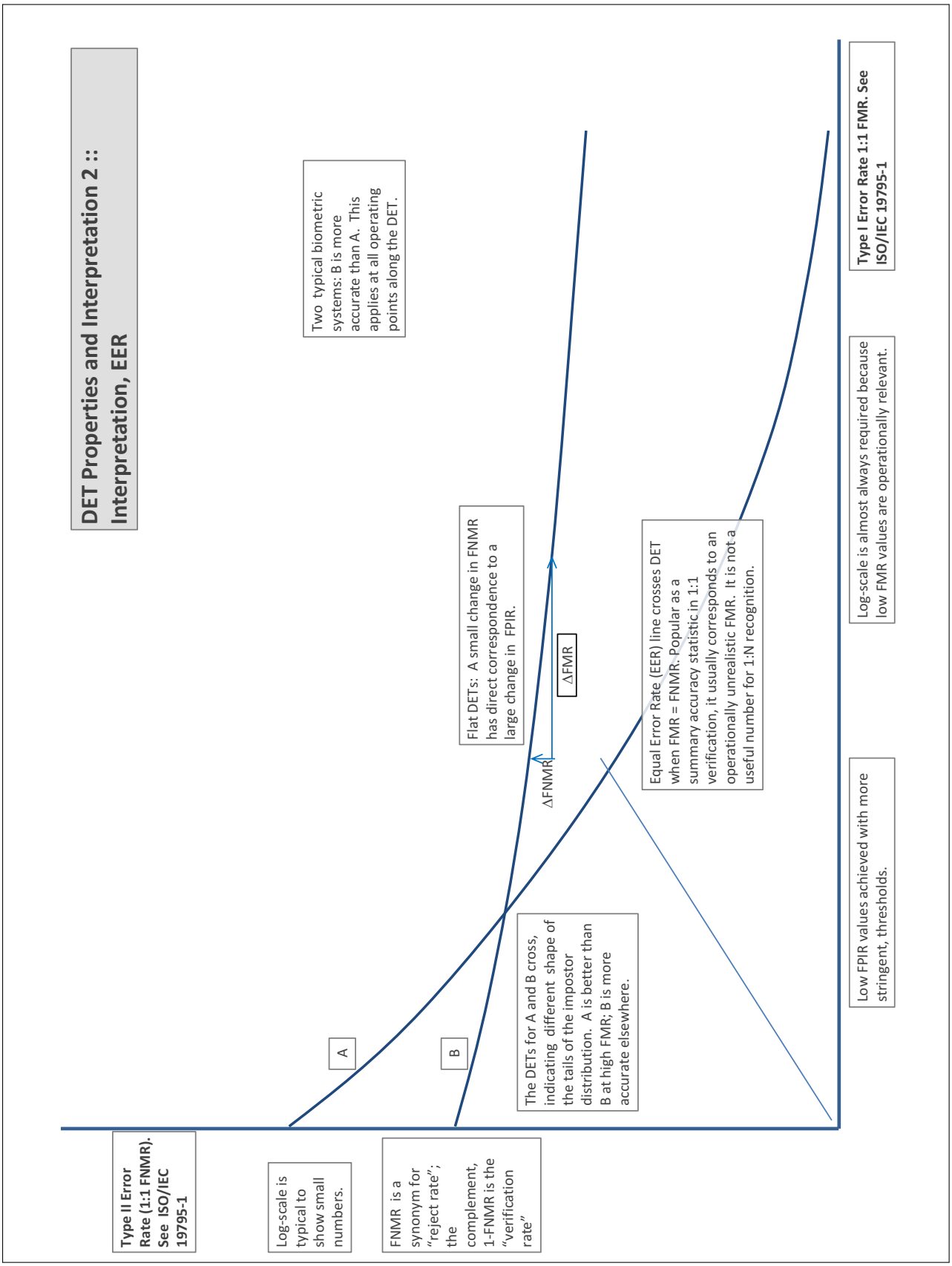
The error tradeoff plots are often called **detection error tradeoff (DET)** characteristics or **receiver operating characteristic** (ROC). These serve the same function but differ, for example, in plotting the complement of an error rate (e.g. TMR = 1 – FNMR) and in transforming the axes most commonly 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

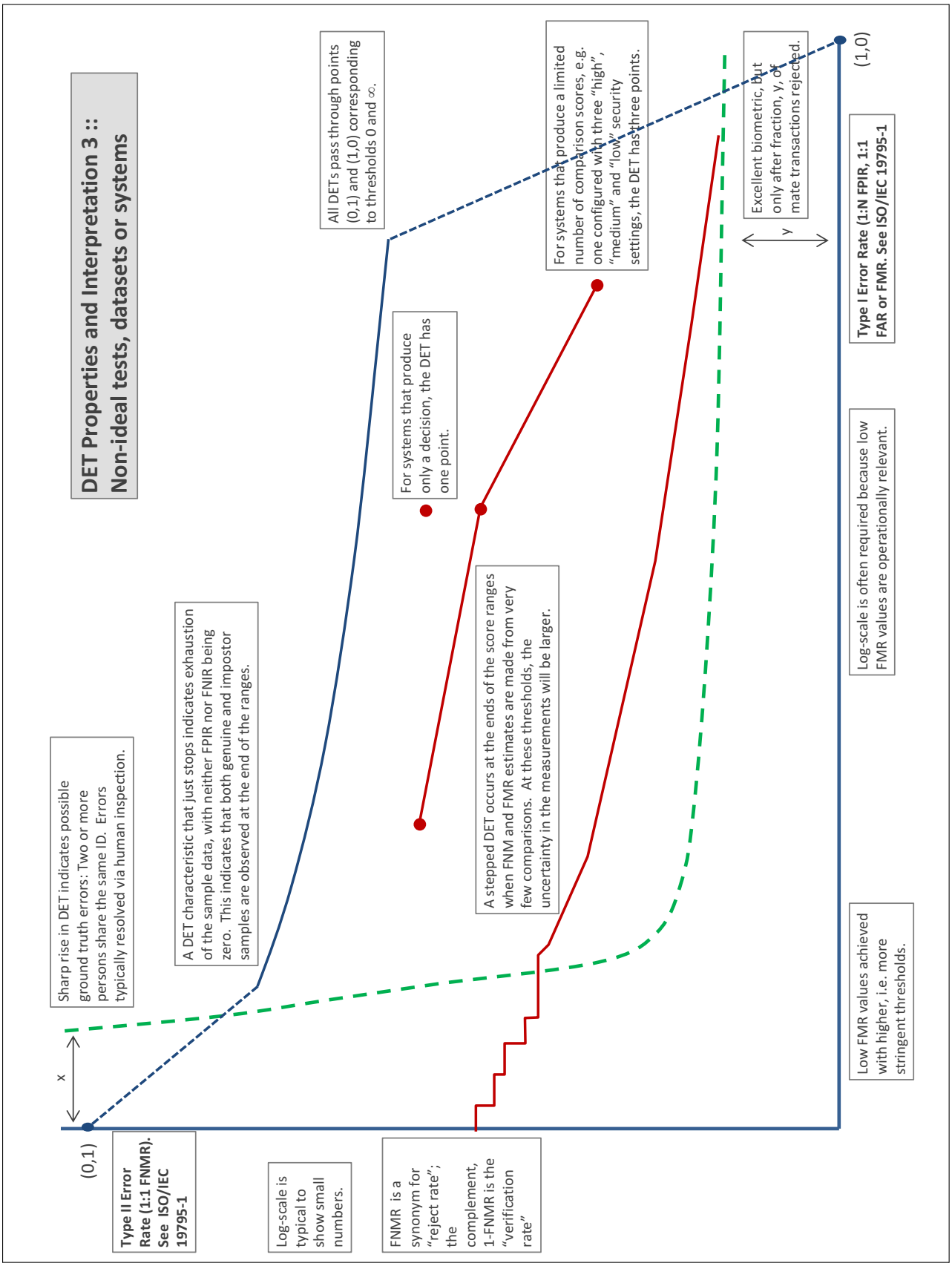
FNMR(T) "False non-match rate"
FMR(T) "False match rate"



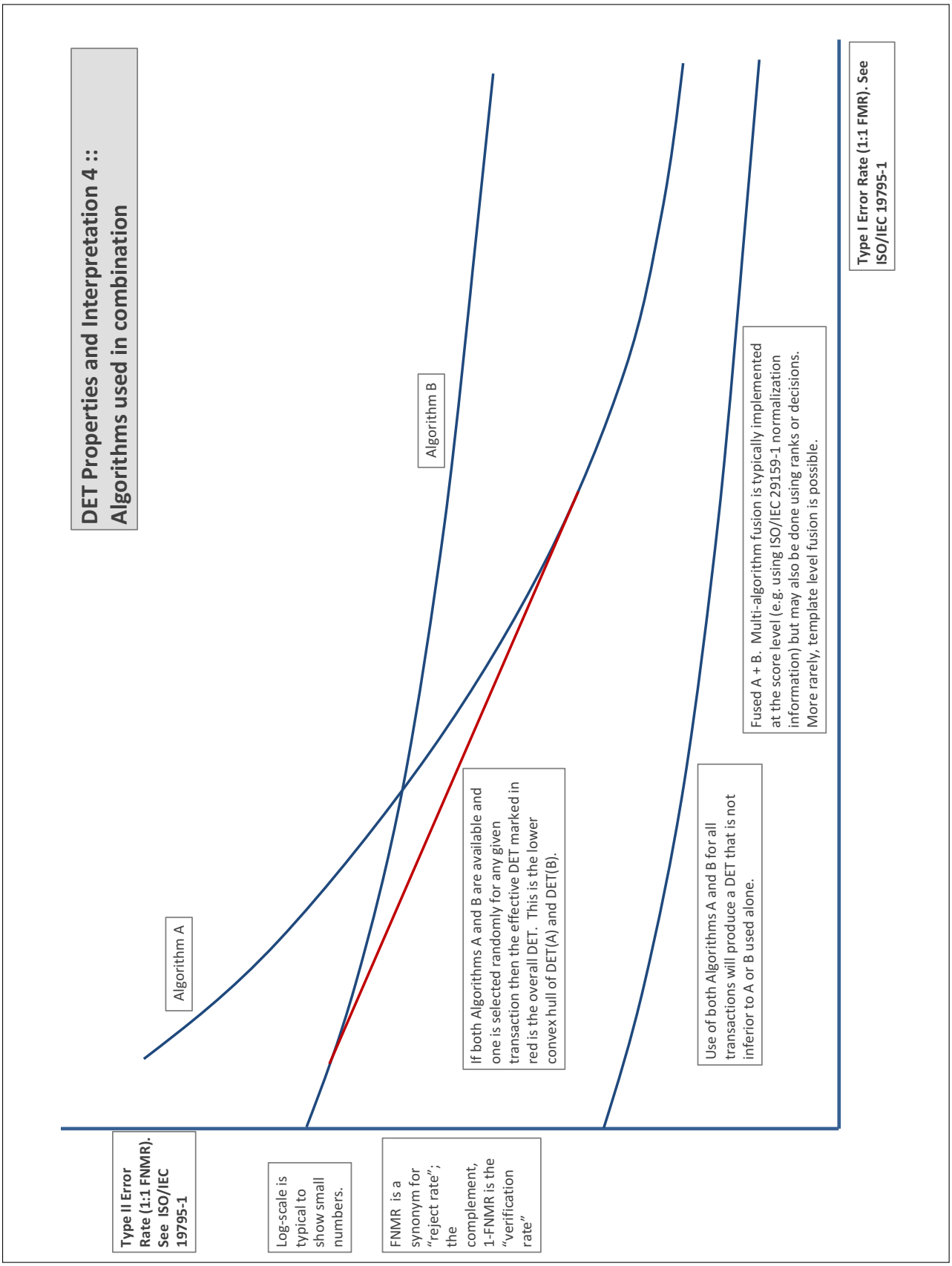
FNMR(T) "False non-match rate"
FMR(T) "False match rate"



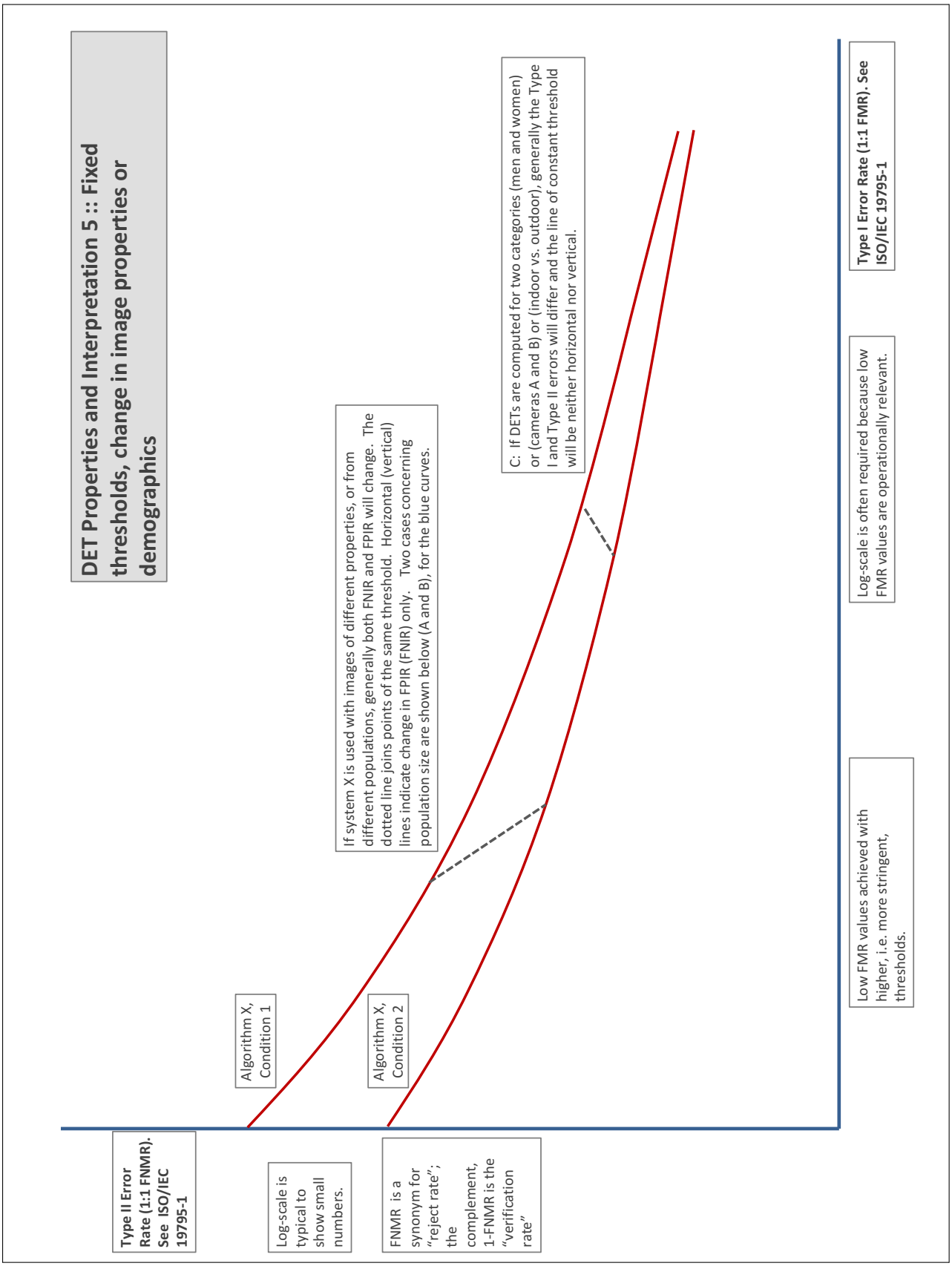
FNMR(T) "False non-match rate"
FMR(T) "False match rate"



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



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



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