RIPA2TAN epidemiological study
Modelling pipeline

Alejandro Navarro Martínez

Computational Biology Group - Life Sciences Department

RIPA2TAN project meeting   15th May, 2024

BSC **Barcelona Supercomputing Center**
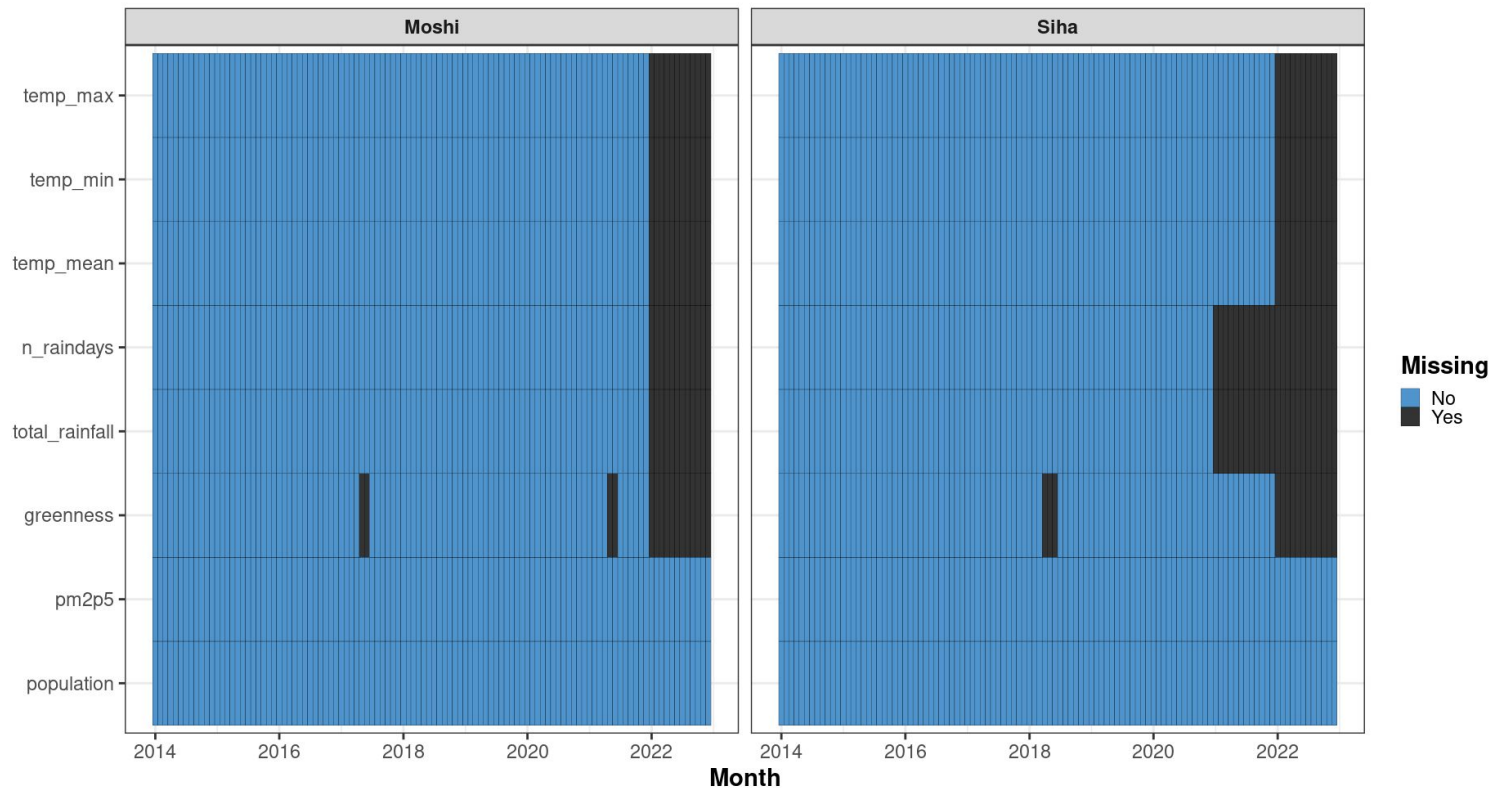*Centro Nacional de Supercomputación*

Goal

- **Epidemiological** and **environmental** data collected for Siha (rural) and Moshi (urban)

- Describe **patterns** in disease incidence

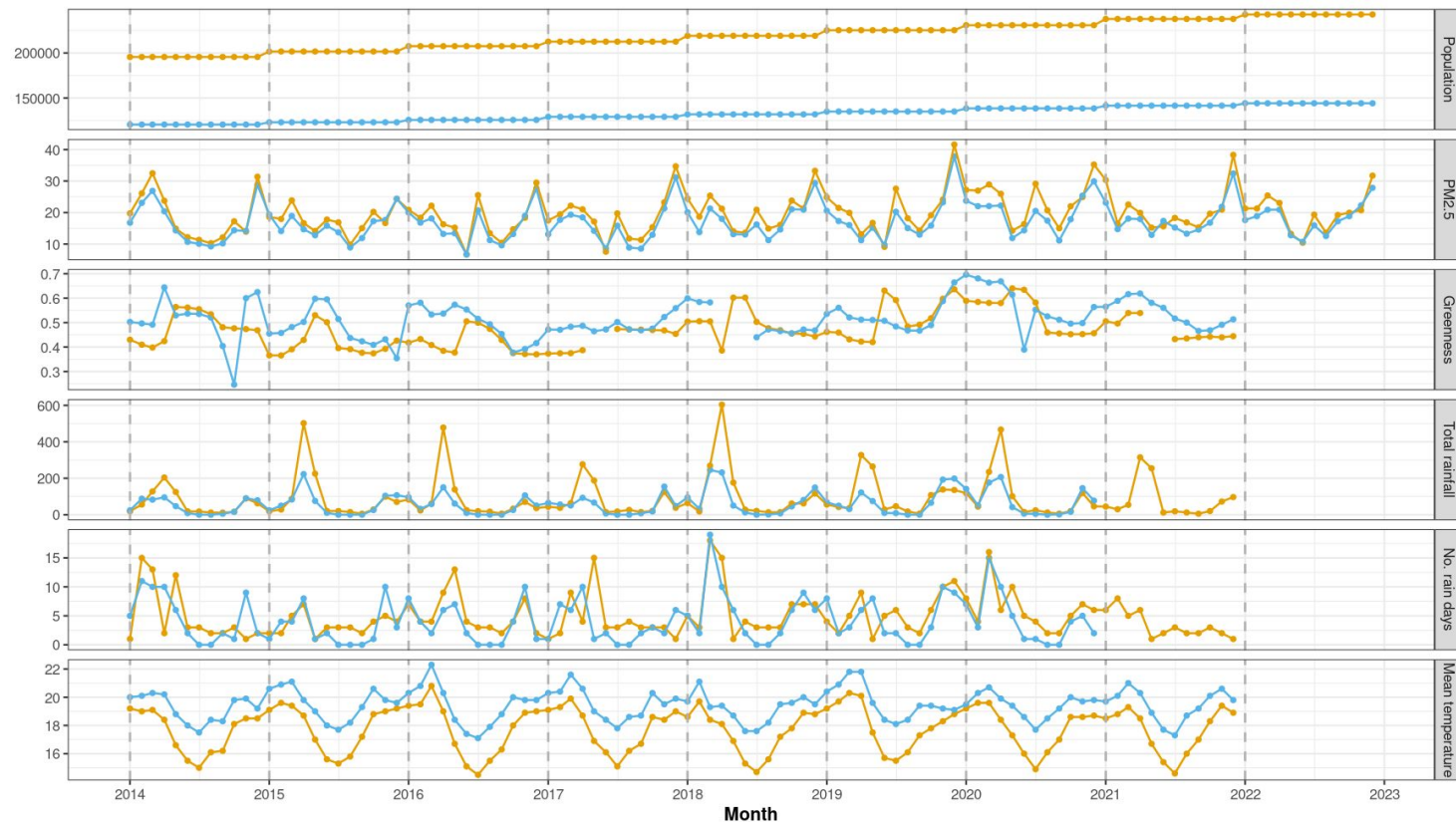- Evaluate **disease risk** from exposure to environmental factors
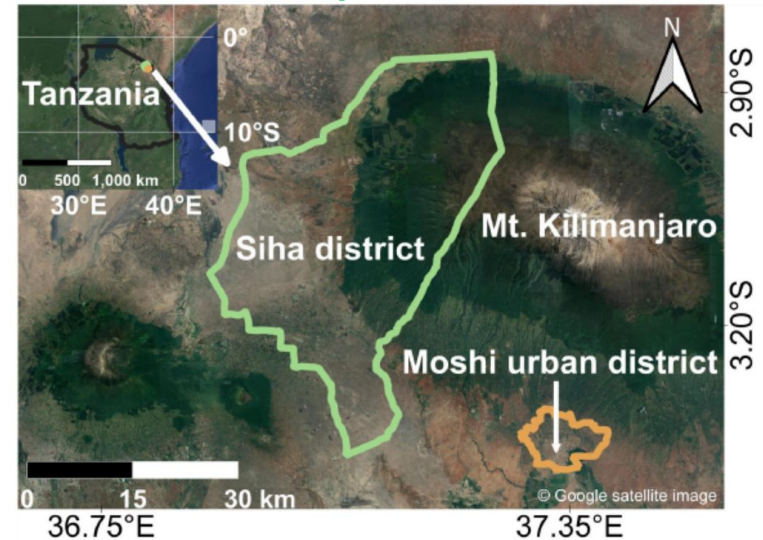
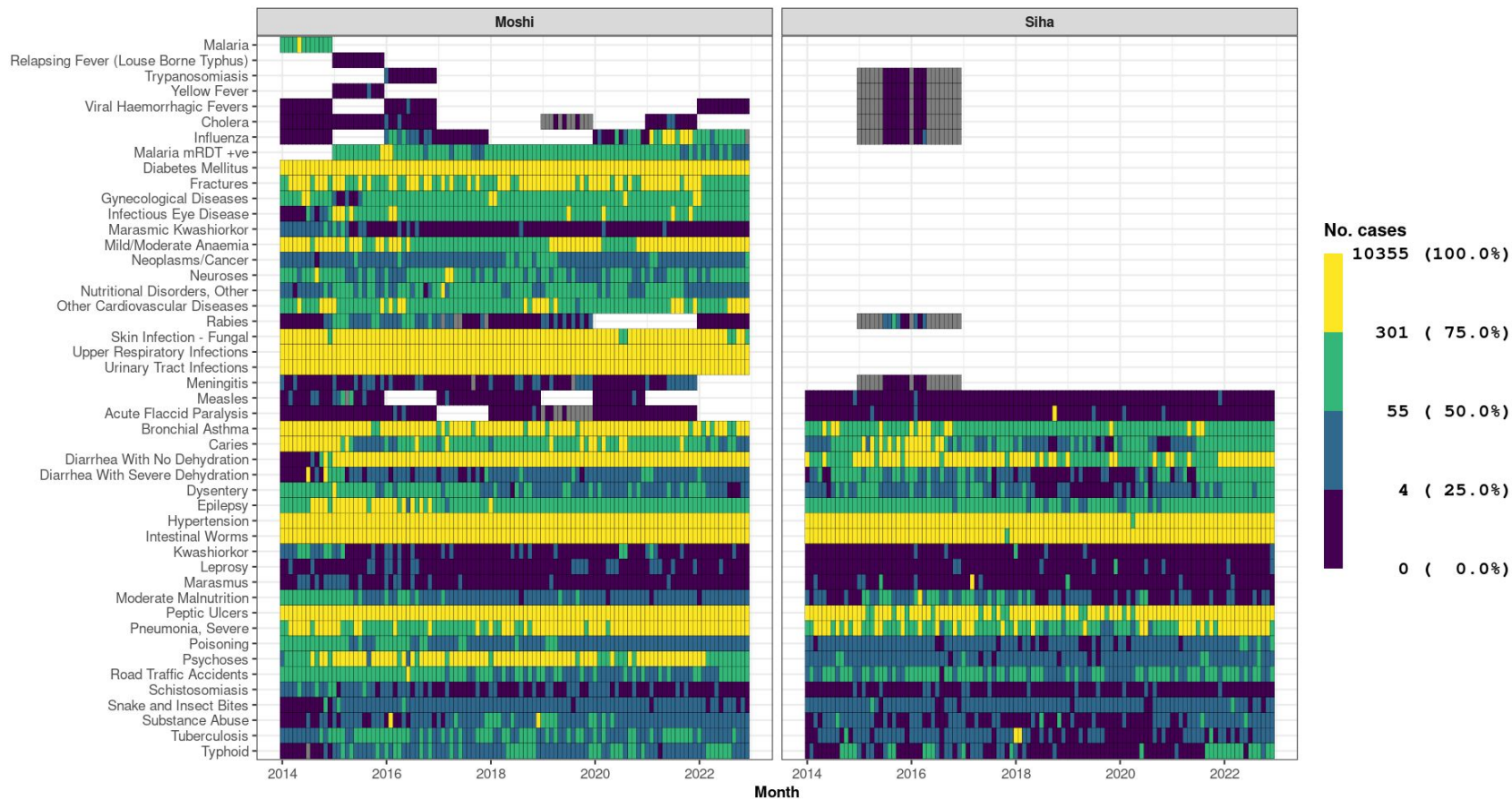*By Netra Bhandari

Environmental data

Environmental data

Working with No. rain days because Total rainfall depends on area

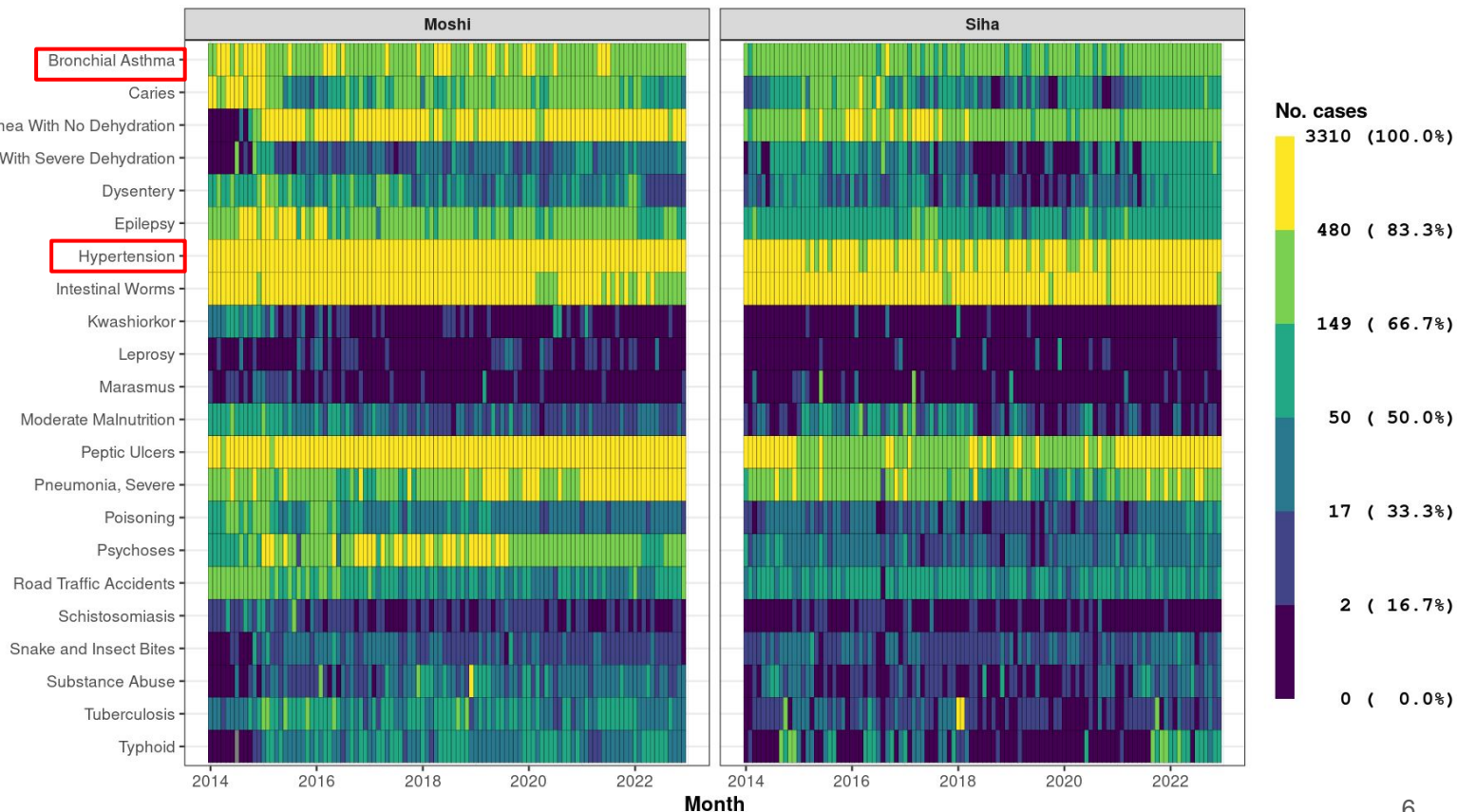Removed Min and Max because correlated

4

Disease data

Disease data

Focusing on diseases
available at both districts

*Associated with PM2.5

Modelling plan

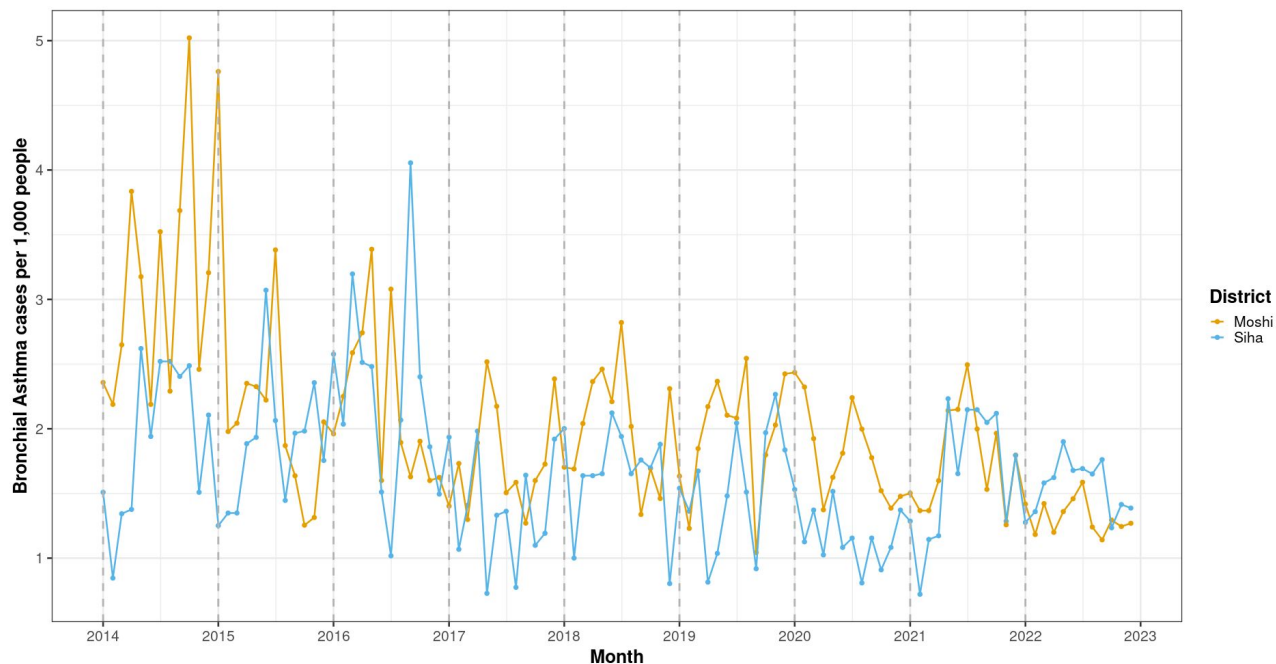# Model 1: Trend and District

Bronchial asthma

glm.nb(n_cases ~ offset(ln_population) + district + time_ind + district:time_ind)

*For incidence rate   *Baseline rate   *Linear   *Different slope
change for Siha   time trend   for both districts

Model 1: Trend and District

Bronchial asthma

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-5.8932954	0.0549917	-107.167	< 2e-16	***
districtSiha	-0.3337873	0.0783267	-4.261	2.03e-05	***
time_ind	-0.0061194	0.0008772	-6.976	3.03e-12	***
districtSiha:time_ind	0.0028985	0.0012491	2.320	0.0203	*

exp_estimate	exp_conf.low	exp_conf.high
<dbl>	<dbl>	<dbl>
0.00276	0.00248	0.00307
0.716	0.615	0.834
0.994	0.992	0.996
1.00	1.00	1.01

Model 2: Environmental effects

Bronchial asthma
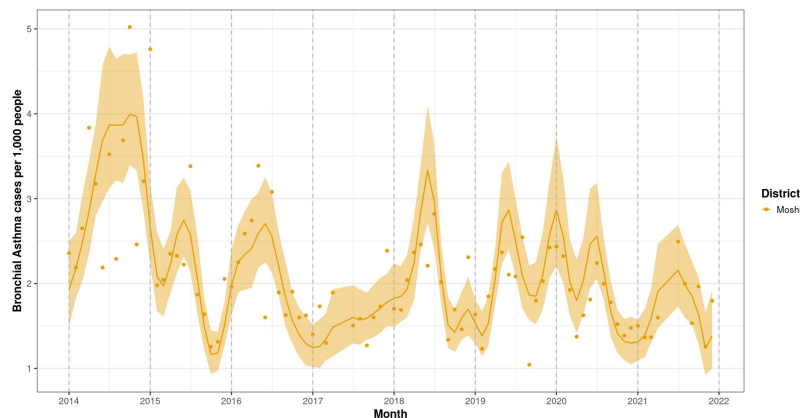
Separate models for Moshi and Siha

n_cases ~ offset(ln_population) + pm2p5 + greenness + n_raindays + ns(time_ind, df=degfree)

*Environmental exposures as linear functions

*Basis of natural cubic splines of time 5 df per year (40 df)

Moshi

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-5.6716678	0.2782840	-20.381	< 2e-16	***
pm2p5	0.0118898	0.0038143	3.117	0.001826	**
greenness	-1.7444773	0.5252980	-3.321	0.000897	***
n_raindays	0.0008652	0.0060631	0.143	0.886526	
ns(time_ind, df = degfree)1	0.7708211	0.2506473	3.075	0.002103	**
ns(time_ind, df = degfree)2	0.6312272	0.2464541	2.561	0.010430	*

exp_estimate	exp_conf.low	exp_conf.high
<dbl>	<dbl>	<dbl>
0.00344	0.00200	0.00593
1.01	1.00	1.02
0.175	0.0623	0.489
1.00	0.989	1.01
2.16	1.32	3.55
1.88	1.17	3.02

Model 2: Environmental effects

Bronchial asthma

Separate models for Moshi and Siha

n_cases ~ offset(ln_population) + pm2p5 + greenness + n_raindays + ns(time_ind, df=degfree)
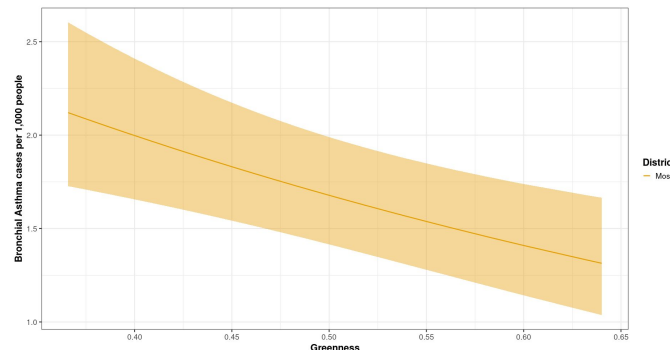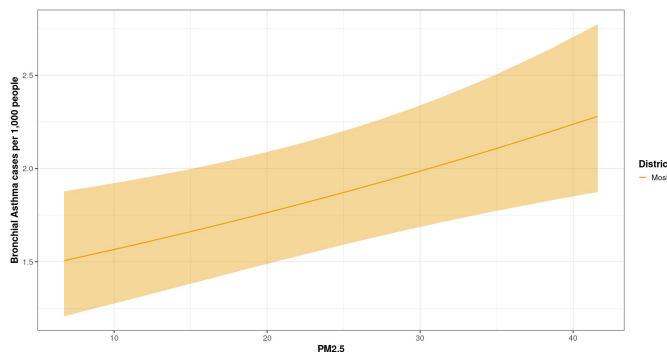
*Environmental exposures as linear functions

*Basis of natural cubic splines of time 5 df per year (40 df)

Moshi

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-5.6716678	0.2782840	-20.381	< 2e-16	***
pm2p5	0.0118898	0.0038143	3.117	0.001826	**
greenness	-1.7444773	0.5252980	-3.321	0.000897	***
n_raindays	0.0008652	0.0060631	0.143	0.886526	
ns(time_ind, df = degfree)1	0.7708211	0.2506473	3.075	0.002103	**
ns(time_ind, df = degfree)2	0.6312272	0.2464541	2.561	0.010430	*

exp_estimate	exp_conf.low	exp_conf.high
<dbl>	<dbl>	<dbl>
0.00344	0.00200	0.00593
1.01	1.00	1.02
0.175	0.0623	0.489
1.00	0.989	1.01
2.16	1.32	3.55
1.88	1.17	3.02

Model 2: Environmental effects

n_cases ~ offset(ln_population) + pm2p5 + greenness + n_raindays + ns(time_ind, df=degfree)
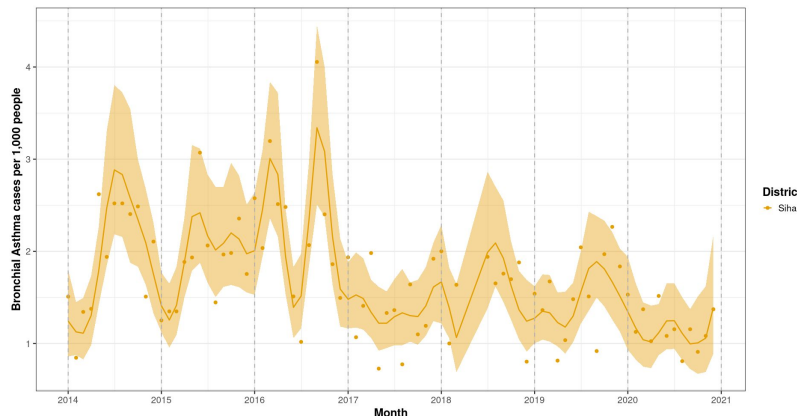
*Environmental exposures as linear functions

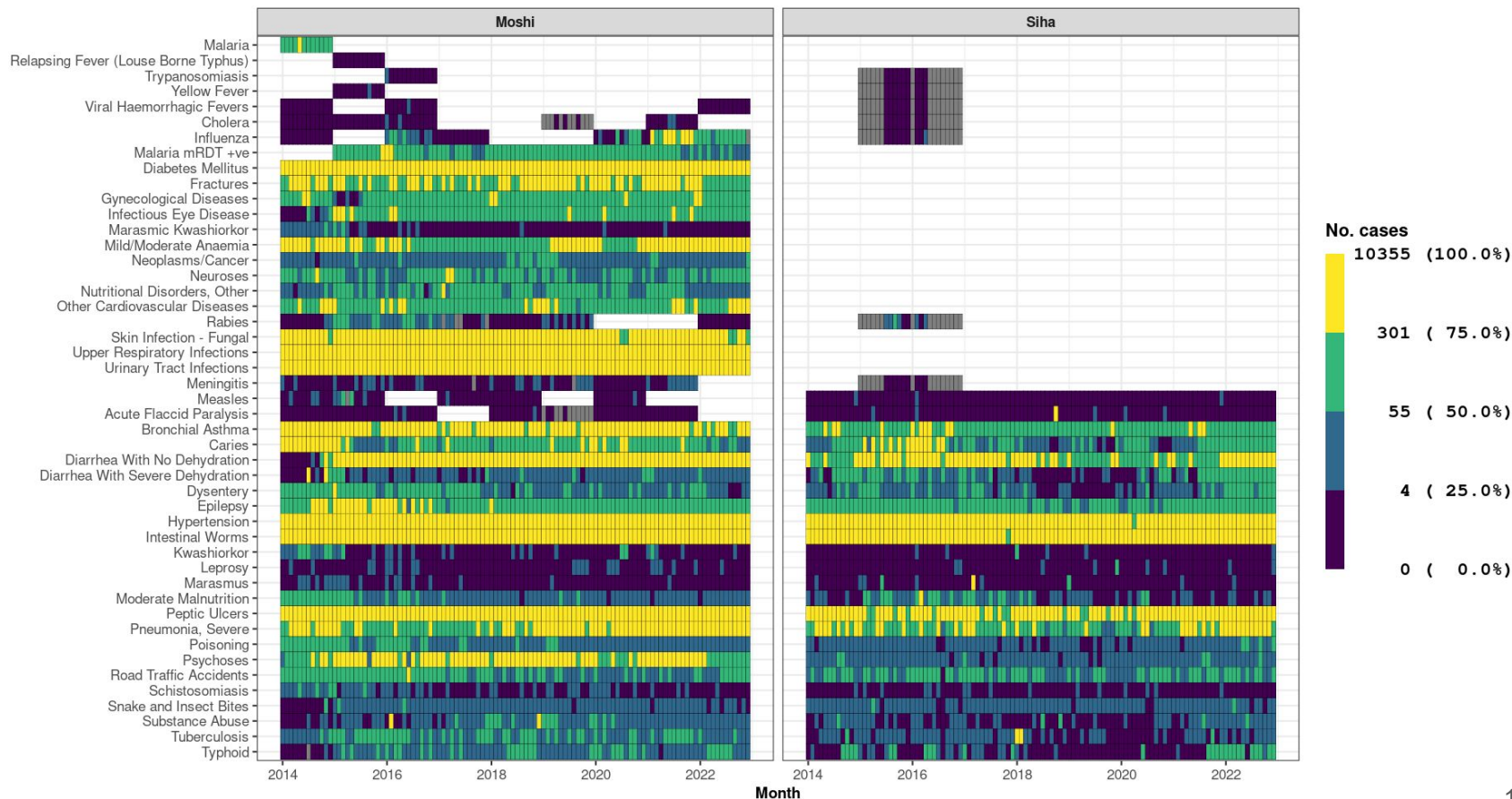*Basis of natural cubic splines of time 5 df per year (40 df)

Siha

No significant coefficients

	Estimate	Std. Error	z value	Pr(>|z|)	
(Intercept)	-6.980e+00	3.607e-01	-19.351	< 2e-16	***
pm2p5	3.848e-04	6.672e-03	0.058	0.95400	
greenness	3.974e-01	5.955e-01	0.667	0.50456	
n_raindays	1.960e-02	1.030e-02	1.904	0.05695	.
ns(time_ind, df = degfree)1	1.049e+00	3.425e-01	3.064	0.00218	**
ns(time_ind, df = degfree)2	6.982e-01	3.921e-01	1.781	0.07494	.

exp_estimate	exp_conf.low	exp_conf.high
<dbl>	<dbl>	<dbl>
0.000931	0.000468	0.00187
1.00	0.987	1.01
1.49	0.471	4.65
1.02	0.999	1.04
2.86	1.46	5.66
2.01	0.941	4.31

Further analyses

RIPA2TAN epidemiological study
Modelling pipeline

Alejandro Navarro Martínez

Computational Biology Group - Life Sciences Department

RIPA2TAN project meeting   15th May, 2024