

## Research appointments

- 2022 – present **Schmidt Futures AI2050 Early Career Fellow**, Department of Computer Science, University of Oxford
- 2022 – present **Junior Research Fellow**, Kellogg College
- 2022 – present **Postdoctoral researcher**, Department of Computer Science, University of Oxford
- 2021 – 2022 **Research Associate in Mathematics**, Department of Mathematics, Statistics Section, Imperial College London
- ▶ Modelling spatial data by leveraging deep generative models (VAEs) to power faster Bayesian inference.
  - ▶ Developing adaptive survey design approach allowing to increase efficiency of surveys, lower the cost of data collection and bias (in collaboration with US Census).
  - ▶ Performing applied epidemiology work to inform public policy as a response to the COVID-19 pandemic: (i) analysis of global genomic surveillance disparities; (ii) investigating paediatric disease severity
- 2019 – 2021 **PostDoc Fellow in Bayesian Machine Learning**, AstraZeneca, Cambridge, UK
- ▶ Developing a Bayesian neural network for toxicity prediction to communicate uncertainty and prevent overfitting of biological data.
  - ▶ Predicting drug toxicity by incorporating scientific knowledge into Bayesian ML models.
  - ▶ Fitting flexible concentration-response curves of a novel drug modality using Gaussian Processes.
  - ▶ Developing models of COVID-19 testing regimes to inform internal company's testing policy.

## Qualifications

- 2014 – 2019 **PhD in Epidemiology** (*summa cum laude*), Swiss Tropical and Public Health Institute, University of Basel, Switzerland: *Bayesian modelling of large spatio-temporal disease surveillance and environmental data*.
- 2003 – 2008 **Diploma in Mathematics** (first class honours), Moscow State University, Moscow, Russia

## Publications

My research interests include Bayesian inference, spatiotemporal statistics, deep generative modelling, epidemiology, drug discovery and mathematics. I have published 16 papers; in disciplines where order matters (i.e. not mathematics), I am first (or joint first) author on 4 and corresponding author on 3 (\* denotes joint authorship and † denotes corresponding author).

## Peer reviewed publications

1. A. Brito\*, **E. Semenova**\*, G. Dudas\*, G. Hassler, C. Kalinich, M. Kraemer, S. Hill, O. Pybus, C. Dye, S. Bhatt, *et al.*, Global disparities in SARS-CoV-2 genomic surveillance, *Nature Communications*, 2022.
2. J. W. Firman, M. T. Cronin, P. H. Rowe, **E. Semenova**, and J. E. Doe, The use of Bayesian methodology in the development and validation of a tiered assessment approach towards prediction of rat acute oral toxicity, *Archives of toxicology*, 1–14, 2022.
3. **E. Semenova**\*†, Y. Xu\*, A. Howes, T. Rashid, S. Bhatt, S. Mishra, and S. Flaxman†, Priorvae: Encoding spatial priors with variational autoencoders for small-area estimation, *Journal of the Royal Society Interface*, **19**, (191), 20220094, 2022.
4. L. Grinsztajn, **E. Semenova**, C. C. Margossian, and J. Riou, Bayesian workflow for disease transmission modeling in Stan, *Statistics in Medicine*, 1–26, 2021.
5. L. H. Mervin, S. Johansson, **E. Semenova**, K. A. Giblin, and O. Engkvist, Uncertainty quantification in drug design, *Drug discovery today*, **26**, (2), 474–489, 2021.
6. S. Mishra, S. Mindermann, M. Sharma, C. Whittaker, T. A. Mellan, T. Wilton, D. Klapsa, R. Mate, M. Fritzsche, M. Zambon, J. Ahuja, A. Howes, X. Miskouridou, G. P. Nason, O. Ratmann, **E. Semenova**, G. Leech, J. F. Sandkühler, C. Rogers-Smith, M. Vollmer, H. J. T. Unwin, Y. Gal, M. Chand, A. Gandy, J. Martin, E. Volz, N. M. Ferguson, S. Bhatt, J. M. Brauner, S. Flaxman, and The COVID-19 Genomics UK (COG-UK) Consortium, Changing composition of SARS-CoV-2 lineages and rise of delta variant in england, *EClinicalMedicine*, **39**, 101064, 2021.
7. L. Scott, N.-Y. Hsiao, S. Moyo, L. Singh, H. Tegally, G. Dor, P. Maes, O. G. Pybus, M. U. Kraemer, **E. Semenova**, *et al.*, Track Omicron's spread with molecular data, *Science*, **374**, (6574), 1454–1455, 2021.

8. **E. Semenova**<sup>†</sup>, M. L. Guerriero, B. Zhang, A. Hock, P. Hopcroft, G. Kadamur, A. M. Afzal, and S. E. Lazic, Flexible fitting of PROTAC concentration-response curves with changepoint Gaussian Processes, *SLAS Discovery*, **26**, (9), 1212–1224, 2021.
9. P. Vel'misov, Y. A. Tamarova, and **E. Semenova**, Asymptotic equations of gas dynamics: Qualitative analysis, construction of solutions, and applications, *Journal of Mathematical Sciences*, **259**, (3), 309–325, 2021.
10. S. E. Lazic, **E. Semenova**, and D. P. Williams, Determining organ weight toxicity with Bayesian causal models: Improving on the analysis of relative organ weights, *Scientific reports*, **10**, (1), 1–12, 2020.
11. **E. Semenova**<sup>†</sup>, D. P. Williams, A. M. Afzal, and S. E. Lazic, A bayesian neural network for toxicity prediction, *Computational Toxicology*, **16**, 2020.
12. D. P. Williams, S. E. Lazic, A. J. Foster, **E. Semenova**, and P. Morgan, Predicting drug-induced liver injury with Bayesian machine learning, *Chemical research in toxicology*, **33**, (1), 239–248, 2019.
13. G. Crippa, **E. Semenova**, and S. Spirito, Strong continuity for the 2d Euler equations, *Kinetic and Related Models*, (4), 685–689, 2018.
14. P. A. Velmisov, U. J. Mizher, and **E. P. Semenova**, Asymptotic study of nonlinear viscous gas flows, in *AIP Conference Proceedings*, AIP Publishing LLC, vol. 2048, 2018, pp.040012.
15. P. A. Velmisov, A. V. Ankilov, and **E. P. Semenova**, Mathematical modeling of aeroelastic systems, in *AIP Conference Proceedings*, AIP Publishing LLC, vol. 1910, 2017, pp.040010.
16. P. A. Velmisov, A. V. Ankilov, and **E. P. Semenova**, Dynamic stability of deformable elements of one class of aeroelastic constructions, in *AIP Conference Proceedings*, AIP Publishing LLC, vol. 1789, 2016, pp.020021.

## Under revision

1. **E. Semenova**, S. Dalmini, and P. Vounatsou, Modelling Log-Gaussian Cox Processes on fine spatio-temporal scale, *Spatial and Spatio-temporal epidemiology (under revision)*, 2022.

## Pre-prints

1. S. Flaxman, C. Whittaker, **E. Semenova**, T. Rashid, R. Parks, A. Blenkinsop, H. J. T. Unwin, S. Mishra, S. Bhatt, D. Gurdasani, *et al.*, Covid-19 is a leading cause of death in children and young people ages 0-19 years in the united states, *medRxiv*, 2022.

## Book chapters

1. S. K. Ashenden, A. Bartosik, P.-M. Agapow, and **E. Semenova**, Introduction to artificial intelligence and machine learning, *The Era of Artificial Intelligence, Machine Learning, and Data Science in the Pharmaceutical Industry*, 15–26, 2021.

## Grants and awards

- 2022 Schmidt Futures AI2050 Early Career Fellowship (\$ 270'000).
- 2022 The International Congress of Toxicology, Travel grant to deliver a lecture (£ 700).
- 2022 Nordic ProbAI summer school, Travel grant to deliver a lecture (£ 800).
- 2022 UK-China collaborative workshop: AI for Climate, Environment and Sustainability, Travel grant (£ 500).
- 2021 Google Season of Docs (GSoD) bursary to contribute Julia documentation (\$ 3000)
- 2020 AstraZeneca R&D, Postdoc of the Year award

## Invited talks and lectures

- 2022, Nov 18 Talk at Autonomous Intelligent Machines and Systems seminar, Oxford (upcoming)
- 2022, Oct 25 "How can deep generative modelling help with spatial statistics?" (talk), Infectious Disease Epidemiology seminar series, Imperial College London
- 2022, Sep 18 "Capturing uncertainty in toxicity profiling models", XVIth International Congress of Toxicology (talk), Maastricht, Netherlands
- 2022, Aug 21 "Mathematics and Statistics for Machine Learning" (lecture), Deep Learning Indaba, Tunis, Tunisia
- 2022, Aug 19 Uniq+, DeepMind social lunches in Oxford
- 2022, Aug 5 "Encoding spatial priors with VAEs for geospatial modelling" (talk), 5th Workshop on Tractable Probabilistic Modeling at Uncertainty in Artificial Intelligence (UAI) conference, Eindhoven, Netherlands
- 2022, Jun 26 Oxford Computational Statistics and Machine Learning (OxCSML) seminar
- 2022, Jun 13 "Bayesian workflow" (lecture), Nordic Probabilistic AI (ProbAI) summer school, Helsinki, Finland
- 2022, Apr 25 "Bayesian modelling of spatial data" (lecture), Cambridge's first machine learning research event for undergraduates, University of Cambridge, UK
- 2022, Feb 10 "Encoding spatial priors with VAEs for geospatial modelling" (talk), *University of Lancaster, CHICAS seminar*, remote

## Contributed talks and poster presentations

- 2022 **E. Semenova**, S. Mishra, S. Bhatt, S. Flaxman, H. J. Unwin "Mapping malaria prevalence in Kenya by reconciling changes in administrative boundaries using MCMC and Deep Learning", *American Society of Tropical Medicine & Hygiene*, annual meeting (remote)
- 2021 L. Grinsztajn, **E. Semenova**, C. Margossian, J. Riou "Bayesian workflow for disease transmission modeling in Stan", *Berlin Bayesians Meetup*, (remote)
- 2020 **E. Semenova**, "Building an ordered logistic regression model for toxicity prediction" (talk), *PyMCon conference*, remote
- 2020 **E. Semenova**, A. Afzal, D. Williams, S. Lazic, "Bayesian Neural Networks for toxicity prediction" (talk), *Applied Machine Learning Days (AMLDD) conference*, Lausanne
- 2019 **E. Semenova**, D. Williams, S. Lazic, "Predicting severity of drug-induced liver injury with uncertainty" (poster), *AwayDay conference*, University of Cambridge
- 2019 **E. Semenova**, D. Williams, S. Lazic, "A Bayesian multi-layered model to predict mechanisms, types, and severity of drug-induced liver injury" (poster), *StanCon-2019 conference*, Cambridge
- 2017 **E. Semenova**, P. Vounatsou, "Fitting large-scale Log-Gaussian Cox Process model with an efficient sampling algorithm" (poster), *46th SpeedUp Workshop on "Uncertainty Quantification and HPC"*, Bern
- 2016 **E. Semenova**, "Point Patterns Analysis for Malaria Surveillance" (talk), Malaria Modelling Group's seminar, Basel,
- 2016 **E. Semenova**, P. Vounatsou, "Early Warning Platform for Malaria Surveillance and Response" (talk), *Swiss Academy for development, Research Fair "Transcending the Silos to Prevent Crises"*, Bern
- 2014 **E. Semenova**, G. Crippa, "The Euler equation: stability and continuity in 2D" (talk), *Analysis seminar*, Geneva

## Organisation, outreach and public engagement

- Co-organizer of a (workshop at NeurIPS )(2022)
- Co-organizer of Gaussian Process seminar series (<https://gp-seminar-series.github.io/>)
- Data Science ambassador of the Faculty of Natural Sciences at Imperial College London (2022)
- Guest at the podcast "Learning Bayesian Statistics"; among the top popular episodes on the channel (2020)
- Organising committee of PyPharma conference (2019)
- Basel Computational Methods for Research Community, Meetup organiser (2017-2019)
- Speaker at Women Techmakers Leads Summit Europe (Google), Madrid (2018)

## Esteem indicators

- Reviewer for PharML conference and journals 'Geospatial Health', 'Parasite epidemiology and control', 'Tropical medicine and international health', 'Journal of the Royal Statistical Society: Series A (JRSSA)'
- Member of European Women in Mathematics (EWM), International Society for Bayesian Analysis (ISBA)
- Hackathons: invited jury member (Basel Hack), winner (Hack Zurich & Crack the Hack) (2017-2018)

## Teaching, supervision and mentoring

- Research Supervision
  - Co-supervising a PhD student on malaria mapping in sub-Saharan Africa using modern computational tools (2022).
  - Supervising a PhD mini-project (3 months) on identifiability of Bayesian statistical models (2022).
  - Co-supervising an MSc project estimating the number of orphans in Kenya due to Covid (2022)
  - Co-supervising an MSc project using RKHS for GP to estimate mobility patterns in the US (2022)
  - Co-supervising an MSc project on PriorVAE and graph representation learning (2022)
- Academic teaching
  - Substitution lecturer in Bayesian modelling and biostatistics. A course taught to Epidemiology and Infectious biology MSs: theory and hands-on practicals. Groups of about ~20 people (2014-2019).
  - Weekly marking of problem sheets and discussion of solutions for two graduate courses: Harmonic Analysis and Partial Differential Equations, groups of ~15 students (2012-2014).
  - Bi-weekly grading and in-person presentation of solutions in Calculus to mathematics and physics undergraduates, groups of ~30 (2012-2014)

### ► Workshops at conferences and meetups

- "Bayesian modelling of spatial data", Cambridge University, guest lecture and mentorship round table (2022)
- "Hands-on Bayesian Machine Learning: embracing uncertainty" (in Julia). Over 100 participants; the most popular workshop at the conference. Applied Machine Learning Days (AMLDD), Lausanne, Switzerland (2020)
- "Bayesian Inference with Python", PyPharma conference, Basel (2019)
- "Machine Learning for news (using NLP): theory, applications and visualization in Python", Lausanne (2018)
- "Bayesian Statistics with R and Stan", R-Lausanne Meetup (2018)

### ► Mentorship

- AISTATS2022, mentoring PhD applicants (2022)
- Mentorship round tables for undergraduate students in mathematics and CS, Cambridge University (2022)

## Skills

|                           |   |
|---------------------------|---|
| General programming       | R, Python, Git, Julia   |
| Probabilistic programming | Stan, PyMC3, NumPyro, Turing.jl   |
| Languages                 | Russian (native), English & German (working proficiency), Italian & French (beginner) |

## Technical writing

1. L. Badillo, **E. Semenova**, and L. Kilpatrick, *Julia ecosystem contributor's guide*, 2021.

## PhD Thesis

My PhD thesis "*Bayesian modelling of large spatio-temporal disease surveillance and environmental data*" is under embargo until March 2023. The chapters (beyond Introduction and Discussion) which it contains are as follows:

1. **E. Semenova**, S. Dalmini, and P. Vounatsou, Modelling Log-Gaussian Cox Processes on fine spatio-temporal scale, 2019.
2. **E. Semenova** and P. Vounatsou, A comprehensive modelling approach for Bayesian analysis of marked point patterns, 2019.
3. **E. Semenova** and P. Vounatsou, Delineating hotspots on the map of schistosomiasis prevalence in Togo, 2019.
4. **E. Semenova** and P. Vounatsou, Restoring missing data in remote sensing imagery via scalable spatio-temporal kriging, 2019.
5. **E. Semenova** and P. Vounatsou, Wombling methodology to identify hotspots on gridded disease surfaces, 2019.

## Previous work experience

2017 – 2018 **Board Member and Educator**, women++ ([www.womenplusplus.ch](http://www.womenplusplus.ch)), Zurich, Switzerland

- Ideated and organized the first female-friendly hackathon in Switzerland for over 100 participants. In 2019 it was included into the list of 80 most important diversity-related initiatives in Europe.
- Designed and tutored a series of programming workshops; organized tech- and career-related professional training, and networking events to foster entrepreneurship and leadership.

2016 – 2016 **R Developer**, remote

- Created R-scripts and web services enabling statistical analysis of clinical data.

2012 – 2014 **Teaching Assistant in Mathematics**, University of Basel, Switzerland

- Harmonic Analysis, Partial Differential Equations, Calculus (graduate and undergraduate levels)

2010 – 2011 **Actuarial Consultant**, PriceWaterhouse Coopers, Moscow, Russia

2008 – 2010 **Actuary**, Zurich Insurance Company, Russia / Germany / Switzerland

- Statistical data analysis, data mining, risk modelling, pricing tool development, due diligence