François COLLIN, PhD

Statistical consultant

Bażantów 41G/3
40-668 Katowice, Poland
0048 787 906 750
FCACollin.github.io/Latarnia
31 years old

Statistician with experience in biostatistics
– Medical sciences, genetics, plant physiology, geostatistics –

Experience

Today
April. 2020 **Senior R programmer**, *FCA Collin for Roche*, Remote/international team (from Warsaw to San Francisco) following **Agile process**, involved in R package development from design to execution, 6 months (ongoing).

- Programming R-base suite for statistical analysis of clinical data.
- From clinical standards STDM/ADaM and analysis to Tables, Listings and Graphs (TLG).
- R for visualisation: base graphs, ggplot2, code new grid elements, Shiny Apps for interactive applications and dashboards, interactive html/js graphs.
- R for statistics: function development for standard outputs (e.g. mixed model for repeated measurements, Cox PH regressions, binary response analysis).
- R in a project: package development with continuous integration and developement with Jenkins (CI/CD), git-branching workflow, team convention and good practices for robust code.
- Large autonomy granted, proactive attitute to create value was encouraged: new ideas, identification of risks/opportunitites, generate new designs and prototypes.

Today
Jan. 2020 **Independent consultant**, *FCA Collin for the Medical University of Bialystok*, Katowice Poland, 9 months (ongoing).

- **Offer and experience in the area of:** observational studies, real world evidence, study design and statistical analysis plans (SAP), biomarker identification, predictive algorithms, publication support, reporting automation.
- **Portfolio**: `https://fcacollin.github.io/Latarnia/Portfolio.html`
- **Recent misc. assignements**:
  - Clinical study - Investigate the relations between baryatric surgery, endocannabinoid systems and diabetes. Scope: SAP and execution for lab test / vital signs data and efficacy evaluation.
  - Clinical study - micro-RNA and diabetes. Scope: SAP for Next Generation Sequencing data and execution covering differential expression, classification, patient profiling, building biomolecular networks and research of genetic candidate biomarkers.
  - Real-World Evidence - Multiple Sclerosis (MS): analysis of public eletronic health records for the evaluation of MS therapies in Poland from 2014 to 2018. Two recent developments: the JCV seroconvertion for patient receiving natalizumab, the disease progression independent of relapse activity.

Dec. 2019
Aug. 2018 **Human Genetics/Expert in Sciences and Technology**, *the Medical University of Białystok (Poland)*, Centre for bioinformatics and data analysis, 1.5 years.
Research support for data analysis, machine-learning models and publication.

- **Lung cancer - Next Generation Sequencing (NGS)** for prognosis, **biomarker**, transcriptome for the prediction of type and stage of tumor implemented in an interactive and reactive prototype.
- **Multiple Sclerosis - Observational study** for post-registration therapy evaluation.
- **Diabetes - NGS/Clinical study**: research of therapeutic targets via analysis of RNA-seq data.
- **Teaching**: 8-hour practical for **RNA-seq data analysis**, PhD student co-supervisor.
- **Pharmacogenomics**: PGx-based contraindication prediction.
- **Personalised Medicine:** statistics, product development and prototyping for Imagene.me.

2020-10-08: I hereby agree for processing my personal data, included in my application for a job, for the purposes of recruitment process, according to the General Data Protection Regulation.

Mar. 2014 **R&D Researcher/Stat. modeler**, *Arvalis (Orsay, France)*, 1.5 years.
Sept. 2012 Improvement of market-leader Farmstar: satellite-image based management of wheat crops.

Education

Jan. 2018 **PhD, Dual Degree**, *The University of Nottingham/AgroParisTech (**UK/France**)*, 4 years.
Apr. 2014
- **Statistics:** Tolerance of wheat crops to STB. Study, protocol design, Statistical Analysis Plans.
- **Project management:** 1 meta-analysis + 3 experiments during 3 years in 2 countries with 3 institutes. Outcome: international conferences, peer-review articles, English-written thesis.

Sept. 2012 **Master of Applied statistics**, *Agrocampus Ouest (Rennes, France)*, 2 years.
Sept.2010 Content: 1.5-year learning program for applied statistics + 6-month graduation internship.

Skills

Methods **Statistics** from usual inferential techniques to state-of-the-art predictive/machine-learning algorithms, exploratory strategies, structure/unstructured data. **Statistical programming** (10 years experience). Relational **database**. Automated **reports and dashboards**.

Code/Soft.
- Senior programmer: **R, Shiny, rmarkdown, markdown** for statistics, reactive prototypes, automated reporting, reproducible science and R package.
- Daily user of **git, github** (branching workflow), CI/CD (Jenkins).
- Good knowledge of **SQL (PostgreSQL), LaTeX**, bash, html/css, HPC and much more.

Languages **Native French**; **Fluent English**; **Polish** (in progress).

Scientific productions (ORCID 0000-0003-0524-5755)

2020 **Article (peer reviewed)**, *K. Kapica-Topczewska, **F. Collin**, J. Tarasiuk, A. Czarnowska, M. Chorąży, A. Mirończuk, J. Kochanowicz, A. Kułakowska*, The JCV status, seroconversion rate and the risk of progressive multifocal leukoencephalopathy in Polish JCV seronegative patients with relapsing-remitting multiple sclerosis, European Neurology, accepted, soon published.

2020 **Article (peer reviewed)**, *K. Kapica-Topczewska, **F. Collin**, J. Tarasiuk, M. Chorąży, A. Czarnowska, M. Kwaśniewski, W. Brola, H. Bartosik-Psujek, M. Adamczyk-Sowa, J. Kochanowicz, A. Kułakowska*, Clinical and epidemiological characteristics of multiple sclerosis patients receiving disease-modifying treatment in Poland, Neurol Neurochir Pol, 54:161–168, `https://doi.org/10.5603/PJNNS.a2020.0020`.

2020 **Article (peer reviewed)**, *M. Niemira, **F. Collin**, A. Szalkowska, A. Bielska, K. Chwialkowska, J. Reszec, J. Niklinski, M. Kwasniewski, A. Kretowski*, Molecular Signature of Subtypes of Non-Small-Cell Lung Cancer by Large-Scale Transcriptional Profiling: Identification of Key Modules and Genes by Weighted Gene Co-Expression Network Analysis (WGCNA), MDPI, 12:37, `https://doi.org/10.3390/cancers12010037`.

2020 **Article (peer reviewed)**, *F. Padilla-Martínez, **F. Collin**, M. Kwasniewski, K. Adam*, Systematic Review of Polygenic Risk Scores for Type 1 and Type 2 Diabetes, MDPI, 21:1703, `https://doi.org/10.3390/ijms21051703`.

2019 **Article (peer reviewed)**, *K. Kapica-Topczewska, J. Tarasiuk, **F. Collin**, W. Brola, M. Chorąży, A. Czarnowska, M. Kwaśniewski, H. Bartosik-Psujek, M. Adamczyk-Sowa, J. Kochanowicz, A. Kułakowska*, The effectiveness of interferon beta versus glatiramer acetate and natalizumab versus fingolimod in a Polish real-world population, PLOS ONE, 14:1–12, `https://doi.org/10.1371/journal.pone.0223863`.

2020-10-08: I hereby agree for processing my personal data, included in my application for a job, for the purposes of recruitment process, according to the General Data Protection Regulation.

2019 **Conference: ESHG 19**, *Gothenburg (Sweden)*, **F. Collin**, M. Niemira, A. J. Krętowski, J. Nikliński, M. Kwaśniewski, Poster: User-friendly and machine learning-empowered platform for classification of NSCLC based on RNA-seq profiling, `https://fcacollin.github.io/Latarnia/doc/poster_190603.pdf`, Awarded fellowship for European Countries.

2018 **Ph.D Manuscript**, ***F. Collin***, The tolerance of wheat (*Triticum aestivum* L.) to Septoria tritici blotch, University of Nottingham/AgroParisTech, `https://pastel.archives-ouvertes.fr/tel-02443529/document`.

2018 **Article (peer reviewed)**, ***F. Collin***, *P. Bancal, J. Spink, P. Kock Appelgren, J. Smith, N.D. Paveley, M.O. Bancal, M.J. Foulkes*, Wheat lines exhibiting variation in tolerance of Septoria tritici blotch differentiated by grain source limitation, Field Crops Research, 217:1–10, `https://doi.org/10.1016/j.fcr.2017.11.022`.

2018 **Conference: Phloème 2018 - 1ère biennales de l'innovation céréalière**, *Paris (France)*, **F. Collin**, P. Bancal, M.J. Foulkes, M.O. Bancal, La tolérance du blé à la septoriose, awarded "Promising thesis".

2016 **Conference: ESA 14**, *Edinburgh (UK)*, **F. Collin**, P. Bancal, J. Foulkes, M.O. Bancal, Poster: A statistical analysis of GxE contribution to leaf senescence during grain filling in wheat, `https://fcacollin.github.io/Latarnia/doc/Collin2016_ESA_poster.pdf`.

2016 **Conference: 9th ISSDC**, *Paris*, **F. Collin**, D. Gouache, M.O. Bancal, P. Bancal, Poster: Tolerance of wheat to Septoria tritici blotch: genetic vs environmental variations of key traits.

2015 **Article (peer reviewed)**, *P. Bancal, M.O. Bancal,* ***F. Collin****, D. Gouache*, Identifying traits leading to tolerance of wheat to Septoria tritici blotch, Field Crops Research, 180:179–185, `https://doi.org/10.1016/j.fcr.2015.05.006`.

2012 **Master thesis**, ***F. Collin***, Modelling the Nitrogen Use Efficiency of an inorganic nitrogen fertilizer application on winter wheat, `https://dumas.ccsd.cnrs.fr/dumas-00741001/en/`.

2012 **Conference: R meeting**, *Bordeaux*, G. Bessigneul, **F. Collin**, M. Gauthier, M. Gérard, S. Lê, Lightning talk: A method for ecological data mining, `r2012.bordeaux.inria.fr/recueil_resumes_R2012.pdf`.

2011 **Misc: Sensometric school project**, *Rennes*, **F. Collin**, M. Gauthier, Sensorial analysis methods: the 'tablecloth' method, `sensominer.free.fr/desc-napping.html`.

2011 **Article (peer reviewed)**, *F. Montfort, S. Poggi, S. Molière,* ***F. Collin****, E. Lemarchand, D.J. Bailey*, Opportunities to reduce *Rhizoctonia solani* expression on carrots by biofumigation with indian mustard, Acta Horticulturae (ISHS), 917:149–157, `https://doi.org/10.17660/actahortic.2011.917.19`.

2020-10-08: I hereby agree for processing my personal data, included in my application for a job, for the purposes of recruitment process, according to the General Data Protection Regulation.