# Connie Nakasi Musisi

Date of birth: 19/02/1996 │ Nationality: Ugandan │ Phone: (+32) 467666240 (Mobile) │ Email address: conniemusisi6@gmail.com │ Work permit: Belgium │ Website: https://github.com/Connie-Musisi │ Address: Helbeekplein 6, bus 7, 3500, Hasselt, Belgium (Home)

## About me

I am a Biostatistician passionate about the application of statistical methods to improve daily living with a focus on ethical, impact-driven pharmaceutical research. I turn clinical data into clear, reproducible analyses that stand up to regulatory review using statistical programming tools while following clinical data standards (CDISC/SDTM/ADaM). I take initiative, communicate plainly, and work well both independently and with cross-functional teams, translating complex results into practical actions for diverse stakeholders. My research interests and focus areas;

Statistical methods & modelling: linear, logistic & Poisson regression, ANOVA/ANCOVA, mixed-effects modelling, Design of Experiments (DOE), Dose-Response modelling, survival data analysis, longitudinal & Clustered, multivariate regression, Bayesian hierarchical models, Non-linear & spline/GAM models, Causal inference, Data Integration Techniques, Simulations, multiplicity control, missing data, specification setting (confidence & tolerance intervals), comparability assessments, control charts.

Infectious Disease Modelling: Compartmental and Stochastic modelling, Time Series Analysis in Infectious Diseases, Geospatial Analysis.

Statistical Computing: Development of Statistical Software with R, RShiny, SAS (Macros), JMP, JAGS and Python.

Bioinformatics: High Dimensional Data Analysis, Big Data analytics, Data Integration and Multi-Omics, Statistical methods for Bioinformatics.

## Work experience

**PhD researcher** │ GSK Biologicals │ 06/12/2021 - 30/06/2022 | Rixensart, Belgium

- Contributed to protocol development, providing input on endpoints/estimands, objectives, analysis sections) and Case Report Form (CRF) alignment.

- Developed Statistical Analysis Plan (SAP) components, including analysis specifications and sensitivity analyses.

- Integrated longitudinal clinical and molecular data from COPD patients, applied time-to-event and mixed-effects alongside Machine Learning tools for feature selection

- Produced ADaM-aligned analysis structures with traceability to raw data

- Delivered tables, listings and figures (TLFs) and model summaries for study teams and contributed to quality control (QC)

- Explained methods and results to cross-functional partners

**Junior Statistician** │ Uganda Virus Research institute │ 01/09/2017 - 31/08/2018 | Entebbe, Uganda

- Ran Genomewide Association Studies (GWAS) for immunerelated diseases including eczema, chronic kidney disease among others

- Built models for infectious disease research

- Performed literature reviews, prepared SAPs and analysis summaries

**Teaching assistant** │ Uganda Mathematical Society │ 15/03/2017 - 31/08/2017 | Kampala, Uganda

- Delivered undergraduate lectures (Differential Equations, Dynamical Systems and Biomathematics), assessments and tutorial support

## Education & Training

**Ph.D. Biostatistics** │ Hasselt university │ 01/07/2021 - Current │ Hasselt, Belgium

- Development of adaptive methods for differential abundance analysis in Microbiome studies

- Development of statistical programming tools (R packages) for Microbiome research

- Interdisciplinary research, research communication and presentations, and reproducible workflows

- Consultancy on COPD project for GSK Biologicals

- Teaching (Linear models and SAS Programming, Introduction to Mathematics for Statisticians)

**Field of study** Statistics | **Level in EQF** 8

**Master of Statistics: Biostatistics** │ Hasselt university │ 13/09/2018 - 04/09/2020 │ Hasselt, Belgium

**Level in EQF** 7 | **Thesis:** Longitudinal data analyses to investigate the effect of phloretin in reducing inflammation in a EAE mouse model

**Bachelor of science (Mathematics and Statistics)** │ Makerere University Kampala │ 10/08/2014 - 31/05/2017 │ Kampala, Uganda

**Level in EQF** 6

## Projects

### An Adaptive Test for Differential Abundance in Microbiome Studies │ 01/01/2024 - 28/02/2025

Developed a novel adaptive test for differential abundance in microbiome studies. Unlike traditional approaches, our method constructs a training dataset to label taxa as differential abundant or not, and then fits a linear model using ordered normalized counts. P-values are obtained from a permutation null distribution and adjusted for FDR, resulting in improved FDR control and competitive sensitivity compared to existing methods. This method is published in the R package ADATEST.

https://github.com/Connie-Musisi/ADATEST.git

### ANCOM-LL: a Log-Linear Variant of the ANCOM-BC Method │ 01/02/2022 - 30/06/2023

Developed a novel method to test for differential abundance in microbiome studies. This method is an extension of ANCOM-BC, a popular R package for analysis of microbiome data. The new method corrects for bias due to differential sampling fractions between groups in microbiome studies, using a log-linear model to better handle sparse microbiome datasets, which showed improved FDR control.

### Analysis of multi-omics COPD data │ 06/12/2021 - 30/06/2022

A consultation for GSK Biologicals, where we integrated longitudinal multi-omics datasets (microbiome, gene expression, clinical data) to develop prediction models for disease status and exacerbations in patients. We also implemented methodologies for feature selection, prediction model evaluation, and latent association analysis using statistical and machine learning techniques.

### Adaptive benchmark and choose method for Microbiome differential abundance analysis │ 01/10/2025 - Current

Due to the nature of microbiome data, there is no single method that can account for all the characteristics. However, how do we know what the best method is for a particular dataset. This is where my project came in and we developed a method that selects the most reliable differential abundance analysis method to use for a given dataset. It simulates realistic variations of the Original data, evaluates competing methods against rigorous error-control criteria, and then applies the best-performing method while correcting for multiple testing to control the FDR. This method is published in the R package ADAM.

## Honours and Awards

**Bijzonder Onderzoekfonds** │ Hasselt University Belgium

**ICP Scholarship to pursue Master's degree** │ VLIR-UOS

**Government Scholarship to pursue Bachelor's degree** │ Government of Uganda

## Language Skills

Mother tongue(s): **Luganda**

| | Understanding | | Speaking | | Writing |
|---|---|---|---|---|---|
| | Listening | Reading | Spoken production | Spoken interaction | |

| | | | | | |
|---|---|---|---|---|---|
| **English** | C2 | C2 | C2 | C2 | C2 |
| **Dutch** | A2 | A2 | | A2 | A1 |
| **French** | A1 | A1 | | A1 | A1 |

## Hobbies and Interests

### Leisure
Watching crime series and travelling

### Sport
Running, swimming and ice skating

## Trainings & Certifications

**SAS Base Programming for SAS 9, SAS Institute (2025)**

**Medical Writing, Emtex (2023)**

**R programming, Johns Hopkins University on Coursera (2022)**

**Introduction to High Performance computing, Vlaamse Supercomputers, Hasselt University (2022)**

**Introduction to Python, UDEMY (2022)**

## Software

### ADATEST
An R packages designed for testing for differential abundance of microbial taxa between two groups.

https://github.com/Connie-Musisi/ADATEST.git

## Skills

### Mathematics and Statistics
Longitudinal and clustered data analysis

Mixed effects modelling

Survival data analysis

Non-linear modelling

Infectious disease modelling

Bayesian data analysis

Data visualization

Machine Learning techniques

### Programming Skills
R (RMarkdown, tidyverse, ggplot2, Shiny)

JAGS (Bayesian Analysis)

SAS (Base, Clinical)

Python (pandas, NumPy)

Latex

STATA

SQL

Linux

**AI & Machine Learning tools**

ChatGPT

**Communication, Interpersonal and Organizational skills**

Public speaking and good presentation skills, team player, good time management, adaptive learner and goal oriented