



**SSACAB**

SUB-SAHARAN AFRICA CONSORTIUM  
FOR ADVANCED BIOSTATISTICS TRAINING



SEP 2025 TO JAN 2026

# Sub-Saharan Africa Consortium for Advanced Biostatistics (SSACAB) II Newsletter.

Strengthening Africa's brain  
capital through data

.....  
Researchers and policy  
makers engagement workshop

.....  
Biostatistics career event

# Newsletter

## Sep 2025 to Jan 2026

As we reflect on the past few months and have already sprinted into 2026, we continue our abiding mission: grow Africa's data capabilities to solve the continent's most pressing health challenges.

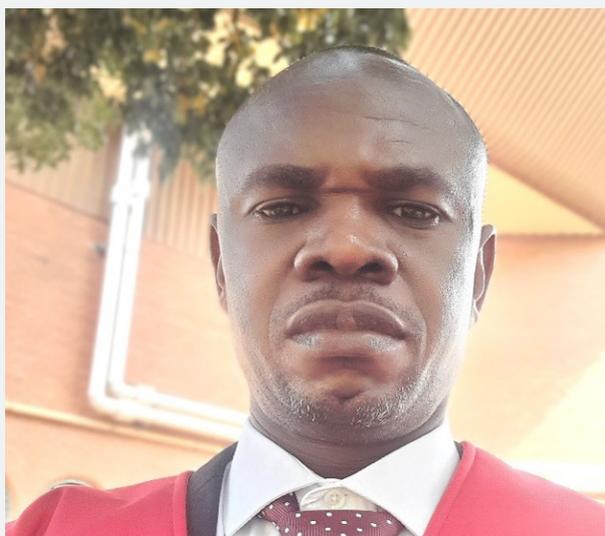
**O**ur highlight reels continue. A significant milestone was reached when SSACAB 2 recruited its 69<sup>th</sup> fellow! We are growing in our network of African universities.

By supporting early career researchers, SSACAB continues to build a pipeline of African biostatisticians who are not only technically skilled but also embedded in local health challenges and research priorities. This investment strengthens long-term research capacity across the continent.

## Spotlight on SSACAB post docs

Dr Mohammed is a SSACAB fellow from UKZN. Under Professor Henry Mwambi as his supervisor, Dr Mohammed's research focus is on "Advanced imputation methods for missing data in complex high-dimensional correlated health data using deep learning and dimensionality reduction techniques."

Dr Mohammed has previously dived into data to decode the complexities of the stock market. Starting with a foundational degree in Statistics and Computer Science, he progressed through rigorous academic training to now researching at the cutting edge of statistical and deep learning.



*Dr Anas Eisa Abdelkreem Mohammed*

Whether presenting findings on global stages or mentoring students in various statistical methodologies, he thrives on challenge and continuous learning. Over the years, he has developed a robust skill set in data analysis, excelling in SPSS, Excel, R, and Python. This technical proficiency, combined with his role as a lecturer, allows him to translate complex statistical concepts into understandable terms and actionable insights.

# Conferences and Thematic areas Of excellence

Health, Environment and Agronomy using  
Structural Equation Modelling:  
A workshop held in Benin.



As part of its partnership with the Sub-Saharan Africa Advanced Consortium for Biostatistics (SSACAB), LABEF organised a training course on Structural Equation Modelling for applications in health, environment and agronomy.

More than 300 young researchers from 18 African countries took part in the advanced statistical and methodological training.

The first day focused on methodological principles and illustrated through case studies in agriculture, health, and environmental sciences and ran a practical component to apply R to define latent variables using the EFA model. This session

was delivered by Dr Souand Tahi (SSACAB Postdoctoral Fellow), Midokpè Merveille Essetcheou (SSACAB PhD Fellow), Kangela Matazi Alain (SSACAB PhD fellow) and Dr Salako.

Day 2 combined theoretical foundations with practical demonstrations in R, and participants were introduced to Structural Equation Modeling.

Structural Equation Modeling allows researchers to rigorously analyse complex, real-world relationships across health, agriculture, and environmental systems therefore, this short course is of great importance to researchers.

# Bayesian Modelling Using R at a workshop in Tanzania



**P**rofessor Eustasius Musenge, Deputy Director and co-principal investigator of the SSACAB and Professor Innocent Maposa a SSACAB collaborator based at Stellenbosch University, served as facilitators at the recent “Bayesian Modelling using R” workshop in Tanzania. Hosted by the Catholic University of Health and Allied Sciences (CUHAS) in Mwanza, the workshop

presented advanced statistical methods to over 30 participants from across Africa.

Bayesian statistics is a way of combining what we already know with new data to get a better, more realistic answer. Traditional statistics only looks at the data collected in a study, but Bayesian methods blend the information from the data with any useful prior knowledge.

# Advanced Machine Learning

## Short course in South Africa



We welcomed Prof. Ruurd Kuiper from Utrecht University in the Netherlands. Utrecht is a northern partner institution of SSACAB's. He spent two impactful weeks in Johannesburg, South Africa, first assisting the Wits team (Professor Musenge, Dr Awol Ebrie, Dr Okechinyere Achilonu and Mr Sam Chikasha) on the Advanced Machine Learning course and then went on to teach, the Large Language Models (LLMs) for Healthcare course. This innovative course explored the full lifecycle of LLMs

in healthcare, from the inner workings of transformers to training and inference pipelines, while examining their unique capabilities and potential risks. Participants also engaged in hands-on coding sessions and a final project where they trained, fine-tuned, and implemented their own models. Prof. Ruurd Kuiper expressed deep gratitude for the opportunity to teach, connect, and collaborate with such a motivated and supportive group of participants and colleagues.

# Illustrious Visits

It was a pleasure to welcome Professor Ann Mwangi, one of the Consortium's Co-Principal Investigators, during her visit to Wits from Moi University in Kenya. Prof Mwangi was in South Africa as part of the Consortium for Advanced Research Training in Africa (CARTA) facilitating team, in November 2025 meeting. CARTA supports the development of a vibrant African academy able to lead world-class multidisciplinary research that impacts positively on public and population health. During her visit, she highlighted the importance of sustained and wide-ranging



training programmes—short courses, seminars, network exchanges, and collaborative supervision—designed to prepare early-career biostatisticians to work confidently with complex datasets.

The SSACAB secretariat hosted Professor David MacLeod, Associate Professor at the London School of Hygiene of Tropical Medicine (LSHTM) and one of the consortium's Northern partners. LSHTM is one of SSACAB's two Northern partners and plays an essential role in helping the consortium expand advanced biostatistics training, co-supervision, and mentorship across sub-Saharan Africa.



# Achievements

SSACAB PhD Fellow,  
Sarah Ogutu, recognised  
at two prestigious national  
research events.

**S**arah Ogutu, a SSACAB PhD Fellow at the University of KwaZulu-Natal, has recently distinguished herself on two major academic stages, earning recognition for both the quality of her research and her growing presence within South Africa's statistical community. Her achievements at the 2025 Postgraduate Research and Innovation Symposium (PRIS) and the 66th Annual Conference of the South African Statistical Association (SASA) underscore her promise as an emerging biostatistician.

At PRIS 2025, hosted by the College of Agriculture, Engineering and Science at UKZN, Sarah was awarded First Prize for Oral Presentation in the School of Mathematics, Statistics, and Computer Science, aligned with the theme: "International Year of Quantum Science and Technology". The award acknowledges the impact of Sarah's PhD research on advancing statistical and machine-learning approaches for health data. "This award marks a significant milestone in my career, bringing further depth to my work and motivation to pursue additional collaborations and high-impact contributions to public health research," she said.

Shortly after her success at PRIS, Sarah presented her doctoral work at the South African Statistical Association conference, which combined plenary lectures, specialised research streams and dedicated sessions for young statisticians.

As an oral presenter, Sarah shared advances from her PhD work on longitudinal and survival modelling. The feedback she received from



experts deepened her understanding of the methodological strengths and refinements needed in her study. The engagement that followed her talk strengthened her scientific communication skills and extended the analytical direction of her work. She was exposed to current trends in statistical modelling and encouraged to think creatively about how these methods can be applied in public health contexts. "A memorable highlight of the SASA programme was interacting with Hadley Wickham, one of the world's leading data science researchers and chief scientist at Posit PBC. It will open up new mentorship and networking opportunities."

Sarah attended Wickham's workshop on large language models, where she learned about his newly developed R-based interface, ellmer, designed to integrate LLMs seamlessly into data science workflows. Meeting Wickham, whose tidyverse ecosystem has shaped much of modern R-based data analysis, was a formative moment that strengthened her sense of belonging within the global data science community. Together, these two achievements reflect a period of significant professional growth for Sarah.

## Strengthening Africa's brain capital through data: SSACAB at the G20 Brain Health Side Meeting



**A**t the G20 Ministerial meeting on Brain Health in November 2025, co-hosted by the Davos Alzheimer's Collaborative and the Science for Africa Foundation, SSACAB was represented by

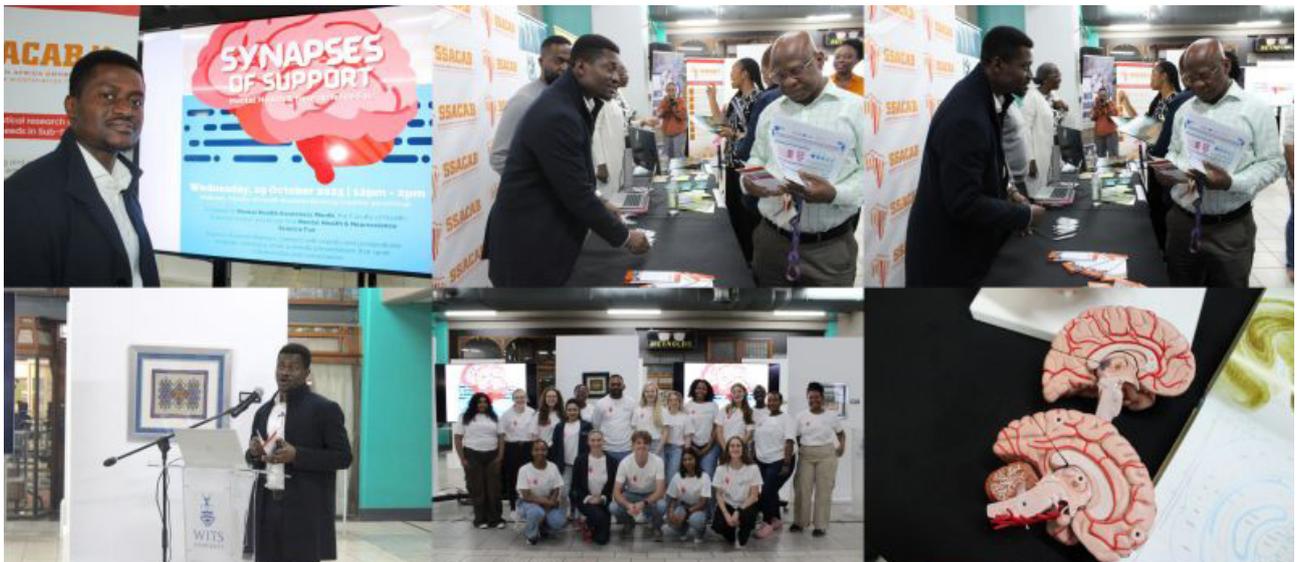
Professor Eustasius Musenge and Caroline Vika. Their participation underscored the central role of biostatistics and data science in shaping Africa's next generation of health and economic strategies.

## AI in Health– Uganda

**M**utale Sampa, a SSACAB PhD Fellow, attended the AI in Health Conference 2025 in Uganda. She was inspired by the incredible work taking place across Africa in the AI and health innovation space, and by the collaboration that continues to drive impactful change. Together, these efforts are reshaping the narrative around technology and healthcare in Africa, building a future where innovation and inclusion go hand in hand.



## Mental Health Matters



**S**SACAB II was proud to participate in the Wits Faculty of Health Sciences's Mental Health and Neuroscience Fair. This was a dynamic space connecting students, researchers, and practitioners dedicated to advancing mental well-being. Our PhD Fellow, Godwin Kalu, presented on how biostatistics can illuminate mental-health challenges across Sub-Saharan Africa. By analysing structured and unstructured

health data, biostatistical models reveal patterns, risk factors, and intervention outcomes. This helps design evidence-based, community-driven mental-health solutions. Godwin's own PhD work explores how AI text-generation tools can deliver personalised mental health and alcohol education messages to South African university students.

# Malawi researchers and policy makers engagement workshop

In October 2025, SSACAB was part of a momentous moment in Malawi's health history: the country's first-ever Health Policy and Biostatistics Research Engagement Workshop. Traditionally, policymakers and researchers have missed critical opportunities to collaborate on implementing policies that address the country's persistent health challenges; however, they have now come together to strengthen their relationships for meaningful collaboration. SSACAB cofounded the event which brought together biostatisticians, researchers and policymakers with the aim

of creating a platform to collaborate on health policy formation and implementation. The workshop was attended by Professor Tobias Chirwa, Director of SSACAB; Caroline Vika, Programme Manager; Dr Halima Twabi and Dr Fiskani Kondowe, PhD both SSACAB alumni; Dr Tsirizani Kaombe, a SSACAB collaborator; and Geoffrey Singini, a SSACAB PhD Fellow. In attendance were key academics from University of Malawi - UNIMA, the National Planning Commission representatives, Malawi Ministry of Health representatives and key Malawian researchers. This event truly marked the



first structured effort to bring together these diverse actors in dialogue. On Day 1, Professor Chirwa delivered the keynote address on the discipline of biostatistics and its relevance to the field of health.

Prof Chirwa's message underscored that the discipline is not only technical but also strategic—central to how societies plan, prioritise, and protect public health. Across the discussions, a clear message emerged. Policymakers noted that research findings are often produced but not translated into actionable insights that can guide programmes or legislation. Researchers, on the other hand, observed that governments tend to work in silos, pursuing distinct objectives without shared mechanisms for collaboration. This disconnect, they

agreed, limits the national impact of critical evidence being used in policy making and ultimately, in health settings. Yet there was a broad consensus that this could change. Biostatistics provides a crucial bridge between evidence and policy. When properly interpreted and communicated, data can illuminate where systems are working effectively, where gaps persist, and how interventions can be adapted to maximise public benefit. This inaugural workshop marks a significant milestone, which SSACAB is proud to have been a part of. It signals a growing recognition that partnerships between universities, policymakers, and technical experts are essential for developing responsive and equitable health systems.

## News and press

### *SSACAB has been prolific in the press recently*

Click the links below to here to read and listen to our feature articles and radio interviews on diverse topics

Edson Mwebesa from Moi University, Kenya, a SSACAB Fellow knows children who died from late-diagnosed malaria. His research has also revealed that malaria is more prevalent in pregnant Ugandan women than in any other population. Malaria is also endemic across Africa, impacting all aspects of social and economic life.



<https://www.wits.ac.za/news/latest-news/research-news/2025/2025-11/is-malaria-messaging-working-african-biostatisticians-have-the-data.html>

Dr Geoffrey Singini, A University of Malawi SSACAB fellow led on an interesting research. The research show cases that African biostatisticians have offered a powerful, data-driven alternative that can accurately inform HIV policy, reduce healthcare costs, and save more lives. HIV and tuberculosis (TB) treatment programmes have always relied on clinical trials to understand whether treatment interventions are actually working, but they're costly and cumbersome. Meanwhile, African biostatisticians have offered a powerful, data-driven alternative that can accurately inform HIV policy, reduce healthcare costs, and save more lives.



<https://www.wits.ac.za/news/latest-news/research-news/2025/2025-12/real-world-data-gives-africa-a-smarter-path-to-fighting-hiv-and-tb.html>

Professor Romain Gelele Keke a Co PI of SSACAB based at the University of Abomey Calavi, Benin, led on a study that found that flies can transmit cholera-causing pathogens. Africa carries the largest cholera burden, which is why an oral cholera vaccine trial is timeous. The study focussed on the role that flies play in spreading cholera – a bacterial disease that causes severe diarrhoea and dehydration – has been underestimated in Africa, which carries the largest burden of the disease.

Professor Romain Glele Kakaï is interviewed on ENCA:



<https://www.enca.com/news/study-finds-houseflies-accelerate-spread-cholera>



<https://www.wits.ac.za/news/latest-news/research-news/2026/2026-01/flies-and-the-hidden-drivers-of-cholera.html>

Caroline Vika interviewed by Benin public TV about public engagement

At the Public Engagement with Science event at the University of Abomey-Calavi in Benin in late January, SSACAB Programme Manager, Caroline Vika was interviewed by UAC Web TV about the importance of raising awareness of biostatistics to young people. Biostatistics is an important career in the context of pressing and emerging health issues on the continent. Caroline spoke about the international network of biostatisticians working towards a robust data future for Africa.



<https://youtu.be/nDa3UH4dPJO>

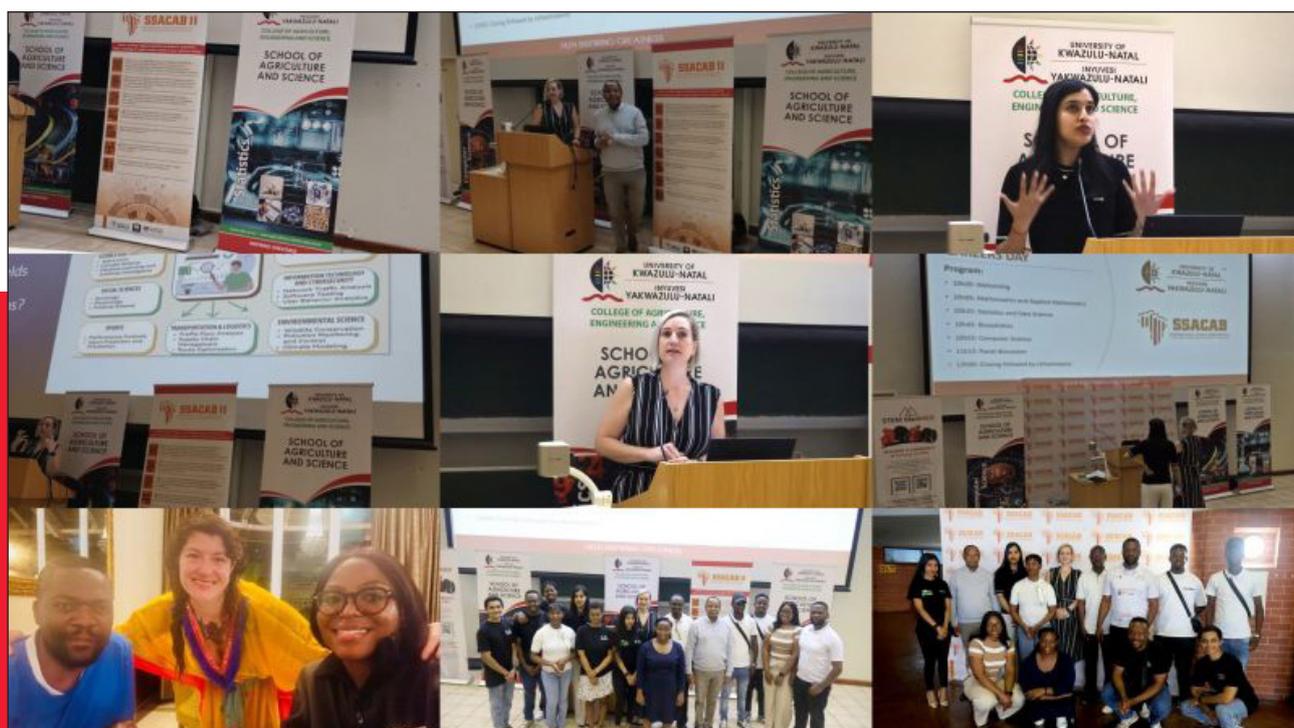
# Public Engagement with Science (PES)

Sub-Saharan Africa Consortium for Advanced Biostatistics (SSACAB) has been involved in various PES events, including monthly workshops. The activities were as follows:

- In September 2025, SSACAB conducted a PES monthly workshop, whereby the concept of PES guidelines was reintroduced to the Masters fellows, PhD, and Postdoctoral fellows, as well as their supervisors, principal investigators, and the broader SSACAB community. The workshop also reviewed how science communication is integrated into fellows' research projects, including an examination of specific activities related to community engagement, public engagement, and policy engagement.
- In October 2025 SSACAB in partnership with the University of KwaZulu-Natal (UKZN) and the College of Agriculture,

Engineering and Science, facilitated an exciting STEM career awareness event for learners in Grades 9–11 held in Durban. The goal was simple but powerful: to inspire young people to make informed subject choices that open pathways to careers in Science, Technology, Engineering, and Mathematics (STEM). Speakers demonstrated how mathematics and statistics underpin almost every aspect of our daily lives – from music, gaming, and artificial intelligence to cosmology, GPS, and satellite technology. Through SSACAB's involvement, learners were introduced to the emerging and impactful field of biostatistics – a discipline that applies statistical reasoning to solve real-world health challenges.

Here are some of the highlights in the picture below:



In November 2025, during the monthly PES workshop, Individual Policy Exchange Programme (iPEP) fellows presented progress updates and key insights from SSACAB policy-related projects, highlighting the importance and impact of the iPEP fellowship.

In addition, the SSACAB team attended the pre-launch event of a clinic digital booking

system for linkage to care. This community engagement activity took place at Stanza Bopape Health Care Centre in Pretoria, Gauteng Province, South Africa. The purpose of the event was to engage community members within the Stanza Bopape Health Care Centre catchment area and introduce them to the new digital system.



A public engagement seminar themed “Biostatistics: Concepts, Applications, and Professional Perspectives” was held at the Idriss Déby Amphitheatre of the University of Abomey-Calavi in Benin. Organized by LABEF in collaboration with SSACAB on the 31st of January, 2026. The event brought together approximately 200 participants from academic, scientific, and professional communities.

The objectives of the seminar were to: (i) deepen understanding of biostatistics as a discipline; (ii) highlight the diversity of available academic and professional pathways; and (iii) foster meaningful connections among learners, students, and professionals.

The opening of the seminar was marked by keynote addresses delivered by the SSACAB Programme manager (Caroline Vika), the Director of LABEF (Prof. Romain Glele Kakai), the Dean of the Faculty of Agronomic Sciences (Prof. Nadia Fanou), and the Vice-Rector in Charge of Scientific Research (Prof.





Aliou Saidou). Interactive discussions and testimonials from alumni of the master's programs in Biostatistics and the PhD program in Biometry at the University

of Abomey-Calavi further enriched the exchanges and stimulated strong interest in the fields of Biostatistics and Biometry.

# Number of SSACAB II Fellows: 69

**University of Abomey-Calavi (SSACAB) has enrolled 8 fellows**

Masters: 4    PhD: 3    Post Doc: 1    Completed: 3

**University of Pretoria (SSACAB) has enrolled 7 fellows**

Masters: 3    PhD: 3    Post Doc: 1    Completed: 4

**Moi University (SSACAB) has enrolled 6 fellows**

Masters: 2    PhD: 3    Post Doc: 1    Completed: 2

**University of KwaZulu-Natal (SSACAB) has enrolled 9 fellows including 3 fellows from Kilimanjaro Christian Medical University College**

Masters: 4    PhD: 3    Post Doc: 2    Completed: 3

**KEMRI Wellcome Trust Research Programme (SSACAB) has enrolled 6 fellows**

Masters: 3    PhD: 2    Post Doc: 1    Completed: 3

**University of Witwatersrand (SSACAB) has enrolled 33 fellows including partially funded fellows in collaboration with the following institutions: University of Namibia; University of Nairobi; University of Makerere; University of Namibia and Utrecht University**

Masters: 23    PhD: 7    Post Doc: 2    Completed: 15 (includes partial funding fellows)

# Publications

2025

---

Adéoti, O.M., Diop, A., Hounkonnou, M.N. & Glèlè Kakaï, R. (2025) 'Bayesian flexible multilevel nonlinear models in infectious disease modeling using non-informative priors', *Biostatistics & Epidemiology*, 9(1), e2505290. <https://doi.org/10.1080/24709360.2025.2505290>

---

Adéoti, O.M., Diop, A., Hounkonnou, M.N. & Glèlè Kakaï, R. (2025) 'Bayesian flexible multilevel nonlinear models in infectious disease modeling using non-informative priors', *Biostatistics & Epidemiology*, 9(1), e2505290.

---

Anteneh, L.M., Hounkonnou, M.N. and Glèlè Kakaï, R., 2025. A stochastic continuous-time Markov chain approach for modeling the dynamics of cholera transmission: Exploring the probability of disease persistence or extinction. *Mathematics*, 13, p.1018.

---

Anteneh, L.M., Zanzo, S.D., Traore, K. & Glèlè Kakaï, R. (2025) 'Modelling the impact of vaccination on cholera transmission dynamics under stratified populations and seasonality', *Infectious Disease Modelling*, 10(4), pp. 1138–1152.

---

Beh Mba, R. & Glèlè Kakaï, R.L. (2025) 'Bayesian modeling of geostatistical survival data with misaligned covariates: A simulation-based study of malaria risk in Gabon', *Journal of Probability and Statistics*, 2025, Article ID 1277268, 13 pages.

---

Bere, A., Maposa, I., Matsena-Zingoni, Z. et al., 2025. Modeling timing of sexual debut among women in Zimbabwe using a Geoaddditive Discrete-Time survival approach. *BMC Women's Health*, 25, p.137. Available at: <https://doi.org/10.1186/s12905-025-03590-w>

---

Glèlè Kakaï, R. & Djomatin, L.C. (2025) 'An age structured mathematical model to assess the combined effects of vaccine and non vaccine interventions on malaria transmission and burden', *Infectious Disease Modelling*, 11(1), pp. 355–376.

---

Hyle, E.P., Maphosa, T., Rangaraj, A., Feser, M., Singini, G.C., Shrestha, P., Shroufi, A., Reddy, K.P., Matiya, E., Dambe, R., Talbot, V.R., Chamanga, R., Horsburgh, C.R., Weinstein, M.C., Nyirenda, R.K., Ford, N., Tiam, A., Phillips, A. & Freedberg, K.A., 2025. Clinical impact and cost-effectiveness of the WHO-recommended advanced HIV disease package of care. *The Lancet Global Health*. Available at: [https://doi.org/10.1016/S2214-109X\(25\)00190-1](https://doi.org/10.1016/S2214-109X(25)00190-1)

---

Kallah-Dagadu, G., Mohammed, M., Nasejje, J.B. et al., 2025. Breast cancer prediction based on gene expression data using interpretable machine learning techniques. *Scientific Reports*, 15, p.7594. Available at: <https://doi.org/10.1038/s41598-025-85323-5>

---

---

Mabasa, V., Seheri, M.L. and Magwira, C.A. (2025) 'Expression of salivary hepcidin and its inducer, interleukin 6 as well as type I interferons are significantly elevated in infants with poor oral rotavirus vaccine take in South Africa', *Frontiers in Immunology*, 16, 1517893, pp. 1–8. <https://doi.org/10.3389/fimmu.2025.1517893>

---

Makonokaya, L., Ngwira, P., Kalitera, L.U., Singini, G.C., Khumbanyiwa, M., Maphosa, T. and Bhatt, N.B., 2025. Tuberculosis preventive treatment uptake among patients initiating antiretroviral therapy in Malawi: Children left behind. *Southern African Journal of HIV Medicine*, 26(1), a1760. Available at: <https://doi.org/10.4102/sajhivmed.v26i1.1760>

---

Makonokaya, L., Ngwira, P., Kalitera, L.U., Singini, G.C., Khumbanyiwa, M., Maphosa, T. & Bhatt, N.B. (2025) Tuberculosis preventive treatment uptake among patients initiating antiretroviral therapy in Malawi: Children left behind. *Southern African Journal of HIV Medicine*, 26(1), a1760.

---

Mwebesa, E., Agbi, D.M., Iha, D.T., et al. (2025) 'Effect of exposure to malaria messages on insecticide-treated net use among women and under-five children in Uganda: a propensity score matched analysis', *Malaria Journal*, 24, article 386.

---

Mwebesa, E., Musinguzi, B., Legason, I.D., Opoke, R., Agaba, B.B., Kananura, R.M. & Mwangi, A. (2025) 'Pooled prevalence and factors associated with insecticide-treated net use among pregnant women in malaria high-burden countries in sub-Saharan Africa: a multilevel mixed-effects analysis', *Research*, 53, article 166.

---

Ogutu, S., Mohammed, M. and Mwambi, H. (2025) 'Deep learning models for the analysis of high-dimensional survival data with time-varying covariates while handling missing data', *Discover Artificial Intelligence*, 5, 176. <https://doi.org/10.1007/s44163-025-00429-z>

---

Twabi, H.S. & Manda, S.O.M. (2025). Methods Using Propensity Scores to Estimate Causal Effects in Observational Longitudinal Studies. In: Kazembe, L.N., Kaombe, T.M. (eds) *Statistical Modelling for Evidence-Based Public Health*. Emerging Topics in Statistics and Biostatistics. Springer, Cham.

---

Voogdt-Pruis, H.R., Grobbee, D.E., Peters, S.A.E., Hehakaya, C., Jambroes, M., Klipstein-Grobusch, K. and Janssens, A. (2025) 'How do policymakers involve citizens in advancing health? A mixed-method qualitative study in municipalities in the Netherlands', *Frontiers in Public Health*, 13, p. 1708209.

---

Wamae, K., Magudha, J., Kakungu, A., Aricha, S., Langat, D., Kinyanjui, S., Mokaya, J., Thomson, N.R., Agoti, C. & Githinji, G., 2025. Cholera in Kenya: a scoping review of current research, evidence gaps and future directions. *bioRxiv*. <https://doi.org/10.1101/2025.09.16.25335755>

---

# SSACAB STATISTICAL CORNER: A BLOG

*Featuring Bongani Ncube (Masters in Biostatistics Wits University).*

## Introduction to Logistic Regression (LR)

Logistic regression is a simple yet powerful model to predict binary outcomes like (yes/no) scenarios. That is, whether something will happen or not. It's a type of classification model for supervised machine learning. One of the primary applications of logistic regression in health research is in disease prediction and risk assessment.



## Logistic Regression overview

Logistic regression is often used to predict the likelihood of an individual developing a specific disease based on various risk factors. For example, it has been extensively used in predicting heart disease, diabetes, and cancer. These conditions are influenced by multiple factors such as age, gender, lifestyle, family history, and underlying medical conditions. By incorporating these variables into logistic regression models, researchers and clinicians can estimate the probability of a patient developing these diseases, enabling early intervention and preventive care. In general, Logistic regression is characterized by research questions with binary (yes/no or success/failure) or binomial (number of yesses or successes in trials) responses for instance:

- a. is dying from influenza related to certain variables?
- b. Is exposure to a particular chemical associated with a cancer diagnosis?

The outcome variable is also known as a “target variable” or a “dependent variable”. There are many variables that could influence the outcome such as ‘Age’, ‘Gender’, ‘lifestyle’ etc. The influencing variables are known as features, independent variables, or predictors all these terms mean the same thing.

Other examples include whether a patient has diabetes or not and whether a drug will cure a patient or not.

## How does logistic regression work?

Logistic regression finds the best possible fit between the predictor and target variables to predict the probability of the target variable belonging to a labelled class/category. Linear

regression tries to find the best straight line that predicts the outcome from the features and uses optimization to try and find the best possible values of intercept and slope. Logistic regression works similarly, except it performs regression on the probabilities of the outcome being a category. It uses a sigmoid function (the cumulative distribution function of the logistic distribution) to transform the right-hand side of that equation. Again, the model uses optimization to try and find the best possible values of intercept and slope. Since the algorithm for logistic regression is very similar to the equation for linear regression, it forms part of a family of models called “generalized linear models”. This is why logistic regression has “regression” in its name, even though it is a classification model.

The sigmoid function resembles an S-shaped curve in the image below. It takes the real numbered input values and converts them between 0 to 1 (by shrinking from both sides, i.e., the negative values to 0 and very high positive ones to 1). Furthermore, the cut-off threshold is the overlaying deciding factor that bins the output into categories or classes when applied on top of these probabilities.

The complex concepts are best understood when explained with examples, so let us pick up an analogy to register the working of the LR algorithm. Let’s assume that the LR model is tasked to identify diagnosis of heart disease by looking at multiple indicators, such as the age, Cholesterol level, chest pain type etc. The objective is to determine the likelihood of being diagnosed of heart disease- which constitutes the target variable. The model assigns weights to the predictors based on how they impact the target variable and combines them to calculate the normalized score or probability of diagnosis. Read the step by step tutorial on building logistic model in R [Logistic regression in detail](#).



Please feel free to see what else we have been up to on our website and linkedin page. [SSACAB > Home](#)



QR for Website



QR for LinkedIn

For any queries, please contact Promise Majozi SSACAB administrator at [pmajozi@wits.ac.za](mailto:pmajozi@wits.ac.za)

