MSC BIOSTATISTICS

Our Journey:

2017-2021

DIVISION OF EPIDEMIOLOGY AND BIOSTATISTICS

UNIVERSITEIT
iYUNIVESITHI
STELLENBOSCH
UNIVERSITY
100
1918·2018

Acknowledgements

DELTAS

DELTAS Africa Initiative SSACAB (Grant No. 107754/Z/15/Z). The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS) Alliance for Accelerating Excellence in Science in Africa (AESA) and is supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust (Grant No. 107754/Z/15/Z) and the UK government.

NIH

Fogarty International Center of the National Institutes of Health under Award Number D43 TW010547.

Contents

Biostatistics

Biostatistics is the branch of statistics concerned with how we ought to make decisions when analyzing biomedical data. It is the evolving discipline concerned with formulating explicit rules to compensate both for the fallibility of human intuition in general and for biases in study design in particular.

The course offers rigorous training for those with a background or experience in quantitative or health-related discipline who wish to pursue a career in biostatistics. The MSc Biostatistics consists of structured modules and a research assignment.

Meet the MSc Biostatistics teaching staff

2017 - 2021

Dr Maxwell Chirehwa
Data Management and Statistical Computing

Mr Lovemore Sigwadhi
Data Management and Statistical Computing (Teaching Assistant)

Linear Models

Dr Alfred Musekiwa
Mathematical Statistics
Epidemiology & Biostatistics

Dr Merga Feyasa
Mathematical Statistics

Prof Taryn Young: HOD
Fundamentals of Epidemiology

Dr Chris Muller
Linear Models and Statistical Inference

Dr Elphas Okango
Linear Models and Longitudinal Data Analysis

Dr Innocent Maposa
Principles of Statistical Inference

Mr Lovemore Mapahla
Principles of Statistical Inference (Teaching Assistants)

Prof Rhoderick Machekano
Observational Data Analysis

Prof Jim Todd
Categorical Data Analysis and Generalized Linear Models

Prof Paul Mostert
Survival Analysis

Prof Edmund Njeru-Njagi
Longitudinal Data Analysis

Dr Andrew Abaasa
Longitudinal Data Analysis

Dr Carl Lombard
Design and analysis of Clinical Trials

Prof Lehana Thabane
Biostatistical Collaboration

Prof Tonya Esterhuizen
Biostatistical Collaboration

Dr Faheema Kimmie-Dhansay
Biostatistical Collaboration

Dr Christel Faes
Bayesian Analysis

Memorial

In remembrance of

Prof Birhanu Ayele

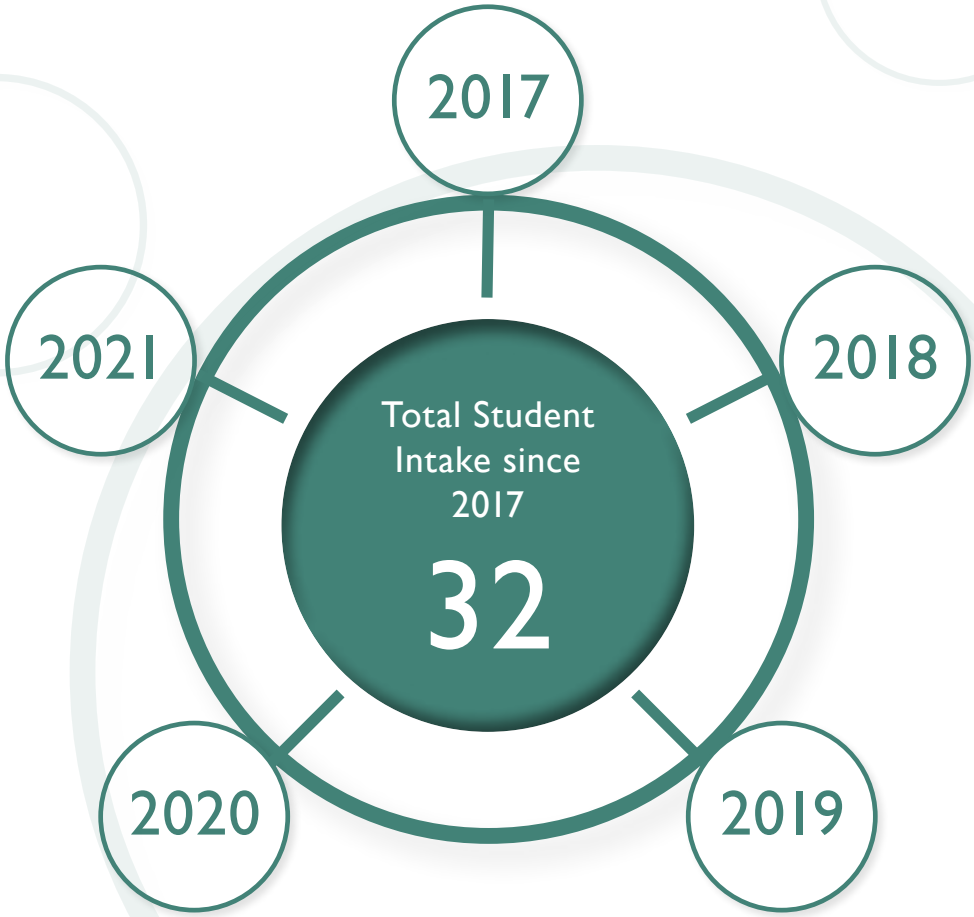
Respected scholar, researcher, collaborator, mentor, colleague, friend

Prof Birhanu Ayele


Birhanu as a Teacher
Dr Faheema Kimmie-Dhansay


Welcome to the
Biostatistics Clinic

Our first intake was in 2017. Many students applied and after several selection meetings 5 students were accepted as our first intake



2017	5
2018	5
2019	7
2020	7
2021	8

Keys

Year

Students

Meet the MSc Biostatistics students

2017 - 2021


2017

James Schmidt

"ALWAYS WORK SMART, NOT HARD"

I am a born and bred Pretoria boy, having attended Pretoria Boys High School. I have an undergraduate degree in Investment management, Honours in Statistics and now am in Biostatistics. I am an enthusiastic DIY-er and home chef, music and outdoor lover, and enjoy spending time at our farm in Tonteldoos, Limpopo. I am married and have two large brown dogs.

Who is your biggest inspiration?

I have no single inspiration but my biggest inspirations and drivers to study and improve myself are without doubt my wife and parents.

Where do you see yourself in 5 years from now?

I try not to do too much forward planning as whenever I do, things seem to always take a dramatic step in a different direction. My current career goals are to find a job in the biostatistics, statistics or data analytics field, preferably in or near the health sector.

The highlight of your 2 years of Masters study

Finally handing in my completed research assignment on the 1st of December 2018.Attending the SSACAB conference in Nairobi, Kenya.

About your fellow students:

We were a small group, being the first cohort to do our MSc in biostatistics at Stellenbosch and so all developed a good bond between us. We had many laughs and I really appreciate the constant support and guidance they provided.

Moleen Dzikiti

"....."

As a trained biostatistician and clinical epidemiologist, my primary research interests includes: design and analysis of observational studies, and clinical trials using both the frequentist and Bayesian approaches.

Who is your biggest inspiration?

My grandmother.

Where do you see yourself in 5 years from now?

My goals are to contribute to a society by generating research evidence for policy formulation and guiding clinical practice decision making, especially in people who are disproportionally affected by HIV/AIDS.

The highlight of your 2 years of Masters study

Giving birth to my son and finishing my MSc studies in real time.

About your fellow students:

It was great meeting you. Let us keep in touch.

Shibe Mhlongo

"DO YOUR BEST AND LET GOD DO THE REST"

I am a very loving down to earth reliable person. Very well organised and fastidious. Active and keep fit with exercising. I am happiest when I make a difference in someone's life.

Who is your biggest inspiration?

My mom. She is a hard worker, selfless and has taught me so much.

Where do you see yourself in 5 years from now?

In 5 year time, I will be 36 and I see myself with more responsibilities. Career wise, I will be in a senior position, working independently and collaboratively in major impactful studies. I will also be completing my PhD and hopefully travelling with my family.

The highlight of your 2 years of Masters study

Travelling to Kenya for the SSACAB conference. It was the second country I visited outside South Africa and my first conference presentation. Had a great experience of being in a different country plus meeting students from other universities.

About your fellow students:

My fellow classmates are very dedicated and focused people with a good sense of humour.

Tichaona Mapangisana

"ALL OUR DREAMS CAN COME TRUE… IF WE HAVE THE COURAGE TO PURSUE THEM."

I have a great interest in research, biostatistics, big data analytics, bioinformatics. I am a person who is positive about every aspect of life. I like reading, writing and listening. I like to be surrounded by smart and committed people.

Who is your biggest inspiration?

I'm inspired by people who make complex things simple. My biggest inspiration is my father. When I looked back and imagine how he was able to sail through difficult situations without showing that the situation was difficult. This always give me the courage to pursue my dreams with determination.

Where do you see yourself in 5 years from now?

Looking back over the past five years, I would never have predicted that at this point in my life I would be where I am. After five years from now I know I will have completed a PHD and I will be working hard on something exciting, alongside smart and committed people. I will be one of the best researchers, biostatistician in Africa and world over.

The highlight of your 2 years of Masters study

My two-year master's programme at Stellenbosch was one of the memorable times in my life. I had a very bad start, I was late for school by almost three weeks but non the less with the help of my classmates, lectures and other staff members I managed to sail through. The chill weather of cape town is one I will not forget. I had gained a lot of good stuff from Stellenbosch, the leaning environment was wonderful, with the good lectures and staff members.

About your fellow students:

It was a great moment to meet and share ideas for two years with Shibe Mhlongo, James Schmidt, Moleen Zunza and Cladnos Mapfumo. You were always there for me guys. Keep it up.

Cladnos Mapfumo

"IF I HAVE SEEN FURTHER THAN OTHER MEN, IT IS BECAUSE I STOOD ON THE SHOULDERS OF GIANTS"

A young and enthusiastic Biostatistician.

Who is your biggest inspiration?
Strive Masiiwa (Econet Wireless Founder).

Where do you see yourself in 5 years from now?
Complete PhD in Biostatistics.

The highlight of your 2 years of Masters study
Two full years of learning statistics and interacting with people of varied backgrounds.

About your fellow students:
Friendly and cooperative.

2018

Lovemore Sigwadhi

"UNLESS ONE IS A GENIUS, IT IS BEST TO AIM AT BEING INTELLIGIBLE"

I am an ambitious and driven person. I thrive on challenge and constantly set goals for myself, so I have something to strive towards. I am passionate about my work. Because I love what I do, I have a steady source of motivation that drives me to do my best and impact the community. I respect God and the community.

Who is your biggest inspiration?

Nick Vujicic.

Where do you see yourself in 5 years from now?

My goal right now is to find a position at a company where I can grow and take on new challenges over time. Ultimately, I'd like to assume more management responsibilities and get involved in product strategy. But most importantly, I want to work for an organization where I can build a career and impact to the community.

The highlight of your 2 years of Masters study

MSc in Biostatistics is a very good program which requires teamwork spirit and can work for long hours. Learning of time management skills is very important to meet deadlines. The MSc develops good thinking power, judgement and decision making. It gives you skills to be a more powerful researcher who can take part in all the research elements and become an effective and efficient collaborator. The program provides learners with knowledge in survival analysis, longitudinal data analysis, observational studies and clinical trials which I would say are the core in medical studies.

About your fellow students:

I love my fellow students there are a great team I ever worked with, so committed, punctual, respectful and joyful.

Ndamona Haushona

"IT DOES NOT HURT TO TRY"

My name is Ndamonaonghenda Haushona, a second year MSc Biostatistics student, originally from Namibia. I graduated with a BSc Honours in Applied Statistics (2017) and a BSc in Applied Mathematics and Statistics (2016) from the Namibia University of Science Technology. I have worked for GIZ-Namibia supporting the Ministry of Industrialisation, Trade and SME development as a junior advisor on economic framework conditions and monitoring and evaluation for two years. My primary research interests lie in Paediatric HIV/AIDS and TB research. My passion is to help women who are living with or at higher risk of HIV to sustain their health in order to halt the mother to child transmission. Furthermore, my research interest includes joint modelling of longitudinal and recurrent-events data.

Who is your biggest inspiration?

My parents are my biggest inspiration, they taught me to be the best I can, never give up in life and to serve others by their example.

Where do you see yourself in 5 years from now?

In five years, I see myself as a successful female biostatistician in Namibia with profound expertise in medical research. And helping to end disease epidemics in Sub-Sahara Africa, through collaborations.

Your favourite thing to do when you are not working/studying

During my free time, I enjoy spending time with my family, catching up with friends and reading novels.

About your fellow students:

My classmates made my Stellenbosch journey worthwhile, I could not have come this far without their assistance and encouragement. Our bond holds up so firmly, making us work together significantly as a team.

Tinashe Mhike

"ZVICHANAKA CHETE, GOD'S TIME IS ALWAYS THE BEST TIME" MEANING 'IT SHALL BE FINE, GOD'S TIME IS ALWAYS THE BEST"

I am a Holder of Honors degree in Statistics from University of Zimbabwe (2008-2012) and currently an MSc in biostatistics student at Stellenbosch University. I am also an author in the field of Mathematics. I am very much interested in medical research and can't wait to practice as a Biostatistician. I am passionate about my work, results oriented and a good communicator. Also, it is my greatest hope that I help those in need.

Who is your biggest inspiration?

Morgen, my brother. I constantly learn from him, how to love the next person, how to help those in need, and being funny when its time. Of course, it won't be just not to mention my mother who has been influential in my academic life.

Where do you see yourself in 5 years from now?

I see my self as a PhD holder of Biostats, and heavily involved in medical research. I intend to be involved in charity work, helping the less privileged since I have been a beneficiary as well.

The Highlights of your 2 years of Masters study

Overall it has been a busy time, but quite interesting. Met highly experienced lecturers in different fields of research and learnt a lot of new things which will be of great help in my career as a Biostatistician.

Your favourite thing to do when you are not working/studying

I am fan of soccer and basketball and I also like going out to places with friends and family. I enjoy writing Mathematical books.

About your fellow students:

Wonderful colleagues with strong sense of team work. Joy to work with and humble. They are also very funny people who play hard and work hard.

Albert Chinhenzva

"I JUST WANTED TO SAY HI"

I am Albert Chinhenzva, husband to one wife and father of three. I am a pharmaceutical scientist, public health specialist and fighting hard to become a Biostatistician. My life has been medicines and medicines from a 5g packet of pills in a clinic to tonnes and tonnes of various medicines in the country's national warehouse. Biostatistics for me is not just numbers, but men, women, born and unborn children who deserve the best from our scientific innovations.

Who is your biggest inspiration?
My Family

Where do you see yourself in 5 years from now?
Lead Scientist in the Biopharmaceutical Industry

The Highlights of your 2 years of Masters study
Driving to Stellenbosch Campus every Tuesday for 14 weeks!!

Your favourite thing to do when you are not working/studying
Drinking plenty of fluids

About your fellow students:
The statement "Too many cooks spoil the broth" has been inapplicable.

Mapahla Lovemore

"COUNTERFACTUAL"

Mapahla Lovemore is an Honours Degree in Statistics graduate with the University of Zimbabwe. His work has been centered on data analysis using different statistical software. He worked as a Research executive, Data analyst and Statistical analyst. Currently, Mapahla Lovemore is employed by the government of Zimbabwe as Advanced Level Mathematics and Statistics teacher.

Who is your biggest inspiration?
Prof Machekano

Where do you see yourself in 5 years from now?
Complete my Biostatistics PHD and help the community, region and whole world using the gained skills and knowledge

The Highlights of your 2 years of Masters study
It was and it is a long walk. I gained skills and experience. I endured and soldiered on. Frequently reminding my spirit to hang in there. I enjoyed the exposure. History written.

Your favourite thing to do when you are not working/studying
Watching soccer, Touring

About your fellow students:
I am luck and delighted to be part of this group. They are a brother's keeper. My colleagues are kind, motivating and considerate. We worked as a team. 8. Wire toy car

2019

Leonard Mwandingi

"PAY ATTENTION TO YOUR THOUGHTS. YOU BELIEVE EVERYTHING YOU SAY TO YOURSELF"

My name is Leonard Mwandingi born 09 January 1994, raised among 4 brothers in Omuthiya, Namibia. I finished my high school in 2011, obtained my Diploma in Applied Statistics in 2014 and then a BSc. Honours degree in Population Studies in 2017 from the University of Namibia. During my studies at the University of Namibia, I served as the chairperson for the Statistics Society. I have more than 3 years working experience in the public health sector of Namibia where I worked as a Monitoring and Evaluation Officer and as a consultant. My research interest includes population health modelling; sexual and reproductive health and rights; monitoring and evaluation of unemployment; well-being and disease mapping

Who is your biggest inspiration?

My biggest inspiration are my parents, they sacrificed everything to make sure that my brothers and I receive the necessary education and grow up to be responsible, caring and independent people.

Where do you see yourself in 5 years from now?

Within five years, I would like to become one of the best Biostatistician in Namibia and equally be a publishing researcher. I want to be involved in the science that tackles problems such as emerging diseases, controlling epidemics. In addition, I would also like to be involved in the development of new vaccines and technologies which is directed to better people's lives. I want to contribute immensely to The Emerging Data Revolution in Africa which is aimed at strengthening the Statistics, policy and decision-making chain. In doing so, I feel I'll be fully prepared to take on any greater responsibilities which might be presented in the long term.

Your favourite thing to do when you are not working/studying

I spend my free time listening to Afrobeat music. I also enjoy spending time with family and friends, and I exercise at least three times a week.

About your fellow students:

Getting to work with such a diverse knowledgeable group is one of the blessings that ever happened to me this year. My fellows are very supportive, selfless, and we work well together as a team.

Mercy Rop

"LIVE SIMPLY AND CARE DEEPLY"

I'm a hard worker, always driven by the need to learn and to always improve steadily. I'm passionate about health research and motivated to contribute to finding solutions to persistent and emerging health problems facing African populations. I am an introvert and mostly able to keep my cool under pressure

Who is your biggest inspiration?

MMy former boss is my biggest inspiration. Her passion for health research and her contribution to improving health among marginalized rural communities motivates me. I've learnt to always stay focused, to be goal oriented and to push myself out of my comfort zone through her.

Where do you see yourself in 5 years from now?

Five years from now, I see myself in Cancer research, probably teaching on a part time basis and doing gardening on the side

Your favourite thing to do when you are not working/studying

Nature walks with friends, watching movies and gardening

About your fellow students:

It's the best team ever. Always working together and never missing something to laugh about even when things are super tight.

Anteneh Yalew

"THERE IS FULL OF CERTAINTY OF UNCERTAINTY IN REAL LIFE IN GENERAL AND IN STATISTICS (RESEARCH) IN PARTICULAR".

I am Anteneh Yalew from Ethiopia, married (with Frewoyni) and have a two-year-old beautiful daughter called Loza. I am obedient, honest, sociable and responsible professional person. My personal ambition is to be a great (bio)statistician. Evidence based decision with certainty of uncertainty motivated me to stay in the field. I am eager to learn.

Who is your biggest inspiration?

My father! He thought me how to stand for the truth and serve the community in practice.

Where do you see yourself in 5 years from now?

Successfully complete my MSc in Biostatistics study and start a PhD in Biostatistics (Statistics).

Your favourite thing to do when you are not working/studying

Reading books (historical, political, spiritual, …), listening spiritual hymns (hymns of Ethiopian Orthodox Tewahedo Church are my favourite), visiting relatives and discussing with friends.

About your fellow students:

We are diverse with respect to where we came from, culturally, linguistically and even background field in some extent and we exchange and learn from one another. We have very strong team spirit that benefitting us to discuss and study as a team for better understanding of our modules.

Diribsa Bedada

"GOD IS GOOD ALL THE TIME."

My name is Diribsa. I am from Ethiopia. I was born in 1986. I have three sisters and five brothers. My dad is a farmer. My parents are my inspiration. They have always stood my side and keep telling me that just focus on your goal not anything else.

Who is your biggest inspiration?

My parents are my inspiration

Where do you see yourself in 5 years from now?

To complete my MSc in Biostatistics and to join my PhD degree in Statistics (Biostatistics) at a well-recognised higher learning institution and to become a senior Biostatistician within the next five years.

Your favourite thing to do when you are not working/studying

I like relaxing with my friends, reading spiritual books, listening spiritual songs and sleeping.

About your fellow students:

My class fellows are very nice, and we find many interesting things about each other and experience new things. I appreciate friendship and people who surround me.

Faheema Kimmie-Dhansay

"NO WAY"

I'm a dentist from Cape Town. I am a wife, mother, daughter, sister and friend. I love meeting new people and learning about new cultures and languages. I love travelling when I'm not studying or working. I enjoy helping others and that's why I love my job.

Who is your biggest inspiration?
My mother

Where do you see yourself in 5 years from now?
Working as a more confident Biostatistician at the Dental Faculty

Your favourite thing to do when you are not working/studying
Worrying that I should be working / studying. Jokes aside, I love to spend time with my family.

About your fellow students:
Our class has really bonded well with one another over the last few months. Our journey has just begun, yet it feels as if we have known each other for years. I look forward to many more laughs through exhausted eyes over lame statistical jokes and pizza from Mr Delivery.

Moses Ouma

LET NOT THE SKY LIMIT YOU FROM MAKING BETTER THE SOCIETY THAN YOU FOUND IT.

I am a passionate community servant with a strong motivation of using the very little knowledge I have acquired in life and education to transform the lives and livelihoods of the young people. I have on many occasions volunteered to organize and participated in services of the community.

Who is your biggest inspiration?

My biggest inspiration is my widowed grandmother. Her struggles to up bring us from the meagre peasant income inspires me a lot more and I therefore always strive to go a notch higher for the best. Her effort and enthusiasm trained me to perceive life positively.

Where do you see yourself in 5 years from now?

MSc in Biostatistics at this reputable institution of learning forms the building block of my career progression upon completion in one and a half year- time. After graduation, I plan to join the Kenya Medical Research Institute to further fortify my expertise in the application of Biostatistics to generating evidence-based research including training upcoming scientists. After garnering a one-year experience at KEMRI I will enrol in a Biostatistics PhD program which will enable me to become a prolific researcher who is able to take the lead in research arena.

Your favourite thing to do when you are not working/studying

My favorite engagements are sport activities and team building.

About your fellow students:

Stellenbosch University has an amazing community of students whom can easily be mingled with and share ideas.

Perseverence Savieri

"NOTHING IS IMPOSSIBLE."
"ATTITUDE IS EVERYTHING."

I am God-fearing, goal oriented, optimistic and persistent. I believe success is centred on attitude and the obstacle is the way. My name always reminds me that nothing is impossible if you truly put your heart and mind to it.

Who is your biggest inspiration?
My mother

Where do you see yourself in 5 years from now?
A PhD holder in Bioinformatics. My ambition is to be a full-time research consultant in the Biostatistics and Bioinformatics fields.

Your favourite thing to do when you are not working/studying
Reading and writing motivational work (and listening to motivational speakers) or watching sports and cooking shows.

About your fellow students:
They're my family away from home. I feel their love and respect, even in our diverse cultures, skills and backgrounds.

Makabongwe Nombula

"NTINGA",
"GO BEYOND"

I am Makabongwe Nombula, I am a talkative, I love people beyond measures because they are the ones who bought me where I am and God as always. I am not a genius, but I work very hard to get things done. I love excellence, perfection, I wish I can always be there best in all times, and I hope one day I will reach that moment. I am always happy to see amazing people like Albert Einstein, Bill Gates and Leonel Messi.

Who is your biggest inspiration?
I am my inspiration

Where do you see yourself in 5 years from now?
It won't be easy, but I need to push hard. In the next five years I am seeing myself owning one of the Biostatistics consultant company or being the head on any company specializing clinical data.

Your favourite thing to do when you are not working/studying
I love gym, I always prefer to gym when I become free. Secondly, I love to talk with people if I am done with my gym and I am still free, I prefer to talk to a person.

About your fellow students:
My fellow students are amazing, of course they are different, but they accommodate the variation among us. Some are genius, and they share their knowledge, so it is wonderful to have them.


2020

Xan Swart

"GOOD MORNING DREAM TEAM!"

Musician and amateur mushroom cultivator with a love for clinical medicine and an unfortunate abhorrence for weekend and night shifts.

Who is your biggest inspiration?
Eugene Marais, naturalist and poet.Roald Dahl and David Attenborough

Where do you see yourself in 5 years from now?
In South African academia with a teaching and research focus.

Your favourite thing to do when you are not working/studying
Making Music with friends, hiking and foraging in the pine forests around cape town.

About your fellow students:
Inspiring group of individuals. I am forever grateful for the department for facilitating the crossing of our paths. I have yet to meet a better group of people.

Your favourite childhood toy and a picture of you as a young child:
Dinosaurs

Belachew Hunde

"WE ALL HAVE A SMART MIND GIFT TO THINK FOR A BETTERMENT, THE DIFFERENCE IS THE WAY WE THINK."

I am Belachew Melese Hunde. I love my profession with passion and dedication. I am always eager to learn new things from anybody. I have been lecturer at Semera and Arsi university from 2011 to 2017 and from 2017-2019 respectively. During my stay at both universities, I like and advised my student to work in team and I collaborated with different professional staff members and did research in team. Currently, I am second year biostatistics student at Stellenbosch university, and I have an interest in research work on longitudinal data analysis, survival time-to-event and joint modelling.

Who is your biggest inspiration?

I find inspiration from what I have learned as education in the class then I started searching myself what I want to be.

Your favourite thing to do when you are not working/studying

I like watching football and reading daily important posts on social media when I am free from any work.

Your favourite childhood toy and a picture of you as a young child:

When I was child, I like playing with mud and my boy and girl friends around our village.

Hillary Kibet Kiprono

"IF YOUR SALARY IS YOUR ONLY SOURCE OF INCOME, YOU ARE ONE STEP AWAY FROM POVERTY."

I am a Kenyan, a finalist student Masters in Biostatistics at Stellenbosch University. My background is in statistics (BSc Applied statistics). I have had previous experience working in a health care research as a data manager. My motivation is to build on existing health care system and best care management of patients through research and using available data to make informed decision.

Who is your biggest inspiration?

My dad. Though without having graduate education, the confidence and faith he has bestowed on me is on another level. That is my driving force each and everyday to make him proud.

Where do you see yourself in 5 years from now?

My ambition is to further my studies. A PhD in Biostatistics is my primary goal. In the next five-year years, I see myself being a senior biostatistician working either in the academic or a not-for profit organization.

Your favourite thing to do when you are not working/studying

I love watching Netflix movies. Especially series action movie.

About your fellow students:

All my fellow students are amazing, dedicated on their studies and enables us to interact and learn diverse cultures and practices.

Meseret
Mamo Bazezew

"NEVER GO INTO THE MATCH WITH AN EXCUSE TO LOSE".

Meseret Bazezew is a statistician involved in different health related research starting from 2010 to 2017. She earned her BSc in statistics from Addis Ababa University and her Master of Public Health from University of Gondar, Ethiopia. The first project she was involved in was Ethiopian Demographic and Health Survey (EDHS) as a data quality controller. Afterwards, she was involved in Nutrition, HIV, Women empowerment and Agriculture surveys.

Who is your biggest inspiration?
Opera Winfrey

Where do you see yourself in 5 years from now?
In five years, my goal is to complete my PhD in Biostatistics. Besides, increasing the number of publications and taking various training that will improve my skills and knowledge are some of my goals that will go hand in hand. I hope to collaborate and network with senior researchers in Biostatistics.

Your favourite thing to do when you are not working/studying
I am nature lover person, especially the ocean. I enjoyed spending time at the ocean.

About your fellow students:
My fellow students are very determined, kind and support each other to achieve our common goal of completing this Msc. Biostatistics program. Coming from different educational background created a good combination of our batch.

Musa Gwala

"MAKE YOUR LIFE A MASTERPIECE; IMAGINE NO LIMITATION ON WHAT YOU CAN BE, HAVE OR DO." – BRIAN TRACY

I am a UKZN double Graduate for BSc degree in Applied Mathematics and Statistics and BSc Honors Statistics. I am a young energetic male from a place called KwaXimba in KwaZulu-Natal, and a fast learner with good communication skills. I have been introduced in variety of Mathematics and Statistics modules and I have been working as an intern in the Division of Epidemiology and Biostatistics at Stellenbosch University Tygerberg Campus, which have given me so much knowledge when it comes to dealing with real life problems.

Who is your biggest inspiration?
My biggest inspiration is to be the best I could be and achieve my goals so that my son will follow my footsteps and see nothing can stop him to do well.

Where do you see yourself in 5 years from now?
I see myself working as one of the best Biostatistician and persuading my PhD.

Your favourite thing to do when you are not working/studying
I like to play video games like FIFA, Call of Duty, Mortal Kombart, etc.

About your fellow students:
I call them "Champions" simply because they want to achieve so much in life, and they are all hard workers with good hearts too. I am grateful to be their Captain. Lol

Your favourite childhood toy and a picture of you as a young child:
Pretty much every toy but a toy car was the best lol.

Emmanuel Chimbunde

OUR DEEPEST FEAR IS NOT THAT WE ARE INADEQUATE. OUR DEEPEST FEAR IS THAT WE ARE POWERFUL BEYOND MEASURE. (MARIANNE WILLIAMSON)

I am an experienced Data Analyst with over two years' experience in analysing and interpreting big data. I have worked with various organisations including tech-companies. I have broad understanding of advanced modelling techniques and modern data visualisation tools.

Who is your biggest inspiration?
Elon Mask

Where do you see yourself in 5 years from now?
Successful Data Scientist, who will be changing people's lives through quality research and analytics combining Artificial Intelligence, Machine learning and Biostatistical modelling skills.

Your favourite thing to do when you are not working/studying
Programming, Watching Movies, Visiting Friends

About your fellow students:
Hard working and caring team.

Your favourite childhood toy and a picture of you as a young child:
car

Yebelay Berehan

"IF THE STATISTICS ARE BORING, THEN YOU'VE GOT THE WRONG NUMBERS." AH, TUFTE WRITES

I am Yebelay Berehan from Ethiopia M.Sc. fellow in Biostatistics at Stellenbosch University South Africa. I have been a teacher at Debremarkos University, Ethiopia since 2014 taught statistics for undergraduate students. I consider myself energetic, hardworking, visionary, and motivated person who believes that I can do something for the benefit of myself and my country.

Who is your biggest inspiration?

My biggest inspiration was my mother. After my father's death, she faced a lot of adversity to feed us as she had to raise eight children by herself. we are working from hand to mouth class citizens, so she worked extremely hard. I was crying always when I saw her hard working as a single mother. Then I developed a particular interest to help her. Then I study hard even it was hard condition to learn.

Where do you see yourself in 5 years from now?

Over the next two years, I want to expand my biostatistical knowledge on the advanced statistical methods and state of the art techniques of communicating statistical findings with public health areas. Then in five years, I want to have gained experience in Biostatistical consulting and I want to find myself as a PhD student.

Your favourite thing to do when you are not working/studying

When I have not study time, I like to read fictions and helping my wife in home works, playing with my child who refresh my mind.

About your fellow students:

I really appreciate my classmates who are very best friends.

Your favourite childhood toy and a picture of you as a young child:

I was born on small rural village. So, our favourite childhood toy was playing up with mud by creating different styles with it.

2021

Oluwaseun Adeyemi

"GOD KNOWS BEST AND WILL SORT IT OUT"

Who is your biggest inspiration?
Jesus Christ
Where do you see yourself in 5 years from now?
5 years from now I hope to have accomplished a PhD
Your favourite thing to do when you are not working/studying
When not studying I like resting
About your fellow students:
My fellow students are goal oriented

Yeshework Amare

"NEVER GIVE UP BECAUSE GREAT THINGS TAKE TIME"

I am a statistician and was working as a lecturer at the Department of Statistics of the Jinka University Since October 2020. I graduated with BSc in Statistics from Jimma University in 2015 and MSc in Applied Statistics from Addis Ababa University in 2019 in Ethiopia. I previously participated in the Ethio-Demographic Demographic and Health Research Project, and later in nutrition, HIV, women's empowerment, and agricultural research. I also worked as a junior researcher in Federal Vital Event Registration Agency in Ethiopia for two years. In addition, I also worked in various positions such as data analyst, research expert, survey field supervisor, qualitative and quantitative data collector for numerous studies that focused on public health, HIV/AIDS, nutrition and agricultural surveys at governmental and non-governmental organizations (on contract bases) for more than two years. I have a keen interest to pursue a career in biostatistics.

Who is your biggest inspiration?

My family is a great encouragement to me. They trusted me and always pushed me to be a good person and I hope to show them that I will be one more day in their interest.

Where do you see yourself in 5 years from now?

My intention in five years is to complete MSc study and also continue studying my PhD in biostatistics. Furthermore, collective the number of publications and taking several practical working exercises to improve my skills and knowledge are my top goals. I look forward to collaborating with top senior investigators in biostatics.

Your favourite thing to do when you are not working/studying

I am curious about new things. I love landscapes and nature, and I especially enjoy spending time in places near the sea and oceans.

About your fellow students:

My fellow students are very resolute, caring, kind and helpful of each other to accomplish our common objective of completing this MSc. Biostatistics Program. Coming from different educational backgrounds, our side was a good fit.

Your favourite childhood toy and a picture of you as a young child:

I enjoyed playing soccer with my friends, swimming in the pool.

Mulugeta Geremew Geleso

"WHEN YOU WANT TO SUCCEED AS BAD AS YOU WANT TO BREATHE, THEN YOU'LL BE SUCCESSFUL".

I graduated with a BSc degree in Statistics from Dilla University, Ethiopia and MSc degree in Mathematical and Statistical Modelling from Hawassa University, Ethiopia. I have more than seven years of working experience as a Lecturer of Statistics. I published more than 6 articles in international peer-reviewed journals and more than 5 conference contributions. My research interest has been focused on Biostatistics, Infectious Disease Modelling, Computational Statistics, Survival Analysis and Causal Inference.

Who is your biggest inspiration?
My father

Where do you see yourself in 5 years from now?
After completing my MSc degree in Biostatistics, I want to pursue my PhD degree in Biostatistics which emanates from my strong desire to be a competent biostatistics expert with a long-term objective of being a university professor in biostatistics, envisioned through a clear objective of helping the rigor, quality, credibility, reliability and reputability of health science researches in the university. I will conduct high-quality researches and publish the findings in international peer reviewed journals.

Your favourite thing to do when you are not working/studying
I love playing and watching football.

About your fellow students:
They are very supportive and friendly.

Hillary Katsabola

"PUT GOD FIRST, STAY HUMBLE AND WORK LIKE NOBODY'S BUSINESS"

My name is Hillary Katsabola. I am a Malawian currently pursuing a master's degree in biostatistics at Stellenbosch University. I completed my bachelor of science degree at the University of Malawi Chancellor College in November 2018, where I majored in Statistics and took mathematics as a minor. Since then, I have worked as a data officer at Baylor College of Medicine Children's Foundation. I am one of those people who believes that anything is possible. I believe that if you truly want to archive something in life, you put forth the effort, you never give up, and the likelihood of succeeding grows exponentially with each attempt until you eventually archive it.

Who is your biggest inspiration?

My parents are my greatest sources of inspiration. They have so much faith in me and are constantly pushing me to be a better person, and I hope to demonstrate to them one day that I am capable of going the extra mile.

Where do you see yourself in 5 years from now?

In five years, I envision myself with a PhD, working at a higher education institution, and running my own scientific research company

Your favourite thing to do when you are not working/studying

When I'm not working or studying, I like to surf the web and learn new computer software that isn't related to my field of study. Chess, which has been my favourite board game since I was a child is another pastime of mine.

About your fellow students:

My fellow students are wonderful, helpful and happy people. I'm always pleased that I met them here.

Your favourite childhood toy and a picture of you as a young child:

I didn't have many toys as a child. We grew up making toys out of clay soil, which was my favourite and a lot of fun.

Mary Tiya Magoya

"YOU HAVE TO LOOK FORWARD TO YOUR DREAM AND YOU CAN'T LET ANYBODY GET IN THE WAY OF IT, NO MATTER HOW TOUGH IT MAY BE, NO MATTER HOW MANY TEARS YOU MIGHT CRY, YOU HAVE TO KEEP PUSHING. AND YOU HAVE TO UNDERSTAND THAT NOTHING COMES EASY. KEEPING YOUR EYES ON THE PRIZE, YOU CAN SUCCEED."

I am Mary Tiyankhulenji Magoya from Malawi. A girl full of dreams and an achiever.

Who is your biggest inspiration?

My inspiration is all women who have attained high positions in science related fields gives me an assurance that one day I will be one of them.

Where do you see yourself in 5 years from now?

It's really important to me to keep growing my skills and adding value to my career in new ways. One thing that drew me to this field was the opportunities for professional development and growth. So, in five years, I plan to master my skills as a biostatistician, so that I have a solid base for taking on managerial responsibilities. I'd love to be further my studies into a PHD either in Pharmacometrics, Epidemiology or Bio-informatics.

Your favourite thing to do when you are not working/studying

I am a big fan of amapiano music and I love listening to it when I am not working, it gives me energy and positive vibes.

About your fellow students:

I really appreciate all my classmates for the support we show to each other as a team. I would be happy to see everyone of us win. Many appreciation to the ones who finished ahead of us, you've supported us in countless ways, assuring us that we will get through. THANK YOU!!!!!

Peter Ndaramu

FORTUNE FAVOURS THE BRAVE

I am an experienced Statistician with over three years' experience analyzing and interpreting complex data. I have worked with various research experts; am very knowledgeable in using statistical tools and methodologies; adept in conducting sampling and validation. I have broad understanding of statistical concepts, methods and models and an expert in using computer applications.

Who is your biggest inspiration?
Chadwick Boseman

Where do you see yourself in 5 years from now?
In 5 years I will be a Biostatistics PHD holder contributing positively towards improvement of human health by training aspiring statisticians and conducting health researchs addressing a specific problem.

Your favourite thing to do when you are not working/studying
Fishing, Watching Movies, Cricket, and Soccer, Visiting Friends, Socialising on Social Media

About your fellow students:
I have had the pleasure of interacting with fellow students who have expert knowledge in different areas. The cultural diversity of the group composition is quite phenomenal

Ivan Nkuhairwe

"TOMORROW PROMISES TO BE BETTER"

I hold a Bachelors in Environmental Health Science from Makerere University Kampala. I have special interest in epidemiology and clinical biostatistics.

Who is your biggest inspiration?
My Dad. (Best 5 years of my Life)
Where do you see yourself in 5 years from now?
Too early to tell.
Your favourite thing to do when you are not working/studying
Football and anything sports.
About your fellow students:
Intelligent and warm.
Your favourite childhood toy and a picture of you as a young child:
I loved toy cars and planes.

Hilary Takawira

SINCE TOMORROW IS NOT GUARANTEED DO WHATEVER YOU ARE SUPPOSED TO DO TODAY. ALWAYS BE ON TOP OF YOUR GOALS.

In 2017, Mr. Takawira graduated with a distinction from Midlands State University, Zimbabwe where he earned a Bachelor of Science in Applied Biosciences and Biotechnology Honours Degree. Hilary Takawira has a strong bias towards bioinformatics and currently aspires to move to a higher level of research work for doctoral studies so as to get a deeper understanding of the link between bioinformatics, statistics and machine learning.

Who is your biggest inspiration?
Dr Strive Masiiwa

Where do you see yourself in 5 years from now?
I envision myself being in the third year of my doctoral studies pursuing PhD in Biostatistics/Bioinformatics. I see myself having published more than two papers of my work. I envision myself being an independent Biostatistician doing statistical consultancy and keeping on building on my data analysis skills and venturing more into bioinformatics and machine learning.

Your favourite thing to do when you are not working/studying
When not working or studying I usually volunteer at NGOs to help giving back to the community. I also enjoy hiking.

About your fellow students:
I am grateful for the numerous contributions in terms of fruitful discussions that we do with my classmate. They make certain concepts easier to grasp and understand.

Your favourite childhood toy and a picture of you as a young child:
I always loved cards

Inside the classroom

"Principles of Statistical Inference" with Dr Chris Muller

"Workshops" with Tonya Esterhuizen

"Bayesian Analysis" with
Isaac Fwemba and the Zambian Students

"Estimation of sample size" workshop with
Dr Lawrence Mbuagbaw

Outside the classroom

"Class of 2019"

"Having fun"

"Exciting outing"

"Studying together"

Enjoying nature

getting to know each other outside of class

Graduation

2018

Dr Moleen Dzikiti

Longitudinal studies investigating changes in a specific outcome that is measured repeatedly over time are common. Such studies are appropriate for investigating the change of the outcome, and the effect of associated covariates. The credibility of study findings depend on the appropriateness of the analysis method used. It is important to assess the sensitivity of the study findings to different possible methods of analysis to ensure appropriate conclusions are drawn from the study.

We performed six models using an empirical data set obtained from a real life non-trial clinical care setting to explore if there exists significant differences in the change of a dichotomous outcome (infection-related hospitalization), over the first year of life, between infants who were predominantly breastfed and those who were not. Data are from the Mother Infant Health study, a longitudinal cohort study that compared infection-related hospitalization among HIV-exposed uninfected infants and HIV-unexposed infants. The primary objective of this study was to assess the sensitivity of the findings to different models used to account for dependency of a binary outcome measured repeatedly over time.

We estimated the effect of predominant breastfeeding on infection-related hospitalization using generalized linear mixed models with (1) a random intercept and (2) a random slope. We used the logit link function and fitted the model using the Gauss-Hermite quadrature approximation. The fixed effects quantify the subject-specific effects of predominant breastfeeding on the odds of infection-related hospitalization, in children who share the same propensity to infection-related hospitalization. The random effects quantify the variation between subjects (random intercept model) and variation between subjects and variation in infant age (random slope model). The final models included fixed effects (predominant breastfeeding, infant HIV-exposure status, mother's educational level, and infant age).

We found a clinically important but statistically insignificant reduced odds of infection-related hospitalization among predominantly breastfed infants compared with non-predominantly breastfed infants, irrespective of the method of analysis used. The distinction between the methods was less important for the longitudinal binary outcome data with insufficient replications of the outcome.

The method of analysis we used had insignificant effect on the parameter estimates and this could have been driven by insufficient replication of infection-related hospitalization outcome in a cluster. Our estimates from the six models are likely to be unstable and need to be interpreted with caution.

My studies were supported through DELTAS Sub-Saharan Africa Consortium for Advanced Biostatistics Training initiative. This study was also partly supported by Fogarty International Center of the National Institutes of Health under Award Number D43 TW010547.

"Exploring association between predominant breastfeeding and infection-related hospitalization, over time: an empirical comparison of models to account for clustering"

Mr Tichaona Mapangisana

"Virus Load Differentiated Care of HIV-1 Infected Children and adolescents is feasible and effective in remote rural Zimbabwe"

There are many challenges to ART delivery to children and adolescents in remote communities of rural Africa. The Zimbabwe Population Based HIV Impact Assessment (ZIMPHIA), found rates of Virologic suppression (VS) <50% among children and adolescents on ART. We hypothesized that community-based virus loads differentiated care (VLDC) can identify virologic failure (VF) and prompt switching to second-line ART to achieve virologic suppression (VS) among children in rural Africa.

A retrospective longitudinal study of 306 children and adolescents on ART was conducted from 2016 to 2018 in rural Hurungwe Province. ART was provided, bimonthly at youth friendly support clinics or by active outreach to 8 rural sites coordinated by Chidamoyo Christian Hospital (CCH). Initial viral load testing by Roche COBAS® Ampliprep®/COBAS® Taqman48® HIV-1 v2.0 was performed in Harare in 2016. At follow-up in 2018 near Point of Care (POC) virus load testing with a Simplified AMplification-Based Assay (SAMBA), (Diagnostics for the Real World, Sunnyvale California) was performed at CCH. Virologic failure (VF) was defined as viral load ≥1,000 copies/ml. We assessed virologic suppression (VS), defined as viral load <1,000 copies/ml. A logistic regression model including demographics, care-givers and treatment regimens assessed risks for VF and loss to follow-up (LTFU) adjusting for the baseline variables. Of 306 children and adolescents, 208 (68%) had viral load <1000 copies/ml. VF was significantly associated with lower CD4 cells (p<0.001) and non-parental caregivers (p=0.04). Follow-up in 2018 demonstrated that 42 (14%) were no longer receiving care through CCH; 23 (7%) had transferred to other ART facilities, 17(6%) were lost to follow up (LTFU) and 2 had died. Older age (age ≥10 years) was significantly associated with LTFU (p=0.030). Fifty-eight (22%) had switched to second line ART. Of the 264 retested in 2018, 212 (81%) had VS. VS rates on first and second line in 2018 were 82% and 72%, respectively. Overall, VS increased from 68% in 2016 to 81% in 2018 (P < 0.001) following the introduction of VLDC and second line treatment.

Second line ART for children and adolescents, VLDC, new health service delivery models and near POC testing are associated with improved virologic suppression rates in this vulnerable population. Provision of VLDC including near POC virus load monitoring in remote rural setting is feasible and effective.

He expressed his sincere gratitude to the staff at Biomedical research training institute (BRTI), staff at University of Stellenbosch, Master in Biostatistics colleagues, Faculty of Medicine and Health sciences at Stellenbosch University for all their support.

He also thanked DELTAS Africa Initiative SSACAB for funding this study. This study was also partly supported by Fogarty International Center of the National Institutes of Health under Award Number D43 TW010547.

He expressed a special thanks to his family for the support throughout the study.

Mr James Schmidt

Patient registry or routine/programme data are commonly available data, created with the use of observational study methods, and are often a statutory requirement or used to answer important medical questions. "Common statistical methods such as descriptive statistics may not be appropriate for such data due to complexities such as imbalance, repeatedness, clustering or missingness", said biostatistician, Mr James Schmidt. "Therefore, more advanced statistical analyses such as linear mixed effects modelling or time-to-event analysis may be required."

James' MSc project argued that Linear Mixed Models (LMM) are more appropriate where the data are repeated or clustered as they incorporate

"The Application Of Linear Mixed Effects Models And Survival Analysis To Determine The Effect Of Waiting Time For Biologic Treatment Therapy Among Rheumatoid Arthritis Patients In South Africa"

random effects for varying intercepts and slopes between clustered observations. Time-to-Event Analysis (TTEA) is used in instances where the outcome of interest is the time taken for an event to occur. In TTEA, censoring is used where the exact time to event is unknown, for instance where the study ends before the occurrence of the event.

As an application example, he applied both LMM and TTEA to determine the effect of the waiting time from biologic treatment therapy amongst Rheumatoid Arthritis (RA) patients using the South African Rheumatism and Arthritis Association (SARAA) patient registry database. He found the relationship between those who received biologic treatment within the same year the diagnosis was made, as compared to those who had a waiting period of greater than one year for biologic treatment, to be insignificant across both analysis techniques.

He concluded that patient registries provide a wealth of data, easily available and accessible, in order to assess and answer disease progression and patient centred outcomes hypotheses and questions, unhindered by complex planning, rigorous criteria and ethical considerations, requiring more advanced statistical techniques to gain the required insights.

He acknowledged the support of his wife, and his family for their never-ending encouragement, support and guidance and he thanked Professor Paul Mostert, University of Stellenbosch, and Professor Khangelani Zuma of the HSRC for their invaluable contribution towards the time-to-event analysis and Dr Elsa van Duuren, Jacaranda Hospital for her vital insights into RA and the SARAA patient registry. He sincerely thanked the Division of Epidemiology, University of Stellenbosch and the course conveners and coordinators for their guidance, patience, support and tuition. He extended his gratitude to the financial support received from the DELTAS SSACAB Africa Initiative. This study was also partly supported by Fogarty International Center of the National Institutes of Health under Award Number D43 TW010547.

"Lastly, I would like to acknowledge and thank my fellow MSc Biostatistics classmates. The support received, friendship and memories made will forever be cherished and never forgotten" he said.

Mr Cladnos Mapfumo

"Estimating effects of smoking and alcohol exposures on oesophageal cancer in a high incidence area in South Africa: A Targeted Maximum Likelihood approach for case control data"

Estimation of causal effects from case-control data has two complications namely confounding since the exposure under study is not randomized, and biased sampling scheme. Logistic regression is the common analysis method for case-control data, from which estimation of the odds ratio is conditional on the covariates. This conditional parameter depends on correct specification model, which is uncertain in practice. Methods that incorporate propensity scores such as the G-formula and targeted maximum likelihood estimation (TMLE) are preferred. Targeted maximum likelihood estimation is a general class of procedures for estimating semiparametric statistical models, that is models with a nonparametric nuisance parameter.

Using case-control data we demonstrate causal inference estimation in a case- control study using TMLE. We demonstrate the application of TMLE using tmle package and logistic regression in R programming software. Education, residence, and sex were confounding variables adjusted in the models.

Six hundred and seventy cases were enrolled and 1188 control patients. Results show that, if the whole population reduced tobacco consumption to below 5 grams per week the odds of oesophageal cancer would be reduced by 46%, Causal odds ratio (COR)=0.54 (95% CI:0.49-0.59). Results further shows that, if the whole population reduced ethanol consumption to ≤135 grams per week the odds of oesophageal cancer would be reduced by 62%. COR=0.38 (95% CI:0.34-0.44).
It was found that low exposures below the 5g/per week(tobacco) and 135g/ week(ethanol) could decrease the odds of developing oesophageal cancer. Education of the population to stop or reduce taking these harmful exposures would be a suggested health intervention.

The authors gratefully acknowledge the entire research team which made this study possible, and study subjects. He also thanked Delta Sub –Saharan Consortium for Advanced Biostatistics (SSCAB) for funding his studies. This study was also partly supported by Fogarty International Center of the National Institutes of Health under Award Number D43 TW010547.

Ms Shibe Mhlongo

Post rape psychological trauma is well described in literature, mostly through cross-sectional studies. The Rape Impact Cohort Evaluation (RICE) study is a comparative cohort study that aims to advance our understanding on the medium-term and long-term health consequences of rape and of particular interest is the mental health impact. Longitudinal studies are common in medical research and consist of measurements taken repeatedly on subjects over a certain period. Usually the objective of such studies is to observe change in the outcome of subjects under study and the rate of change of the outcome for an individual subject relative to others. The correct statistical methods need to be applied to make valid inferences since the repeated observations on the subjects are likely to be correlated over time.

A lot of research has been conducted on the immediate effects of rape through cross-sectional studies, but little has been done on the medium- and long-term effects of rape. The research from the longitudinal comparative cohort study on the impact of rape will add knowledge on the health consequences post rape and especially the mental health status. For illustration purposes the evolution of the continuous depression status of rape exposed women will be compared with non-exposed women. In the paper we considered a mixed linear splines model for comparison of the mean depression score of women that have been raped and compared to women that have not been raped. Methods and estimation procedures used in this paper highlight the practical aspects and challenges that come with analysing longitudinal data. The mixed linear spline model forms the basis for further work in determining valid estimates of depression medium and long term. Rape impacts on women's mental health, more immediate effects than long term effects. The overall conclusions from the model are in line with what is found in literature about recovery rates of depression given recent exposure to a traumatic event.

"Statistical methods to facilitate the comparison of the mental health status of rape exposed and non-exposed women one-year post enrolment: Rape Impact Cohort Evaluation (RICE) study"

I would like to acknowledge my family (Maria Mhlongo, Khulu Mhlongo, Jabes Jaars, Christina Jaars), who have been understanding and supportive in this journey of completing my degree. I would like to thank Dr Carl Lombard who guided me in the research project. I am honoured to have worked with him in this project and have learnt so much more about longitudinal data analysis. I would like to thank Dr Naeemah Abrahams as the Principal Investigator of the RICE study for the encouragement, support and allowing me to use the data for my project. My studies were financially supported through the DELTAS Africa Initiative SSACAB and the South African Medical Research Council. This study was also partly supported by Fogarty International Center of the National Institutes of Health under Award Number D43 TW010547.

This research would not be possible if women did not share part of their lives with us and helped in conducting such research. This study will hopefully advance our understanding on the impact of rape and improve women's life's post rape and identify key areas in planning services to mitigate and prevent rape and its consequences.

Graduation

2019

Lovemore
Sigwadhi

Data plays an important role in every aspect of statistical modelling. In this study, we used a cross-sectional study from the African European Tuberculosis Consortium (AETBC)[2] which was approximated to follow a logistic population.

The prevalence of tuberculosis (TB) was approximately 25.2% among the adults who were enrolled in the study between November 2010 and November 2012[2]. In epidemiological studies, the risk ratio (RR), prevalence ratio (PR) and the odds ratio (OR) are widely used measures of association in cross-sectional studies[3]. PR and RR are mathematically equivalent and are recommended to be reported for cross-sectional and cohort studies. Although our data was assumed to be from a logistic population, the odds ratio may not be the best measure of association to use because of its interpretation[3]. Espelt et al (2017) recommended the use of PR instead of OR in cross-sectional studies when the prevalence is high[4]. Furthermore, Coutinho et al (2008) noted that if the OR is interpreted as RR it will overestimate the associations for outcomes with low, intermediate and high prevalence by 13%, almost by 100% and fourfold, respectively[5]. Therefore, in this study we explored the best effect measure to use with the appropriate statistical method to estimate the measure.Logistic regression is the most popular model for binary outcomes, and it expresses the effect measure on the outcome as an odds ratio [6]. Since, the PR and RR are the recommended measure of effect[4] when the outcome is rare the odds ratio is a good estimate when used to approximate the RR. However, for common outcomes, the odds ratio overestimates the association when used to estimate the RR or PR [7]. Frequently, in cross-sectional and cohort studies where PRs and RRs are the effect measures of primary interest, ORs are reported instead [8]. Furthermore, the log-binomial regression has been recommended to estimate PRs directly [4] [9].

Since we were unsure which effect measure and regression model to use when dealing with common outcomes from a logistic population. Logistic regression is not the best model and the ORs are not good effect measures when the outcome of interest is moderate to high. The comparison for model performance in estimating the effective measure (RR, OR, PR) and regression model was done using a simulation study based on the AETBC data. A continuous covariate and binary covariate were considered for the simulation. The binary covariate had 0.5 probability of success for all scenarios and the continuous covariate had a mean of 35 and variance of 1. Simulations for low, moderate and high prevalences were considered for low, moderate and high sample sizes. The nine scenarios had a baseline risk (_0 = 2:2), (_1 = 1:5) and [_2 = 0:04; 0:048(0:047); 0:09] which were used for data simulation. One binary independent variable and one continuous independent variable and different coefficients were used in the simulation study. Let Yi denote a binary outcome for subject i for (i = 1; _ _ _ ;N) and _i = P(Yi = 1jXi) be the probability of success, where Xi = (X1;X2). Ten thousand datasets were simulated and logistic regression, log-binomial and robust Poisson were used to analyse the datasets. The Bayesian logistic could not provide results with 10000 datasets because of its computer intensiveness and 1000 datasets were considered for the Bayesian model.

"Performance of Effect Measures Estimation methods in Studies with Common Outcomes: Simulation and AETBC study"

Ndamona
Haushona

Competing risks arise when subjects are exposed to more than one mutually exclusive cause of failure and failure due to one cause precludes the occurrence of failure from other causes. A competing risk is defined as an event that prevents the observation of an event of interest or that in principle modifies the probability of occurrence of the outcome of interest in a study 12. Death is one of the most common competing risks in health studies, as treatment effects or occurrence of an event of interest will not be observed because the participant is dead.

For instance, in our study, death precludes the occurrence of other unfavourable retention in care outcomes such as loss to follow-up or stopped treatment among People Living with Human Immune Virus (PLHIV) who were on Isoniazid preventive therapy (IPT) and antiretroviral therapy (ART). The death toll competes with observing the events under investigation, and it prevents us from knowing when an individual would have experienced another unfavourable retention in care outcome had they been alive. Even if the duration of follow-up is extended, a subject will never be observed as loss to follow up or stopped treatment once he or she has died. In literature, many researchers estimate the treatment effect or occurrence of an event in time to event studies without accounting for competing risks 3. Failures from other causes are treated as censored events and survival probabilities are estimated using standard survival methods, such as Kaplan Meier and Cox proportional hazard models. These methods assume censored events are independent of the outcome of interest. However, this is not the case in the presence of competing risks unless it can be demonstrated that competing risks are independent. Failure to account for competing events reduces statistical power and results in biased inferences, as the treatment effect or occurrence of an event of interest, may be distorted by the competing event 2,4,5.

Consequently, various techniques to deal with competing risks were developed, such as the ones by Fine and Gray 6 and Prentice at al. 7. These techniques yield unbiased results in the presence of competing risks and do not relax on the assumption of independent censoring.

Most studies also report composite outcomes with the aim of accounting for competing risks 28. The composite outcome is a combination of multiple individual outcomes into a single endpoint. Combining individual outcomes into composite outcomes can increase the overall event rate and the statistical power for a study 910. Ideally, a composite outcome comprises of outcomes that in principle are assumed to have similar importance, similar relative frequencies of occurrence, similar underlying etiology, similar precision of measurement, and similar magnitude and direction of treatment effects 10 8. The use of composite outcomes is deemed appropriate when the treatment effect across individual outcomes within a composite outcome is homogeneous 2,5. However, if a composite outcome is designed to quantify risk benefits or capture competing risks, there is no assumption of homogeneous treatment effect across individual components.

Standard survival techniques do not account for competing risks, and failure to account for the latter may lead to incorrect interpretations of the probability of the outcomes of interest. Therefore, the main aim of this paper is to reanalyse the data from 11 accounting for death as competing risk, with a focus on the composite of unfavourable retention in care outcome comprising of loss to follow up, stopped treatment and death. We will use Fine and Gray model to determine baseline covariates that are associated with the composite of unfavourable retention in care outcomes among HIV patients who were on IPT plus ART, while accounting for death as a competing risk and allowing for within hospital clustering. We will perform sensitivity analysis using two additional competing risk models and cox proportional hazards regression model to assess the robustness of the findings.

"Empirical comparison of time-to-event models to analyze a composite outcome in the presence of death as a competing risk"

Tinashe Mhike

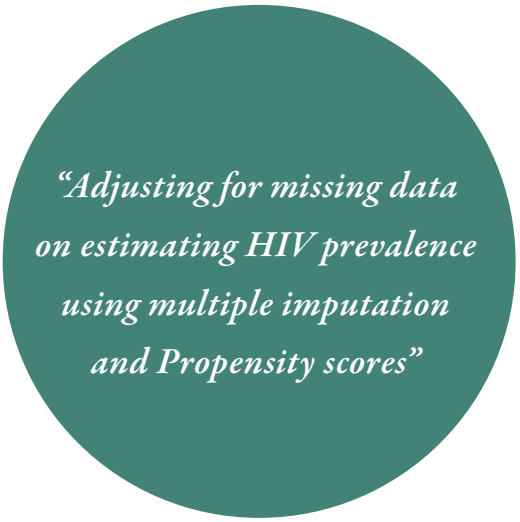
Population surveys and demographic studies are the gold standard for estimating national HIV prevalence. However, participation in these surveys is of major concern. Individuals may not participate because the interviewers could not contact them for interview or because they refuse to give consent to an HIV test or overall participation (McGovern et al., 2015). Complete case analysis is the most common method used which may result in biased estimates. To obtain better estimates of HIV prevalence where there is non-response, more advanced methods must be used to obtain unbiased estimates.

Population surveys and demographic studies are the gold standard for estimating national HIV prevalence. However, participation in these surveys is of major concern. Individuals may not participate because the interviewers could not contact them for interview or because they refuse to give consent to an HIV test or overall participation (McGovern et al., 2015). Complete case analysis is the most common method used which may result in biased estimates. To obtain better estimates of HIV prevalence where there is non-response, more advanced methods must be used to obtain unbiased estimates.

Prevalence measures the burden of disease in a population in a given location and at a particular time, as represented in a proportion of people affected. Estimates of HIV prevalence are frequently used to monitor and study the determinants of HIV epidemic, identify groups at high risk of HIV infection, and to assess the need for HIV prevention and treatment(Bärnighausen, Bor, Wandira-Kazibwe, & Canning, 2011). Population surveys and demographic studies have become the gold standard for estimating national HIV prevalence. However, participation in these surveys is of major concern. Individuals may not participate because the interviewers could not contact them for interview or because they refuse to give consent to an HIV test (McGovern et al., 2015). Non-response can bias population-based estimates of HIV if it is systematically associated with HIV status. This could occur for two reasons due to refusal to participate in HIV testing if an individual knows his status or individual involved in high sexual risk behavior (Marston, Harriss, & Slaymaker, 2008). Missing data in research can be classified into three mechanisms. Data can be missing completely at random (MCAR), which means that missingness is independent of the outcomes (observed or unobserved). Data can be missing at random (MAR), that is missingness can be dependent only on observed covariates. MAR assumes that the missingness of the data may depend on the observed data, but is independent of the unobserved data. Therefore, testing MAR is in general impossible since it requires unavailable information about the missing data (Li, 1988). Lastly, data can be missing not at random (MNAR), which means that data is neither missing completely at random nor missing at random. When missing data depends on both the observed and unobserved data, data are MNAR (Li, 1988).

An example of MNAR data occurs if an eligible study participant does not come for testing because they already know their HIV status, and because the missingness is directly related to the outcome, the data are missing not at random. If, however, a patient misses a test, because he had a long way to walk, then data would be MAR, because although missingness is not related to their HIV status, it may be related to their residence or other observed covariates. Finally, if the patient gave a blood sample but the sample was destroyed before it was tested then the missingness is not associated with their HIV status or any other observed covariates, then data will be missing completely at random (MCAR). When observations of a variable are missing completely at random, the missing observations are a random subset of all observations; the missing and observed values will have similar distributions. Missing at random means there might be systematic differences between the missing and observed values, but these can be entirely explained by other observed variables. For example, if HIV status is missing at random, conditional on age, residence, marital status and sex, then the distributions of missing and observed HIV status will be similar among people of the same age, residence, marital status and sex (Bhaskaran & Smeeth, 2014).

Most researchers use conventional methods such as the complete case or available case analysis where the assumption used is that data are missing completely at random, that is only the available data is analysed and assumed to be a random sample of the target population. Using these methods in presence of missing data results in loss of information and may lead to biased estimates of prevalence (Chinomona & Mwambi, 2015b).
There has been development of statistical methods that can be applied to adjust for missing data such as inverse probability weighting (IPW), maximum likelihood estimation and imputation methods.
The IPW methods rely on the intuitive idea of creating a pseudo-population of weighted copies of the complete cases to remove selection bias introduced by the missing data. However, different weighting approaches are required depending on the missing data pattern and mechanism (Robins, 2014). Maximum likelihood estimation and multiple imputations are the other methods used to adjust for missing data (Allison & Horizons, 2012). However, these methods are merely used either due to its complexity and software availability. Depending on the pattern and mechanism assumed some techniques are superior than other. In this study we will assume data to be missing at random and obtain prevalence of HIV adjusting for missing data using multiple imputation and propensity scores weighting.



Albert
Chinhenzva

Randomised controlled trials provide best evidence in terms of effectiveness of treatment. Randomisation allows estimation of causal effect by balancing characteristics between treatment groups. In longitudinal studies differential dropouts between treatment groups is one factor that is inherent and can negate treatment effects. Clinical trials are not often analysed using methods that account for missing outcomes and this may lead to biased estimation of causal effects. We used marginal structural models, a class of causal models that allow improved adjustment of confounding and selection bias introduced by differential dropout between study arms. We applied method of inverse probability weighting to account for dropout by calculating the propensity of being retained in the study based on covariates, treatment and outcomes. This was applied in a longitudinal

study assessing neurodevelopment outcomes in HIV infected and uninfected children. The application of inverse probability weighting in the longitudinal analysis produced different estimates from the intention to treat model. This demonstrates effectiveness of inverse probability weighting in accounting for dropouts.

The application of IPW in the longitudinal analysis showed an impact in the estimates of causal effects of different HIV treatments regimens to neurodevelopment in children in comparison to ITT and adjusting for covariates. Results from all three analysis strategies lead to the same conclusion, however we saw that failing to handle dropouts could potentially lead to reporting of biased estimates of causal effects in RCTs10. The IPW estimate was approximately 3% greater than ITT estimate which is not a large difference, however the confidence interval for IPW appeared wider than ITT because we used estimated weights which themselves may have some imprecision. Implementation of IPW allowed estimation of causal effects in contrast to associations which were established in the CHER study and ITT analyses6. Handling of dropouts is not limited to the IPW method but can also be done using likelihood based procedures and multiple imputation. Multiple imputation involves replacement of missing values by simulated numbers followed by creation of several versions of complete dataset which is analysed using complete case analysis methods to obtain estimates18. The other method used is Linear Increments (LI) which is based on assumption of hypothetical complete dataset paying particular attention to outcomes that would have been observed in the event of no dropout19,20.

Constructions of inverse probability weights is a much more thoughtful process which requires understanding of predictors of dropouts in children13. This analysis could be enhanced if more variables which predict dropouts in paediatrics are used in the propensity score model21. If the dropout model was misspecified then estimates from IPW methods were biased. Application of doubly robust methods such as augmented inverse probability weighting (AIPW) and targeted maximum likelihood estimation (TMLE) can be applied to reduce chance of bias caused by model misspecification15,22. The doubly robust methods use both dropout and outcome model thereby guaranteeing unbiased estimates if one of the models is not misspecified15. Simulation studies when used as standard of truth have shown that AIPW models perform better than other methods20. The method of choice for handling dropout must be based on understanding the underlying bias that dropouts may have created. If several methods are used to handle dropouts it is worthy to report all estimates and provide more information to readers21.

We recommend application of at least IPW method in handling dropouts for future longitudinal studies to allow unbiased estimation of causal effects4.

"Statistical methods for causal inference in the presence of dropouts in longitudinal studies"

Mapahla
Lovemore

The study aimed to review the imputation methods for missing data in cluster randomised control trials, review the software programmes that have been implemented for imputation and analysis and to apply some of the methods to a cluster randomised trial conducted in the primary clinics of the Eden district, in the Western Cape Province. Blood pressure measurements performed on all participants of the Eden trial were analysed using imputed data. This entailed using multilevel multiple imputation and performing an intent to treat analysis. Joint model and sequential regression multilevel imputation strategies (level 1: participants; level 2: clinics) were done and compared to single level multiple imputation (level 1 only) as well as the complete case analysis. Jomo and hmi package in R were used to do imputations from joint model and sequential

regression respectively. The results were pooled using the Rubin rules into final treatment effect estimate. The results based on the multilevel imputed data confirms the published results for the blood pressure outcomes (based on complete case analysis) showing no significant intervention effect.

This study gave me the opportunity to investigate how to execute an intention to treat analysis for a cluster randomised controlled trial and understand how to implement an appropriate imputation strategy to deal with incomplete data in cluster randomised trial. I managed to do analysed- as-randomised analyses after imputing data without ignoring the cluster effect.

Hmi and jomo (with mitml as interface) R packages were successfully applied to do the multiple imputations with and without considering the effect of clusters. However, the sequential modelling multilevel imputation using hmi took more time (at least 2 hours) to converge as compared to the joint model imputation using jomo which took roughly 30 mins to do multilevel multiple imputation. I hardly manage to save the imputed dataset implying that imputation process was repeated every time I restart my computer. Multilevel multiple imputation was justified in this study as compared to complete case analysis. We realised an increase in error margin due to a decrease in sample size when complete case analysis was applied. This implies that multilevel multiple imputation was slightly more precise as compared to complete case analysis in this study.

I hardly managed to obtain between patients within cluster variability estimates after sequential regression using the hmi package. This means that only between cluster variability were compared to the joint model imputation estimates

"Joint modeling and sequential regression multilevel multiple imputation of cluster randomised trial data"

Graduation

2020

Mercy
Rop

Mercy Rop completed her MSc. Biostatistics in the Division of Epidemiology and Biostatistics, Stellenbosch University in December 2020. She also holds a Masters' and a Bachelor's degree in Statistics from 2 universities in Kenya. She has close to 6 years' experience in management and analysis of health research data. Currently, Mercy is a teaching assistant for MSc. Biostatistics, Longitudinal Data Analysis Module. Her research interests include Longitudinal Causal inference, Survival Analysis and Clinical Trials.

Migration and the risk of HIV acquisition in Western Kenya: A Targeted Maximum Likelihood Estimation approach to a population-based data. Migration has long been identified as an independent risk factor and an important driver of HIV acquisition. Studies on HIV have often established a strong positive link between migration and individual's HIV status but whether migration is a cause or consequence of HIV infection in unclear. In this study, we investigated the causal relationship between migration and the risk of HIV acquisition using population-based pre and post migration data. Data on 13,646 subjects whose past HIV status was known to be negative and were between 15-60 years of age was extracted. Predictors of migration were explored; regression analyses were done to check associations between migration and HIV as well as migration and sexual behaviour in the past 12 months. Akaike Information Criteria was used for variable selection and incidence rate ratios and odds ratios were reported. Causal relationship between migration and HIV acquisition was then investigated using Targeted Maximum Likelihood Estimation method. This method was utilized because it has desirable statistical properties but has not yet been widely implemented in applied epidemiological research as other causal inference methods. Results from this study indicate that education, age, and socio-economic status were important predictors of migration in the surveillance area and that migrants had higher odds of engaging in casual sex than non-migrants. The risk of acquiring HIV infection was found to be about 2 times higher in migrants than in non-migrants but this risk was statistically insignificant at 5% level. The causal relative risk of HIV acquisition was about 2, this implies that if everyone in the population had migrated, on average, the risk of HIV acquisition would almost double as compared to if everyone had not migrated. This effect was however statistically insignificant at 5% level. The main challenge in this study was nonresponse in sensitive questions about sexual behaviour. Although the results from this study were statistically insignificant at 5% level, it is important to continuously explore the relationship between migration and HIV acquisition to gain insights on possible interventions to prevent new HIV infections.

Diribsa
Bedada

Handling of Missing data in HIV studies with Survey Weights Missing data are inevitable in most applied researches, including HIV surveys. Despite the increasing number of population-based surveys that provide testing and counselling for HIV over the past decade, understanding the impact of nonresponse in these surveys is still limited. HIV prevalence is usually obtained from nationally representative population studies, such as demographic health surveys (DHS). However, surveys often have a problem of missing data, which can bias the results and reduce study precision. Accurately

estimated HIV prevalence is key for the prevention and treatment of HIV, monitoring and evaluation of the ongoing programs such as sustainable development goals (SDGs), achieving 90-90-90 targets and the allocation of resources. The 2016 Ethiopian Demographic and Health Survey (2016EDHS) data which consists of 15,683 women aged 15-49 was used in this study. We applied widely used and advanced methods (complete case analysis, inverse probability weighting, standard multiple imputations, consent to HIV adjusted inverse probability weighting and multilevel multiple imputations) to handle missing data problems in HIV survey research. HIV prevalence was six times higher in urban than in rural areas (3.22% vs 0.56%). All statistical methods predicted an overall HIV prevalence of approximately 2%, while the reported national HIV prevalence was 1.20%. The 95% confidence interval of the predicted prevalence of the heteroscedastic multilevel multiple imputations was the narrowest, while that of complete case analysis and inverse probability weighting were the widest. Age, place of residence, religion, region, and education level were significantly associated with HIV status. Multilevel multiple imputation is preferable to handle missing data in HIV surveys, and to predict the prevalence of HIV as it accounts for survey weights correctly. The predicted prevalence obtained using all statistical methods implies that the reported national HIV prevalence is underestimated.

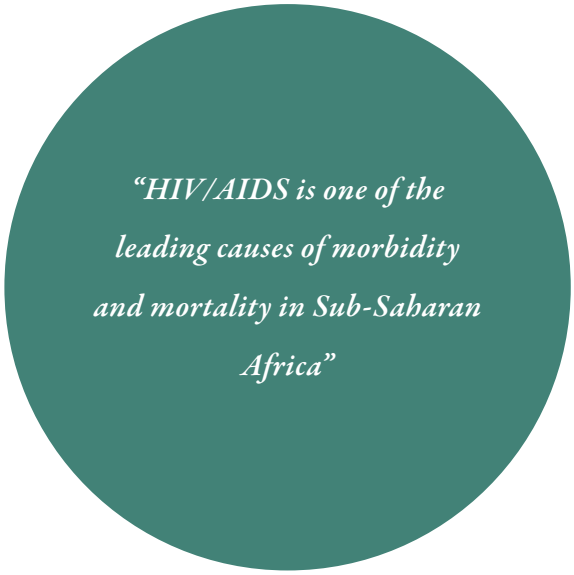
With a correctly specified imputation model, including a linear effect of weights (as a continuous covariate), and a linear interaction with the other variables, allows Rubin's rule to give a valid estimate of the variance. With correct model for consenting, consent to the HIV adjusted IPW will give unbiased estimates. Inverse probability weighting combined with an appropriate HIV prediction model can be a useful tool to correct for non-response to HIV testing, especially if the number of tested individuals is very minimal at subnational level.

Keywords: Survey weights, Inverse probability weighting, Multilevel multiple imputation, Missing at random, Demographic Health Survey

Moses
Ouma

HIV/AIDS is one of the leading causes of morbidity and mortality in Sub-Saharan Africa. Recent studies have shown that women are at higher risk of HIV, physical and sexual violence and correlates of HIV are associated with experiencing physical and sexual violence. However, little has been done on their joint mapping. Understanding HIV interrelationships can provide a basis for intervention. We, therefore, focus on assessing spatially the risk of HIV and experiencing violence against reproductive-aged women in the Kenya counties.

Methods: The study uses data from KENPHIA 2018-2019 national household survey in the 47 counties of Kenya. The study provides the descriptive statistics for each covariate by each outcome and strength of association. We also implemented Shared Component model (SCM) to assess the joint spatial distribution of HIV, physical and sexual. Further, the structural equation model was implemented to assess the interrelationship between the latent violence and HIV.



Results: The national prevalence of HIV, physical and sexual violence among women aged 15-49 years were 6.51 %, 12.48 % and 10.96 % respectively. HIV was most prevailing in Homa Bay (28.07 %), Siaya (20.84 %), Migori (19.90 %), Kisumu (17.88 %) with Wajir, Mandera and Garissa having 0 % prevalence. Physical violence was prevalent in Busia 24.44 %, Siaya (21.50 %), Kisumu (22.82 %), Migori (19.23 %), Trans-Nzoia (20.03 %) and Homa Bay (19.96 %). Counties with prevalence of sexual violence against larger than 15 % are Kisumu (21.28 %), Homa Bay (18.22 %), Siaya (21.04 %), Tharaka (15.17 %), Migori (15.53 %), Kwale (16.95 %). The scaling of HIV, physical and sexual violence from the SCM were 1.01[95% Cred. Int=0.47; 2.06], 0.98[95% Cred. Int=0.46; 2.08] and 0.90[95% Cred. Int=0.44; 1.87] respectively.
Conclusion: Physical and sexual violence show similar magnitude of association with HIV transmission as scaling estimates were close to 1.

Perseverence

Savieri

"Microarray transcriptomic analysis of tuberculosis relapse"

A longitudinal study of pulmonary tuberculosis (PTB) was undertaken to study molecular biology underpinning the differences in clinical outcomes. The outcome being measured was tuberculosis relapse and it was determined after a follow-up period of two years to determine whether any patients relapsed. The criteria for relapse were based on completing treatment successfully and being culture negative on two consecutive tests and presenting with a

new episode of TB resulting from the same strain as the first. Analysis of the data could provide valuable insight into the pathobiology of tuberculosis and the host that leads to relapse.

We investigated data from thirty-seven cured TB patients and thirteen patients who relapsed within two years after the initial cure. The analysis was conducted using blood samples from 214 Affymetrix U133 Human Genome Plus 2.0 GeneChips. These ex-vivo samples from PTB patients were taken at diagnosis and after 1, 2, 4, and 26 weeks of successful treatment. A Bayesian hierarchical model was used for the analysis of successful treatment of TB and adjustments were made for multiple comparisons using the Benjamini and Hochberg method. We observed a significant ≥ 1-fold change in expression of 97 genes during treatment. These changes comprised 69 genes with a down-regulated expression in week 2. This was followed by slower changes in expression at week 4 and the end of treatment week 26. Pathway enrichment analysis identified fifteen biological processes and genes that were differentially expressed between cured and relapse patients were mostly involved in activation and transmission of signals to the nucleus.

This study shows the biological processes underpinning the differences in clinical outcomes in a longitudinal study of pulmonary tuberculosis (PTB). However, the samples taken at different PTB treatment visits were affected by variability emanating from processing grants. Although missingness is a rare issue is microarray data analysis after performing quality control, our results cannot be generalized to all PTB patients as only 38% of the patients had complete data after follow-up. Additional studies are required to explore multiple imputation as most of the downstream analyses for microarray data require complete datasets. This means accurate algorithms for missing value estimation are needed for improving the performance of microarray data analyses. Furthermore, microarray data analysis requires higher computing power therefore we recommend using a server for analysing the data as this was a limitation in this study. Also, a less restrictive FDR cut-off of 0.1 and fold change cut-off of 2 can be used to get more differentially expressed genes on some of the time points. Other baseline characteristics that were not included in the design matrix and the contrasts matrix could also be incorporated to explore different conclusions.

Leonard
Mwandingi

Multiple imputation approach for non-responses in survey data: An application to the 2017 South Africa HIV nationally representative population-based survey

Shaalula Leonard Mwandingi1*, Alfred Musekiwa1,2, Khangelani Zuma3 and Musawenkosi Mabaso3
1Department of Global Health, Faculty of Medicine and Health Sciences, Division of Epidemiology and Biostatistics, Stellenbosch University, Francie van Zijl Drive/ Rylaan, Tygerberg, 7505 Cape Town, South Africa.
2University of Pretoria, School of Health Systems and Public Health, 31 Bophelo Road, Prinshof 349-Jr, 0002 Pretoria, South Africa.
3Human Sciences Research Council (HSRC), Human and Social Capabilities (HSC) Research Division, 134 Pretorius Street, 0002 Pretoria, South Africa.

Abstract

Background: Missing data poses a pervasive challenge in health research, especially in population-based surveys where there is a risk of non-respondents being systematically different from respondents. Despite the current software resource capabilities to do advanced analysis for handling missing data, complete case procedures that are only valid under missing completely at random mechanism are still commonly used in surveys. These procedures result in biased estimates, loss of information, and reduced efficiency.

Methods: In this study, we motivated for the use of the multiple imputation approach as an appropriate method for handling missing data as opposed to complete case analysis. The South Africa's 2017 National HIV Prevalence, Incidence, Behavioural, and Communications Survey was used as an application example, which was a population-based household survey with a complex sampling design. The pattern of missing data in this study followed an arbitrary or non-monotone missing pattern hence a multivariate imputation using chained equations (MICE) was considered. Also, a survey logistic regression was applied to identify the determinants associated with HIV status. All analyses were done in Stata/IC 15.1.

Results: The analysis included 24 966 participants aged 15 years or older who completed the survey interview. About 17 347 (69.5%) of the participants who completed the survey interview provided a blood specimen for HIV testing, of which 2 828(16.3%) participants were HIV positive of which 2 016 (19.5%) were women and 809 (11.7%) were men. Both approaches elicited an overall estimated HIV prevalence of 18.8% (95% CI: 17.5%, 20.1%) and females had the highest HIV prevalence of 23.1% as compared to males (14.2%). The results showed that the odds of HIV positive status were over three times higher for females as compared to males (OR 3.72, 95% CI: 2.114, 6.546) under the complete case approach and less than three times (OR 2.602, 95% CI: 1.857, 3.646) under the multiple imputation approach.

Conclusion: Despite no notable statistical difference in results between the two approaches in this study, it was evident that the multiple imputation approach yielded reliable results compared to the complete case approach. Its 95% confidence intervals were narrower and standard errors were smaller. This can be attributed to the ability of the multiple imputation method to account for uncertainty during the imputation process.

Keywords: missing data; multiple imputation; complex design; HIV/AIDS

"the multiple imputation approach yielded reliable results compared to the complete case approach"

Makabongwe
Nombula

Topic: Impact of Missing Data on the Modelling HIV/AIDS Progression in Subjects Receiving Antiretroviral Therapy (ART)
Background: Multistate Markov models (MSM) offer great tools for modelling disease progression. In addition, MSM offers a flexible tool for gaining a deeper understanding of how subject characteristics influence disease progression and process.

Methods: The data used for this study was collected for the HIV Centre for the AIDS Programme of Research in South Africa (CAPRISA), (CAPRISA) cohort study, running from June 2004 to August 2013. Four World Health Organization (WHO) severity states were considered: Stage 1 is where CD4+ T cells ≥ 500 CD4+ T cells/mm3; Stage 2,is 350 ≤ CD4+ T cells/mm3≥499; Stage 3 is 200≤ CD4+ T cells/mm3≥ 349; Stage 4 is <200 CD4+ T cells/mm3. Participants were allowed to transition between all the WHO severity states, and a semiparametric time-homogenous multistate Markov model was used to obtain the estimates of interest.
Results: This study consisted of 4014 subjects which contributed to 26 048 observations, 24469 of these observations had time recorded transitions and contributed 12684 transitions. According to the participant transition, factors such as environmental site, gender, TB and WHO staging at baseline were associated with both delay and acceleration of immune recovery and immune deterioration. Subjects with tuberculosis (TB) had a high immune deterioration rate, males had a high immune recovery, and subjects at WHO stage 2 were associated with immune recovery while subjects at WHO stage 4 transitioned very slowly from this state.

Conclusions: MSMs are the best tool for studying HIV/AIDS trajectories. It gives several different estimates that support each other for both immune recovery insight and immune deterioration. MSM also offers one to adjust for covariates to study these process effects are affected by one's characteristics. MSM can identify risk factors associated with HIV progression and disease processes. Not only providing treatment can help but finding population that is vulnerable is known to be significant in public health, this include teaching young girls and boys on how to protect themselves or how effective the treatment is especially one can be committed attend all the visits that are desired to be attended.