ASSAf
ACADEMY OF SCIENCE OF SOUTH AFRICA

South African Journal *of* Science

120 years

Shane Redelinghuys
National Institute for Communicable
Diseases, South Africa

Submissions
Submissions should be made at
www.sajs.co.za

Cover caption
Portulacaria afra bush in Worcester, South Africa. *Portulacaria afra* – commonly known as spekboom or elephant bush – is a medicinally important plant species endemic to South Africa. Basson and colleagues investigated the effect that rising atmospheric CO_2 levels may have on the medicinal properties of this plant.

Photo by Dean Erasmus on Unsplash

The university of global excellence

In many ways, and this is as it should be, the *South African Journal of Science* is at the heart of what may be termed the 'establishment' of academic and research life in South Africa. As a journal soon to celebrate 120 years of operation, we are part of our country's science history and present, marked by the intertwining of science with instruments of power and influence. In recent years, we have been very active in trying to increase access to our pages and to broaden our readership, as evidenced, for example, by our Inclusive Language Policy. But we are not a journal which sets up to disrupt, and our Vision and Mission are cast in very general terms, with a wish to include and diversify, rather than to take a specific political stand as a journal. Although we do not have the data to support this view, we imagine that among our authors and readers there are wide differences of opinion on what the statement "excellent South African research for the local and global academic community," part of our Vision, does and should mean in practice.

This said, the heart of good science (and of good academic practice more broadly) is constructive, robust debate, and we are pleased that we are able to showcase debates in our journal. No academic, however much a part of the 'establishment' they may be, should shy away from critique and contestation. Dogma is the enemy of sound academic practice. It is in this spirit that when we came across Sioux McKenna's poem, 'Welcome to the University of Global Excellence', we asked her if we could include it in an editorial in our journal. We are delighted she agreed. Sioux McKenna is the Director of the Centre for Postgraduate Studies at Rhodes University in South Africa, a globally known expert on higher education and a disciplined and thoughtful scholar on a range of topics – including the neoliberal university. We like her poem for a number of reasons; for example, it is great fun, and it speaks directly to experiences that many academics globally (and especially in South Africa) will recognise. Everything is there in the poem: the corporatisation and branding of universities, the peddling of 'wellness' in a context of increasing academic demands and precaritisation of the academic workforce; the chasing of numbers and rankings when true quality in context is much harder to measure; the universities in the Global South still being in thrall to metrics designed from and for the Global North; and even the peddling of 'decolonisation' as a marketable product. And to top it all, this poem is meticulously referenced, with pointers to very important scholarly works.

Welcome to the University of Global Excellence

Dear New Staff Member,

Congratulations!
You are joining a prestigious brand,
ranked in the Top Five[1]
hundred or so[2].

Most of our permanent academics[3]
rated working for us to be satisfactory[4].

Student evaluations indicate that our lecturers[5],
are four and a half on a scale of 5.

Please ensure that you sign your performance agreement.[6]
KPIs are publications.[7]
Terms and conditions are not negotiable.

Please join the Division of People's Wellness and Culture[8]
for online Meditations-at-Midday[9]
to enhance productivity in afternoon meetings[10].

Our Division of Customer Services is here to serve our students[11],
who are reminded to sign the disciplinary code[12].
Every assignment must include a statement of originality,
and be added to Turnitin's® database.

Our Division of Teaching Excellence and Policy Compliance[13]
drives throughput in regulation time.
Please join their monthly NAFFE workshops:
Necessary Acronyms For Full Engagement.

The Division of Quality Enhancement, Assurance, Reporting,
Monitoring and Evaluation[14]
kindly requests that templates be completed quarterly[15].
They must be signed off by the
Executive Head of Department,
Executive Dean[16],
and the Executive Director of the Division of People's
Wellness and Culture.

Welcome to the University of Global Excellence!
We are an African University committed to decolonised education.
Sioux McKenna

Not everyone will agree with all the sentiments in the poem, and not everyone will agree with our decision to publish it, but that is the nature of debate. We enjoyed reading the poem and learned things at the same time – and we hope that our readers do as well.

References

1. Hazelkorn E. Reshaping the world order of higher education: The role and impact of rankings on national and global systems. Policy Rev High Educ. 2018;2(1):4–31. https://doi.org/10.1080/23322969.2018.1424562

2. Marginson S. Global stratification in higher education. In: Slaughter S, Taylor B, editors. Higher education, stratification, and workforce development. Higher Education Dynamics vol. 45. Cham: Springer; 2016. p. 13–34. https://doi.org/10.1007/978-3-319-21512-9_2

3. Shihaam S, Du Plessis M. Temporary academics in South African higher education institutions: Their lived precarity experiences. J Psychol Afr. 2023;33(5):522–529. https://doi.org/10.1080/14330237.2023.2258001

4. Loveday V. The neurotic academic: Anxiety, casualisation, and governance in the neoliberalising university. J Cult Econ. 2018;11(2):154–166. https://doi.org/10.1080/17530350.2018.1426032

5. Darwin S. What contemporary work are student ratings actually doing in higher education? Stud Educ Eval. 2017;54:13–21. https://doi.org/10.1016/j.stueduc.2016.08.002

6. Kairuz T, Andriés L, Nickloes T, Truter I. Consequences of KPIs and performance management in higher education. Int J Educ Manag. 2016;30(6):881–893. https://doi.org/10.1108/IJEM-05-2015-0067

7. Kenny J. Academic work and performativity. High Educ. 2017;74:897–913. https://doi.org/10.1007/s10734-016-0084-y

8. Badr S. Re-imagining wellness in the age of neoliberalism. New Sociol J Crit Praxis. 2022:3–9. https://doi.org/10.25071/2563-3694.66

9. Carvalho A, Grácio R. The dark side of mindfulness: Workplace socialization, neoliberalism and the self. Commun Lang Work. 2022;8(2):63–77. https://doi.org/10.7146/claw.v8i2.135105

10. Deem R. New managerialism in higher education. In: Teixeira PN, Shin JC, editors. The international encyclopedia of higher education systems and institutions. Dordrecht: Springer; 2020. p. 2083–2088. https://doi.org/10.1007/978-94-017-8905-9_308

11. Bunce L, Baird A, Jones SE. The student-as-consumer approach in higher education and its effects on academic performance. Stud High Educ. 2017;42(11):1958–1978. https://doi.org/10.1080/03075079.2015.1127908

12. Naidoo R, Williams J. The neoliberal regime in English higher education: Charters, consumers and the erosion of the public good. Crit Stud Educ. 2015;56(2):208–223. https://doi.org/10.1080/17508487.2014.939098

13. Alvesson M, Spicer A. (Un)Conditional surrender? Why do professionals willingly comply with managerialism. J Organ Chang Manag. 2016;29(1):29–45. https://doi.org/10.1108/JOCM-11-2015-0221

14. Shore C, Wright S. Audit culture revisited: Rankings, ratings, and the reassembling of society. Curr Anthropol. 2015;56(3):421–444. https://doi.org/10.1086/681534

15. Churcher M, Talbot D. The corporatisation of education: Bureaucracy, boredom, and transformative possibilities. New Form. 2020;100:28–42. https://doi.org/10.3898/NewF:100-101.03.2020

16. De Boer H, Goedegebuure L. The changing nature of the academic deanship. Leadership. 2009;5(3):347–364. https://doi.org/10.1177/1742715009337765

HOW TO CITE:

AUTHORS:
Hannelie Human[1]
Robin Crewe[1,2]
John Sharp[2]

AFFILIATIONS:
[1]Department of Zoology and Entomology, University of Pretoria, Pretoria, South Africa
[2]Centre for the Advancement of Scholarship, University of Pretoria, Pretoria, South Africa

CORRESPONDENCE TO:
John Sharp

EMAIL:
john.sharp@up.ac.za

Sue Nicolson (1950–2023): Internationally recognised insect physiologist and dedicated mentor

The South African scientific community lost a biologist of rare quality with the death of Professor Sue Nicolson on 27 April 2023. Her passing is a very sad loss to her family, friends, colleagues, and the many students she mentored over a long and influential career, first at the University of Cape Town and subsequently at the University of Pretoria.

Sue was born in Dargaville, New Zealand in 1950, and completed a BSc (Hons) in Zoology, with distinction, at the University of Auckland. She was encouraged by her Auckland supervisor to study further at Cambridge University, and was one of the first women to be admitted to King's College in 1972. She completed a PhD degree in Insect Physiology, based on research on osmoregulation in several different insect species. Over time she developed an international reputation in this field of study, leading to an important review article in 1993 on the ionic basis of fluid secretion in insect malpighian tubules.

In 1975, Sue married a South African who was also at Cambridge – they were the second King's couple to marry in the centuries-long history of the college. In 1978, they moved to the University of Cape Town, where Sue held a teaching position in the Department of Zoology for a couple of years before becoming a research fellow in Professor Gideon Louw's laboratory. She later returned to the teaching staff, and her expertise and experimental skill led to her undertaking a series of studies on Namib Desert beetles that furthered scientific understanding of their adaptation to the harsh environment they inhabited.

Sue's interest in osmoregulation also stimulated her to explore the way in which nectar-feeding carpenter bees deal with excess water in their food and conserve ions. This resulted in her focusing most of her subsequent research, at the University of Cape Town and then at the University of Pretoria, on various nectar-producing plants and the insects and birds that forage on them. In later years, before and after retirement, Sue extended these studies with research on the feeding preferences of sunbirds and the metabolism of the nectar sugars on which they feed, as well as on the nutritional physiology of honeybees. Some of her work on bees was undertaken in collaboration with colleagues at universities around the world, including Oxford and leading universities in Italy, Israel, and China. Sue often said that her bee research marked a return to questions with which she had grappled in her doctoral research at the start of her career – she examined how forager bees succeeded in reducing flight energy cost, for individual bees and their colonies as a whole, by getting rid of most of the water in nectar at the flower and in flight.

Sue's bee research was timely, given the ongoing threat to the world's bee population, and she and several international colleagues used their extensive knowledge to develop and patent a feed supplement that could potentially help bees survive a considerable crisis.

Sue and her husband moved to the University of Pretoria in 2001, with the university creating a chair in Zoology for her as an inducement to leave Cape Town. At both universities she made very significant contributions to undergraduate teaching and the mentoring of postgraduate students and postdoctoral fellows. She had a profound impact on many of her postgraduate students, who acknowledge that the opportunity to work with her shaped the course of their subsequent careers. She had a particular liking for mature postgraduates who showed willingness to put in long hours in the field and the laboratory. Sue could be hard on students who did not display sufficient enthusiasm for their work, but was kind and accommodating to those who did, and made a point of visiting them at their field sites in order to offer encouragement and guidance. She published extensively in collaboration with postgraduates who showed promise, and these publications are a legacy of her influence on them.

In addition to regular funding from South Africa's National Research Foundation, she received funding from the Insect Pollinator initiative of the UK's Biotechnology and Biological Sciences Research Council (BBSRC) from 2011 to 2014, and collaborated in the BBSRC's study of the macronutrient regulation of adult worker honeybees from 2017 to 2019. She was on the national steering committee of the Global Pollinator Project, South Africa from 2009 to 2014. Sue received the Exceptional Achiever Award from the University of Pretoria each year from 2003 to 2016.

Sue was the author of some 185 scientific publications that established her as an internationally recognised research scientist (NRF B1 rating). With co-author Steven Chown, she wrote *Insect Physiological Ecology: Mechanisms and Patterns* (Oxford University Press, 2004), which won the prestigious Bill Venter/Altron Literary Award for the best book in science in 2008. The book reflected the depth of their insights into physiological ecology and their comprehensive grasp of the field. The widespread and growing interest in pollination biology prompted Sue to co-edit the book *Nectaries and Nectar* (Springer, 2007) with Massimo Nepi and Ettore Pacini, which has become a key text for researchers interested in nectar worldwide.

Sue published both books during her term as Head of the Department of Zoology and Entomology at the University of Pretoria (2003–2011), and maintained all her teaching commitments at undergraduate and postgraduate levels throughout this period. Colleagues recall that she was a brave head of department – the first woman to hold the position in the department's history – who could also be impatient and blunt if she believed others were barking up the wrong tree on important issues. Sue's years in office were ones in which Zoology and Entomology was the most research-intensive academic department in the university, and was recognised as a role model for research and teaching excellence.

Sue was a fellow of the Royal Society of South Africa, of the Royal Entomological Society, and of the Cambridge Philosophical Society. She received the Gold Medal of the Zoological Society of Southern Africa, as well as the Chancellor's Award for Research from the University of Pretoria in 2013.

In addition to her academic work and achievements, Sue played a major part in nurturing her two children (twins). She encouraged them through school and university, and into highly successful careers, without ever being prescriptive. She trusted their judgement from an early age, and delighted in their achievements. In her spare time, she was a gifted artist who produced many paintings on New Zealand and South African landscapes.

Sue was a warm and engaging, and at the same time a challenging, colleague who will be sorely missed in the Department of Zoology and Entomology, and in the University of Pretoria's Centre for the Advancement of Scholarship where she was a senior research fellow after her formal retirement.

Sue is survived by her husband and children, and by a granddaughter who was born shortly before she died. Knowing how advanced her illness was, Sue counted herself fortunate to have the opportunity to see her granddaughter and hold her in her arms before the end.
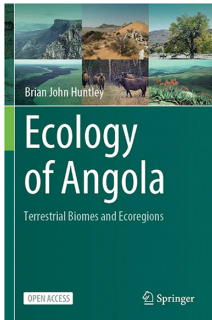
AUTHOR:
Brian J. Huntley

REVIEWER:
Brian W. van Wilgen[1]   iD

AFFILIATION:
[1]Centre for Invasion Biology, Department of Botany and Zoology, Stellenbosch University, Stellenbosch, South Africa

EMAIL:
bvanwilgen@sun.ac.za

New ecological textbook for Angola

Angola is one of the most ecologically diverse countries in the world, with terrestrial ecosystems ranging from the tropical rainforests of the Congo Basin to the hyper-arid Namib Desert. Although there is a rich and expanding literature on the terrestrial ecology of southern and eastern Africa, almost no work has been conducted in Angola until fairly recently. Consequently, there is very little in the way of relevant academic resources for training and supporting Angolan students and researchers in the ecological sciences. In addition, the scarcity of literature in Portuguese has been a critical barrier to the development of a new generation of Angolan scientists and ecosystem managers. Although this book is in English, a Portuguese translation is nearing completion and will be published shortly. In addition, this book is available online as an open access resource that can be downloaded free of charge at https://link.springer.com/book/10.1007/978-3-031-18923-4. Both of these features mean that the book will be widely accessible to academics, students and ecosystem managers in Angola.

This volume is primarily a textbook for students. It deals broadly with ecological concepts, and goes on to illustrate the relevance of these concepts in Angola's biomes, which include rainforests, grassland-forest mosaics, arid and mesic savannas, the Namib Desert, mangroves, and an escarpment zone. Each of these is treated in its own chapter in the book. In so doing, it is more than just a textbook as it contains up-to-date assessments of the state of knowledge pertaining to each biome. In the opening chapter, Huntley explains that the book covers selected ecological concepts, illustrated with local examples and supported by references to many other sources. A feature of the book is its focus on ecological concepts and terminology, with each term or concept printed in bold type where it occurs throughout the text, and defined at first mention. These definitions are collected in a glossary of ecological terms at the end of the book for ease of reference.

Brian Huntley is well positioned to write this volume, having worked as an ecologist in Angola during his early career, and having maintained contact with colleagues and undertaken many visits and expeditions to the country since then. He also consulted numerous leading experts while compiling this volume to ensure that the concepts and the content were accurate and up to date. The result is a detailed and extremely useful overview of the ecology of a little-known but important part of Africa. The book is also very well illustrated with colour photographs, maps and diagrams, which will enhance its appeal. It is perhaps worth noting that this book deals with the ecology of undisturbed ecosystems, how they came about, and the forces that maintain and shape them. It provides brief synopses on emerging environmental problems relevant to Angola, focusing on land degradation and transformation, invasion by alien species, climate change, and resource overuse and pollution, all of which will require urgent attention from ecologists. It thus provides an introduction to the baseline understanding vital to addressing these elements of rapid global change.

Open access publishing comes at a cost. Funding for this venture arose from strong collaboration between the Angolan and Portuguese ministries of science, technology and higher education. This admirable relationship between an African country and its former coloniser has already resulted in several important publications. This book is a very welcome addition, and I would highly recommend it to students of ecology, not just in Angola, but elsewhere in Africa and beyond. For Angolan students and lecturers, I would like to think that having a resource that so clearly illustrates the country's enormous diversity and value will engender a sense of pride and ownership in a country that has suffered from war and neglect for too long.

REVIEWER:
Gabriel Konayuma[1,2]

AFFILIATIONS:
[1]Senior Vocational Education and Training Officer, Ministry of Technology and Science, Lusaka, Zambia

[2]Country Representative (Zambia), e/merge Africa, Cape Town, South Africa

EMAIL:
gkonayuma@gmail.com

The missing link in modern teaching? A review of *Co-Teaching and Co-Research in Contexts of Inequality*

The subject of co-teaching and co-research in contexts of inequality is timely and useful for transforming teaching and learning in higher education and vocational education. As we emerge from the COVID-19 pandemic, co-teaching and co-research have become very important.

The chapters in this book originated from a call for contributions that was shared before the pandemic and that invited colleagues to share their experiences and reflections on teaching and researching across institutional and geographical boundaries (p. xxxiii). This edited collection consists of 15 contributions from more than 40 international authors from Africa, Europe, the USA, South America and Australia. Generally, most books in the educational field are authored by men, so it is refreshing that more than 70% of the contributors to this book are women. Given that the editors of the book are based in South Africa, the majority of African authors are from South Africa. The book would have been made richer with an inclusion of authors from West Africa, Central Africa and Asia.

The book focuses on how networked learning can be used to connect Africa and the world when co-teaching and co-research are undertaken. This is especially useful in addressing the contexts of inequality that characterise the higher education landscape of Africa. Shangase et al. note that the power of networked learning to connect staff and students globally has never been more evident than during the last few years, during which higher education globally pivoted to online learning (p. xxxiii). Hodgkinson-Williams states that the particular perspective being explored is the value of networked technologies to enable, broaden and sustain team teaching and/or collaborative research within and beyond Africa (p. xxxv).

The book is divided into two main sections: 'Connecting Africa through co-teaching and co-research' and 'Connecting Africa and the world through co-teaching and co-research'. Section 1 consists of six chapters, connecting institutions both within South Africa and the African continent.

Chapter 1 is an introduction to the book by the editors. The editors' definition of co-teaching has much in common with what the literature terms 'Collaborative Online International Learning', which connects classrooms across geographical locations. The book aims to encourage readers to enter co-teaching and co-research opportunities and to contribute to a much-needed conversation around collaborating across unequal contexts and to add to the growing body of knowledge in this field (p. 9).

Chapter 2 provides insights into the interplay that occurred between structure, culture and agency when course facilitators from different higher education backgrounds collaboratively facilitated a postgraduate course. Challenges in collaborating across such differently placed institutions are honestly addressed by the authors. One challenge noted is the inequalities observed within all of the represented higher education institutions and manifested to various degrees of prominence (p. 17). This chapter will be useful to facilitators and managers planning to implement co-teaching as they navigate institutional differences. The results arising from the authors' reflections on co-designing and co-facilitating a module involving structure, culture, agency, curriculum and pedagogy (p. 23–29) are useful for the reader who is experienced or new to co-teaching. For example, on agency, "a strong suggestion for a shared passion for social justice and acute sensitivity for inequalities" is reported (p. 26). This aligns very well with the aim of the book which is to contribute to the conversation around collaborating across unequal contexts.

Chapter 3 shares the reflections of colleagues from three institutions who began collaborating during a one-year Postgraduate Diploma in Educational Technology course at the University of Cape Town. One of the factors reported to have enabled collaboration in the co-research among the authors is group cohesion established by the course design, co-teaching/co-learning approach and block format (p. 46). Researchers seeking to collaborate on research will find useful the reflections on how the authors managed challenges in collaborating across geographical boundaries and the factors that enabled collaboration.

Chapter 10, which is part of the second section of the book, explores how connected co-learning and co-teaching can be designed for inter-institutional collaboration, in online and blended global studios across cultural boundaries – in this case the Global South and North. The conversations of the collaborators were guided by seven themes (p. 181). The chapter ends with four design principles for the conceptualisation and implementation of connected co-learning and co-teaching in online and blended global architecture studios:

1. Employ relevant technologies and techniques through learning design.

2. Acknowledge students as partners to promote student agency and well-being.

3. Consider flexibility through multiple interlinked learning settings and modes.

4. Recognise humanity, humour, culture and community (p. 194–195).

If the reader wanted only an overview of this chapter, I would recommend reading these learning design principles on their own.

Chapter 11 discusses emerging principles for online cross-cultural collaborative research. These principles are based on what the researchers learnedabout their own practices, to ensure committed and sustained engagement in collaborative research online in South Africa, Uganda and the USA. These principles are divided into two main types: inter-personal principles and process-based principles. Inter-personal principles are divided into relationship focused, safe, supportive spaces, collaborative and shared passion. Flexibility, evolving, applied, reflective and contextual are discussed under process-based principles. The reader can find more detailed discussion of these principles on pages 216–218. The emerging principles discussed in this chapter provide views and perspectives on co-teaching and co-research using networked learning, and as such achieve the overall purpose of the book as outlined in the preface. However, the chapter would have been enriched by a collaborator from another continent.

There are very few books, if any, that deal with co-teaching and co-research in contexts of inequality in the way that this book does. *Co-Teaching in Higher Education: From Theory to Co-Practice* by Daniel Jarvis and Mumbi Kariuki (University of Toronto Press; 2017) comes close to this book, but falls short in only dealing with co-teaching and not with co-teaching and co-research online. I would recommend *Co-Teaching and Co-Research in Contexts of Inequality* as a handy reference guide to which one can refer again and again in their co-teaching and co-research journey.
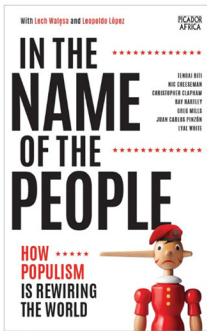
AUTHORS:
Tendai Biti, Nic Cheeseman, Christopher Clapham, Ray Hartley, Greg Mills, Juan Carlos Pinzón and Lyal White

REVIEWER:
Johannes Machinya[1] iD

AFFILIATION:
[1]Department of Sociology, University of the Witwatersrand, Johannesburg, South Africa

EMAIL:
joemachinya@gmail.com

A comprehensive analysis of populism: *In the Name of the People*

In the Name of the People: How Populism is Rewiring the World is a collaborative work from a diverse group of contributors with different expertise and perspectives on populism: Tendai Biti is a prominent politician and former finance minister of Zimbabwe, Nic Cheeseman is an accomplished scholar of African politics, Christopher Clapham is an esteemed expert on political history, Ray Hartley is a seasoned journalist, Greg Mills is an accomplished author on political economy, Juan Carlos Pinzón is a former ambassador of Colombia to the United States and a member of the Brenthurst Foundation's Advisory Board, and Lyal White is an independent researcher specialising in political economy issues in Africa, Asia, and Latin America. The book offers a comprehensive exploration of how populism, which at some point in time was associated with dysfunctional political formations in Latin America, has morphed into a global political phenomenon.

In the Name of the People draws on different case studies across diverse continents and historical contexts – from the old-style populist movements in Latin America, to Africa's liberation-movement populist formations, to new-order populist movements around the world – to underscore the diverse and common properties of populist movements in different contexts, highlighting what constitutes populism, its origins and causes, its strategies, and impact on society. As a result, the book gives a rich and widely researched account of populism that not only demonstrates its complexity but also its paucity of generalisations. This is because, as one understands from the book, populism does not embody a distinctive political regime and is not ideologically confined, or a mere transient political trend. The book shows that, despite some populists leaning towards one or another style of populism, populism in general adopts a language, style, or content that is shaped by the political culture of the society in which it emerges. This makes populism an adaptable political style for political parties and actors on all sides of the political spectrum, from the right to the left, giving it no particular ideological content. But despite populism's lack of conceptual and definitional clarity, the book shows that the political phenomenon is deeply rooted and possesses the power to reshape the political landscape in many countries around the world. With this offering, the contributors of this book avoid falling into the trap of oversimplifying the populist political phenomenon. Instead, they give a nuanced analysis and acknowledge the variations of populism that emerge in different contexts.

While highlighting that there are different flavours of populism, *In the Name of the People* also provides valuable insights into the strategies employed by populist leaders to gain and maintain power, and the consequences that such actions have on democracy and the economy. The authors show that populism's political rhetoric is less/not grounded in ideology but rather gives reference to 'the people' whereby it evokes latent grievances and appeals to emotions that resonate with broad segments of the population. This political rhetoric mobilises the idea of 'the people' that separates 'the people' from an enemy 'other', whereby the 'other' is constructed as posing an existential threat to 'the people', thus creating a Manichean divide between 'us' and 'them'. The classic populist strategy is rooted in anti-elitism, which sets the people against an enemy elite. Depending on the context, the enemy of the people can be constructed as a corrupt political or ethnic elite, opposition political formations, or former colonial powers, as in the case of liberation-movement populist formations. However, substantial populist energy today in Europe, North America, and other parts of Africa is being directed against a demonised foreign 'Other', the ('illegal' migrant). By looking into this populist strategy – which in many cases evokes a politics of fear, fear of the enemy 'Other' – the book is able to point to the reason why populism can be so appealing to large sections of society. The book further underlines the deleterious effects that populism has on society: it results in the hollowing of democratic institutions and abuse of human rights; corruption thrives under populism, and the economy ultimately crumbles, plunging many people into political and economic misery.

Beyond its captivating analytical approach, *In the Name of the People* is an accessible text which makes the book an important and comprehensible resource beyond the academic and policymaking circles.

All in all, *In the Name of the People: How Populism is Rewiring the World* makes an invaluable contribution to the subject of populism that is accessible to a wider readership. The book provides a thorough analysis of the origins of populism, its strategies, and its impact, not only on democratic governance but also on the wider society. The book is highly recommended to anyone curious to understand the phenomenon of populism and how it is reshaping the political landscape around the world.
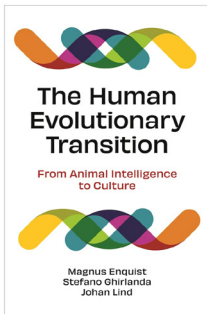
AUTHORS:
Magnus Enquist, Stefano Ghirlanda
and Johan Lind

REVIEWER:
Sarah Wurz[1,2]

AFFILIATIONS:
[1]School of Geography, Archaeology
and Environmental Studies, University
of the Witwatersrand, Johannesburg,
South Africa
[2]SFF Centre for Early Sapiens
Behaviour (SapienCE), University of
Bergen, Bergen, Norway

EMAIL:
sarah.wurz@wits.ac.za

Becoming human: Expert sequential and flexible thinking led to cumulative culture

When, why and how humans started to think and behave in typically human ways is a topic that has long fascinated and vexed scholars of hominin evolution. The origins of 'behavioural modernity' are discussed and debated in a variety of fields – it is, for example, close to the hearts of archaeologists and evolutionary anthropologists working on the material cultural legacy of *Homo sapiens* and their ancestors in Africa.[1,2] This book provides a bold new theory on becoming human that considers why human intelligence did not evolve in other animal species. It is a refreshing perspective on the topic as it comprehensively reviews animal behaviour in a wide range of species, from rats to pigeons, octopuses, and elephants. Humans represent a small part of the animal kingdom, yet most animal behaviour, apart from that of higher primates, is never discussed in accounts of human behavioural evolution. This new theory emphasises how mental mechanisms relate to behaviour rather than the brain; this avenue of research is preferred as much more is known about behaviour than is known about the brain. It is hoped that this theory would avoid the disagreement and fragmentation characterising cognition and learning research by not relying on concepts such as cognition, intelligence, planning, and reasoning that are 'inherently slippery' to describe human and animal behaviour. However, a degree of conceptual ambiguity is probably unavoidable, as illustrated by the focus on 'intelligence' in this book.

A 'strong' mathematical theory – in the sense that it can generate empirical predictions – is put forward. A model of sequence learning based on information, learning and decision is created and a simulation environment is developed to specify experimental designs and mental mechanisms. Combinatorial dilemmas stand in the way of productive behavioural sequences as flexible solutions are costly and time-consuming. Enquist et al. suggest that natural selection would have restricted mental flexibility to decrease learning costs. Therefore, a cost–benefit analysis is an important component of the model that is used to evaluate whether a mental mechanism would have been a viable evolutionary solution to a problem.

An impressive number of animal studies conducted under controlled conditions, from the early 1900s until recently, are reviewed. The authors conclude that chaining through association is the mental mechanism accounting for learning in animals. Animals tackle combinatorial dilemmas by learning short sequences and simple algorithms, favoured by genetic predispositions and the ecological and social environment. The view that associative learning is frequently superior to human-type thinking is a welcome reminder of animal cognitive genius in a research field dominated by anthropocentrism.

Human thinking happens when information from different events is recombined, and causal relationships between events are remembered. Humans became experts in sequential information processing and mental flexibility. Mental flexibility involves a series of mental procedures and intermediate memory states that retain the results of mental processes – these components are collated into a behaviour. Adaptive filtering or the preferential retention of useful information is a crucial aspect of human intelligence. In human thinking, the combinatorial cost incurred by mental flexibility is offset by using culture as a "reservoir of mental and behavioural skills". In this way, individuals can learn skills far beyond what they can learn and discover for themselves, thus setting the scene for cumulative culture. This strategy enables humans to consider increased possibilities and overcome the combinatorial dilemmas that hamper the discovery of productive behavioural sequences.

Genetic evolution of abilities linked to faithful sequence representation and mental flexibility started the evolutionary transition towards human thinking. These changes would have been domain-general (not modular) and initially "almost behaviorally silent." Its full potential would have been realised through cumulative cultural evolution. Gene-culture coevolution would have been involved in changes related to inborn specific mental skills, for example those used to support language. A stable environment most likely created the ideal conditions for cumulative cultural evolution. Stable environments would have provided the opportunity for long sequences of behaviour to develop, which could be scaffolded into further complex behaviours. In modelling language, a stable environment would have provided a favourable entry pattern for the evolution of faithful sequence representation or sequential abilities. A further essential aspect in the evolution of human cumulative culture is longer and protected childhoods which supported active teaching and sharing of information. However, humans and animals share the inborn motivational and emotional systems that drive much of behaviour.

A crucial question is when the genetic grounding for mental flexibility and long sequence behaviour might have been in place. It is suggested that the last common ancestor of Neanderthals and *Homo sapiens* would have already evolved the brain capacity for cumulative culture. This is at odds with one of the major ideas that had previously dominated thinking on the evolution of modernity – that human cultural capacities developed as recently as 40 000 years ago in modern humans and spread through migration and contact. This idea is not as widely rejected as suggested in the book, but indeed the "rubicon expectations"[3] that assume that there would be clear archaeological markers for modern behavioural and cultural processes have proved to be less tenable. The computational theory put forward here agrees with models and theories that emphasise the importance of the demographic and social environments in the development of complex material culture.

This book was not intended to be a theory of everything, and the testable predictions could be fruitfully investigated further. This engaging account explicitly shies away from correlating behaviour with brain functioning. The fundamental general domain capabilities on which human culture rely remain a black box and there is thus ample scope for integrating biological brain evolution research to a larger degree. There is also much potential to link the hypotheses on flexible sequential thinking more extensively to archaeological material culture. The book focuses

mostly on impressive recent cultural expressions, such as playing chess, musical notation and monumental architecture. However, this represents only the most recent manifestation of cumulative cultural evolution, leaving fertile ground for further exploration of the archaeological cultural reservoir.

References

1. Beyin A, Wright DK, Wilkins J, Olszewski DI. Handbook of Pleistocene Archaeology of Africa: Hominin behavior, geography, and chronology. London: Springer; 2023. https://doi.org/10.1007/978-3-031-20290-2

2. McBrearty S, Brooks AS. The revolution that wasn't: A new interpretation of the origin of modern human behavior. J Hum Evol. 2000;39(5):453–563. https://doi.org/10.1006/jhev.2000.0435

3. Meneganzin A, Currie A. Behavioural modernity, investigative disintegration & Rubicon expectation. Synthese. 2022;200(1):1–28. https://doi.org/10.1007/s11229-022-03491-7

AUTHORS:
Marlo Möller[1] iD
Eileen Hoal[1] iD
Brenna M. Henn[2,3] iD

AFFILIATIONS:
[1]DSI-NRF Centre of Excellence for Biomedical Tuberculosis Research, South African Medical Research Council Centre for Tuberculosis Research, Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa

[2]Department of Anthropology, University of California, Davis, California, USA

[3]Genome Center, University of California, Davis, California, USA

CORRESPONDENCE TO:
Marlo Möller

EMAIL:
marlom@sun.ac.za

The power of contemporary African DNA: Exploring models of human evolution and health in Africa

Significance:

It is generally accepted that humans evolved in Africa, but several opposing conceptual models representing our origins have been proposed. We shed light on the divergence of human populations on the African continent and challenge traditional models, suggesting a new framework – represented by a tangled vine with offshoots – in which stem populations separated but continually exchanged genetic material. This work would not have been possible without sequencing the most genetically diverse human genomes in the world; contemporary African DNA is not only key to understanding deep human history, but is also central to answering other health-related questions.

Genetic models of human evolution

The study of human evolution has been a subject of great interest as we seek to understand the origins and history of our species. One of the fundamental organising concepts in biology is the 'tree of life', which was the predominant model of human evolution for many years. This model of human population divergence from a single ancestral population in Africa was supported by the existing genetic data, but was difficult to align with the evidence of *Homo sapiens* fossils and archaeological sites across the continent.[1]

Fossils and archaeological records indicate the presence of anatomically modern humans across Africa between 300 000 and 100 000 years ago. Several key fossils, such as those found at Jebel Irhoud in Morocco, Herto in Ethiopia, and Klasies River in South Africa, demonstrate that anatomical features that originated in *Homo sapiens* were present throughout the continent during this period. Additionally, archaeological sites associated with *Homo sapiens*, particularly from the Middle Stone Age, are widely distributed across Africa, including the site with the oldest footprint identified.[2] Whether these populations represent the direct ancestors of contemporary humans or represent isolated local populations remains unanswered. Due to recent advances in population genetic tools, more complex modelling and inference using larger data sets have become possible.[3]

The discovery of Neanderthal admixture in Eurasian populations prompted research that suggested that an archaic hominin 'ghost' population contributed to African populations.[4] These studies indicated that such a model could better explain the observed genetic data, particularly in western, southern, and central African populations. However, these studies contrasted only a single-origin model with an archaic hominin admixture model, leaving out other plausible models.[1]

Challenging the genetic models: The study and its methodology

We aimed to discriminate between a broader set of demographic models by analysing the genomes of contemporary populations from diverse regions in Africa.[1] We considered four main models: single-population expansion, single-population expansion with regional persistence, archaic hominin admixture, and multi-regional evolution.[5] By including genetically and geographically diverse populations, we were able to infer demographic models that better explain the observed genetic diversity.

Critical to this study, we obtained whole-genome sequencing data for four African populations. The Nama from South Africa are an indigenous population that form part of a larger group of geographically close and culturally related individuals known collectively as the 'Khoe-San'. The Khoe-San are reported to have the most divergent lineages of any living population grouping and it is believed that they have largely remained isolated until ~2000 years ago. Participants from the Mende from Sierra Leone, the Gumuz from Ethiopia, and eastern African agriculturalists (Amhara and Oromo) were also included. British individuals were included as a representative of back-to-Africa gene flow and recent colonial admixture in South Africa. To account for gene flow from Neanderthals in Europeans, a high-coverage ancient Neanderthal genome was included in the analysis.

The results of our analyses confirmed the inadequacy of tree-like models and provided insights into more complex population structures. Rather than a simple single-origin model or archaic hominin admixture model, the best-fit models involved population reticulation, migration between early hominin populations, and divergence followed by merger events – a tangled vine with weakly separated offshoots instead of a tree of life (Figure 1).[6] These 'weakly structured stem' models indicated major stem lineages in southern, eastern, and western/central Africa during the late Middle Pleistocene, followed by subsequent fragmentation and combination of subpopulations. The inferred models suggested that the Middle to Late Pleistocene was a critical period of change, with merger events between divergent stems likely influenced by shifts in wet and dry conditions across the African continent. The findings also highlighted the ongoing contribution of one of the stems to western Africans during the Last Glacial Maximum, indicating gene flow in western and/or central Africa. In addition to distinguishing between models of early human history, by including many present-day individuals and groups, we can better predict and describe genetic variation for people living today, making these models more broadly useful.
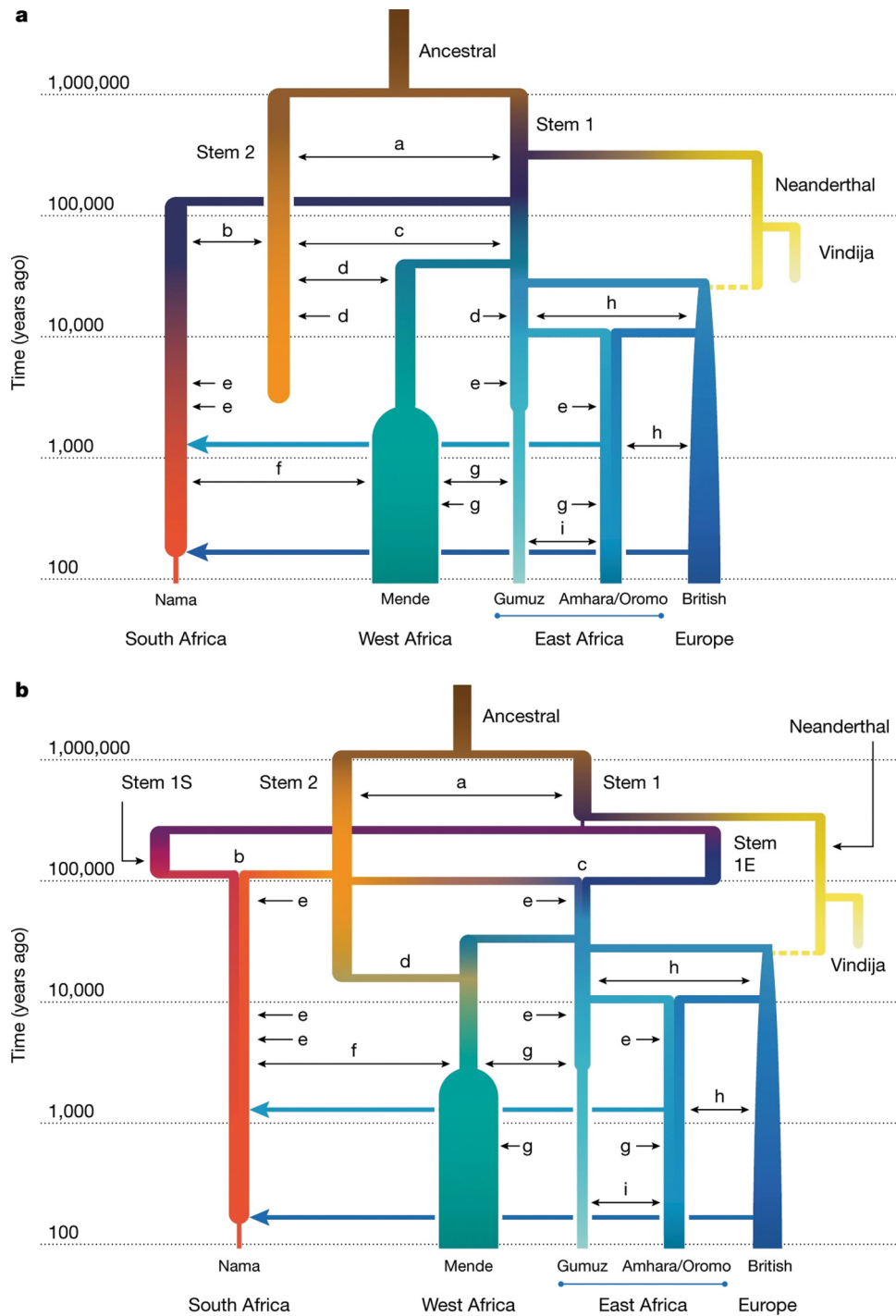
Source: Ragsdale et al.[1]; ©The Authors, under exclusive licence to Springer Nature Limited; reproduced with permission

Figure 1: New 'weakly structured stem' models of human evolution with (a) continuous migration and (b) multiple mergers: where stem populations separated, but continually exchanged genetic material.

It is often stated that "All models are wrong, but some are useful" (quote attributed to the statistician George Box). This quote highlights how constructing detailed models of human history is challenging, as model misspecification is inherent in such studies. It is also difficult to explore all plausible models, including the possibility of more complex models that involve additional stems or hybrid scenarios. The study's interpretations were also subject to uncertainties in estimating divergence times and migration rates, emphasising the need for further research and testing with ancient DNA samples (though rarely found in Africa) and additional populations.

Why do we need to sequence additional African genomes?

Africa is home to the greatest level of genetic diversity in the world. Historically, genomic research has been biased towards populations of European and Asian ancestry, leading to an underrepresentation of African populations in genomic studies.[7] Results from these investigations may not be applicable to individuals with more diverse ancestries. To prevent further disparities, and to improve knowledge about human history, health and disease, it is imperative to include diverse genetic data

sets. However, efforts have been made to address this imbalance and increase the representation of African genomes in genomic databases. Initiatives such as the African Genome Variation Project (AGVP) and the H3Africa (Human Heredity and Health in Africa) Consortium have been instrumental in collecting genomic data from diverse African populations.[8,9] A recent study which sequenced 180 individuals from 12 indigenous African populations identified millions of unreported variants, many with functional consequences.[10] Even so, there are definitely not enough African genomes sequenced (from modern or archaic samples) to represent the genetic diversity on the continent, but there is a drive by scientists to correct this underrepresentation.[11]

The establishment of local sequencing facilities on the African continent is another important step, as some African countries do not allow the export of DNA samples. Locally in South Africa we are fortunate to have access to several sequencing facilities, such as the South African Medical Research Council Genomics Platform, as well as the Centre for Epidemic Response and Innovation (CERI) Genomics Centre at Stellenbosch University, to name but two.[12] While significant progress has been made in recent years, it is important to note that the sequencing of African genomes is an ongoing process, and there is still much work to be done. Access to African genomes will advance our knowledge of human evolution and improve our analyses of genomic variation linked to complex health traits.

How can African genomes contribute to health?

African genomes are not only key to deciphering human history, but can advance our understanding of how to use population-specific variants in health and disease so as to implement precision and preventative medicine on the continent. Many African countries face a significant burden from infectious diseases, as well as non-communicable diseases such as diabetes, cardiovascular disorders, and certain types of cancer. African genomes can aid in understanding the genetic basis of these diseases, identifying high-risk individuals, and developing targeted prevention and treatment strategies. These genetic data can also shed light on host–pathogen interactions and contribute to the development of vaccines and therapeutics.

However, to make precision and preventative medicine a reality for all, investigations of population genomics in diverse populations from across the world are crucial. As an example, a study involving South Africans that included Nama DNA as a reference population, identified ancestry-specific expression quantitative loci (eQTLs) associated with tuberculosis and type-2 diabetes that could potentially guide the search for new therapeutic targets for these diseases in African populations.[13] It is troubling that three African-specific eGenes would have been missed if we did not consider the genetic ancestry of the study participants. This was also the case in a genome-wide association study of tuberculosis.[14]

It is unfortunate and concerning that the genetic data currently available are clearly not representative of the world's population. These data sets may also not be relevant to understudied groups and could even be unhelpful or misleading when determining genetic risk profiles for diseases in these settings, as a clear understanding of local population genomics is needed.[7] At the same time, developing economies, especially those in Africa, bear the brunt of socio-economic inequalities, poor living conditions and disease. Primary health care incorporating genomics could assist developing economies in not only diagnosing and treating diseases, but by helping to prevent diseases, resulting in cost savings.

Societal impact

As mentioned above, a number of consortia have made significant contributions to include more diverse populations that have historically been underrepresented, and contributed to the training and development of African genetic researchers. Clearly, global cooperation is key in these consortia, as our collaboration has also shown. We have worked together on a number of studies involving population genetics as well as genetic susceptibility to tuberculosis since 2009.[15-19] This longstanding collaboration has resulted in many co-authored manuscripts, grants and the training of PhD students. However, none of this work would have been possible without the study participants. We have made several trips to provide feedback on the research, and build on community participation and collaboration. As part of this collaboration, local clinics were better equipped and community members were trained to assist with recruitment (Figure 2). Such efforts can help to promote science education and a culture of science in communities.

Conclusion

Our study highlighted the critical role that DNA from contemporary Africans can play in understanding deep human history. By integrating the genetic data of a large number of contemporary individuals and groups, we can better anticipate and explain genetic variation in present-day individuals, allowing us to apply these models to health-related research concerns. There is clearly still much to be learnt by focusing on genetic data from contemporary individuals, especially when ancient DNA – crucial in revealing intriguing history and answering important concerns – may not exist for the relevant time periods, as is typically the case in Africa. Overall, our results contribute to a better understanding of human, and specifically African, population history and highlight the limitations of simplistic models, encouraging the re-evaluation of previous interpretations of genomic and fossil data.

Figure 2: Human genetics research is built on community participation and collaboration. As part of our research, local clinics were better equipped and community members were trained to assist with recruitment. Pictured here are study co-investigators, the study nurse and community members involved in recruitment.

Acknowledgements

Competing interests

We have no competing interests to declare.

References

1. Ragsdale AP, Weaver TD, Atkinson EG, Hoal EG, Möller M, Henn BM, et al. A weakly structured stem for human origins in Africa. Nature. 2023;617(7962):755–763. https://doi.org/10.1038/s41586-023-06055-y

2. Helm CW, Carr AS, Lockley MG, Cawthra HC, De Vynck JC, Dixon MG, et al. Dating the Pleistocene hominin ichnosites on South Africa's Cape south coast. Ichnos. 2023;30(1):49–68. https://doi.org/10.1080/10420940.2023.2204231

3. Scerri EML, Thomas MG, Manica A, Gunz P, Stock JT, Stringer C, et al. Did our species evolve in subdivided populations across Africa, and why does it matter? Trends Ecol Evol. 2018;33(8):582–594. https://doi.org/10.1016/j.tree.2018.05.005

4. Hammer MF, Woerner AE, Mendez FL, Watkins JC, Wall JD. Genetic evidence for archaic admixture in Africa. Proc Natl Acad Sci USA. 2011;108(37):15123–15128. https://doi.org/10.1073/pnas.1109300108

5. Henn BM, Steele TE, Weaver TD. Clarifying distinct models of modern human origins in Africa. Curr Opin Genet Dev. 2018;53:148–156. https://doi.org/10.1016/j.gde.2018.10.003

6. Scerri EML. One species, many roots? Nat Ecol Evol. 2023;7:975–976. https://doi.org/10.1038/s41559-023-02080-2

7. Martin AR, Teferra S, Möller M, Hoal EG, Daly MJ. The critical needs and challenges for genetic architecture studies in Africa. Curr Opin Genet Dev. 2018;53:113–120. https://doi.org/10.1016/j.gde.2018.08.005

8. Choudhury A, Aron S, Botigué LR, Sengupta D, Botha G, Bensellak T, et al. High-depth African genomes inform human migration and health. Nature. 2020;586(7831):741–748. https://doi.org/10.1038/s41586-020-2859-7

9. Gurdasani D, Carstensen T, Tekola-Ayele F, Pagani L, Tachmazidou I, Hatzikotoulas K, et al. The African Genome Variation Project shapes medical genetics in Africa. Nature. 2015;517(7534):327–332. https://doi.org/10.1038/nature13997

10. Fan S, Spence JP, Feng Y, Hansen MEB, Terhorst J, Beltrame MH, et al. Whole-genome sequencing reveals a complex African population demographic history and signatures of local adaptation. Cell. 2023;186(5):923–939.e14. https://doi.org/10.1016/j.cell.2023.01.042

11. Wonkam A. Sequence three million genomes across Africa. Nature. 2021;590(7845):209–211. https://doi.org/10.1038/d41586-021-00313-7

12. Glanzmann B, Jooste T, Ghoor S, Gordon R, Mia R, Mao J, et al. Human whole genome sequencing in South Africa. Sci Rep. 2021;11(1):606. https://doi.org/10.1038/s41598-020-79794-x

13. Swart Y, Uren C, Eckold C, Cliff JM, Malherbe ST, Ronacher K, et al. cis-eQTL mapping of TB-T2D comorbidity elucidates the involvement of African ancestry in TB susceptibility [preprint]. bioRxiv. 2022:2022.10.19.512814. https://doi.org/10.1101/2022.10.19.512814

14. Swart Y, Uren C, Van Helden PD, Hoal EG, Möller M. Local ancestry adjusted allelic association analysis robustly captures tuberculosis susceptibility loci. Front Genet. 2021;12:716558. https://doi.org/10.3389/fgene.2021.716558

15. Smith MH, Myrick JW, Oyageshio O, Uren C, Saayman J, Boolay S, et al. Epidemiological correlates of overweight and obesity in the Northern Cape Province, South Africa. PeerJ. 2023;11:e14723. https://doi.org/10.7717/peerj.14723

16. Reynolds AW, Grote MN, Myrick JW, Al-Hindi DR, Siford RL, Mastoras M, et al. Persistence of matrilocal postmarital residence across multiple generations in southern Africa. Hum Nat. 2023;34:295–323. https://doi.org/10.1007/s12110-023-09452-4

17. Schurz H, Kinnear CJ, Gignoux C, Wojcik G, Van Helden PD, Tromp G, et al. A sex-stratified genome-wide association study of tuberculosis using a multi-ethnic genotyping array. Front Genet. 2018;9:678. https://doi.org/10.3389/fgene.2018.00678

18. Van Eeden G, Uren C, Pless E, Mastoras M, Van der Spuy GD, Tromp G, et al. The recombination landscape of the Khoe-San likely represents the upper limits of recombination divergence in humans. Genome Biol. 2022;23(1):172. https://doi.org/10.1186/s13059-022-02744-5

19. Uren C, Kim M, Martin AR, Bobo D, Gignoux CR, Van Helden PD, et al. Fine-scale human population structure in southern Africa reflects ecogeographic boundaries. Genetics. 2016;204(1):303–314. https://doi.org/10.1534/genetics.116.187369

AUTHOR:
Martin Bekker[1]

AFFILIATION:
[1]School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg, South Africa

CORRESPONDENCE TO:
Martin Bekker

EMAIL:
martin.bekker@wits.ac.za

Large language models and academic writing: Five tiers of engagement

Significance:

Against a backdrop of the rapidly expanding use of large language models (LLMs) across diverse domains, this discussion breaks LLM usage into tiers of use, offering practical guidance to cautiously embrace the benefits of this significant new tool.

Introduction

2023 will be remembered as the year of large language models (LLMs), which, led by their brash poster child, ChatGPT, have changed the world forever. LLM-assisted writing will indelibly alter many writing tasks, offering speed and efficiency, and even automating-away many tasks.

However, academic, scientific, and intellectual integrity are at risk: not only due to mistakes that may creep in via the automation of writing, but also – and more importantly – owing to the loss of the ability to construct well-crafted arguments, ostensibly through the dulling of scholars' reasoning via the outsourcing of thinking that LLMs could engender. Moreover, ethical principles around intellectual process and ownership ought to be protected against the vague accountability of black-box algorithms (termed 'algorithmic opacity' by Eslami et al.[1]) with respect to published or submitted work.[2,3]

Academic journals, along with university and high-school curricula developers and assessment setters, need immediate yet thoughtful guidelines (rules and standards) for using LLMs and AI in the scientific process. In this Commentary, I propose a five-tier system that stipulates permissions and prohibitions around the use of LLMs in the academic writing process. I recapitulate what has changed, what – amid all the apparent changes – has stayed the same, and introduce the five tiers of LLM-assistance to academic writing, motivating the lines of distinction and suggesting appropriate uses.

What has changed?

In existence for over 40 years, language models are probabilistic models of a human language that can generate likelihoods of a series of words, based on text corpora on which they have been trained.[4] Over the last decade, the size of the training text corpora and the number of weights between concepts held within the models have increased, necessitating affixing 'large' to recent models, now known as LLMs.[5] ChatGPT was released in November 2022, combining the then-most advanced LLM with a chatbot-interface, simplifying the prompting (requesting) and serving (receiving responses) process. With their promise of speed and efficiency, ChatGPT and other LLMs have had an immediate impact on the academe; demonstrating the ability to automate the writing of reports, research and literature papers, exams, and computer code, among others, to various degrees of human ability, with each iteration showing improvement.

Contemporary LLMs represent a break with the past, based on (1) the speed and scale of information processing, (2) an unprecedented function of research process assistance (which includes research summary and data manufacture), and (3) the potential for the outsourcing of thought. The first of these three innovations often accompanies new technological tools. However, the scale and speed at which LLMs perform information processing tasks have now surpassed the human performance of certain tasks within the so-called information economy[6], suggesting a quantum leap in functioning and utility.

The second change is tied to the very nature of the way LLMs can process information. The ability of transformer models – the deep-learning architecture behind LLMs – to manipulate text (or language, or indeed anything that can be represented as a language) has made LLMs superlative at summarising texts, style transfers (ranging from translation to mere tone tweaks) and spelling and grammar corrections. Moreover, in addition to LLMs, several other AI-related tools that assist in the research process have recently been introduced (with many more to follow in their wake). With these, two particular functions spring to mind: first, research summarising tools (e.g. Elicit, Perplexity and Consensus) that can 'find' work (published but unknown to the scholar), 'understand' discourses, identify research gaps, and assist with literature reviews. The second are those that can manufacture artificial (or synthetic) data. Here, a scholar might give instructions regarding what a data set should contain, and, in the absence of this being available (as secondary data) or impossible to gather (for, say, ethical reasons), such data can be 'created' instantly.

The third change relates to the potential for the wholesale contracting out, or 'outsourcing' of thought to a (non-human) algorithm. While cheating is nothing new – academics have long been aware of student essays drafted by 'essay mills', computer code written by friends or copied from online repositories, or even a human double sitting for an exam on behalf of a less-prepared student – LLMs present the academe with a new level of concern. That is, from students to scientists, writers are now able to turn to LLMs to author academic work from conceptualisation through research and writing[7], putting in peril the principle of scientific advancement through human reasoning (or sound thinking as articulated through writing).

Perils notwithstanding, the immediate benefit of such a tool – one that can fix the register, grammar, and punctuation of a text in seconds and apparently at no cost to the user – is immediate and clear. This is especially the case for those writing in a second language, which case applies to the majority of academic scholars, who must publish in English. This 'levelling of the playing field' is to be welcomed by the academic community.[2,3,8] Academics often recite that good writing is indistinguishable from good thinking; the corollary of this is that clear, high-quality writing helps not only non-native speakers get the recognition they deserve, but benefits humanity if readers access knowledge in clear, correct, and accessible language.

The immediate drawback accompanying this new class of technology is that, sadly, many things that the academy has long battled to counter – dishonesty, cheating and plagiarism – have almost instantly become much harder to detect, owing to the increased volume and sophistication of the breaches. Moreover, LLMs are known to routinely produce credible untruths ('hallucinations', 'simulated authority'[8] or 'compelling misinformation'[9]) and omit attributions of their source or training data (plagiarism). Combined with the outsourcing of thought itself, such concerns render the use of LLMs for academic work potentially disingenuous at best, and at worst, in violation of the norms of scientific research. With all of this in mind, the need for practical guidance through the ethical minefield of LLMs is clear.

What has remained constant?

It is comforting that, despite the impressive and daunting changes wrought by the advent of LLMs, in reality, most scientific principles endure. First, the three values of beneficence, autonomy, and justice, all tied to non-maleficence and the avoidance of suffering[10] stand firm; in this sense, right is still right, and wrong remains wrong. Similarly, cheating and plagiarism remain anathema to the spirit of science, while openness, reproducibility, and the sharing (non-obscuring or gatekeeping) of data, keeping in mind all the caveats of harm, still stand as ideals.

Also holding firm are peer review as a well-established principle before publication, and the less-formal review-by-peers; that is, the shaping and improving of ideas (and writing) based on conversations, correspondence, arguments and counsel. Technical help – whether in the form of word processors, spell checkers, pocket calculators and software programs, or human editors and proofreaders – remains accepted and welcomed.

The five-tier system: An ethical guide to using LLMs

The present scramble to incorporate LLM use in academic work (or to find ways to ban it) implies that conversations about ethical guidelines and AI-use standards are timeous and valuable. Given the promise and peril of the new, but guided by the three values (justice, autonomy and beneficence), I propose a five-tier system to simplify thinking around permissions and prohibitions related to using LLMs for academic writing. While representing increasing 'levels' of LLM support that progress along a seeming continuum, the tiers in fact represent paradigmatically different types of mental undertakings.

Tier 1: Use ban

The first level comprises a complete ban on LLM-based support. This means that no LLM tools may be used in the preparation of the academic text. This tier therefore implies the highest level of human authorship and research authenticity.

Given its Draconian nature, coupled with the likelihood of inadvertent violations (e.g. the spelling and grammar checks employed by 'ordinary' word processors use a form of AI, and common word processors will soon be incorporating several other dimensions of LLMs), this tier is the most inviting to be flouted. Difficulty to enforce, lack of benefit and abundance of drawbacks (e.g. a step backwards in terms of present practice, given, for example, the ubiquity of automatic spelling and grammar checks) make such a tier likely only to be used in very specific circumstances, such as proctored university examinations or other formal testing conditions.

Tier 2: Proofing tool

Here, human-written text may be submitted to an LLM, accompanied by a prompt instructing the model to fix spelling, grammar, register, tone, and style (in the manner that products such as Grammarly might do). A proofing tool can be instructed to catch (and recommend remedies for) tone and style variations, identify problematic or misused words, and highlight direct translations.

The point here, of course, is that the work is presented at the end of the writing process – once the experimental and argumentative thinking is complete. Tier 2 does not outsource the thinking (or pain) that goes into the drafting process; rather, it takes fleshed-out thoughts and cosmetically enhances (or translates) them in much the same manner as would a private or in-house proofing team (or academic translation service).

Facilitating word-perfect (or near enough) text before submission, this level of usage can 'level the playing field' for non-native speakers and early-career researchers; that is, allowing those authors to display their arguments and findings to best advantage, smoothing over linguistic objections (tacit or implicit), and thus allowing work to be judged on merit alone. Of course, it may also enable 'lazy' writing (perhaps on the part of native English writers, knowing that sloppy writing will be fixed by an algorithm). Nonetheless, clarity and universality are significant in matters of standards, and thus, under this tier, any author might make use of this LLM proofing assistance.

Tier 3: Copyediting tool

Editing and proofing are distinguished by their sequential place in the preparation of texts, and also in the tasks that they perform: this distinction is echoed in the differences between Tiers 2 and 3. Given Tier 3 permissions, an author may ask for an LLM to alter text beyond correcting linguistic mistakes and aligning stylistic requirements. Shortening wordy text (e.g. reducing an abstract from 300 to 150 words), expanding for clarity, and rephrasing for precision are tasks every academic writer has laboured over and would likely welcome assistance with. Other areas of assistance at this tier would be checking citations for accuracy, style, and appropriateness – all things an LLM editorial function can do, and, in most cases, that paid in-house editors do to ensure the quality of publications (see Table 1 for prompt examples).

An editorial function can ensure a text is well organised, making changes to a text's structure by reordering paragraphs, or highlighting missing arguments. This tier will also include the assistance that Tier 2 permits,

Table 1: Examples of initial prompts per Tiers 2 and 3 (these are intentionally kept straightforward and unsophisticated)

Tier	Sample prompts
2	a) The following is section [x] from my paper, which I aim to submit to [name of journal]. Proofread the section and suggest corrections for spelling, grammar, tone, and style errors. Ensure the text is clear and free from any direct translations. Also, review the section for tone and style variations, and note any such pointwise. Finally, identify problematic or misused words, list each, and provide a recommendation for replacement.
	b) Make recommendations to improve the overall readability and coherence of the text. "[text here]"
3	a) Edit the below to shorten to 2000 words while preserving content, intention and clarity.
	b) Check the citations in the following document for accuracy, style, and appropriateness. List any necessary corrections pointwise.
	c) Make suggestions for the reorganising of paragraphs in the document below to improve the overall structure and flow of the argument. Briefly motivate each modification.

including spelling, grammar, tone and style corrections, flagging unusual words, nonsensical or confusing text, and guiding smooth transitions between paragraphs.

This tier would best be used during and after the writing process, and would likely be used iteratively, perhaps once a substantial section has been written.

Efficient editing (which can be a tedious and expensive journey) and the clarity of resultant texts are among the chief benefits of this tier. If correctly applied, critical thinking (and laboratory work or experimentation) would have preceded this step that in principle simply allows for a near-flawless write-up. Nonetheless, as with Tier 2, intellectual thoroughness and writing rigour may be compromised, ostensibly making this a compromise that must be accepted. This tier raises the point that copyediting is, and should be, regarded as an intellectual contribution (although copyeditors are seldom credited in scholarly journal articles in the way they may be in the publication of books), underscoring the observation that perhaps all forms of support ought to be acknowledged in academic writing.

Tier 4: Drafting consultant

Tier 4 speaks to a process of human–LLM 'co-creation' ('augmented writing'[7] or 'coauthoring'[2]). In addition to the support permitted under Tiers 2 and 3, this tier permits an iterative back-and-forth of ideas as one might do with a coauthor, up to the point of (and including) the LLM suggesting the omission of certain arguments, suggesting alternative 'interpretations', or requesting that one rerun experiments or check back to confirm previous findings.

In this tier, the author can interact with an LLM to plan a research write-up and shape and develop an argument, including requesting sample lines (e.g. instructing the LLM to 'compose an opening line'). Thus, this is not merely a tool offering a 'substantive edit', but a tool that can ensure one's evidence backs up one's argument (that is, where an LLM might even contribute to shaping that argument at earlier stages), and, where this is not the case, can provide suggestions on how to remedy such gaps.

Unlike the previous tiers, Tier 4 implies LLM engagement before the writing process commences, followed by iterative 'reporting back' sessions with the LLM as the writing advances. Potential benefits include the support such a routine would lend to early-career researchers: 'handholding' and sense-checking their writing process and arguments, while also offering suggestions and criticisms throughout the process to ensure quality. While the approach allowed by this tier is likely to significantly speed up the writing process, it does appear to tip the scales in terms of potential risks. Hallucinations and biases (subtle or not), both artefacts of LLMs, are more likely to manifest in co-created works. Also, LLMs may establish ways to obscure poor research behind apparently brilliant writing. It also follows that an overreliance on LLMs – here operating well beyond argumentative and stylistic considerations – would

constitute a loss of skills, the mastery of which would be expected in professional scientists or academics.

Tier 5: No limits

The fifth tier allows any LLM assistance at any stage. This tier includes brainstorming avenues of research, discussing and suggesting hypotheses and ideas, and even allowing the LLM to write text on the author's behalf. Such a no-holds-barred approach also includes interpreting results, summarising other scholarly work, and suggesting the implication of findings. In other words, Tier 5 permits the outsourcing of thought.

This tier is likely to be useful for instructing students on AI literacy and AI usage, and more generally demonstrating the dangers of stochastic models. This tier has limited value to scholarly, peer-reviewed publications, as the principle of authorship and the requisite of originality (at least as currently conceived) would likely be violated by, for example, a systematic review conducted solely by an AI agent.

Overview of tiers

With their growing levels of permissions, the tiers represent not only increasing degrees of LLM support, but also increasing levels of LLM dependence. Table 2 illustrates the tiers, and typical use-case points of entry, alongside the most obvious advantages and disadvantages of each.

As the tiers 'progress', so do the apparent speed and efficiency of tasks, as well as the dangers of LLM hallucination and manipulation – both significant and sensible concerns, given the opaque nature of LLM's stochastic processes and governance, not to mention the monopolising tendencies among Big Tech in general. Added to this is the arguably less immediately pernicious loss of academic integrity that would attend the outsourcing of thinking, and may come to represent a threat to humanity's overall ability to undertake quality scientific research, in the unlikely scenario in which LLM reliance is left unchecked and unregulated. Moreover, real and unconscionable human exploitation[11] and high environmental costs[12], both present in current LLM models, cannot be discounted or wished away.

Other AI tools available to the would-be academic author

While not part of the writing process, and thus adjacent to the proposed tiers, other academic tools draw on new advances in LLM or other recent machine-learning technologies. The ethical and intellectual considerations of their use overlap with those of LLMs, as do the efficiencies they offer researchers and writers.

The first is a class of **research summarising tools** already introduced in this paper. These can perform scans of literature on a topic (e.g. Perplexity), provide one-line summations of research papers (e.g. Elicit), give the apparent scientific consensus on a topic (e.g. Consensus),

Table 2: Summary of large language model (LLM) permission tiers, where they come into the writing process, and their most obvious benefits and risks

Tier	Effect / type of tool	Place in the writing process	Most obvious benefit	Most obvious risk
1	Ban	n/a	Ensures 100% human authorship and does not compromise academic integrity	Inevitably flouted except under exam conditions
2	Proofing	After	Increases efficiency, reduces cost	Might subtly alter meaning or obscure intentions
3	Editing	During	Produces well-organised, word-perfect writing	Language may be bland, may foster laziness
4	Co-creating	Start	Offers an alternative to human partnership–making interpretative and instructive suggestions, error checks	Authorship is opaque, high risk of introducing hallucinations and biases, inexplicability, loss of autonomy, loss of critical reasoning, outsourcing of thought
5	All	All	Enables high-speed academic writing, maximum support for inexperienced academics	As per Tier 4, but more extreme

17

or even create a literature review 'at the touch of a button'. Such tools can provide excellent assistance in exploring new fields of research or fields related to one's own work (see Jansen et al.[13] for a discussion of areas in which LLMs might support survey-based research), but are not substitutes for the process of critically reading to inform and order one's own intuitions and conclusions, gradually bringing new ideas into relation with one's own thought-scape. Moreover, such models may play a role in reifying conventional wisdoms, and in so doing, drown out marginal voices (the latter which may also be thought of as 'majority voices', considering that most people, including academics, are non-Western, non-white, and non-anglophone, despite the outsized influence of Western universities on the global scientific community).

The second non-writing LLM-based tool is '**automatic data analysis**' (e.g. Langchain), whereby data sets can be loaded up to an LLM for statistical analysis. In one dimension, the use of such technologies is equivalent to that of a pocket calculator: a logical time-saver, provided the user understands what the LLM is doing. For example, for at least the past three decades, scientists have routinely used multiple regressions, typically executed by statistical software or a coding routine in a software library. Use of statistical software (not least the writing up of results) requires a basic knowledge of statistics and data analysis. Subject to new developments, danger enters when the process of statistical analysis is not understood by the scientist, but regarded, crudely, as magic (i.e. it cannot be explained). In all, the use of LLM interfaces for statistical analysis will likely become commonplace, to the benefit of science in general, with the qualification that scientists will still be required to understand at least 'the bare bones' of statistical analysis.

The third application of AI tools, now in the form of machine learning, is the **creation of synthetic data** (e.g. Statice). Here, one can ask for data containing certain characteristics and of a given (potentially vast) size, or create data for teaching or illustrative purposes (including instances of data unavailability or cases where ethical or legal considerations proscribe the gathering of such data, such as sensitive healthcare data with patient identifiers). Such data sets will play an increasing role in teaching and testing, and so long as users keep in mind that the data in hand are fabricated, will be of great advantage in several domains (and are currently used in the training of self-driving cars, models for financial service security, and the pharmaceutical industries).

It is reasonable to expect that more AI-based tools will join the arsenal of the scientist. While caution (based, as ever, on the beneficence, autonomy, and fairness and the avoidance of maleficence principles) regarding the tools' creation, application and implications must be exercised, many of these tools will provide important and progressive support to scientific advancements[8], and should be embraced.

LLM tools and safety principles

Returning to LLM tools and how they can support the academic writing process, two inviolable principles merit further reflection: ownership and transparency. AI raises complex questions about these two principles; however, for the time being, the five-tier system, plus the appropriate use of supplementary material, may help to clarify questions around authorship, responsibility, and where we should stand on the place of the thinking process.

Ownership is the tenet that a submitted or published work and all of its contents remain the responsibility of a human author, and that the author is the only accountable party for mistakes or other consequences emanating from the work.[3] Any academic will be familiar with examples of authors choosing to omit their names from a publication (despite having contributed to the scholarship) when they do not fully agree with the contents, or feel unable to be held responsible for arguments contained in the work. Yet the broader point here is that the owner of ideas, the agent bearing the risk, and the agent deciding to be listed as author of a work is, so far, a human one. In our scientific pursuit, in 'advancing on Chaos and the Dark' it is the human – often individual – thinker who toils, who weighs, who risks.[14] Intellectual progress and the advancement of human thinking assumes the hitherto generally unspoken assumption that human thought ought not to be outsourced to non-human entities. Differently put, and evoking Chesterton[15], you

cannot make science without a soul. Even in the scenario in which a human and an LLM 'co-create' a work, the responsibility for the content still must rest somewhere, and, in the spirit of science, this would most obviously be the human author. Similarly, a recent US court ruling heard that 'the guiding human hand' is a 'bedrock requirement' to authorship.[16] The idea of blaming an LLM for mistakes appears disingenuously evasive and points to a concerning ambiguity over authorship. Certainly, under the first three tiers, as suggested here, authorship, and therefore ownership, rests with the human writer. The corollary of this is that the scientific value of papers, books and artefacts produced under Tiers 4 and 5 are de facto limited, and are likely to remain so: should this change, we will have to rethink authorship and academic credit anew.

While it is possible that humans will not forever be regarded as the sole culpable parties of their publications, recent work drawing on AI advances continues to confirm the principle of human authorship: despite the success of AlphaFold being based on Google DeepMind's algorithms, the celebrated *Nature* paper[17] lists only the human authors. Of course, acknowledgement of AI tools used in research (and acknowledgements in general) is categorically different from named authors of a work; that is, contribution does not imply attribution.

The second principle is that of **transparency**. Transparency – that is, showing one's work and thought process – lies at the heart of scientific accountability, reproducibility, peer accessibility, and public trust.[2] LLMs are by their nature opaque[1,5]; that does not mean they cannot be used, but rather that we must be open about *when* and *how* we use them. The alternative scenario is that readers have to guess whether LLMs have been used in the production of text[3] – here, mistrust is a solvent of credibility.

I suggest that authors who use LLM assistance include a way for their readers to see the prompts (and responses) they used in preparing a text. While the American Psychological Association referencing standard has issued a protocol for referencing ChatGPT, possibly a better way of referencing LLM contributions would be to provide a way for readers to access the entire series of human–LLM interactions, including every prompt and response. A hyperlink to an online repository (such as GitHub or Google documents) may risk being too fleeting a base, while a text file as supplementary material might suffice in 'showing one's working' in the same way as sharing one's routine of commands to statistical software, or sharing a codebase when programming.

Discussion: AI hype and despair

There is nothing emergent – in the complex adaptive systems theory sense – in the working of a pocket calculator. Calculators' answers are consistent, predictable, replicable, and regular. In this sense, LLMs are not like calculators: their massive size and black-box nature appear to have given them, at least by most accounts, emergent capabilities. The general public (and technical!) discourse has tended to label this not as 'emergent' but rather as 'generative' – which is, all told, succumbing to the language of AI hype. 2023 may well be marked as a high point of AI optimism, not least owing to the abilities and potential of LLMs. However, the ability to distinguish helpful innovation from unhelpful hyperbole (whether AI saviourism or AI catastrophising) is important in order to keep humankind's present problems and struggles in perspective, recognise our immediate moral duties, and rationally analyse the extent to which a new class of tools can help (or hinder) scientific progress and human betterment.

For scientists across every branch of knowledge, mental panic and ossification remain our nemeses. We gain most by seeing neither cataclysmic doom nor total redemption in technology, but instead recalibrating a new technology's value based on what it can change, and what it can't.

Among the proposed tiers, Tier 1 is techno-pessimistic. This tier assumes that technology per se represents a threat to knowledge production and human capabilities. This position is, to my mind, untenable for an academic journal already heavily reliant on LLMs (e.g. in the form of spell-checkers), not to mention calculator-like technology. In contrast, Tiers 2 and 3 may be regarded as technology-embracive. Optimistic about the

efficiencies LLMs bring to human knowledge and scientific advancement, these tiers advocate for adoption, remove the first-language barriers so often inhibiting the global dissemination of great ideas, and may even expedite the writing-up process. If permitted some rumination, one might suggest that Tiers 4 and 5 are leaning towards AI hype: both of these supposing that LLMs are either on the path to true cognitive supremacy and should thus be employed at all costs (the slave will soon become the benign master), or, alternatively, taking up the despairing position that LLMs will soon be so ubiquitous that any resistance to their use is bound to fail. One might argue that King[2] and Jansen et al.[13], on whose insights this note draws, lean towards the possibility of a Tier 5 future, where AI will become a colleague and coauthor.

Conclusion

Five tiers of LLM support for academic writing have been introduced, each offering a different level of writing support, and each entering the writing (and thought) process at a different stage. With some intentionality, the principles of ownership (plus responsibility) and transparency (sharing of prompts) can, and ought to be maintained.

Competing interests

I have no competing interests to declare.

References

1. Eslami M, Vaccaro K, Lee MK, Elazari Bar On A, Gilbert E, Karahalios K. User attitudes towards algorithmic opacity and transparency in online reviewing platforms. In: Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems; 2019 May 4–9; Glasgow, Scotland. New York: Association for Computing Machinery; 2019. p. 1–14. https://doi.org/10.1145/3290605.3300724

2. King MR. A place for large language models in scientific publishing, apart from credited authorship. Cell Mol Bioeng. 2023;16:95–98. https://doi.org/10.1007/s12195-023-00765-z

3. Li H, Moon JT, Purkayastha S, Celi LA, Trivedi H, Gichoya JW. Ethics of large language models in medicine and medical research. Lancet Digit Health. 2023;5(6):e333–e335. https://doi.org/10.1016/S2589-7500(23)00083-3

4. Rosenfeld R. Two decades of statistical language modeling: Where do we go from here? Proc IEEE. 2000;88(8):1270–1278. https://doi.org/10.1109/5.880083

5. Bender EM, Gebru T, McMillan-Major A, Shmitchell S. On the dangers of stochastic parrots: Can language models be too big? In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency; 2021 March 3–10. New York: Association for Computing Machinery; 2021. p. 610–623. https://doi.org/10.1145/3442188.3445922

6. Bubeck S, Chandrasekaran V, Eldan R, Gehrke J, Horvitz E, Kamar E, et al. Sparks of artificial general intelligence: Early experiments with GPT-4 [preprint]. arXiv;arXiv:2303.12712; 2023. https://doi.org/10.48550/arXiv.2303.12712

7. Kulesz O. The impact of large language models on the publishing sectors: Books, academic journals, newspapers [master's thesis]. Sweden: Linnaeus University; 2023. Available from: https://urn.kb.se/resolve?urn=urn:nbn:se:lnu:diva-119366

8. Rillig MC, Ågerstrand M, Bi M, Gould KA, Sauerland U. Risks and benefits of large language models for the environment. Environ Sci Technol. 2023;57(9):3464–3466. https://doi.org/10.1021/acs.est.3c01106

9. Spitale G, Biller-Andorno N, Germani F. AI model GPT-3 (dis) informs us better than humans [preprint]. arXiv; arXiv:2301.11924; 2023. https://doi.org/10.1126/sciadv.adh1850

10. US National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. The Belmont report: Ethical principles and guidelines for the protection of human subjects of research. Washington DC: US Department of Health and Human Services; 1979. Available from: http://www.hhs.gov/ohrp/regulations-and-policy/belmont-report

11. Perrigo B. Exclusive: OpenAI used Kenyan workers on less than $2 per hour to make ChatGPT less toxic. Time. 18 January 2023. Available from: https://time.com/6247678/openai-chatgpt-kenya-workers/

12. Strubell E, Ganesh A, McCallum A. Energy and policy considerations for deep learning in NLP. In: Korhonen A, Traum D, Màrquez L, editors. Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics; 2019 July 28 – August 02; Florence, Italy. Kerrville, TX: Association for Computational Linguistics. p. 3645–3650. https://doi.org/10.18653/v1/P19-1355

13. Jansen BJ, Jung SG, Salminen J. Employing large language models in survey research. Nat Lang Process J. 2023;4:100020. https://doi.org/10.1016/j.nlp.2023.100020

14. Emerson RW. Self-Reliance (1841). In: Porte J, editor. Essays and lectures. New York: Library of America; 1983. p. 261.

15. Chesterton GK. Orthodoxy. San Francisco, CA: Ignatius Press/John Lane Company; 1908.

16. Thaler Sv, Perlmutter S, Register of Copyrights and Director of the U.S. Copyright Office, et al., Civil Action No. 22-1564 (BAH) (18.08.2023). US District Court for the District of Columbia; 2023.

17. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021;596(7873):583–589. https://doi.org/10.1038/s41586-021-03819-2

AUTHOR:
Charles H. Simpson[1]

AFFILIATION:
[1]Institute of Environmental Design and Engineering, University College London, London, UK

CORRESPONDENCE TO:
Charles Simpson

EMAIL:
charles.simpson@ucl.ac.uk

Comment on Havenga et al. (2022): Standard heat stress indices may not be appropriate for assessing marathons

Significance:

An article in the July/August 2022 issue (Havenga et al., S Afr J Sci. 2022;118(7/8), Art. #13118) argued that changing the date of the Comrades Marathon from May to August would result in increased heat stress for participants. Heat stress was estimated using the Universal Thermal Climate Index (UTCI), which is designed to represent a person walking, not running. In this Commentary, I argue that using the UTCI may lead to an underestimation of heat stress for the Comrades Marathon, and that the conclusion that August has worse heat stress than May depends on the assumptions in the estimation of heat stress.

Introduction

While Havenga et al.[1] are right to examine the thermal environment of the Comrades Marathon, the Universal Thermal Climate Index (UTCI) might not be an appropriate metric. When the thermal environment is simplified into a single index, choices about the relative importance of temperature, humidity, wind, and radiative temperature are codified. Choice of thermal index can reverse the conclusion of a study in some contexts[2], thus it is important to identify cases where choice of thermal index is a critical assumption.

The UTCI has some advantages in that it has a strong thermo-physiological basis, and that it accounts for radiation. However, the derivation of the UTCI contains assumptions about activity and preferred clothing that are not true for a distance running event, which may distort the results.[3] In this Commentary, I aim to identify the effect of these assumptions.

Havenga et al.[1] justified their use of UTCI with reference to other studies, but these other studies do not provide a strong justification for using UTCI. One reference related to the thermal comfort of spectators, rather than competitors.[4] Brocherie and Millet[5] were cited for the statement "the Universal Thermal Comfort Index (UTCI) is regarded to be a better measure to model sports heat stress," but this reference does not actually test this and only proposes that newer indices might improve on the deficiencies of wet bulb globe temperature (WBGT). Honjo et al.[6] used the UTCI alongside WBGT, and noted the limitation that UTCI does not allow variations in metabolism or clothing. Gasparetto and Nesseler[7] used UTCI alongside WBGT but did not note these limitations. None of these studies demonstrates that UTCI is uniquely appropriate for thermal evaluation of distance running, and other research has highlighted these limitations for the sports context.[3] The limitations of UTCI are acknowledged by its developers[8] but Havenga et al.[1] do not discuss how these limitations affect their results.

The UTCI operational procedure is based on a person walking, with a metabolic rate of 2.3 MET.[8] Running involves higher metabolic rates than this: studies of Comrades Marathon participants found metabolic rates ranging from 6.6 to 10.6 MET (23–37 mL O_2/kg/min).[9] The bodies of distance runners need to dissipate much more internal heat than assumed in the derivation of the UTCI. Therefore, UTCI may underestimate heat stress or may not correctly identify the conditions with the highest heat stress.

The UTCI clothing model is based on the assumption that clothing preference is determined by air temperature[10], with clothing insulation reaching a minimum around 35 °C. Air temperatures on historical and proposed race days range from 5 °C to 32 °C with a mean of 19 °C. Applying Equation 3 from Havenith et al.[10] (with an assumed minimum of 0.25 clo), clothing values for the distribution of race temperatures have a mean of 0.75 clo and a maximum of 1.32 clo. The clo unit is defined as the estimated amount of clothing for a person at rest indoors at 21 °C to maintain thermal equilibrium: trousers, long-sleeved shirt, long-sleeved sweater and a t-shirt are 1.0 clo, sweat pants and a sweat shirt would correspond to 0.74 clo, while walking shorts and a short-sleeved shirt would correspond to 0.36 clo.[11] The UTCI clothing model is based on surveys of people going about ordinary daily activities[10], but it is simply wrong to assume that this is also the amount of clothing that runners wear. While runners do vary their level of clothing, metabolic production of heat needs to be included in any prediction of clothing level. The expected effect of this is that clothing levels are lower than assumed by the UTCI clothing model, generally distorting the pattern of heat stress predicted by the UTCI, as higher temperatures will be partly compensated by lower clothing insulation.

In the following sections, I demonstrate how the metabolic heat assumption affects the heat stress calculation and estimate the effect on this study.

Method

To estimate the combined effect of these assumptions, some calculations were performed. The full computer code for the underlying thermo-physiological model of the UTCI is not public, so it is not possible to directly test the effect of these assumptions. Physiological equivalent temperature (PET) operates in a similar way to UTCI but contains less physiological detail; PET code is publicly available and allows activity and clothing assumptions to be changed directly.[12] PET and UTCI do not have the same reference conditions, and are intended to represent slightly different things (heat stress vs heat strain). This is intended only to be an example, and I am not arguing that PET is necessarily the best index to assess thermal conditions for marathons in general.

PET was calculated with two sets of assumptions: (1) metabolic rate of 2.3 MET (based on the UTCI assumptions[8]) and clothing of 0.4 clo, and (2) metabolic rate of 8.6 MET and clothing of 0.4 clo. The chosen metabolic rate of 8.6 MET is the middle of the range observed in Comrades Marathon runners by Byrne et al.[9]

PET was calculated using the 'pythermalcomfort package' (https://py thermalcomfort.readthedocs.io accessed 2023-07-07)[13], which uses the Walther and Qoestchel 2018 specification[12]. When calculating PET, wind speed at 10 m height was transformed to wind speed at a height of 1.1 m using the same logarithmic scaling specified for the UTCI, and wind speed at 10 m height was limited to a minimum of 0.5 m/s for both UTCI and PET.[8] Limits specified in Brode et al.[8] were applied to the UTCI calculation – a step which appears to not have been applied in the ERA-HEAT supplied UTCI, which appears to overestimate heat stress at low wind speeds.

Temperature and humidity were taken from ERA5[14], with radiant temperature from ERA5-HEAT[15]. Hourly PET was calculated at locations for the start, halfway point, and end of the race. Following Havenga et al., the calculation was performed for the last 10 days of May and August. Only data between 03:00 and 16:00 UTC were included, to match the time of the race.

The distribution (across years) of maximum PET and UTCI, and total hours of heat stress categories according to PET and UTCI, were compared during the last 10 days of May and August 1980–2019 to determine if heat stress would typically be higher on August dates or May dates. This calculation was repeated with different metabolic heat assumptions to demonstrate its importance.

Results and discussion

Firstly, I note that the UTCI and PET produce very similar results when calculated with similar assumptions. Figure 1 shows UTCI plotted against PET calculated with the low metabolism assumption. The coefficient of determination between these two quantities is 0.96, i.e. 96% of the variance in the UTCI is explained by the PET. The main difference between UTCI and PET seems to be in sensitivity to wind speed. I argue, therefore, that making the analogy of UTCI and PET is justified for the purposes of this calculation. However, there are individual times when there is a large amount of disagreement about the level of heat stress, as shown by Table 1.

Secondly, I note that, by definition, the PET always increases with the metabolic rate. Figure 2 shows the extent to which PET is decreased by the low metabolism assumption. Changes in magnitude ranged from −13.8 °C to −3.8 °C, with a mean of −7.6 °C, with the largest changes occurring at high values of PET.

Table 2 shows the number of heat stress hours according to UTCI, PET(1) and PET(2). Using UTCI, there are more days in August than in May on which the maximum UTCI indicates 'strong' or 'very strong' heat stress. There are no days when the UTCI indicates 'extreme' heat stress. Using PET(1), there are more days in August than in May on which PET indicates 'strong' heat stress, and no days when the PET indicates 'extreme' heat stress. Using PET(2), there are more days in May than in August on which the PET indicates 'strong' or 'extreme' heat stress. Therefore, UTCI and PET(1) indicate that May has lower heat stress, but PET(2) indicates that August has lower heat stress.

Repeating the calculation with wind speed fixed at 2 m/s, August has higher heat stress using UTCI and PET(2), as shown by Table 3. The heat stress predicted by UTCI and PET is highly sensitive to wind speed, especially at low wind speed, and the two models have different sensitivity to wind speed. ERA5 indicates that wind speed is higher in August, as shown by Figure 3: the lower heat stress in August compared to May indicated by PET(2) is largely the result of wind speed being higher in August. This is problematic as near-surface wind speeds in the actual race environment are likely to differ considerably from the wind speed at a height of 10 m and horizontal resolution of 31 km in ERA5 in ways not well represented by logarithmic scaling. Furthermore, at low wind speed, the effect of the runners' body movements will become a substantial source of air movement, which is not properly taken into account in either the UTCI or PET calculations.

Conclusion

In this Commentary, I have demonstrated how the assumptions of low metabolic heat production used in the UTCI distorts thermal assessment for athletic events. PET calculations indicate that the assumption of low metabolic heat production leads to underestimation of heat stress. Furthermore, PET calculations using a higher metabolic heat assumption can indicate the opposite conclusion to PET with a lower metabolic heat assumption, i.e. August dates for the Comrades Marathon have lower heat stress. However, there is a strong dependence on low wind speeds
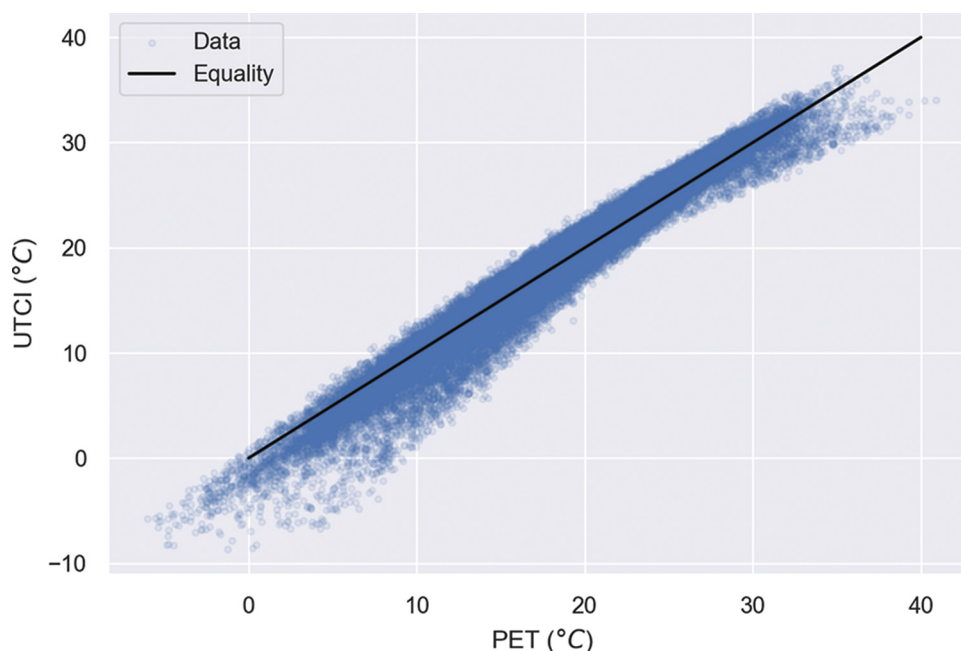
Figure 1: Universal Climate Thermal Index (UTCI) versus physiological equivalent temperature (PET). PET is calculated here with a metabolic rate matching the UTCI.

Table 1: Cross-tabulation of physiological equivalent temperature (PET) and Universal Thermal Climate Index (UTCI) categories calculated from hourly ERA5 data for the Comrades Marathon. Counts are hours for each pairing of categories.

	UTCI heat stress categories				
	No	**Moderate**	**Strong**	**Very strong**	**Extreme**
PET heat stress categories					
No	11 238	0	0	0	0
Slight	8516	0	0	0	0
Moderate	5137	2415	0	0	0
Strong	399	2889	300	0	0
Extreme	35	603	148	0	0
PET heat stress categories (low metabolism)					
No	21 837	0	0	0	0
Slight	3476	4015	0	0	0
Moderate	12	1844	385	0	0
Strong	0	48	63	0	0
Extreme	0	0	0	0	0
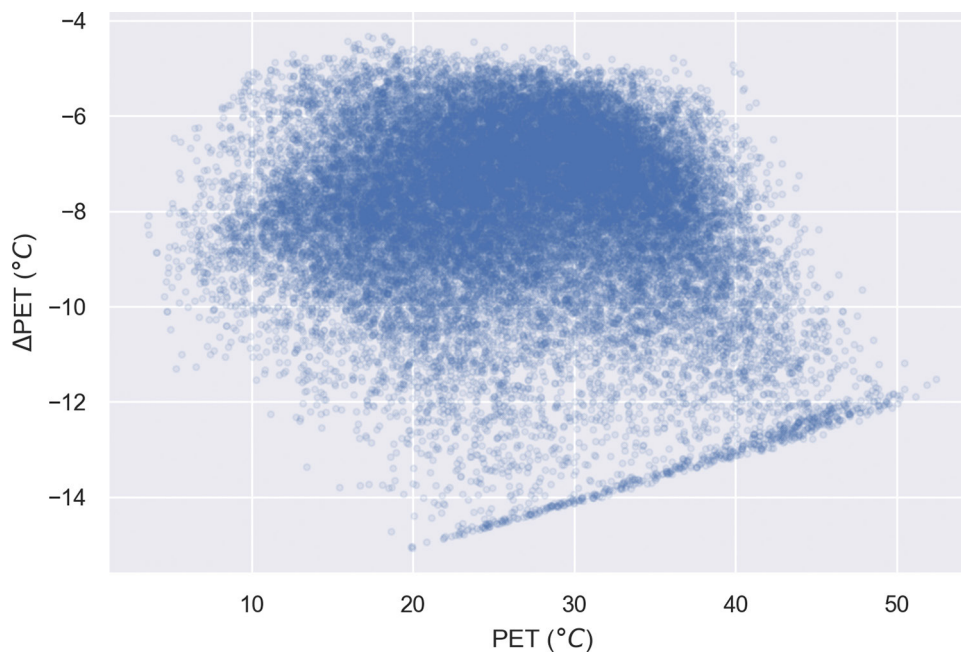
Figure 2: Difference in physiological equivalent temperature (PET) calculated with different metabolic assumptions. The *x*-axis shows the PET calculated with high metabolic rate (8.6 MET), whereas the *y*-axis shows the difference in the PET when calculated with a lower metabolic rate (2.3 MET).

in both PET and UTCI, and ERA5 wind speed at 10 m might not well represent the real race environment.

Grundstein and Vanos[3] argued that none of WBGT, UTCI, or PET are ideal for monitoring heat strain in athletes (although they refer to a PET implementation that did not allow for changes to metabolic rate or clothing). There may be demand for an equivalent of the UTCI with modified clothing and metabolism in the future, which would be useful for sport and occupational contexts. The ability to modify clothing and metabolic rate assumptions is vital in this context, and would point towards using an implementation of PET which allows these modifications, or another model of heat balance.

Table 2: Counts of hours in different heat stress categories defined by Universal Thermal Climate Index (UTCI) and physiological equivalent temperature (PET)

	May (hours)	August (hours)	Difference (hours)
UTCI			
No	12 456	12 869	−413
Moderate	3280	2627	653
Strong	104	344	−240
Very strong	0	0	0
Extreme	0	0	0
PET(1) (low metabolic heat)			
No	10651	11186	−535
Slight	3977	3514	463
Moderate	1160	1081	79
Strong	52	59	−7
Extreme	0	0	0
PET(2) (high metabolic heat)			
No	5501	5737	−236
Slight	3974	4542	−568
Moderate	3872	3680	192
Strong	1979	1609	370
Extreme	514	272	242

Table 3: Counts of hours in difference heat stress categories defined by Universal Thermal Climate Index (UTCI) and physiological equivalent temperature (PET), with windspeed fixed at 2 m/s

	May (hours)	August (hours)	Difference (hours)
UTCI with fixed wind speed			
No	11 777	11 393	384
Moderate	3937	3946	−9
Strong	126	501	−375
Very strong	0	0	0
Extreme	0	0	0
PET(2) with fixed wind speed			
No	4738	4442	296
Moderate	4719	4459	260
Slight	3949	3966	−17
Strong	2417	2808	−391
Extreme	17	165	−148
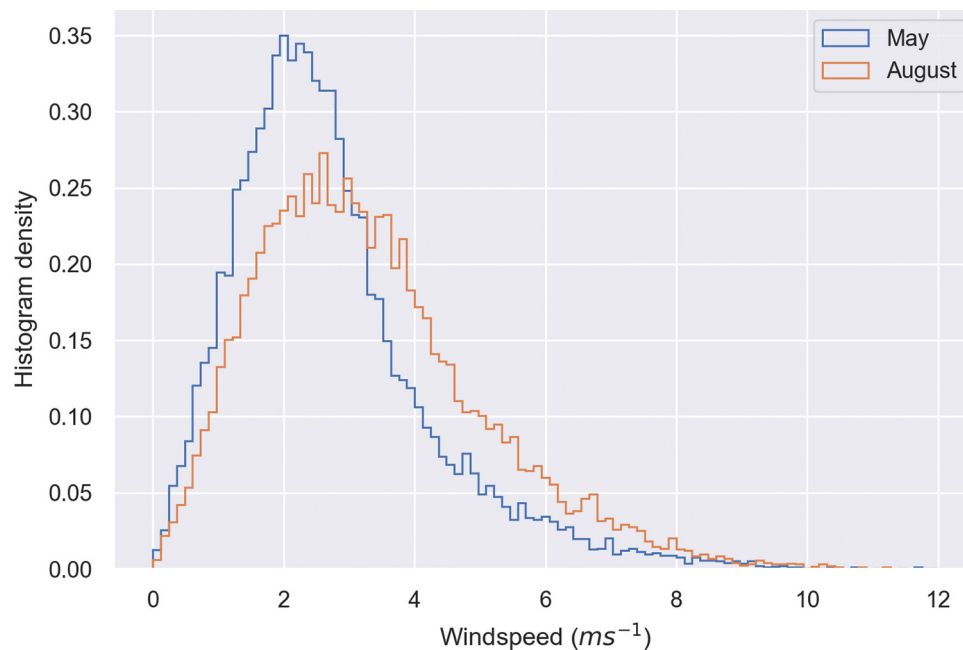
Figure 3: Histogram of wind speed in August and May at race locations. Windspeed is typically higher in August than in May.

Acknowledgements

Competing interests

I have no competing interests to declare.

Data availability

All code and data supporting the results of this study are archived and freely available for download from https://doi.org/10.5281/zenodo.8348335. ERA5 is freely available from the Copernicus Climate Service https://cds.climate.copernicus.eu/.

References

1. Havenga H, Coetzee B, Burger RP, Piketh SJ. Increased risk of heat stress conditions during the 2022 Comrades Marathon. S Afr J Sci. 2022;118(7/8), Art. #13118. https://doi.org/10.17159/sajs.2022/13118

2. Simpson CH, Brousse O, Ebi KL, Heaviside C. Commonly used indices disagree about the effect of moisture on heat stress. npj Clim Atmos Sci. 2023;6, Art. #78. https://doi.org/10.1038/s41612-023-00408-0

3. Grundstein A, Vanos J. There is no 'Swiss Army Knife' of thermal indices: The importance of considering 'why?' and 'for whom?' when modelling heat stress in sport. Br J Sports Med. 2021;55(15):822–824. https://doi.org/10.1136/bjsports-2020-102920

4. Vanos JK, Kosaka E, Iida A, Yokohari M, Middel A, Scott-Fleming I, et al. Planning for spectator thermal comfort and health in the face of extreme heat: The Tokyo 2020 Olympic marathons. Sci Total Environ. 2019;657:904–917. https://doi.org/10.1016/j.scitotenv.2018.11.447

5. Brocherie F, Millet GP. Is the wet-bulb globe temperature (WBGT) index relevant for exercise in the heat? Sports Med. 2015;45(11):1619–1621. https://doi.org/10.1007/s40279-015-0386-8

6. Honjo T, Seo Y, Yamasaki Y, Tsunematsu N, Yokoyama H, Yamato H, et al. Thermal comfort along the marathon course of the 2020 Tokyo Olympics. Int J Biometeorol. 2018;62(8):1407–1419. https://doi.org/10.1007/s00484-018-1539-x

7. Gasparetto T, Nesseler C. Diverse effects of thermal conditions on performance of marathon runners. Front Psychol. 2020;11, Art. #1438. https://doi.org/10.3389/fpsyg.2020.01438

8. Bröde P, Fiala D, Błażejczyk K, Holmér I, Jendritzky G, Kampmann B, et al. Deriving the operational procedure for the Universal Thermal Climate Index (UTCI). Int J Biometeorol. 2012;56(3):481–494. https://doi.org/10.1007/s00484-011-0454-1

9. Byrne C, Cosnefroy A, Eston R, Lee JKW, Noakes T. Continuous thermoregulatory responses to a mass-participation 89-km ultramarathon road race. Int J Sports Physiol Perform. 2022;17(11):1574–1582. https://doi.org/10.1123/ijspp.2022-0043

10. Havenith G, Fiala D, Błazejczyk K, Richards M, Bröde P, Holmér I, et al. The UTCI-clothing model. Int J Biometeorol. 2012;56(3):461–470. https://doi.org/10.1007/s00484-011-0451-4

11. American Society of Heating R and ACE Inc. 9.4.3.1 Thermal insulation. 2017 ASHRAE® Handbook – Fundamentals (SI Edition). Peachtree Corners, GA: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc. (ASHRAE); 2017.

12. Walther E, Goestchel Q. The P.E.T. comfort index: Questioning the model. Build Environ. 2018;137:1–10. https://doi.org/10.1016/j.buildenv.2018.03.054

13. Tartarini F, Schiavon S. pythermalcomfort: A Python package for thermal comfort research. SoftwareX. 2020;12, Art. #100578. https://doi.org/10.1016/j.softx.2020.100578

14. Hersbach H, Bell B, Berrisford P, Hirahara S, Horányi A, Muñoz-Sabater J, et al. The ERA5 global reanalysis. QJR Meteorol Soc. 2020;146(730):1999–2049. https://doi.org/10.1002/qj.3803

15. Di Napoli C, Barnard C, Prudhomme C, Cloke HL, Pappenberger F. ERA5-HEAT: A global gridded historical dataset of human thermal comfort indices from climate reanalysis. Geosci Data J. 2021;8(1):2–10. https://doi.org/10.1002/gdj3.102

AUTHORS:
Henno Havenga[1] [iD]
Ben Coetzee[2] [iD]
Roelof P. Burger[1] [iD]
Stuart J. Piketh[1] [iD]

AFFILIATIONS:
[1]Unit for Environmental Sciences and Management, North-West University, Potchefstroom, South Africa
[2]Physical Activity, Sport and Recreation Research Focus Area, Faculty of Health Sciences, North-West University, Potchefstroom, South Africa

CORRESPONDENCE TO:
Henno Havenga

EMAIL:
henno.havenga@nwu.ac.za

Response to Simpson (2024): Standard heat stress indices may not be appropriate for assessing marathons

Significance:

We value comments on our research paper in a Commentary in this issue (Simpson, S Afr J Sci. 2024;120(1/2), Art. #16445). Acknowledging the Universal Thermal Comfort Index (UTCI)'s limitations in capturing individual physiological responses remains important; however, we argue for its appropriateness based on recent thermophysiology and heat exchange advancements during its development and broader alignment with standardised indexing efforts. Our original research paper set out with these considerations in mind, and our conclusions remain valid. We further argue for refinement of the UTCI for specific activities instead of using the PET. Finally, future efforts should focus on monitoring data in real-world scenarios to validate and improve thermal indices.

We thank Simpson[1] for his insightful commentary on our research paper[2], which critically examines the appropriateness of a *standard* heat stress index for a specific event. In our research, we chose the Universal Thermal Comfort Index (UTCI) to evaluate heat stress during a major ultra-marathon event. We acknowledge that while UTCI provides a broad overview of heat responses in the average person, and may not capture the specific physiological reactions and microclimatic conditions experienced by each athlete, this limitation is inherent to all heat-related stress indices.[3] We argue that the UTCI is the appropriate index to use, as it is based *on the most recent scientific progress in both thermophysiology and heat exchange theory.*[4,5] The validation of the physiological equivalent temperature (PET) vs UTCI (or other indices) was beyond the scope of our paper, because it has been done numerous times.[6,7] Our conclusions, as we set out in our objectives and data analysis, remain valid and are supported by Simpson's own calculation of the UTCI, despite his criticism of the index – which will hold true in many other comparisons as well.[3,8]

However, to verify our decision to use the UTCI, we will briefly examine the history of this index within the greater bioclimatic community. The UTCI was developed to consolidate the wide range of thermal comfort indices – by 2015, more than 162 indices were documented in the scientific literature.[9] Subsequently, our selection aligns with the mission of the International Society of Biometeorology Commission 6 and the European COST Action 730, which aimed to standardise such indices.[3,10,11] Our decision to use the UTCI for assessing marathon running was also informed by other peer-reviewed studies.[12-14] The choice of an indicator that has been used in previous peer-reviewed studies also allows for comparisons between studies. The UTCI has also been applied to a range of sporting events[15], and physical activities[16]. Despite the limitations of the UTCI, the vision of the larger research community and published peer-reviewed articles directed our use of the index. According to us, there is no other peer-reviewed index that was specifically developed for a ultra-marathon (during which runners cover 89 km, over complex topography, with water points every 3 km, over various landscapes) which accounts for individual heat adaptations, training history, injury history, sweat rate, evaporative cooling, and race day hydration strategies.

Simpson rightly points out that the fixed metabolic rate (MET) that is prescribed by the UTCI might have underestimated the METs of the runners. However, the metabolic rate parameters are derived from Fiala's multi-node human physiology and thermal comfort model, which is widely accepted as the most advanced thermophysiological model to date.[17-19] The UTCI-Fiala contains substantially more physiological parameters than the simple PET-Munich Energy-balance Model for Individuals (MEMI). A comparison between the MEMI model (and other basic metabolism models) and the UTCI-Fiala model[5] indicated that the UTCI-Fiala model was more accurate in representing the physiological responses in a variety of environmental conditions. A possible reason for this finding is the fact that the UTCI-Fiala model was built on studies which included athletes.[5,19] Just because the PET allows the user to change the MET, does not mean that the underlying model is fundamentally correct.

A similar argument can be used for the clothing model query. The UTCI clothing model dynamically[20] adjusts for increases and decreases in temperature, whereas the simple PET model assumes uniform clothing[21]. Due to the Comrades Marathon's early morning start, and especially when the event starts in Pietermaritzburg, runners typically dress in warm attire and remove clothing as the event continues throughout the day. However, we agree that the clothing can be tailored for specific activities, which could in some cases alter the conclusions drawn. Nevertheless, the PET will lead to less accurate findings than a model in which the UTCI is modified. Therefore, we suggest that future research must rather focus on refining the UTCI to also consider individual changes in MET and clothing.

Despite the above arguments, we reiterate that the goal of our original research paper was not to evaluate all 162 indices, but to specifically apply the UTCI (as used in other peer-reviewed studies) to examine the possible prevalence of heat stress among runners and predict the possible influence of a change in climatological conditions on the incidence of exertional heat illness among athletes due to a change in the event date.

According to Grundstein and Vanos[22], who were referenced in our paper, researchers who investigate thermal indices can be classified into two primary groups: the bioclimate and the physiology communities. Our approach aligns with the bioclimate perspective, focusing on mostly generalised and predicted physiological variables combined with advanced meteorological data. The comments of Simpson[1] suggest a more balanced perspective by contending for a more tailored approach. However, we suggest that this modification should be based on the UTCI.

Alternatively, and drawing from the climate modelling community, we recommend the ensemble approach for future research. This would require that future efforts are based on a variety of heat stress indices that are tailored for specific scenarios. Such an approach can assist researchers to find possible variations in outcomes and indices. Similar to the Coupled Model Intercomparison Project Phase 6 (CMIP6) and Coordinated Regional Climate Downscaling Experiment (CORDEX) initiatives[23,24], the ensemble approach can help future researchers to strengthen the accuracy of conclusions when using heat stress indices.

We value Simpson's Commentary as it compels researchers to expand on existing research approaches, pointing towards more nuanced and detailed investigations into heat stress during physical activities. We also recognise that the choice of the applied heat stress index can significantly impact study outcomes, as was also the case with our study. Therefore, despite the value of the UTCI in providing researchers with a broader understanding of the thermal environment, the ideal would be to supplement the UTCI with local, micro-meteorological, and physiological data that are specific to the conditions and participants involved in a specific event. However, since the publication of the paper, we have engaged in more than six thermic-related investigations in collaboration with medical and sports scientists, especially focusing on high-level track and field, and road running events. These efforts aim to enhance our understanding of the correlation between environmental conditions and athletes' physiological responses to these conditions. We hope that these investigations will further contribute to the refinement of heat stress indices for use in real-life scenarios. Ultimately, the aim is to improve our understanding of the role that environmental factors play in sports performance and injury risk.

Competing interests

We have no competing interests to declare.

References

1. Simpson C. Comment on Havenga et al. (2022): Standard heat stress indices may not be appropriate for assessing marathons. S Afr J Sci. 2024;120(1/2), Art. #16445. https://doi.org/10.17159/sajs.2024/16445

2. Havenga H, Coetzee B, Burger RP, Piketh SJ. Increased risk of heat stress conditions during the 2022 Comrades Marathon. S Afr J Sci. 2022;118(7/8), Art. #13118. https://doi.org/10.17159/sajs.2022/13118

3. Jendritzky G, de Dear R, Havenith G. UTCI – why another thermal index? Int J Biometeorol. 2012;56:421–428. https://doi.org/10.1007/s00484-011-0513-7

4. Broede P, Blazejczyk K, Fiala D, Havenith G, Holmer I, Jendritzky G, et al. The Universal Thermal Climate Index UTCI compared to ergonomics standards for assessing the thermal environment. Ind Health. 2013;51(1):16–24. https://doi.org/10.2486/indhealth.2012-0098

5. Havenith G, Fiala D. Thermal indices and thermophysiological modeling for heat stress. Compr Physiol. 2011;6(1):255–302. https://doi.org/10.1002/cphy.c140051

6. Fröhlich D, Matzarakis A. A quantitative sensitivity analysis on the behaviour of common thermal indices under hot and windy conditions in Doha, Qatar. Theor Appl Climatol. 2016;124:179–187. https://doi.org/10.1007/s00704-015-1410-5

7. Pantavou K, Lykoudis S, Nikolopoulou M, Tsiros IX. Thermal sensation and climate: A comparison of UTCI and PET thresholds in different climates. Int J Biometeorol. 2018;62:1695–1708. https://doi.org/10.1007/s00484-018-1569-4

8. Binarti F, Koerniawan MD, Triyadi S, Utami SS, Matzarakis A. A review of outdoor thermal comfort indices and neutral ranges for hot-humid regions. Urban Clim. 2020;31, Art. #100531. https://doi.org/10.1016/j.uclim.2019.100531

9. De Freitas CR, Grigorieva EA. A comprehensive catalogue and classification of human thermal climate indices. Int J Biometeorol. 2015;59:109–120. https://doi.org/10.1007/s00484-014-0819-3

10. Jendritzky G, Höppe P. The UTCI and the ISB. Int J Biometeorol. 2017;61(Suppl 1):23–27. https://doi.org/10.1007/s00484-017-1390-5

11. Jendritzky G, Havenith G, Weihs P, Batchvarova E, DeDear R. The universal thermal climate index UTCI goal and state of COST action 730. Environmental ergonomics XII. Ljubljana: Biomed; 2007. p. 509–512.

12. Gasparetto T, Nesseler C. Diverse effects of thermal conditions on performance of marathon runners. Front Psychol. 2020;11, Art. #1438. https://doi.org/10.3389/fpsyg.2020.01438

13. Honjo T, Seo Y, Yamasaki Y, Tsunematsu N, Yokoyama H, Yamato H, et al. Thermal comfort along the marathon course of the 2020 Tokyo Olympics. Int J Biometeorol. 2018;62:1407–1419. https://doi.org/10.1007/s00484-018-1539-x

14. Hollander K, Klöwer M, Richardson A, Navarro L, Racinais S, Scheer V, et al. Apparent temperature and heat-related illnesses during international athletic championships: A prospective cohort study. Scand J Med Sci Sports. 2021;31(11):2092–2102. https://doi.org/10.1111/sms.14029

15. Konefał M, Chmura P, Zacharko M, Baranowski J, Andrzejewski M, Błażejczyk K, et al. The influence of thermal stress on the physical and technical activities of soccer players: Lessons from the 2018 FIFA World Cup in Russia. Int J Biometeorol. 2021;65:1291–1298. https://doi.org/10.1007/s00484-020-01964-3

16. Young E, Kastner P, Dogan T, Chokhachian A, Mokhtar S, Reinhart C. Modeling outdoor thermal comfort along cycling routes at varying levels of physical accuracy to predict bike ridership in Cambridge, MA. Build Environ. 2022;208, Art. #108577. https://doi.org/10.1016/j.buildenv.2021.108577

17. Blazejczyk K, Epstein Y, Jendritzky G, Staiger H, Tinz B. Comparison of UTCI to selected thermal indices. Int J Biometeorol. 2012;56:515–535. https://doi.org/10.1007/s00484-011-0453-2

18. Fiala D, Havenith G, Bröde P, Kampmann B, Jendritzky G. UTCI-Fiala multi-node model of human heat transfer and temperature regulation. Int J Biometeorol. 2012;56:429–441. https://doi.org/10.1007/s00484-011-0424-7

19. Psikuta A, Fiala D, Laschewski G, Jendritzky G, Richards M, Błażejczyk K, et al. Validation of the Fiala multi-node thermophysiological model for UTCI application. Int J Biometeorol. 2012;56:443–460. https://doi.org/10.1007/s00484-011-0450-5

20. Havenith G, Fiala D, Błażejczyk K, Richards M, Bröde P, Holmér I, et al. The UTCI-clothing model. Int J Biometeorol. 2012;56:461–470. https://doi.org/10.1007/s00484-011-0451-4

21. Matzarakis A, Mayer H, Iziomon MG. Applications of a universal thermal index: Physiological equivalent temperature. Int J Biometeorol. 1999;43:76–84. https://doi.org/10.1007/s004840050119

22. Grundstein A, Vanos J. There is no 'Swiss Army Knife' of thermal indices: The importance of considering 'why?' and 'for whom?' when modelling heat stress in sport. Br J Sports Med. 2021;55(15):822–824. https://doi.org/10.1136/bjsports-2020-102920

23. Deser C, Phillips A, Bourdette V, Teng H. Uncertainty in climate change projections: The role of internal variability. Clim Dyn. 2012;38:527–546. https://doi.org/10.1007/s00382-010-0977-x

24. Tél T, Bódai T, Drótos G, Haszpra T, Herein M, Kaszás B, et al. The theory of parallel climate realizations: A new framework of ensemble methods in a changing climate: An overview. J Stat Phys. 2020;179:1496–1530. https://doi.org/10.1007/s10955-019-02445-7

AUTHORS:
Filistéa Naudé[1]
Jan H. Kroeze[2]

AFFILIATIONS:
[1]Personal Librarian: College of Science, Engineering and Technology, University of South Africa, Johannesburg, South Africa
[2]School of Computing, College of Science, Engineering and Technology, University of South Africa, Johannesburg, South Africa

CORRESPONDENCE TO:
Jan Kroeze

EMAIL:
jan.kroeze@gmail.com

South African research contributions to *Lecture Notes in Computer Science*, 1973–2022

Lecture Notes in Computer Science (LNCS) is a globally recognised publication outlet for the field of Computer Science, including in South Africa. In this study, spanning from 1973 to 2022, we investigated the research participation of South African based authors in LNCS. The publication output and citation impact of these authors were compared to the global Computer Science and LNCS output. The authorship patterns and collaborative behaviour of South African LNCS papers were explored, and a keyword or topic analysis also conducted. Of the total of 518 662 LNCS papers published globally between 1973 and 2022, South African based researchers contributed 1150 papers (0.22%). The LNCS papers from South Africa exhibit a strong collaborative publication culture, with 1043 (91%) co-authored and 107 (9%) single-authored works. Local LNCS researchers prefer institutional collaboration (43%), followed by international (37%) and national collaboration (11%). Europe emerged as the most significant collaboration partner for LNCS researchers in South Africa. Of the 1150 papers, 836 (73%) had received citations, while 314 (27%) had not. On average, papers published by South African based authors received 6.05 citations, compared to the global LNCS average of 9.49 citations per paper. A keyword analysis revealed that the majority of papers by South African authors focus on artificial intelligence. The results indicate that, although LNCS serves as a reputable dissemination platform for Computer Science research output both globally and locally, South African authors should consider publishing more journal articles to build and improve their researcher profiles.

Significance:
- The study shows that LNCS is the most frequent publication outlet for Computer Science researchers, globally and in South Africa.
- The study offers insight into the publication output, authorship patterns, collaborative behaviour and citation impact of South African based Computer Science researchers.

Introduction

Lecture Notes in Computer Science (LNCS), a conference proceedings book and e-book series, is the most prevalent and prominent publication outlet for Computer Science globally[1] and in the Republic of South Africa[2,3]. The series is well established and highly respected. It was founded in 1973[4,5] and is published by Springer, the world's largest academic publisher[6]. Established in 1950, Springer is a reputable publisher of influential scholarly publications[7], which adds to the credibility of LNCS and makes it an attractive venue for researchers. In this article, the term 'Computer Science' is used as an umbrella term for the field encompassing all the sub-disciplines of Computer Science, Information Systems and Information and Communication Technology. This use of the term is in line with its use by Scopus and Springer. Some authors use the terms 'Computing' and 'Computer Sciences' as alternatives.

LNCS established itself as a primary dissemination channel for the publication of the latest developments in Computer Science research. The series includes the sub-series *Lecture Notes in Artificial Intelligence* and *Lecture Notes in Bioinformatics*, both of which were also included in the empirical work of this study. When LNCS commenced publication in 1973, it rapidly attracted attention because of its unprecedented publication turnaround times. In the late 1990s, Springer developed a systematic approach to publishing LNCS as a full-text electronic version in parallel to printed books. Original research results reported in pre-conference peer-reviewed proceedings and revised post-proceedings remain the core of LNCS.[4] The formats within its 2021 e-book collection underline the importance of conference proceedings in Computer Science, with these works comprising 78% of the collection, followed by monographs (8%), contributed volumes (4%), textbooks (3%) and other materials (7%).[8]

Over the years, the classification of LNCS as a journal, conference proceedings or book series has varied. The LNCS series offers comprehensive coverage and indexing in academic databases such as Elsevier's Scopus and Clarivate's Web of Science (WoS). Clarivate indexed LNCS from 1981 onwards in the Web of Science Core Collection.[9] Between 1999 and 2005, Clarivate listed LNCS in the Journal Citation Reports and assigned a journal impact factor. When Clarivate established the Conference Proceedings Citation Index (CPCI) database in September 2008, it moved LNCS to the new index.[9]

LNCS is classified in Scopus under the source type 'book series' and document type 'conference paper'. The source types covered in Scopus are either serial publications that have an ISSN (International Standard Serial Number) (e.g. journals, book series and conference series) or non-serial publications that have an ISBN (International Standard Book Number) (e.g. one-off book publications or one-off conferences). Document types are sub-units or components of source types and indicate the type of publication (e.g. journal article, conference paper or book).[10]

LNCS is readily available through various online platforms and academic databases, ensuring that researchers globally can access and cite LNCS. This availability enhances the discoverability, e-visibility, archiving, citability and research impact of the publication and contributes to the popularity of LNCS as a publication medium. However,

the Springer LNCS e-book series is a commercial, paywalled resource available for purchase as part of the Springer Nature Computer Science eBook Collection, which limits its accessibility compared to open-access publications. Authors are allowed to upload author-accepted versions of their papers on publicly available repositories, thus mitigating the paywall issue to some extent.[11] Springer also allows authors to publish substantially revised and extended versions of their LNCS papers in other outlets, provided that the original LNCS publication is acknowledged.[11]

For conferences to be considered for publication in the LNCS series, they must meet the following criteria: have an international programme committee, focus on a coherent set of topics that are of international relevance, and ensure a minimum of three reviewers per paper.[12] Furthermore, Springer proceedings should contain original research that has not been published or submitted elsewhere.[12]

Although LNCS is subsidy-bearing and accredited as a publication by the South African Department of Higher Education and Training (DHET)[13], its academic status and research significance as a discipline-specific publication venue for South African based Computer Science researchers are unclear. The problem statement is that it is uncertain to what extent Computer Science scholars in South Africa select LNCS as a publication channel for the dissemination of their research results and if it will continue to provide a suitable outlet for their future publications. Furthermore, it is also uncertain what the citation impact of LNCS is compared to other Computer Science journals in South Africa.

The main research question was: What are the publication output and authorship patterns of South African based authors who publish in LNCS? The sub-questions were:

- Which South African institutions contribute the most to the global publication output of LNCS?

- What are the most popular subjects and topics of LNCS papers from South Africa?

- What is the collaboration behaviour of South African based authors who publish in LNCS?

- What institutions, countries and regions feature as the most significant collaboration partners of South African based LNCS authors?

- How does the publication output of South African based LNCS authors compare to the global publication rate in LNCS and the global rate in the category of 'Computer Science' and sub-category 'General Computer Science' (GCS) in Scopus? (LNCS is classified in the GCS sub-category.)

- How does the citation impact of LNCS papers from South Africa compare to the global citation rate of LNCS and the global rate in the category of 'Computer Science' and sub-category 'General Computer Science' in Scopus?

- What is the reputation and status of LNCS, and how do they impact the series as a future publication outlet of choice for South African based authors?

Answering the questions above could reveal the academic status and research significance of LNCS as a publication outlet for Computer Science researchers in South Africa.

We therefore aimed to explore whether LNCS is a popular research outlet for South African authors in Computer Science. To ascertain if LNCS provides a quality platform for future South African publications, it is necessary to provide a holistic picture of the status and reputation of LNCS as a publication venue.

The scope of the article is to determine the involvement and relationship of South African authors with LNCS for the period 1973 to 2022. The number of South African papers is compared with those from other countries. To determine the impact of the South African papers, the citation metrics were also explored and compared to a DHET-accredited South African Computer Science journal.

The article makes a contribution by positioning LNCS in terms of other Computer Science journals. Determining the prominence and status of LNCS – globally and locally – gives an indication of the publication's standing as an outlet for Computer Science research. The research clarifies the contributions of South African based authors to the publication, in terms of both quantity and impact.

Literature review

Traditionally, Computer Science has had a conference-centric publication culture.[14-16] Conferences offer a fast publication cycle and immediate dissemination of cutting-edge findings in a rapidly evolving field of emerging technologies. In a historical investigation into the development of modern Computer Science publications and conference-based publication practices, Bouma-Sims[17] observes that there was an increase in conference publications in the 1980s, with publications in reputable conference proceedings being valued as much as (or even more than) articles in journals. This situation, however, complicates the application of traditional performance metrics in the research assessment of Computer Science researchers.[18,19] Zhang and Glänzel[20] confirm the importance of LNCS as a core Computer Science publication channel. Although proceedings papers receive fewer citations and the citation impact is lower than that of research articles[20], proceedings papers have remained the main publication channel in Computer Science over the past 10 years. The latest Scopus SciVal data show that from 2013 to 2022 there has been a total publication output of 4 462 139 with 41 819 764 citations and an average citation rate of 9.4 citations per paper. There were 2 473 808 (55%) conference papers compared to 1 658 367 (37%) journal articles in Computer Science and 8% were other publication types (see Supplement A[21]). In the General Computer Science subcategory, there was a total of 818 095 publications, with 159 501 citations and an average of 7.5 citations per paper. For GCS, there were 413 574 conference papers (50%) and 295 198 journal articles (36%), and 14% were other publication types (see Supplement A[21]).

However, Fortnow[22] argues that Computer Science is a mature discipline that needs to change its publication behaviour from a conference-based to a journal-based system. Halpern and Parkes[23] discuss the problems associated with the conference-based practice of publication in Computer Science. According to the latest 2022 Scopus SciVal Computer Science category data, the number of conference papers peaked in 2019 at 310 347 and declined to 281 906 in 2022, while the number of journal articles has been growing since 2013 and peaked in 2022 at 277 050, just slightly behind the conference papers. Looking at the sub-category of General Computer Science, the number of conference papers peaked in 2020 at 53 556 and declined to 36 152 in 2022, while the number of journal articles has been growing since 2013 and peaked in 2019 at 46 941 and declined to 37 823 in 2022 (see Supplement A[21]). These findings suggest that the conference-centric publication behaviour has already started to change into a journal-centric publication culture.

Scientometric and bibliometric analyses of Computer Science research output conducted in Argentina[24], Mexico[25], the Republic of Moldova[26], India[27], the Netherlands[28], China[29] and Malaysia[30] emphasise the importance of conference proceedings and LNCS as a core publication outlet for Computer Science research. Subject-specific reviews in the field of Geographical Information Systems[31], the Digital Economy[32], Living Labs and Human-Computer Interaction[33] conclude that LNCS is among the most popular publications for these subject domains.

Fiala and Tutoky[1] conducted a bibliometric assessment of 1.9 million Computer Science journals and conference papers indexed in the WoS and the CPCI for the time frame 1945 to 2014 – an extensive investigation which highlighted the reliance on conference publications in the field of Computer Science. The results indicated that 56% of papers were published as proceedings papers and 35% as journal articles – the latter received 75% of citations, compared to 11% for the former.[1] Journal articles received on average 13.4 citations per article compared to 1.2 for conference papers.[1] The researchers found that most Computer Science papers were published in LNCS and established that the average number of citations per paper for LNCS was 3.6 for the period under study (1945–2014).[1] The most productive subject areas were Artificial Intelligence (31.8%), Theory and Methods (30.3%) and Information

Systems (26.6%), while the United States of America (USA) (24.8%), China (13.7%) and the United Kingdom (UK) (5.7%) were the countries with the most publications.[1] The study showed that LNCS papers (11 259) comprised 0.6% of the total papers (1 922 625) published in Computer Science for the period 1945 to 2014.

In their study of Chinese publications in LNCS, He and Guan[29] analysed 5916 conference papers by Chinese authors between 1997 and 2005: more than half of the papers in Computer Science from China were published in LNCS. He and Guan's analysis also showed a trend of many publications with few citations and little impact. As their results showed, the number of papers by Chinese authors published in LNCS had increased in the time frame investigated, especially since 2004, but citations of the papers remained very low.[29] Authorship patterns indicated a preference for jointly authored papers (97%) over single-authored papers (3%). Also, the Chinese researchers preferred domestic collaboration or co-authoring with fellow Chinese researchers to international collaboration.[29]

South African research studies into the Computer Science research landscape are sparse[34-36], but two recent South African Computer Science bibliometric assessments notably emphasise the popularity of LNCS with South African scholars[2,3].

Parry[3] conducted a comprehensive scientometric investigation into Computing research in South Africa, using the Elsevier Scopus citation-enhanced database. The data set of 11 180 records included journal articles, books, book chapters and conference proceedings of researchers affiliated with South African universities for the period 2008–2017.[3] Parry's study revealed that Computing research in South Africa had increased by 172.61% over 10 years. Conference papers were the most popular publication format (61.40%), followed by journal articles (36.10%), book chapters (2.33%) and books (0.23%).[3] LNCS ranked third in the 25 most prominent publications for South African Computing research. Parry[3] calculated that the mean citation rate per publication for South African Computing research was 4.67 and that 39.19% of Computing publications had no citations. Conference papers accounted for 26.04% of the citations and journal articles for 73.18%, compared to books at 0.78%.[3]

A scientometric assessment of Computer Science in South Africa by Mouton et al.[2] for the period 2005–2020 concluded that LNCS is the predominant publication choice for local Computer Science researchers. They analysed 3441 papers in 472 publications retrieved from the SA (South African) Knowledgebase database of DHET subsidy-earning publications (see Supplement B and C[21]). The adapted data (Table 1)[2] underscore the popularity of conference proceedings as a publication outlet for this discipline, with 6 of the top 20 publication outlets being conference proceedings and 14 being journals.

In South Africa, the DHET funding that a university receives is partially based on the research output units produced by a university and is guided by the 2015 DHET Research Output Policy.[37] LNCS is subsidy-bearing and accredited as a publication by DHET.[13,38] At present, there are seven DHET-approved journal lists. The DHET Scopus journal list includes LNCS. The inclusion of the Scopus list as a DHET journal list was approved from 2016. Before 2016, LNCS papers were submitted to DHET as publications in conference proceedings.

The global publication behaviour in the field of Computer Science prioritises conferences over books or journals. However, this poses significant challenges for South African Computer Science researchers who are required to adhere to the journal-centric subsidy model set by DHET. The South African National Research Foundation (NRF) acknowledges the significance of conference papers as valuable research outputs in the field of Computing, although DHET places a higher emphasis on journal papers and considers them superior.[37] For each article published in a DHET-accredited journal, subsidy is almost guaranteed when claimed, compared to conference papers and book chapters where subsidy may be awarded if DHET is convinced of the standard of the peer-review process based on the details in the portfolio of evidence that should accompany the claim.[37] Conference papers are worth half of a journal article in terms of DHET subsidy. A research paper published in an accredited publication is subsidised as a single unit (1 research output

Table 1: Top 20 popular publication outlets for South African based researchers in Computer Science

Publication name (Publisher)	Document type	Paper count
Lecture Notes in Computer Science (Springer)	Proceedings	337
South African Computer Journal (SAICSIT)	Journal	201
IEEE Access (IEEE)	Journal	161
Communications in Computer and Information Science (Springer)	Proceedings	132
Advances in Intelligent Systems and Computing (Springer)	Proceedings	76
Scientometrics (Springer)	Journal	59
IFIP Advances in Information and Communication Technology (Springer)	Proceedings	54
Discrete Mathematics and Theoretical Computer Science (Maison de l'informatique et des mathematiques discretes) (Springer)	Journal	51
Bioinformatics (Oxford)	Journal	47
Computers and Security (Elsevier)	Journal	47
Mathematical and Computer Modelling (Elsevier)	Journal	39
Computers and Chemical Engineering (Elsevier)	Journal	34
Electronic Journal of Information Systems Evaluation (Academic Conferences International)	Journal	34
Lecture Notes in Artificial Intelligence (Springer)	Proceedings	34
Theoretical Computer Science (Elsevier)	Journal	33
Computers and Education (Elsevier)	Journal	30
Journal of Combinatorial Optimization (Springer)	Journal	30
Journal of Molecular Modeling (Springer)	Journal	29
Lecture Notes in Business Information Processing (Springer)	Proceedings	29
Structural and Multidisciplinary Optimization (Springer)	Journal	28

Source: Mouton et al.[2] with permission

unit), compared to papers published in approved conference proceedings that are allocated a maximum of 0.5 units.[38] DHET defines approved conference proceedings as "those which appear in approved conference lists or other approved indices"[38].

Conference papers that are not published in DHET-approved conference proceedings listed on the DHET-accredited list, must adhere to DHET prerequisites.[38] For instance, a minimum of 60% of contributions published in the conference proceedings should originate from multiple institutions. The primary objective of the conference must be to facilitate the widespread dissemination of original research and new advancements in the relevant field. All submitted papers must undergo rigorous peer review before being accepted for publication. Evidence of the peer review should be provided for subsidy claims. Additionally, DHET mandates that the conference should have an editorial board and/or organising committee, comprising a substantial majority of members from diverse institutions, demonstrating expertise in the respective subject area.[38] The preference for LNCS as a publication outlet among South African researchers may stem from the fact that LNCS is an approved DHET-listed publication[13],

facilitating smoother DHET subsidy claims and rendering a full research output unit.

In order to determine to what extent South African authors have selected LNCS in the past as a publication channel for disseminating their Computer Science research results, and the reasons therefor, this article reviews their publication output numbers in LNCS for the period 1973–2022, as well as the related citation impact, authorship patterns and collaboration behaviour.

Methodology

The article follows a bibliometric approach. Bibliometrics can be defined as "the use of statistical methods to analyse publications with the aim of measuring outputs of individuals, institutions, and countries, and identifying networks between them"[39]. Bibliometric data provide quantitative measures to assess the influence and research impact of academic publications and guide researchers with their publication outlet choices. This study is a quantitative analysis of the South African LNCS contributions for the period from 1973 to 2022. The selected time frame corresponds to the establishment of LNCS in 1973. The data collection occurred in June and July 2023, while the study encompasses the period leading up to 2022.

The Elsevier Scopus citation-enhanced database was selected as the bibliometric data-collection tool, due to the extensive retrospective coverage and availability of LNCS citation data. Scopus, as the largest curated citation database and a reliable high-quality bibliometric data source[40], has better coverage than the WoS when it comes to Computer Science related subjects[41]. Scopus SciVal data were also obtained for the period 2013–2022.

SciVal is an Elsevier research analytics product that uses the publications in the Scopus database as its data set. It provides a wide range of research metrics and has the ability to create reports, and compare and benchmark many different types of entities.[42]

The coverage of Computer Science in Scopus is estimated to be in the region of 60.59%.[42] Scopus uses the ASJC (All Science Journal Classification) subject scheme and classifies all computing-related literature into the broad category of Computer Science.[43] Scopus indexes approximately 2626 sources in the Computer Science subject area.[44] Table 2 lists the number of sources in the 13 sub-disciplines of Computer Science in Scopus.

LNCS is assigned to the subject category 'General Computer Science' and the publication is ranked 126th out of 233 by CiteScore metrics. CiteScore is based on the number of citations to documents (articles, reviews, conference papers, book chapters and data papers) in a journal over 4 years, divided by the number of the same document types indexed in Scopus and published in the same 4 years by that journal[42] (see Supplement D[21]). The 358 General Computer Science sources referred to in Table 2 include active and inactive sources, 233 of which are active (i.e. currently accepting and publishing new articles).

The Scopus database shows that a total of 8 102 462 documents were published globally in the field of Computer Science for the period 1973 to 2022 (Table 3). Of these, 6.40% (518 662) were LNCS papers. A total of 1 450 425 documents were published in the subcategory 'General Computer Science', of which 35.76% were LNCS papers. South African LNCS papers comprised 0.01% (1150) of the total Computer Science papers published globally and 0.08% of the papers published in the subcategory 'General Computer Science'.

A document search was conducted in Scopus using the LNCS Source-ID (25674) to retrieve the LNCS bibliographic data (ISSN:0302-9743 or

Table 2: Number of sources in Computer Science in Scopus

Scopus subject area: Computer Science – Total number of sources			2626
Scopus sub-categories for Computer Science	Number of sources	Scopus sub-categories for Computer Science	Number of sources
Artificial Intelligence	339	General Computer Science	358
Computational Theory and Mathematics	204	Hardware and Architecture	240
Computer Graphics and Computer-Aided Design	133	Human-Computer Interaction	170
Computer Networks and Communications	472	Information Systems	455
Computer Science (miscellaneous)	126	Signal Processing	158
Computer Science Applications	973	Software	561
Computer Vision and Pattern Recognition	121		

Table 3: Publication output in Computer Science in Scopus, 1973–2022

Scopus subject area: Computer Science – Total publication output			8 102 462
Scopus sub-categories for Computer Science	Publication output	Scopus sub-categories for Computer Science	Publication output
Artificial Intelligence (code=1702)	1 119 349	General Computer Science (code=1700)	1 450 425
Computational Theory and Mathematics (code=1703)	430 258	Hardware and Architecture (code=1708)	805 351
Computer Graphics & Computer-Aided Design (code=1704)	333 169	Human-Computer Interaction (code=1709)	439 050
Computer Networks and Communications (code=1708)	805 351	Information Systems (code=1710)	1 003 713
Computer Science (miscellaneous) (code=1701)	173 671	Signal Processing (code=1711)	739 517
Computer Science Applications (code=1706)	2 802 246	Software (code=1712)	1 644 553

Table 4: Top 50 countries contributing to *Lecture Notes in Computer Science*, 1973–2022

Country	Number of papers	Country	Number of papers	Country	Number of papers
USA	86 617	Belgium	7706	Hungary	2376
China	66 050	Portugal	7386	Romania	2113
Germany	56 062	Russian Federation	7258	Malaysia	2023
France	40 422	Israel	7246	Chile	1862
United Kingdom	39 859	Sweden	7235	Tunisia	1680
Italy	28 671	Taiwan	7203	Iran	1516
Japan	28 664	Greece	6523	Colombia	1441
Spain	23 922	Denmark	6100	Vietnam	1360
Canada	19 425	Singapore	5934	Thailand	1251
Australia	16 468	Hong Kong	5891	Bulgaria	1219
Netherlands	16 167	Czech Republic	5840	Luxembourg	1217
South Korea	15 949	Finland	5547	Slovakia	1169
Poland	12 212	Norway	4783	**South Africa**	**1150**
India	11 317	Mexico	4069	Argentina	1122
Austria	10 013	Ireland	3942	Slovenia	1120
Switzerland	9960	Turkey	2814	Saudi Arabia	1064
Brazil	9095	New Zealand	2811		

E-ISSN:1611-3349). The time frame of the search was limited to the period 1973–2022, with a total of 518 662 papers published globally in LNCS during that period.

To analyse and compare the productivity patterns, the results were further refined by limiting the study to papers only from South Africa. There were 1150 papers with a South African affiliation (i.e. 0.22% of the total of 518 662 LNCS papers published internationally). The Scopus data were exported into an MS Excel spreadsheet format for analysis. Using the 'Analyse search results' function in Scopus, the results were analysed according to author, affiliation and country, and then ranked by the top affiliations and countries. The 'Export refine' option in Scopus was used to export the keywords, countries, affiliations and number of papers to MS Excel. The total number of papers per annum for LNCS were exported to MS Excel and the total number of citations per annum was calculated. The average citations per paper per annum were calculated by dividing the citation count (number of citations) by scholarly output (number of papers).[42] The collaboration patterns of South African based authors were also analysed. All LNCS South African papers were assigned a collaboration type by Scopus, based on affiliation information[42]:

- International collaboration: global collaboration between authors from different countries

- National collaboration: collaboration between authors from different institutions within the same country

- Institutional collaboration: collaboration between authors from the same institution

- Single authorship: sole-authored paper with no collaborators

Results

Authors from a total of 159 countries made contributions to LNCS between 1973 and 2020. Table 4 presents the top 50 countries, ranked by their contributions. South Africa ranked 47th out of 159, with a publication count of 1150 research papers (see Supplement E[21]). It seems that this rank is quite stable when compared to more recent periods. In the decade 2003–2012, South Africa retained its 47th place,

but in the decade 2013–2022, it moved up slightly to 44th place (see Supplement F[21]).

Figures 1–6 show the global and South African LNCS output and citation impact (1973–2022). Global publication output in LNCS has grown exponentially since its inception, from 92 in 1973 to 20 372 papers in 2022, while South African publications increased from one paper in 1978 to 57 in 2022 (Figures 1 and 2). South African based authors did not publish in LNCS between 1973 and 1977, nor in 1979, 1981–1989 and 1992. A possible explanation for this finding is the country's isolation and the academic sanctions that were in place during the apartheid era. During the culmination of the era of apartheid in the 1980s, there was widespread support for an academic boycott. This boycott involved various tactics, such as journals and publishers refusing to accept or publish manuscripts originating from South Africa, exclusion of South Africa from participation in international scientific conferences[45], denial of entry visas to academics, international scholars declining to visit South Africa or to collaborate with South African researchers, as well as restricted access to information resources like books, journals, computer software and databases[46]. However, academic sanctions were gradually relaxed from 1990, leading to an increase in South African participation in the international scholarly arena. Figure 2 reflects this increase by showing the exponential growth of South African LNCS publications since 1994.

From 1993 onward, there were regular annual contributions with steady annual growth. Figures 1 to 4 compare the number of global and South African LNCS papers and citations. Figure 5 compares the percentage of South African citations with the percentage of South African papers, and Figure 6 the global and average citations per paper per year. The graphs provide a visual way to gain insight into South Africa's contributions to LNCS. Although all the data available up until 2022 were collected, it should be noted that citations take time to accumulate. Therefore, the data of the last 2 years cannot be used as a reliable indication of any current or future trends.

As shown in Figures 1 and 2, the total number of global LNCS papers grew exponentially until 2006 when it stabilised at between 20 000 and 25 000 p.a., while the South African LNCS papers seem to have gained
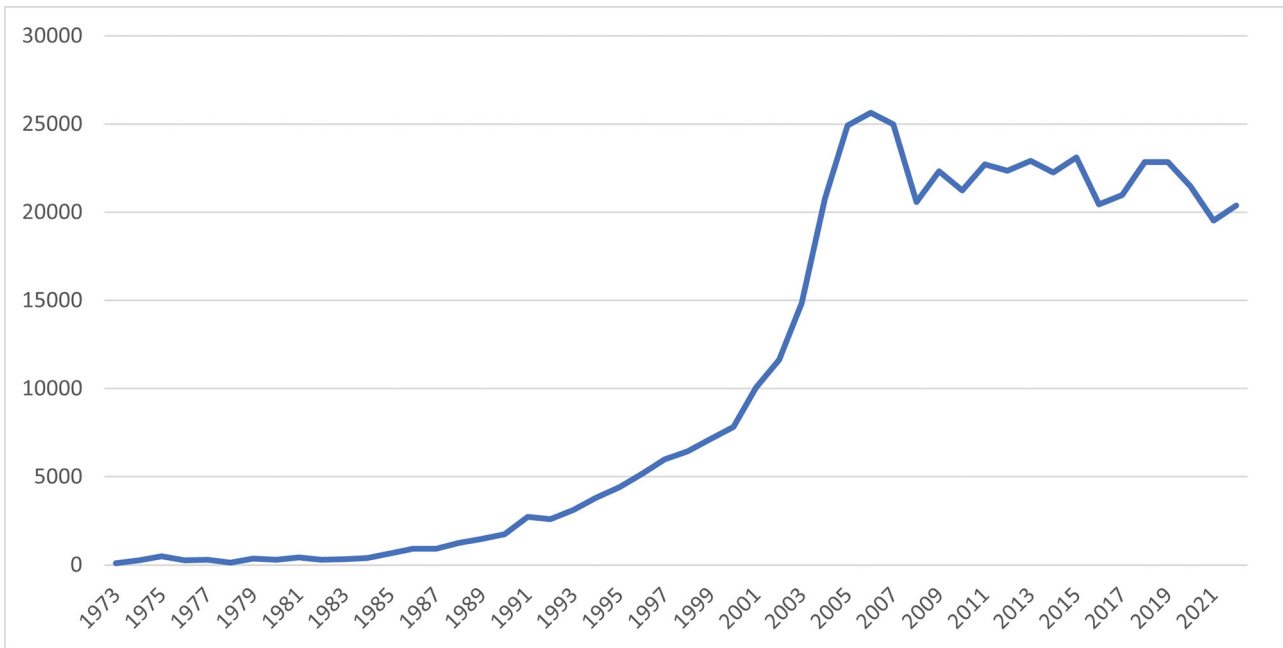
Figure 1: The annual number of global papers published in *Lecture Notes in Computer Science* from 1973 to 2022.
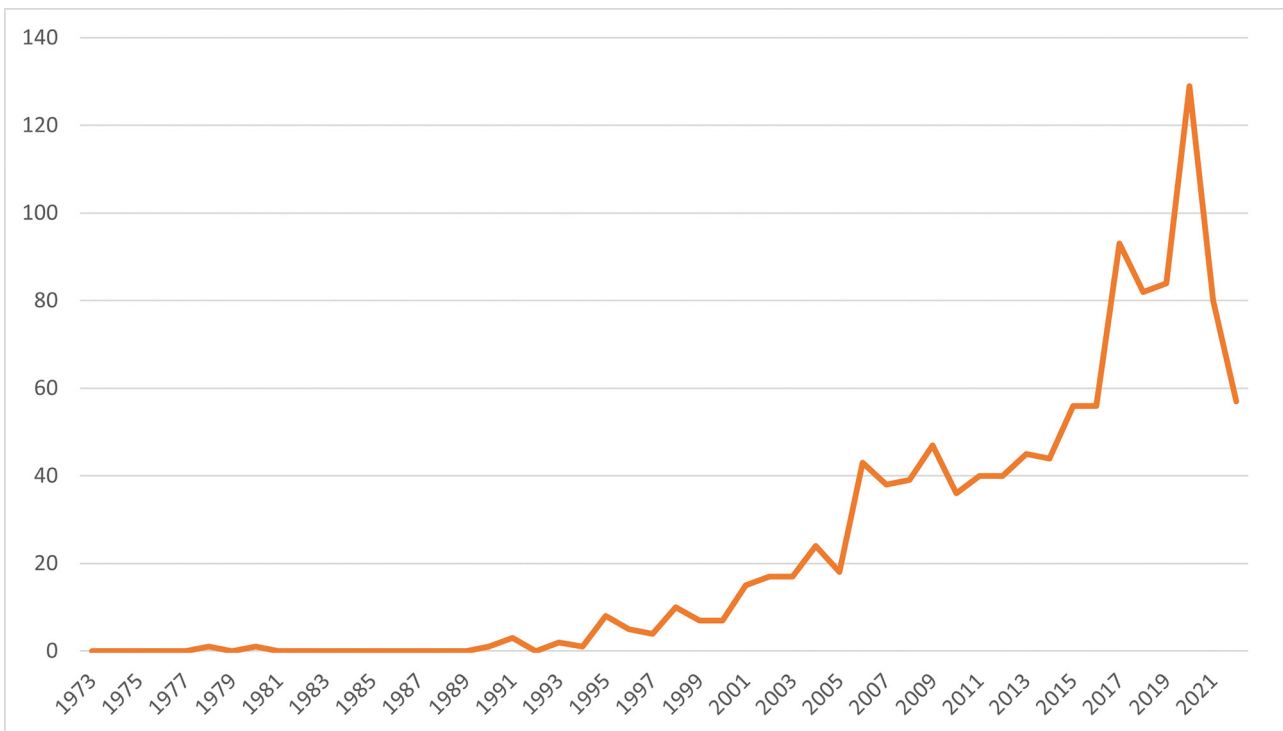
Figure 2: The annual number of papers from South Africa published in *Lecture Notes in Computer Science* from 1973 to 2022.

momentum from around 2005 with spikes in 2006, 2017 and 2020 (129 papers), and then declined again in 2021–2022 (this may be due to the COVID-19 pandemic, and it is not yet clear if the declining trend will continue or if the South African numbers are stabilising around 60 p.a.). The number of annual South African publications peaked 14 years after the global peak. The delayed growth in South African publications may be due to the addition of the Scopus list to the approved DHET list in 2016. This development possibly stimulated a surge in submission of conference papers to LNCS by authors from South Africa from 2016.

The annual number of global LNCS citations grew steadily until 2006 but seems to be diminishing, while the number of South African citations peaked in 2008 and then declined, and seems to be stabilising around 400 p.a. (Figures 3 and 4). It seems that the overall citation impact of LNCS is declining, which prompts a recommendation that authors should also consider other journals with stable or growing citation impact.

The 1150 South African papers were cited in Scopus 6968 times, compared to the global citations of 4 926 601 (i.e. 0.14%). Of the 1150 papers by South African based authors, 836 (73%) had citations and 314 (27%) did not. When comparing the percentage of South African papers in LNCS with the percentage of South African LNCS citations, it is clear that the growth of the number of citations is closely aligned with the number of publications, suggesting that the relationship is quite stable
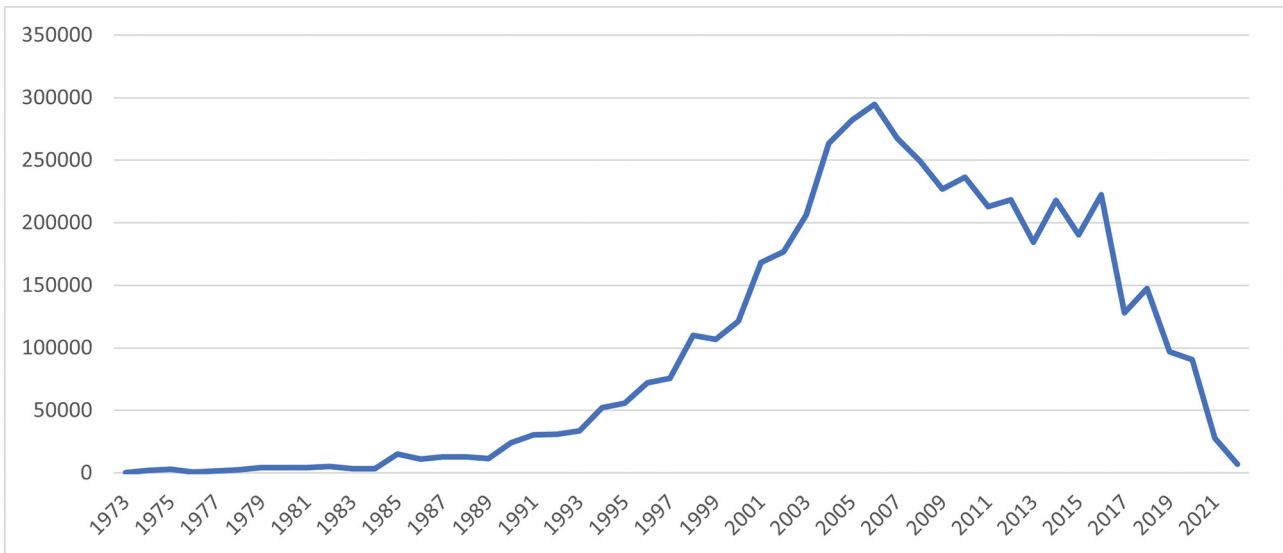
Figure 3: The annual number of citations to global *Lecture Notes in Computer Science* papers.
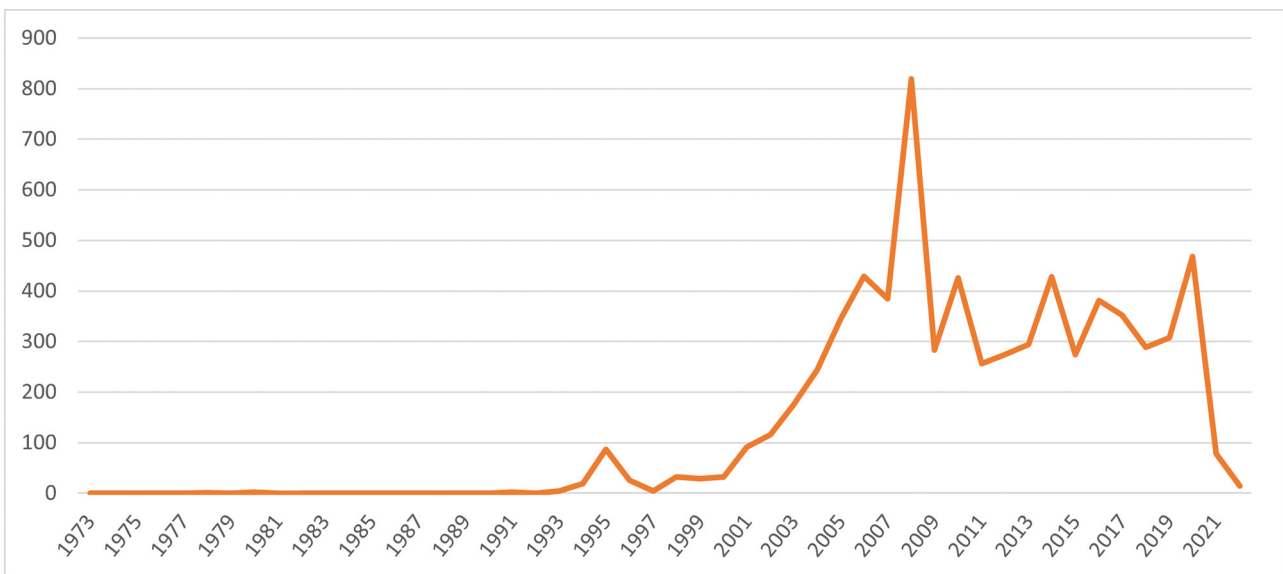
Figure 4: The annual number of citations to *Lecture Notes in Computer Science* papers from South Africa.

(Figure 5). A possible trend that seems to have emerged since 2018 is that South African papers are attracting more citations, almost equalling the percentage of paper numbers in 2020.

When comparing the percentage of the total LNCS average citations per paper with the South African LNCS average citations per paper, it seems that the average number of citations is diminishing but that the South African average is converging to the overall average (Figure 6). This may indicate that the quality of South African papers is becoming comparable to that of global papers and that South African papers are attracting an equivalent number of citations (see Supplements G–J[21]).

Table 5 shows the most productive South African institutions by the number of papers published in LNCS for the period 1978–2022. This information shows that authors affiliated with some of the country's top-ranked universities have published the most in LNCS. This finding may suggest that there has been, and still is, a significant place for LNCS as a Computer Science publication outlet in South Africa.

Figure 7 shows the collaboration behaviour of South African based authors by the number of papers published in LNCS for the period 1978–2022. Of the 1150 papers from South Africa, 492 (43%) represented institutional collaboration, 424 (37%) international and 127 (11%)

national collaboration, while 107 (9%) were single-authored. Overall, LNCS authors from South Africa adopted a collaborative publication culture, with 107 (9%) single-authored and 1043 (91%) co-authored papers. The graph shows that LNCS provides a platform for South African based authors to collaborate with international scholars, as well as with other academics in their home institutions. This may explain the journal's popularity. Surprisingly, national collaboration seems not to be a preferred way of authoring papers in South Africa compared to the Chinese scenario discussed above in which researchers preferred domestic collaboration over international collaboration.[29] Sole authorship is also not popular, which may be explained by a culture in Computer Science of supervisors co-authoring with their postgraduate students (see Supplements K–O[21]).

Scholars from 160 institutions co-authored with South African based LNCS authors (see Supplement P[21]). Figure 8 depicts the international (regional and country) distribution of the LNCS papers contributed by South African scholars and researchers who collaborated internationally. South African based authors collaborated with scholars from 68 countries, with the top five international collaboration partner countries being the UK, Germany, the Netherlands, the USA and Italy (see Supplement Q[21]).
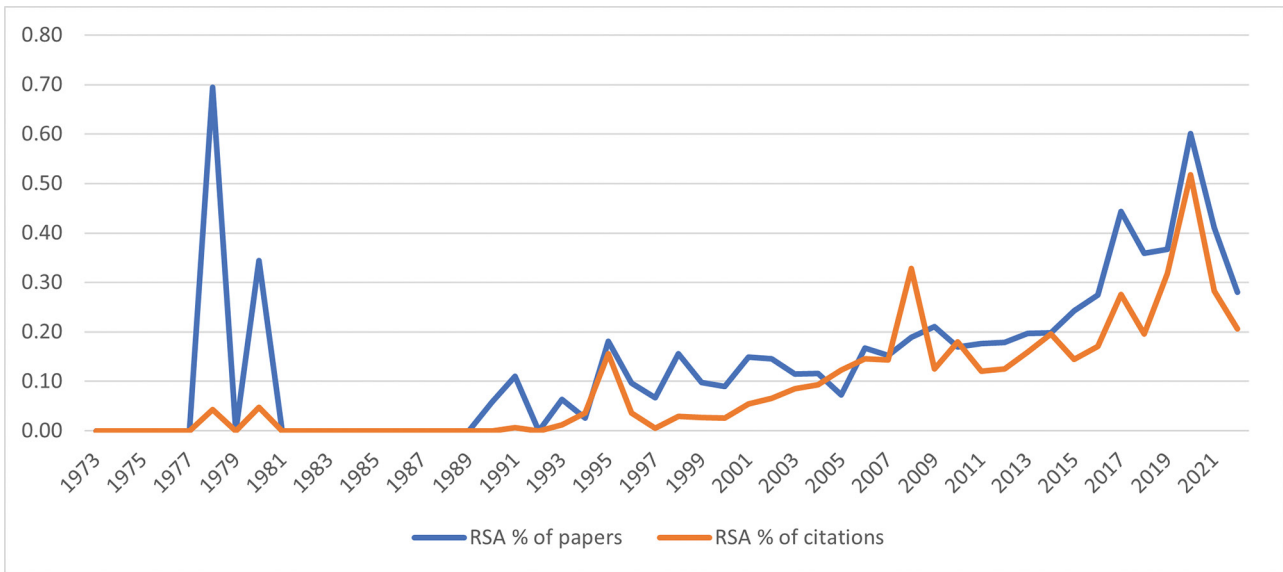
Figure 5: A comparison of the percentage of South African (RSA) papers in *Lecture Notes in Computer Science* (LNCS) with the percentage of South African LNCS citations.
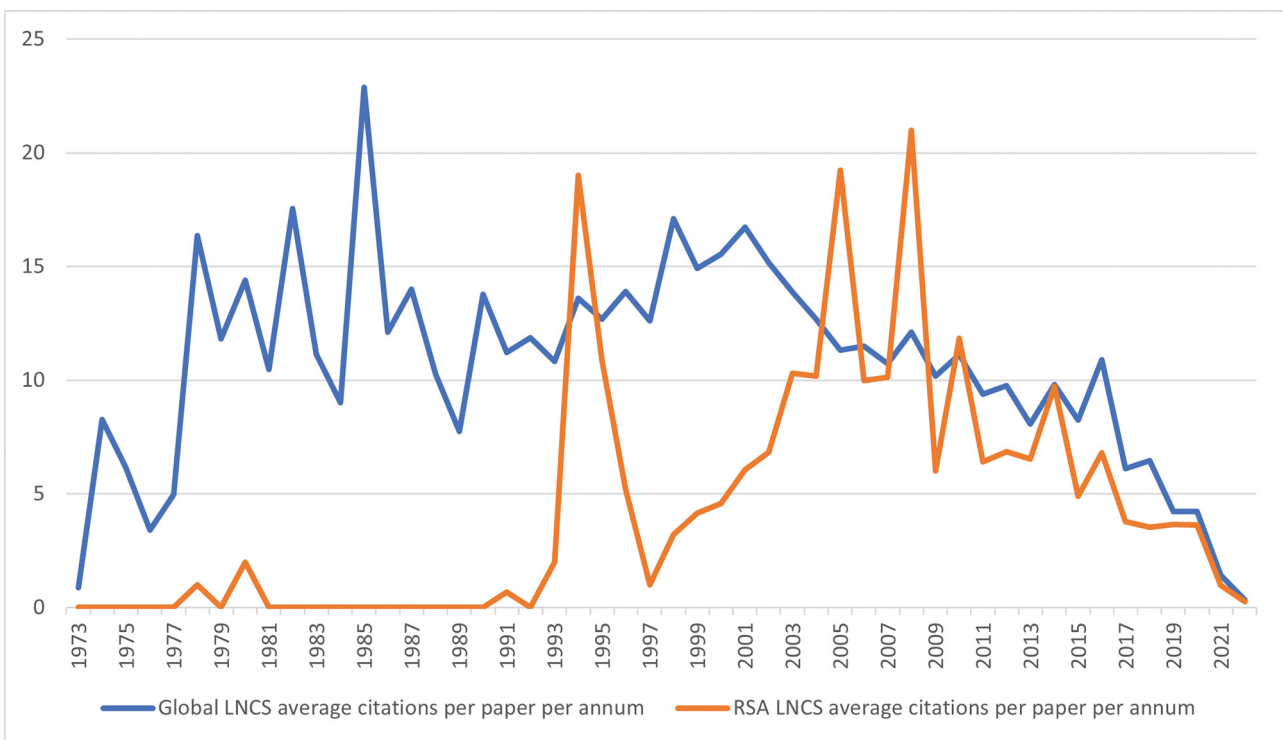
Figure 6: A comparison of the global *Lecture Notes in Computer Science* (LNCS) average citations per paper with the South African (RSA) average citations per paper per annum.

Over the past 44 years, South African based authors have shown a preference to collaborate with scholars in Europe (430), followed by those in North America (64), and less so with scholars in Asia (53), elsewhere in Africa (48), Oceania (41), Latin America (31) and the Middle East (10).

Table 6 ranks the top 50 keywords assigned to papers by the number of papers that include those keywords. A total of 160 keywords were assigned to the 1150 LNCS papers from South Africa. 'Artificial intelligence' was the most used keyword, followed by 'computer science' and 'electronic commerce' (see Supplement R[21]). Artificial intelligence

was also identified by Kotze and Van der Merwe[35], as well as by Fiala and Tutoky[1], as a topic that attracted a significant number of papers.

An integrated comparison of the productivity and citation metrics is useful to gauge South African based authors' contributions and impact. Table 7 provides such an integration. To enable an integrated, fair and accurate comparison, SciVal data for the period 2013–2022 were used. To account for a comparative South African context, data from the *South African Computer Journal* (SACJ) are included in the table. SACJ is the only dedicated Computer Science journal published in South Africa (see Supplements S–W[21]).

Table 5: Top South African contributors to *Lecture Notes in Computer Science* by affiliation (1978–2022)

Affiliation	Number of papers
University of Pretoria	210
University of Cape Town	188
Stellenbosch University	154
University of KwaZulu-Natal	151
University of Johannesburg	125
University of South Africa	91
Council for Scientific and Industrial Research	78
University of the Witwatersrand	74
Tshwane University of Technology	53
Meraka Institute	33
Nelson Mandela University	27
French South African Institute of Technology	20
University of the Western Cape	19
University of the Free State	16
Durban University of Technology	15
North-West University	14
Rhodes University	8
Vaal University of Technology	7
Cape Peninsula University of Technology	5

Discussion

Computer Science researchers could benefit from a shift in publication behaviour, specifically moving towards a more balanced approach by selecting journal papers over conference publications. There are several considerations associated with publishing in LNCS as opposed to journals. Although LNCS is a reputable publication, a paper that appears in this series may not attain the same level of research impact as a paper in a prestigious journal. Conferences typically have tight deadlines and limited timelines for peer review, compared to the longer review cycles of journals and adequate time for detailed feedback and multiple revisions.[47] LNCS papers undergo peer review, but the quality and consistency of the peer review may vary across different LNCS conferences or individual reviewers, making the peer review less rigorous than that of top-tier journals. LNCS has page and space limits and accepts full papers (12–15 pages) and shorter papers (6–11 pages).[12] Researchers can present more comprehensive, detailed and complete work in journals compared to LNCS. Furthermore, journal papers carry more weight than conference papers in individual researcher performance assessments.[19] Compared to publishing in LNCS, publishing in high-impact journals will enhance a researcher's academic reputation, funding opportunities and career advancement. LNCS papers are behind paywalls, which can limit access to researchers without Springer institutional subscriptions. Conference papers tend to become obsolete more quickly than journal articles[48], and the citations they generate are significantly lower[20], which can have a long-term negative impact on a researcher's Hirsch index (h-index), which measures an author's number of publications and citations. While a wide range of journal metrics exists for assessing scientific journals, comparable metrics for scholarly conferences are lacking. Standardised, universally accepted conference citation metrics or curated lists of reputable academic conferences to indicate the quality of conferences are not available.[37] Conference acceptance rates are often used as an evaluation methodology, with lower rates associated with thorough peer-review processes and reputable conferences.[49] However, these acceptance rates are rarely published, making them less accessible. For researchers with a diverse publication record, conference papers can be a complement or supplement to journal papers, but are not a replacement.

In fields where technology develops fast, quick publication turnaround times are important to allow authors to protect their intellectual property on innovative ideas. LNCS provides such a platform: a reliable and established series with short delays between submission and publication and a peer-review process similar to that of journals.

When one compares the average number of publications per author for the whole Computer Science category (all sub-fields and all publication types in the Scopus database), it seems that South African based authors (1.39) are considerably more productive than the global average (0.94) (Table 7). The same is true, and even more so, when one compares the productivity in the sub-field of General Computer Science (GCS) only: South Africa: 1.05 vs global: 0.69. However, South African LNCS authors' productivity (0.64) is slightly below – but very close to – the global GCS (0.69) and LNCS global (0.67) average number of publications per author. This result suggests that South African based authors' use of LNCS as a publication outlet is aligned well with the global pattern, and that LNCS is not overly used by them.

In terms of citations, however, the performance of papers from South Africa is somewhat disappointing. The average number of citations per publication for the whole Computer Science category (all sub-fields and all publication types in the Scopus database) for South African papers (7.5) seems to underperform somewhat compared
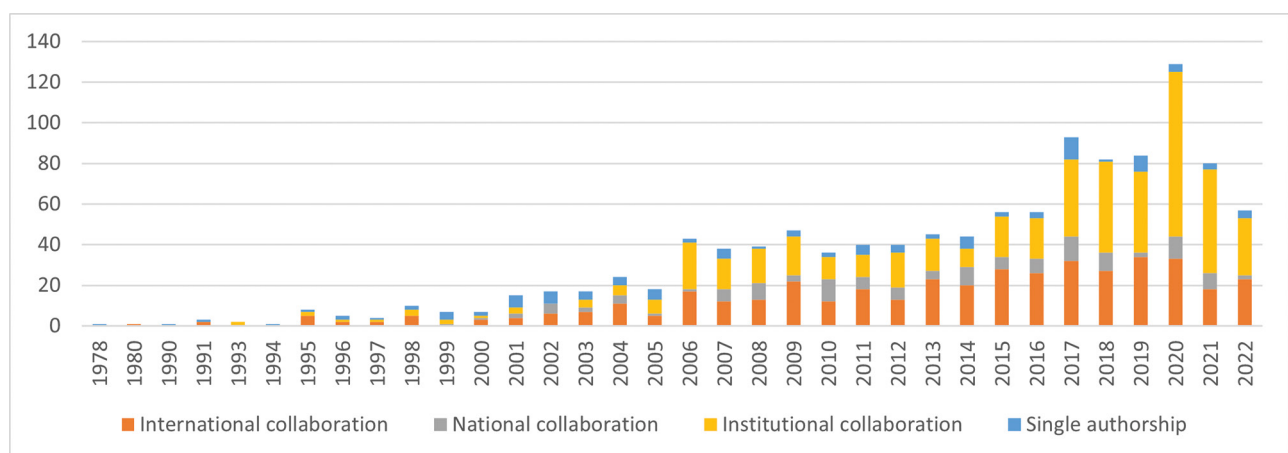
Figure 7: Collaboration behaviour of South African based authors of *Lecture Notes in Computer Science* papers, 1978–2022.
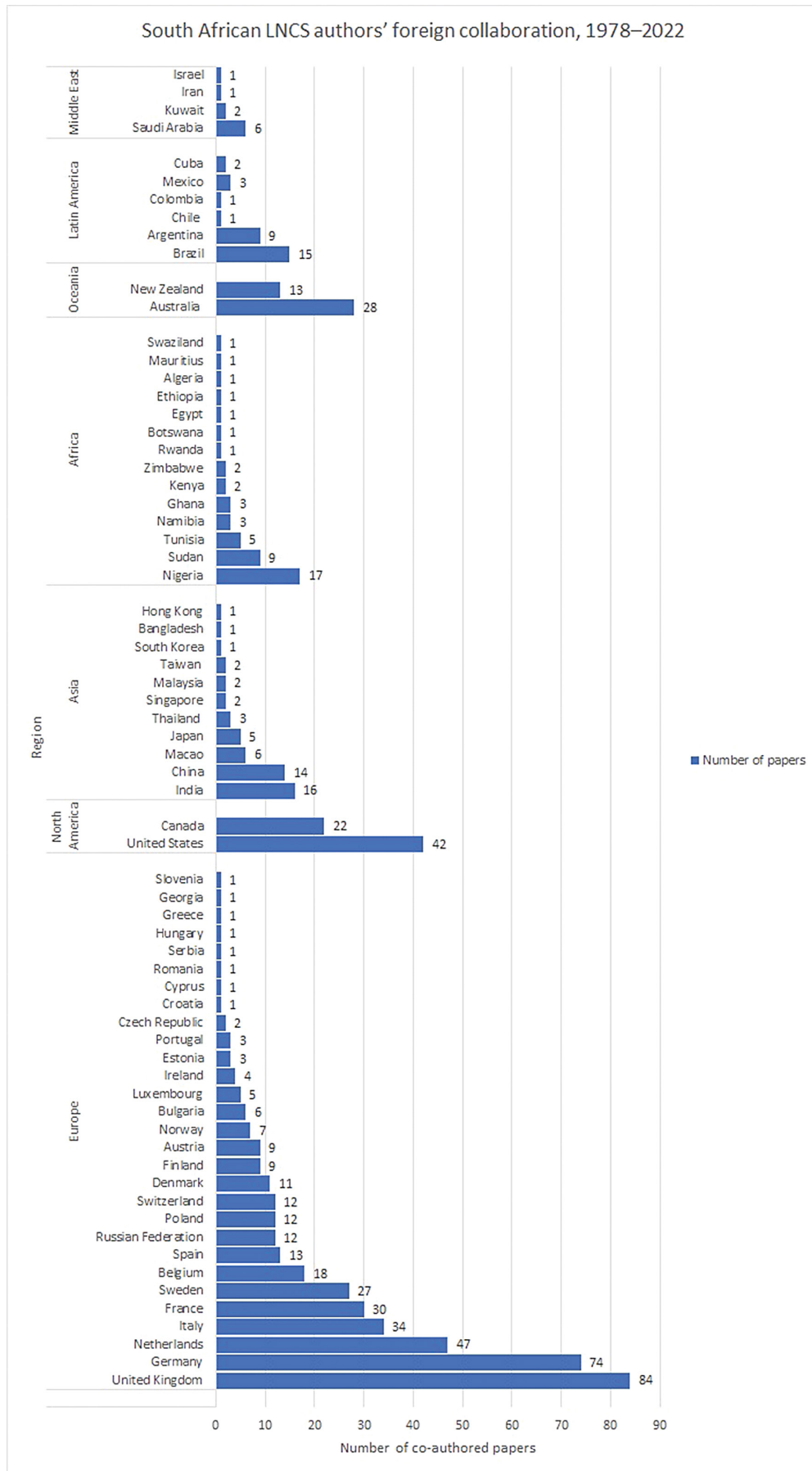
Figure 8: The distribution of international (regional and country) collaboration by South African based authors of *Lecture Notes in Computer Science* papers (1978–2022).

Table 6: Top keywords in *Lecture Notes in Computer Science* papers from South Africa

Keyword	Number of papers	Keyword	Number of papers	Keyword	Number of papers
Artificial Intelligence	172	Digital Libraries	36	Model Checking	26
Computer Science	75	Computer-Aided Instruction	35	Surveys	26
Electronic Commerce	72	Feature Extraction	34	Computer Vision	25
Information Use	72	Students	34	Data Mining	25
Computers	68	Computer Circuits	30	Pattern Matching	25
Algorithms	67	Machine Learning	30	Support Vector Machines	25
Human-Computer Interaction	60	Image Segmentation	30	Developing Countries	24
Semantics	60	Swarm Intelligence	30	Multi-Agent Systems	24
Learning Systems	53	Education	29	Social Networking (online)	24
Particle Swarm Optimization	51	Computational Linguistics	28	Evolutionary Algorithms	23
Optimization	47	Design	28	Learning Algorithms	23
Deep Learning	45	Ontology	28	Convolution	22
E-learning	41	Problem Solving	28	Engineering Education	22
Classification (of Information)	40	Knowledge Management	27	Multi-objective Optimization	22
Neural Networks	40	User Interfaces	27	Websites	22
Information Systems	37	Decision Making	26	Convolutional Neural Network	21
Computation Theory	36	Formal Logic	26	Graphic Methods	21

Table 7: An integrated comparison of publication productivity with citation impact (2013–2022) based on SciVal data

SciVal data 2013–2022	Publication output	Number of authors	Average number of publications per author	Citation count	Citations per publication	FWCI
Global – Computer Science category (all publication types)	4 462 139	4 730 259	0.94	41 819 764	9.4	1.06
South Africa – Computer Science category (all publication types)	21 298 (South Africa ranks 44th out of 220 countries)	15 275	1.39	159 501	7.5	0.92
Global – General Computer Science (all publication types)	818 095	1 183 354	0.69	5 881 007	7.2	1
South Africa – General Computer Science (all publication types)	3829 (South Africa ranks 49th out of 198 countries)	3 651	1.05	26 347	6.9	1.01
Global – LNCS	216 705	324 542	0.67	1 312 423	6.1	1.01
South Africa – LNCS	726 (South Africa ranks 47th out of 159 countries)	1142	0.64	2886	4	0.81
South African Computer Journal	123	195	0.63	469	3.8	0.26

FWCI, Field-Weighted Citation Impact; LNCS, Lecture Notes in Computer Science

to the global average (9.4). South African based authors fare better in the sub-field of General Computer Science (GCS): South Africa: 6.9 vs. global: 7.2. The global LNCS average number of citations per publication (6.1) is somewhat below the global GCS average (7.2), while the South African LNCS average (4) is considerably lower than both. This result suggests that, although South African LNCS authors' production is on par with global LNCS rates, their papers attract considerably fewer citations.

Field-Weighted Citation Impact (FWCI) is the ratio of the total citations actually received by the denominator output and the total citations expected based on the average of the subject field[42]:

- A FWCI of 1.00 indicates that the entity's publications have been cited exactly as would be expected based on the global average for similar publications.

- A FWCI of more than 1.00 indicates that the entity's publications have been cited more than would be expected based on the global average for similar publications.

- A FWCI of less than 1.00 indicates that the entity's publications have been cited less than would be expected based on the global average for similar publications.

The South African FWCI of 0.92 is below the global FWCI of 1.06, showing that South African publications were cited less than would be expected in the Computer Science category (all sub-fields and all publication types in the Scopus database). In the General Computer Science sub-field, the South African FWCI of 1.01 is on par with the global FWCI of 1.00 and equal to the global LNCS FWCI of 1.01. However, the FWCI of South African LNCS papers of 0.81 is below the global LNCS FWCI of 1.01. Although this score is better than the South African average number of citations – which probably says something good about the quality and reputation of South African contributions in LNCS – there is still room for improvement.

The outcome of the comparison above suggests possible reasons for the popularity of LNCS globally and in South Africa. Unfortunately, Scopus citation metrics do not distinguish between journal articles, conference papers and other publication types. It was therefore not possible to compare the various categories with each other, but it was possible to compare LNCS as a conference paper outlet in the GCS field with the trends identified for all publication types in the GCS field. While the average number of publications per author and the number of citations per publication is close to, but somewhat less than, the global GCS rate, LNCS's global FWCI is almost equal to the GCS rate. If one takes into account that LNCS publishes mainly conference papers, while the global GCS category encompasses all publication types, it seems that LNCS can be regarded as a middle-of-the-road outlet for GCS publications. In terms of the FWCI, it is positioned just above the 50th GCS percentile. Therefore, it can be deduced that LNCS probably behaves more like other GCS journals than other conferences. More research is, however, needed to compare the behaviour and performance of various publication types with each other.

Although South African publications in LNCS have attracted fewer citations than the global LNCS publications, the favourable FWCI suggests that LNCS provides an attractive venue for South African publications. The average number of publications per author, which is very close to the global LNCS number, suggests that South African based authors are seizing this opportunity. However, there is still room for improvement in terms of citations and citation impact, but, as suggested by Figure 6, it seems that the number of South African citations may already be converging towards the global LNCS average.

The *South African Computer Journal* (SACJ) is a DHET-accredited journal and the foremost regional publication channel for Computer Science in South Africa, enjoying widespread recognition within the academic community.[35] When comparing the citation metrics of LNCS to SACJ, LNCS outperforms SACJ.[42] In terms of CiteScore, LNCS achieved a score of 2.2 in 2022, whereas SACJ received a score of 0.9. In terms of the SCImago Journal Rank, LNCS obtained a score of 0.320, surpassing SACJ's score of 0.170 in 2022. The SCImago Journal Rank indicator is a measure of the prestige of scholarly journals that accounts for both the number of citations received by a journal and the prestige of the journals from which the citations come.[42] LNCS attained a Source Normalised Impact per Paper (SNIP) score of 0.542 in 2022, while SACJ obtained a score of 0.314. SNIP is a sophisticated metric that intrinsically accounts for field-specific differences in citation practices. It does so by comparing each journal's citations per publication with the citation potential of the field, defined as the set of publications citing that journal.[42] These metrics indicate that LNCS exhibits a higher journal performance than SACJ, which could contribute to the appeal of LNCS among researchers in South Africa.

This study highlights the collaborative nature of South African LNCS researchers, their preference for institutional and international collaboration, and their focus on topics related to artificial intelligence. A trend which emerged, namely that South African based scholars largely prefer to collaborate with European scholars (Figure 8), prompts the recommendation that they should look out for more collaborative research opportunities, not only within their own country but also across the globe, particularly in other parts of Africa. By way of collaborative research outputs in LNCS, Tunisia and South Africa, as the only two African countries among the top 50 countries contributing to LNCS during 1973–2022, can play a leading role in improving other African countries' research footprint. LNCS provides a platform for the publication of papers co-authored with other national and African scholars, as was the case in China (where national collaboration was preferred to international collaboration). These insights regarding collaboration provide valuable directions for future research and emphasise the significance of LNCS in the field of Computer Science. Due to space restrictions, it was not possible to include more detailed groupings broken down by institution or research group or to explore increasing or decreasing collaboration patterns in more depth. This limitation is acknowledged, and further research is suggested to uncover more detail regarding national and international collaboration patterns and trends, as the extent of the relevant data needed in this regard justifies a separate project.

Other related aspects that could be addressed in further research include the following. (1) To help contextualise the results of this article, further research is needed to determine how South Africa's rank in terms of the number of publications in LNCS (Table 4) corresponds with the number of PhDs graduating and the amount of research funding (in general and specific to Computer Science). (2) It may be enlightening to analyse from which specific conferences South African LNCS papers come and how this has changed over time. This will provide further insight into the publication trends discussed above by revealing meaningful patterns regarding the conferences that contribute the South African LNCS papers. (3) To better judge the status of LNCS as a platform for the publication of conference papers, there is a need for a comparative study of South African LNCS papers and South African papers in other conference proceedings or book series such as the Springer series *Communications in Computer and Information Science*. (4) Additional research is warranted to assess the prevalence of Computing researchers in South Africa as first authors in LNCS papers to shed light on the extent of collaborative research spearheaded by researchers in South Africa.

Conclusions

We investigated the involvement and relationship of authors based in South Africa with LNCS for the period 1973–2022. Of the total of 518 662 LNCS papers published globally during the years in question, South African based authors contributed a share of 0.22%, or 1150 papers. Local researchers' publication contributions to LNCS showed a consistent upward trend from 1993 to 2022. Authors from South Africa who published in LNCS preferred to collaborate, with 91% of papers being co-authored. Institutional collaboration (43%) was the most prevalent co-authoring style, followed by international (37%) and national collaboration (11%). The regional analysis showed that Europe was the most significant collaboration partner for local LNCS researchers, followed by North America and Asia. The top five international collaboration partner countries of the LNCS authors from South Africa were the UK, Germany, the Netherlands, the USA and Italy.

The keyword analysis showed that artificial intelligence was the topic of most of the South African LNCS papers.

The citation impact of South African LNCS papers was lower (6.06) than that of global LNCS papers (9.5). Of the 1150 papers from South Africa, 73% had citations and 27% did not. The South African citations represent 0.14% of the global LNCS citations. These metrics compare relatively well to those of journals, and probably much better to other conference proceedings, especially in the South African context.

While the integrated comparison of output numbers and impact showed that South African authors' productivity is on par with that of international scholars, the number of citations per publication is less than that of global scholars. The FWCI confirms that South African publications were cited less than would be expected based on the global average for similar publications.

Overall, the results indicate that LNCS remains a popular publication outlet for Computer Science researchers in South Africa because it has gained and maintained a high scholarly status over the past 50 years. It is a reputable and indispensable publication dissemination platform for Computer Science research, both globally and locally. Proceedings papers remain the most popular publication medium for Computer Science researchers, although there is a trend to find a more equal distribution. LNCS provides a quality platform or outlet for future publications from South African scholars and researchers. However, because LNCS attracts fewer citations than other accredited journals, with the number of citations on the decline, authors are encouraged to explore alternative outlets that offer stronger and more consistent citation impact. Given a seemingly strong reliance on conference publications with shorter papers and much more limited peer review, South African authors should aim and work to compete in a more rigorous and competitive journal-publication world at this stage.

Further investigation is needed to assess the scholarly influence and performance of LNCS (conference proceedings) when compared to other conference proceedings and journals. Conducting a survey among Computer Science researchers to explore the factors influencing their choice of LNCS as a publication venue would also provide valuable insights.

Acknowledgements

Competing interests

We have no competing interests to declare.

Authors' contributions

F.N.: Conceptualisation; methodology; data collection; sample analysis; data analysis; validation; data curation; writing – the initial draft; writing – revisions; project leadership; project management; funding acquisition. J.H.K.: Conceptualisation; methodology; data collection; sample analysis; data analysis; validation; data curation; writing – the initial draft; writing – revisions; project leadership; project management; funding acquisition.

References

1. Fiala D, Tutoky G. Computer science papers in Web of Science: A bibliometric analysis. Publications. 2017;5(4):1–16. https://doi.org/10.3390/publications5040023

2. Mouton J, Blanckenberg J, Van Lill M, Redlinghuys H. A scientometric assessment of the Computer Sciences in South Africa. Paper presented at: 51st Annual Conference of the Southern African Computer Lecturers' Association; 2022 July 21–22; Cape Town, South Africa.

3. Parry DA. Computing research in South Africa: A scientometric investigation. South African Comput J. 2019;31(1):51–79. https://doi.org/10.18489/sacj.v31i1.674

4. Springer. Lecture Notes in Computer Science [homepage on the Internet]. No date. Available from: https://www.springer.com/series/558

5. Gjertsen K. Why are proceedings so important in Computer Science? Springer Nature. 2020 March 22. Available from: https://www.springernature.com/gp/librarians/the-link/blog/blogposts-ebooks/why-are-proceedings-so-important-in-computer-science/17801220

6. Nishikawa-Pacher A. Who are the 100 largest scientific publishers by journal count? A webscraping approach. J Doc. 2022;78(7):450–463. https://doi.org/10.1108/JD-04-2022-0083

7. Springer U. The founding of Springer Publishing Company and its first 20 years, 1950-1970. In: The history of Springer Publishing Company. New York: Springer Publishing Company; 2008. https://doi.org/10.1891/9780826111135

8. Springer. eBooks: Computer Science [document on the Internet]. c2021 [cited 2022 Nov 21]. Available from: https://resource-cms.springernature.com/springer-cms/rest/v1/content/10592782/data/v11

9. Bar-Ilan J. Web of science with the conference proceedings citation indexes: The case of computer science. Scientometrics. 2010;83(3):809–824. https://doi.org/10.1007/s11192-009-0145-4

10. Elsevier Research Intelligence. Scopus Content Coverage Guide [document on the Internet]. c2020 [cited 2021 Jul 22]. Available from: https://www.elsevier.com/__data/assets/pdf_file/0007/69451/Scopus_ContentCoverage_Guide_WEB.pdf

11. Springer. Frequently Asked Questions for Springer Computer Proceedings [webpage on the Internet]. No date [cited 2022 Nov 21]. Available from: https://www.springer.com/gp/computer-science/lncs/faq

12. Hofmann A, Gerstner R, Kramer A, Holzwarth F. Guidelines for Volume Editors of Springer Computer Science Proceedings [webpage on the Internet]. No date [cited 2022 Nov 21]. Available from: https://resource-cms.springernature.com/springer-cms/rest/v1/content/19836488/data/v4

13. Sabinet. Accredited journals [webpage on the Internet]. No date [cited 2022 Nov 21]. Available from: https://journals.co.za/publications-accredited

14. Demetrescu C, Finocchi I, Ribichini A, Schaerf M. On Computer Science research and its temporal evolution. Scientometrics. 2022;127(8):4913–4938. https://doi.org/10.1007/s11192-022-04445-z

15. Vrettas G, Sanderson M. Conferences versus journals in Computer Science. J Assoc Inf Sci Technol. 2015;66(12):2674–2684. https://doi.org/10.1002/asi.23349

16. Franceschet M. The role of conference publications in CS. Commun ACM. 2010;53(12):129–132. https://doi.org/10.1145/1859204.1859234

17. Bouma-Sims E. Learning to drive on the wrong side of the road: How American computing came to rely on conferences for primary publication [thesis]. Raleigh, NC: North Carolina State University; 2021. Available from: http://arxiv.org/abs/2109.06438

18. Meyer B, Choppy C, Staunstrup J, Van Leeuwen J. Research evaluation for Computer Science. Commun ACM. 2009;52(4):31–34. https://doi.org/10.1145/1498765.1498780

19. De Sutter B, Van Den Oord A. To be or not to be cited in Computer Science. Commun ACM. 2012;55(8):69–75. https://doi.org/10.1145/2240236.2240256

20. Zhang L, Glänzel W. Proceeding papers in journals versus the "regular" journal publications. J Informetr. 2012;6(1):88–96. https://doi.org/10.1016/j.joi.2011.06.007

21. Naude F, Kroeze JH. South African LNCS papers 1973–2022 [data set]. Figshare. 2023. Available from: https://doi.org/10.6084/m9.figshare.21397158.v32

22. Fortnow L. Time for computer science to grow up. Commun ACM. 2009;52(8):33–35. https://doi.org/10.1145/1536616.1536631

23. Halpern JY, Parkes DC. Journals for certification, conferences for rapid dissemination. Commun ACM. 2011;54(8):36–38. https://doi.org/10.1145/1978542.1978555

24. Godoy D, Zunino A, Mateos C. Publication practices in the Argentinian Computer Science community: A bibliometric perspective. Scientometrics. 2015;102(2):1795–1814. https://doi.org/10.1007/s11192-014-1450-0

25. Uddin A, Singh VK, Pinto D, Olmos I. Scientometric mapping of Computer Science research in Mexico. Scientometrics. 2015;105(1):97–114. https://doi.org/10.1007/s11192-015-1654-y

26. Țurcan N, Coșuleanu I, Grecu M, Cujba R. Research in computer science in the republic of Moldova: A bibliometric analysis. Rom J Libr Inf Sci. 2019;15(2):27–50. https://doi.org/10.26660/rrbsi.2019.15.2.27

27. Gupta BM, Kshitij A, Verma C. Mapping of Indian Computer Science research output, 1999–2008. Scientometrics. 2011;86(2):261–283. https://doi.org/10.1007/s11192-010-0272-y

28. Visser MS, Moed HF. Developing bibliometric indicators of research performance in Computer Science. Proc ISSI 2005 10th Int Conf Int Soc Sci Inf. 2005;1(February):275–279.

29. He Y, Guan J. Contribution of Chinese publications in computer science: A case study on LNCS. Scientometrics. 2008;75(3):519–534. https://doi.org/10.1007/s11192-007-1781-1

30. Kaur S, Ibrahim R, Selamat A. Research productivity in Computer Science field for Universiti Teknologi Malaysia in 2000–2013. In: 2014 International Conference on Computer and Information Sciences (ICCOINS); 2014 June 03–05; Kuala Lumpur, Malaysia. IEEE; 2014. p.1–6. https://doi.org/10.1109/ICCOINS.2014.6868441

31. De Melo AVF, de Queiroz AP. Bibliometric mapping of papers on geographical information systems (2007-2016). Bull Geod Sci. 2019;25(3), e20190015. https://doi.org/10.1590/s1982-21702019000300015

32. Purnomo A, Susanti T, Rosyidah E, Firdausi N, Idhom M. Digital economy research: Thirty-five years insights of retrospective review. Procedia Comput Sci. 2021;197(2021):68–75. https://doi.org/10.1016/j.procs.2021.12.119

33. McLoughlin S, Maccani G, Prendergast D, Donnellan B. Living Labs: A bibliometric analysis. In: Proceedings of the Annual Hawaii International Conference on System Sciences; 2018 January 03–06; Waikoloa Village, Hawaii. Honolulu, HI: University of Hawaii at Manoa; 2018. p. 4463–4472. https://doi.org/10.24251/HICSS.2018.563

34. Calitz AP. The 50 Year history of SACLA and computer science departments in South Africa. In: Leung WS, Coetzee M, Coulter D, Cotterrell D, editors. ICT Education. SACLA 2021. Communications in Computer and Information Science, vol 1461. Cham: Springer, Cham; 2022. p. 3–23. https://doi.org/10.1007/978-3-030-95003-3_1

35. Kotze P, Van der Merwe A. The research foci of computing research in South Africa as reflected by publications in the South African Computer Journal. South African Comput J. 2010;44(44):67–84. https://doi.org/10.18489/sacj.v44i0.24

36. Kourie DG. The South African computer journal: 1989 to 2010. Trans R Soc South Africa. 2010;65(2):107–111. https://doi.org/10.1080/0035919X.2010.510664

37. Von Solms R, Von Solms B. Publish or perish - but where? South African Comput J. 2016;28(1):44–58. https://doi.org/10.18489/sacj.v28i1.394

38. South African Department of Higher Education and Training. Research outputs policy. Government Gazette. 2015;597(38552):3–31.

39. Abedin B, Jafarzadeh H, Olszak CM. Thirty six years of information systems management: A bibliometric and thematic analysis. Inf Syst Manag. 2021;38(2):151–164. https://doi.org/10.1080/10580530.2020.1781987

40. Baas J, Schotten M, Plume A, Côté G, Karimi R. Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies. Quant Sci Stud. 2020;1(1):377–386. https://doi.org/10.1162/qss_a_00019

41. Cavacini A. What is the best database for Computer Science journal articles? Scientometrics. 2015;102(3):2059–2071. https://doi.org/10.1007/s11192-014-1506-1

42. Elsevier Research Intelligence. Research Metrics Guidebook. 2019;1–68. Available from: https://www.elsevier.com/research-intelligence/resource-library/research-metrics-guidebook

43. Xia W, Jiang Y, Zhu W, Zhang S, Li T. Research fronts of Computer Science: A scientometric analysis. J Scientometr Res. 2021;10(1):18–26. https://doi.org/10.5530/jscires.10.1.3

44. Elsevier. What is the complete list of Scopus subject areas and all science journal classification codes (ASJC)? [webpage on the Internet]. No date [cited 2022 Nov 21]. Available from: https://service.elsevier.com/app/answers/detail/a_id/15181/supporthub/scopus/

45. Benatar SR. An alternative to academic boycott. Nature. 1990;343(6258):505–506. https://doi.org/10.1038/343505a0

46. Haricombe LJ. The effect of an academic boycott on academics in South Africa [dissertation]. Champaign, IL: University of Illinois at Urbana-Champaign; 1992.

47. Gupta S, Kumar N, Bhalla S. Citation metrics and evaluation of journals and conferences. J Inf Sci. Forthcoming 2023. https://doi.org/10.1177/01655515231151411

48. Lisée C, Larivière V, Archambault É. Conference proceedings as a source of scientific information: A bibliometric analysis. J Am Soc Inf Sci Technol. 2008;59(11):1776–1784. https://doi.org/10.1002/asi.20888

49. Chen J, Konstan JA. Conference paper selectivity and impact. Commun ACM. 2010;53(6):79–83. https://doi.org/10.1145/1743546.1743569

AUTHORS:
Martins Nweke[1*] (iD)
Maryjane Ukwuoma[2] (iD)
Ada C. Adiuku-Brown[3]
Adaora J. Okemuo[4]
Princewill I. Ugwu[5] (iD)
Elizabeth Nseka[4]

AFFILIATIONS:
[1]Department of Physiotherapy, University of Benin, Benin, Nigeria
[2]Department of Physiotherapy, University of Nigeria Teaching Hospital, Ituku-Ozalla, Nigeria
[3]Department of Obstetrics and Gynaecology, University of Nigeria, Enugu, Nigeria
[4]Department of Medical Rehabilitation, University of Nigeria, Enugu, Nigeria
[5]Department of Physiology, University of Nigeria, Enugu, Nigeria
*Current: Department of Physiotherapy, University of Pretoria, Pretoria, South Africa

CORRESPONDENCE TO:
Martins Nweke

EMAIL:
Martins.nweke@gmail.com

Burden of postpartum depression in sub-Saharan Africa: An updated systematic review

Postpartum depression (PPD) is a significant public health concern in resource-constrained sub-Saharan Africa (SSA). Efforts to combat this burden are hampered by the region's wide variation in reported prevalence. This review aimed to systematically synthesize up-to-date data on PPD in SSA. The review was structured per the Preferred Reporting Item for Systematic Reviews and Meta-analyses. Included in the review were studies that reported the prevalence of PPD in SSA. A search was undertaken of PubMed, Medline, CINAHL, Academic Search Complete, and PsycINFO. A random-effect model was fitted to estimate the pooled burden of postpartum depressive symptoms in SSA. We conducted subgroup analyses to estimate the distribution of postpartum depressive symptoms based on important study characteristics: sample size, the timing of diagnosis, design, study setting/region, instrument, and income/economy. The prevalence of postpartum depressive symptoms ranged from 3.8% to 69.9%, with a pooled estimate of 22.1% (CI 18.5–26.2; I^2 = 98.2; Tau = 0.848; $p<0.001$). There was a significant variation in postpartum depressive symptoms with sample size ($p<0.001$). The highest prevalence (25.6% CI 21.5–30.1) was obtained within 12 weeks postpartum. The prevalence estimate was highest (23.3%; CI 20.1–26.8) with the Edinburgh Postnatal Depression Scale (EPDS). South Africa (30.6%; CI 23.6–38.7) and Zimbabwe (29.3%; CI 22.2–37.5) reported the highest prevalence rates, while Tanzania (13.5%; CI 10.1–17.9) reported the lowest prevalence estimates. Upper-middle SSA countries presented the highest prevalence rates (30.6%; CI 23.6–38.7). The prevalence was highest within the period 2010–2015. PPD constitutes a significant health burden in SSA and is fast becoming an epidemic in southern Africa.

Significance:

- Given that PPD is a recurring mental health challenge among women in sub-Saharan Africa, there is an urgent need for strategic policy provisions to ameliorate its burden.

- An increase in prevalence of PPD from 2005–2010 to 2015–2021 is indicative of the need for national governments to intensify efforts targeted at achieving the UN Sustainable Development Goals 3 and 5 in the region.

- In SSA, the prevalence of PPD is highest (approximately 30%) in Southern Africa, precisely South Africa and Zimbabwe, where it is fast becoming an epidemic; hence strategies are needed to curtail its growing trend.

- There is a need to characterise and stratify the risk factors of PPD in sub-Saharan to guide policy development of predictive algorithms and implementation strategies.

Introduction

Despite the expanding population of those with postpartum depression (PPD), early detection, treatment, and prevention of postpartum depression remain a challenge, particularly in the resource-scarce region of sub-Saharan Africa (SSA).[1] According to the United Nations, the SSA consists of 46 of the 54 African countries and territories that are fully or partially south of the Sahara, excluding Algeria, Djibouti, Egypt, Libya, Morocco, Somalia, Sudan and Tunisia.[2] PPD results from a complex interplay of physical, mental, and behavioural changes that occur perinatally, especially following childbirth.[3] Those diagnosed with PPD experience extreme feelings of grief, anxiety, or despair that interfere with their ability to undertake daily activities.[4] PPD is a significant medical and psychological condition that predisposes nursing mothers to a low quality of life and ineffective breastfeeding.[5,6] It may impair the care provided to the baby and occasionally exposes babies to physical harm as instances of mothers attempting to harm their infants have been reported in extreme cases of PPD.[7] PPD is often characterised by sadness, disinterest, fatigability, sleep problems, inability to cope with daily activities, and poor appetite.[8] Interestingly, the consequences of PPD are not limited to mothers, with deleterious effects also seen on children's mental development. Children whose mothers had PPD are more likely to have mental health issues and develop ailments in adolescence compared with children of mothers without PPD.[9]

Globally, the prevalence of PPD is estimated to be between 10% and 25%, with sub-Saharan Africa bearing the brunt of the burden.[10,11] In sub-Saharan Africa, varying prevalence rates of PPD have been reported, with Uganda and Zimbabwe having the lowest (7%) and highest rates (33%), respectively.[11] Within a country, prevalence estimates have been found to vary significantly. For example, the prevalence rate in Nigeria ranges from 14.6%[12] to 44.5%[13]. The wide variation in reported estimates of PPD in sub-Saharan Africa may act as a roadblock to strategies aimed at eradicating PPD in the region. In pursuit of universal health coverage and the African Union Agenda 2063, there is a need for a valid and dependable estimate of the PPD burden in sub-Saharan Africa, as we anticipate that policymakers' attention may soon be drawn to the socio-economic implications of PPD, thus catalysing adequate provision for early detection and treatment.

Two systematic reviews[14,15] from Africa on the topic of this review have recently been published. However, one crucial limitation of both reviews was the lack of a detailed search, resulting in relatively few studies contributing to the evidence. For example, Atuhaire et al.[15] reviewed 21 studies, while Dadi et al.[14] reviewed 19 studies. In addition, Atuhaire

et al.[15], who did not employ a meta-analysis, reported a prevalence range of 6.1–44%, while Dadi et al.[14] reported a prevalence range of 3.8–50.3%, with a pooled prevalence of 16.8%. We argue that the discrepancy in the prevalence values achieved in these systematic reviews reflects the difference in the number of articles involved in each of them. The outcome of our preliminary search of five databases yielded over 50 eligible studies, thus our review aimed to systematically synthesise up-to-date evidence on the burden of postpartum depressive symptoms in sub-Saharan Africa.

Methods

Protocol and registration

This was a systematic review of epidemiological studies to systematically summarise the evidence on the prevalence of PPD in sub-Saharan Africa. The protocol was structured using the Preferred Reporting Item for Systematic Reviews and Meta-analyses (PRISMA) checklist.[16] The protocol was registered with the Open Science Framework Registry: https://osf.io/5xzp8/?view_only=f9bb178837474b89828895ed67d5da94

Eligibility criteria

Characteristics of the study

The review included studies that documented the prevalence of PPD in sub-Saharan Africa. Only articles written and published in English were included. Quantitative observational studies were included irrespective of the sample size, sampling technique, and test statistics. The participants in the included studies were postpartum women. Studies were included regardless of whether a control group was employed. The primary outcome was the prevalence of PPD in sub-Saharan Africa. An outcome was included if assessed at least once during the study. Secondary outcomes included clinical, sociodemographic, and study characteristics.

Inclusion and exclusion criteria

In this study, PPD was defined as depression in postpartum women occurring from 10 days to 3 years postpartum. The criteria for inclusion were peer-reviewed observational studies on PPD conducted in sub-Saharan Africa on postpartum women and studies in which PPD was diagnosed 10 days postpartum solely by using a standard instrument (without medical examination/assessment). We excluded studies in which PPD was diagnosed within 10 days solely by the use of a paper-based measure, such as the Edinburgh Postnatal Depression Scale (EPDS) or Patient Health Questionnaire (PHQ). However, we included studies with a mixed period of within and beyond 10 days, provided the average point exceeded 10 days. In addition, we excluded qualitative studies and case studies. The included studies were published between 2006 and 2021.

Sources of information and search strategy

The search strategy was developed, piloted and refined by the primary reviewer (Martins Nweke [MN]). Searches were conducted using a variety of combinations of terms from the medical subject headings (MeSH) and free terms from a selected number of key articles. To begin, a PubMed pilot search was conducted to determine the search strategy's face sensitivity. In the end, the most sensitive strategy was "Depression [All fields] AND (postnatal or postpartum) [All fields] AND Africa [MeSH Terms]", which yielded over 6500 papers. The terms were adapted to the syntax and subject headings of the remaining databases, namely Medline, Academic Search Complete, CINAHL, and PsycINFO. Additional searches for relevant studies were conducted in the references of the identified observational and review articles.

Study records and data management

The results of the literature search were exported to EndNote 8 for removal of duplicates and data management, including the selection of articles for inclusion. Thereafter, the full texts of eligible articles were downloaded. Finally, the included studies' eligibility criteria and screening forms were developed, piloted, and refined.

Procedures for selection and data collection

Two independent reviewers (Adora Justina Okemuo [AJO] and Princewill I. Ugwu [PIU]) conducted an initial screening of the title and abstract concurrently. Conflicting points of view were resolved in consultation with the primary reviewer (MN). The primary reviewer conducted critical cross-checking of the initial screening results and read the full text of selected studies for further screening using the previously defined eligibility criteria. Data were extracted by MN and Maryjane Ukuwoma (MU). Of the eligible studies, 12 (20%) were independently assessed by MN and MU, with an inter-rater agreement of 0.92. We denoted the inter-rater agreement as the ratio of the total number of articles correctly assessed by both raters to the total number of articles assessed. The remaining 46 studies were appraised by MU. Full texts were available for most (56) of the included articles. We contacted two authors to obtain full-text articles, but they did not respond; however, we did not exclude these studies as the abstracts presented important data items. Using the PRISMA diagram, we present details of the flow of studies throughout the selection process, along with the reason for exclusion (Figure 1).

Data items

The following data items were collected from each study: prevalence, authors' identities, and study characteristics such as study design, region, timing of diagnosis, sample size, method of assessment, the instrument used, and sampling technique.

Quality appraisal/risk of bias assessment

The quality assessment checklist for prevalence studies[17] was used to assess the risk of bias. It assesses the methodology's suitability and adequacy, as well as the study's design, participant recruitment, data collection, analysis, and presentation of findings. It is suitable for evaluating most study designs. The tool consists of 10 items, with the 10th item being a corresponding summary score. On a three-point Likert scale, studies were rated as follows: low risk (0–3), moderate risk (4–6), and high risk (7–9). Quality was appraised by two authors (MN and MU). Twelve (20%) of the included studies were independently appraised by MN and MU, with an inter-rater agreement of 0.92. The remaining 48 (80%) were assessed by MU.

Summary measures

Participants' ages were summarised using the mean and standard deviation. Participants' levels of education were summarised in terms of the percentage who attained post-primary education. Prevalence of PPD was summarised using percentages. For longitudinal studies reporting the prevalence of postpartum depressive symptoms at different points, we merged the points and used the average.

Data synthesis and analysis

The pooled prevalence was estimated in the manner described by Wang and Liu[18]. We employed a random effect model throughout.[18,19] The heterogeneity measures, Cochrane's Q statistics, and I^2 were calculated following Higgins et al.[20] I^2 values were interpreted per the Cochrane Handbook for Systematic Reviews of Interventions as follows: 0–40% may indicate low heterogeneity, 30–60% may indicate moderate heterogeneity, 50–90% may indicate substantial heterogeneity and 75–100% may indicate considerable heterogeneity.[19]

Risk of bias across studies and additional analyses

We assessed publication bias using Egger's test. We conducted subgroup analyses to explore the distribution of PPD burden based on study characteristics such as the timing of diagnosis, study design, screening instrument, region, economy, and period. We fitted meta-regression models to identify putative sources of heterogeneity in the burden of PPD in sub-Saharan Africa. Statistical operations were performed using Comprehensive Meta-analysis version 3.

Results

A total of 6861 records were identified from PubMed (326), Medline (598), CINHAL (238), Academic Search Complete (5419), PsycINFO (273), and reference list (7). We eliminated 995 records that were duplicated, leaving 5866 articles for title and abstract screening, of which 5808 were not eligible and 58 studies met the eligibility criteria (Figure 1). In longitudinal studies in which prevalence was reported at different periods, we used a simple average to obtain the summary prevalence
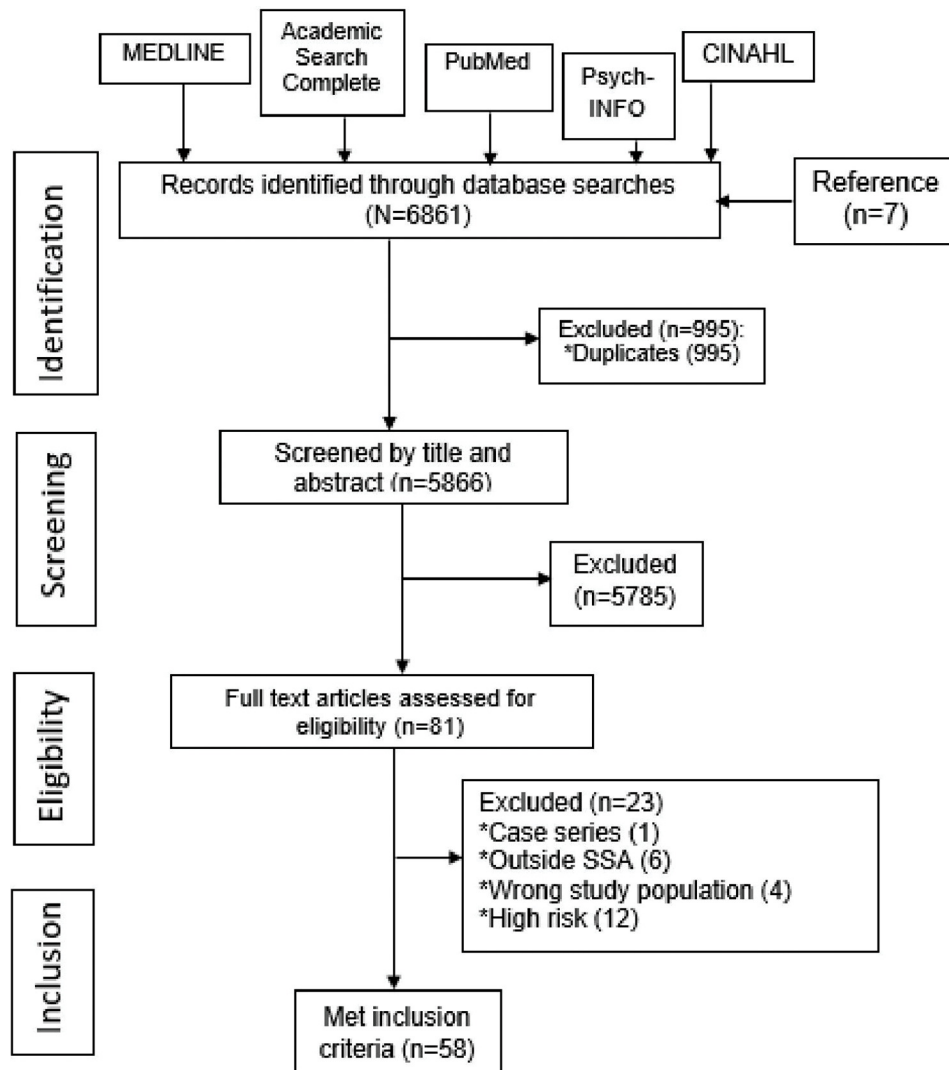
Figure 1: PRISMA flow diagram of the systematic review on burden of postpartum depressive symptoms in sub-Saharan Africa (2004–2020).

point. The 58 studies were also included in the meta-analysis. A sample of 39 090 participants was involved in the prevalence estimate (Table 1).

Study characteristics and study quality

The eligible studies involved 12 of the 46 countries in sub-Saharan Africa. Eleven of the 58 studies were conducted in Ethiopia, and 10 each were conducted in South Africa and Nigeria. More than half (65.8%) of the participants in the included studies had completed at least secondary education. The mean age of the study participants was 20.5 ± 11.8 years. All the included studies had a low risk of bias (Table 1).

Prevalence of PPD in sub-Saharan Africa

The prevalence estimates range from 3.8%[72] to 69.9%[40]. Pooled prevalence was estimated to be 22.1% ([CI 18.5–26.2]; PI=5.0–61.0%), with significant heterogeneity (I^2 = 98.2) and publication bias (Egger's t-test =3.543, p=0.0002) (Figures 2, 3 and 4). Subgroup analyses revealed variation in PPD prevalence due to sample size ($p<0.001$), the timing of diagnosis, study design, study setting/region, instrument, and income/economy. The highest prevalence (25.6% [CI 21.5–30.1]) was obtained less than 3 months postpartum (Table 2). We observed the highest PPD rate in southern Africa (29.5% [CI 24.1–35.6]). South Africa (30.6% [CI 23.6–38.7]) and Zimbabwe (29.3% [CI 22.2–37.5]) reported the highest prevalence rates, while Uganda (11.6% [CI 4.9–]) and Tanzania (13.5% [CI 10.1–17.9]) reported the lowest prevalence estimates. Prevalence was highest within the period 2010–2015. The prevalence estimate was highest (23.3% [CI 20.1–26.8]) with the use of the EPDS (Table 2). A higher rate

of PPD was reported in cross-sectional studies (24.1% [CI 20.5–28.0]) compared to longitudinal studies. Upper-middle sub-Saharan African countries showed the highest prevalence rates (30.6% [CI 23.6–38.7]) (Table 2). A larger sample size was associated with a lower prevalence rate. Incidentally, meta-regression analysis showed only sample size ($p<0.001$) contributed significantly to the study heterogeneity, accounting for 0.40 of the total variance (0.73) in true prevalence (Table 3).

Discussion

The prevalence of PPD varies widely across the globe, with higher rates reported in low- and middle-income countries compared to high-income countries.[11,29] In our study, the prevalence of PPD in sub-Saharan Africa was 21.8% (CI 18.2–25.8) in comparison to 16.8% obtained by Dadi et al.[14], suggesting that sub-Saharan Africa is facing a significant burden of PPD. Clearly, our estimate is higher than the 16.8% reported by Dadi et al.[14] The discrepancy could be because Atuhaire et al.[29] and Dadi et al.[14] employed relatively far fewer studies than were included in our study. In addition, our study was delimited to sub-Saharan Africa, unlike Atuhaire et al.[29] and Dadi et al.[14], which were limited to Africa. PPD remains a public health problem in sub-Saharan Africa, with a prevalence estimate exceeding 10%.[76] Nonetheless, the pooled estimate obtained in this review is within the range reported by Parsons et al.[11] and Dadi et al.[14] for low- and middle-income countries but significantly higher than the values (13.0–13.2%) obtained for higher-income countries.[77] The increasing focus on PPD in industrialised nations may have contributed to increased awareness and policy initiatives to combat PPD, thus reducing

Table 1:   Study characteristics and risk of bias

S/N	Study	Mean (SD) age (years)	Education (% post-primary)	Instrument	Postpartum duration	Country	Quality/ Risk of bias
1	Abadiga et al.[21]	29.6 (9.5)	69.0	EPDS	Within 12 months	Ethiopia	Low risk
2	Abebe et al.[22]	24.3 (3.8)	94.9	EPDS	Within 6 months	Ethiopia	Low risk
3	Abiodun et al.[23]	27.9 (5.9)	81.9	EPDS	6 weeks	Nigeria	Low risk
4	Adewuya and Afolabi[24]	27.5 (11.1)	–	PHQ	Weeks 4, 8, 12, 24 & 36	Nigeria	Low risk
5	Adewuya et al.[13]	28.4 (12.1)	82.6	EPDS	6 weeks	Nigeria	Low risk
6	Adewuya[25]	28.5 (12.2)	–	EPDS	Weeks 4 & 8	Nigeria	Low risk
7	Agbaje et al.[26]	28.9 (6.2)	79.4	EPDS	4–12 weeks	Nigeria	Low risk
8	Anato et al.[27]	28.0 (6.6)	40.9	EPDS	5–10 months	Ethiopia	Low risk
9	Anokye et al.[28]	27.3 (8.31)	96.0	EPDS	Within 12 months	Ghana	Low risk
10	Arach et al.[29]	25.0 (7.0)	77.8	EPDS	Day 50	Uganda	Low risk
11	Atuhaire et al.[15]	27 (–)	98.6	DSM-IV	6 weeks	Ethiopia	Low risk
12	Azale et al.[30]	28.8 (5.2)	20.0	PHQ	1–12 months	Ethiopia	Low risk
13	Bitew et al.[31]	26.8 (–)	61.3	PHQ	4–12 weeks	Ethiopia	Low risk
14	Baggaley et al.[32]	26.0 (7.0)	46	K10	3, 6, 12 months	Burkina Faso	Low risk
15	Chibanda et al.[33]	–	–	DSM-IV	6 weeks	Zimbabwe	Low risk
16	Chinawa et al.[34]	29.5 (4.4)	70.0	EPDS	Within 6 weeks	Nigeria	Low risk
17	Choi et al.[35]	25.0 (5.8)	–	EPDS	Week 6, 6th month & 12th month	South Africa	Low risk
18	Dlamini et al.[36]	31.5 (–)	–	EPDS	Within 6 weeks	Eswatini	Low risk
19	Dow et al.[37]	27.1 (5.3)	86.4	EPDS	Weeks 10 & 14, months 6, 9, 12, 15, 18	Malawi	Low risk
20	Duma and Madiba[38]			EPDS	4–6 weeks	South Africa	Low risk
21	Gebregziabher et al.[39]	27.7 (5.2)	98.7	DSM-IV	2–14 weeks	Eritrea	Low risk
22	Gold et al.[40]	28.0 (11.0)	65.4	PHQ	2 weeks	Ghana	Low risk
23	Guo et al.[41]	29.1 (5.4)	–	PHQ	3 & 12 months	Ghana	Low risk
24	Holm-Larsen et al.[42]	25 (-)	93.0	EPDS	Day 40, 2–3 years	Tanzania	Low risk
25	Hung et al.[43]	26 (-)	–	EPDS	3 months	South Africa	Low risk
26	January et al.[44]	25.4 (5.6)	94.0	EPDS	Within 12 months	Zimbabwe	Low risk
27	January et al.[45]	25.8 (5.4)	55.6	EPDS	6–10 weeks	Zimbabwe	Low risk
28	Kerie et al.[46]	26.1 (5.8)	63.0	EPDS	Within 12 months	Ethiopia	Low risk
29	Madeghe et al.[47]	28.0 (–)	33.5	EPDS	6–16 weeks	Kenya	Low risk
30	Mahenge et al.[48]	27 (–)	25	PHQ	4–36 weeks	Tanzania	Low risk
31	Mbarak et al.[49]	29.2 (6.3)	32.1	EPDS	2–6 weeks	Tanzania	Low risk
32	Nakku et al.[50]	23.4 (4.8)	59.5	SRQ-25, MINI	Within 6 weeks	Uganda	Low risk
33	Nampijja et al.[51]	28.0 (4.8)	98.0	PHQ, MINI plus	1–3 months	Uganda	Low risk
34	Necho et al.[52]	29.85 (6.4)	79.9	EPDS	Within 4 weeks	Ethiopia	Low risk
35	Obindo et al.[12]	28.0 (5.8)	–	EPDS	6–8 weeks	Nigeria	Low risk
36	Odinka et al.[53]	29.65 (4.9)	100	HADS	6–14 weeks	Nigeria	Low risk

Table 1 continued...

S/N	Study	Mean (SD) age (years)	Education (% post-primary)	Instrument	Postpartum duration	Country	Quality/ Risk of bias
37	Okronipa et al.[54]	28.5 (0.3)	67.1	EPDS	Within 12 months	Ghana	Low risk
38	Ongeri et al.[55]	25.5 (–)	73.1	EPDS	6–10 weeks	Kenya	Low risk
39	Owoeye et al.[56]	28.5 (5.3)	82.9	EPDS	Within 4 months	Nigeria	Low risk
40	Pellowski et al.[57]	–	36.5	EPDS	Within 18 months	South Africa	Low risk
41	Peltzer et al.[58]	28.3 (5.7)	78.3	EPDS	6 & 12 months	South Africa	Low risk
42	Pingo et al.[59]	25.1 (6.4)	33.3	EPDS	6 weeks	South Africa	Low risk
43	Ramachandani et al.[60]	26.1 (6.5)	79.2	PHQ	Within 7 weeks	South Africa	Low risk
44	Rogathi et al.[61]	26 (–)	33.9	EPDS	40 days	Tanzania	Low risk
45	Rotheram-Fuller et al.[62]	26.9 (4.5)	63.8	EPDS	2 weeks, months 6, 18, 36	South Africa	Low risk
46	Sefogah et al.[63]	27.1 (4.3)	84.3	EPDS	2 weeks	Ghana	Low risk
47	Shamu et al.[64]	20 (–)	92.9	CES-D	6 weeks	Zimbabwe	Low risk
48	Shitu et al.[65]	30.6 (6.3)	6.4	EPDS	Within 12 months	Ethiopia	Low risk
49	Stellenberg and Abrahams [66]	–	–	EPDS	6, 10, 14 weeks	South Africa	Low risk
50	Stewart et al.[67]	24.4 (–)		DSM-IV	36 weeks	Malawi	Low risk
51	Tomslinson et al.[68]	26.4 (2.9)	45.9	DSM-IV	2 & 18 months	South Africa	Low risk
53	Toru et al.[69]	28.5 (–)	64.3	PHQ	Within 12 months	Ethiopia	Low risk
53	Tungchama et al.[70]	27.0 (6.0)	80.6	EPDS	6–8 weeks	Nigeria	Low risk
54	Turan et al.[71]	24.3 (4.9)	11.9	EPDS	6 weeks	Kenya	Low risk
55	Wemakor et al.[72]	27.9 (8.2)	–	CEDS	Within 5 years	Ghana	Low risk
56	Weobong et al.[73]	22.5 (–)	44.0	PHQ	4 weeks	Ghana	Low risk
57	Wubetu et al.[74]	26.5 (4.5)	66.0	EPDS	Within 6 weeks	Ethiopia	Low risk
58	Yator et al.[75]	31.0 (5.2)	81.3	PHQ	8 weeks	Kenya	Low risk

EPDS, Edinburgh Postnatal Depression Scale; PHQ, Patient Health Questionnaire; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders 4th ed; K10, Kessler Psychological Distress Scale; SRQ-25, Self-Reporting Questionnaire; MINI, Mini-International Neuropsychiatric Interview; HADS, Hospital Anxiety and Depression Scale; CES-D, Center for Epidemiologic Studies Depression Scale; CEDS, Common Education Data Standards

the PPD burden.[29] Interestingly, a higher burden of PPD was observed in upper-middle-income sub-Saharan economies compared to low- and low-middle-income economies. This could be explained by the fact that South Africa, being the only upper-middle-income economy involved in this review, is the epicentre of the HIV/Aids epidemic, which is a risk for PPD.[14,78] It is also possible that South Africa, being an upper-middle economy, might have better reporting of PPD than other African countries with lower economies. Nonetheless, the effects of PPD on mothers are well documented to extend to the partner and child.[79] Manifestations often include crying spells, insomnia, depressed mood, fatigue, anxiety, and poor concentration.[80] Furthermore, severe PPD can result in infanticide and maternal death.[76] Hence, there is a need for persistent and improved support toward ameliorating pregnancy-related mental health challenges in sub-Saharan Africa.

Epidemiologically, the burden of PPD in sub-Saharan Africa varied based on sample size, region, the timing of diagnosis, study design, and test instrument. The variation in PPD rate based on sample size is consistent with Suresh and Chandrashekara[81], and Dadi et al.[14] Because precise

estimation is based on sample size adequacy, studies with larger sample sizes typically estimated lower prevalence than studies with small sample sizes.[14] As the most important source of heterogeneity, attention should be paid to the correct estimation of sample size in sub-Saharan studies, and further reviews should exclude studies that do not meet the country criteria for sample size estimation in terms of the disease burden, age, education, and, where possible, occupation.[16] Low citations of inadequately powered studies will compel journal editors to place a premium on proper sample size estimation. In turn, this will compel mental health researchers in sub-Saharan African countries to use applicable sample sizes. PPD prevalence was highest in southern Africa compared to East and West Africa. This suggests that southern Africa bears the brunt of the PPD burden in sub-Saharan Africa. This may not be unconnected to the fact that most (86%) of the southern African studies were conducted in South Africa, which currently contends with a high burden of HIV/Aids.[77]

PPD prevalence was highest within 12 weeks (25.6% [CI 21.5–30.1]) postpartum, with the use of a cross-sectional design. This result corroborates the finding that PPD diagnosis could be optimised during

Study name	Statistics for each study					Event rate and 95% CI
	Lower limit	Event rate	Upper limit	Z-Value	p-Value	
Abadiga et al. 2019	0.166	0.209	0.260	-9.166	0.000	
Ababe et al. 2019	0.187	0.221	0.259	-11.812	0.000	
Abiodun et al. 2006	0.149	0.186	0.230	-10.806	0.000	
Adewuya & Afolabi, 2005	0.082	0.103	0.129	-16.510	0.000	
Adewuya et al. 2005	0.124	0.146	0.171	-18.456	0.000	
Adewuya et al. 2006	0.082	0.107	0.138	-14.343	0.000	
Agbaje et al. 2019	0.263	0.356	0.462	-2.642	0.008	
Anato et al. 2020	0.179	0.228	0.287	-7.783	0.000	
Anokye et al. 2018	0.045	0.070	0.108	-10.581	0.000	
Arach et al. 2020	0.192	0.211	0.230	-22.779	0.000	
Atuhaire et al. 2021	0.223	0.271	0.324	-7.529	0.000	
Azale et al. 2018	0.111	0.122	0.134	-36.228	0.000	
Bitew et al. 2019	0.177	0.223	0.276	-8.611	0.000	
Baggaley, 2007	0.322	0.440	0.566	-0.935	0.350	
Chibanda 2010	0.270	0.330	0.396	-4.826	0.000	
Chinawa et al. 2016	0.178	0.229	0.290	-7.463	0.000	
Chol et al. 2019	0.328	0.411	0.500	-1.965	0.049	
Dlamini et al. 2019	0.384	0.474	0.565	-0.562	0.574	
Dow et al. 2014	0.085	0.110	0.141	-14.513	0.000	
Duma & Madiba, 2020	0.313	0.400	0.494	-2.083	0.037	
Gebregziabher et al. 2020	0.051	0.074	0.105	-12.802	0.000	
Gold et al. 2013	0.622	0.699	0.767	4.788	0.000	
Guo et al. 2013	0.095	0.118	0.145	-16.600	0.000	
Holm-Larsen et al. 2018	0.086	0.102	0.121	-22.111	0.000	
Hung et al. 2014	0.262	0.317	0.378	-5.628	0.000	
January et al. 2015	0.290	0.342	0.398	-5.320	0.000	
January et al. 2017	0.227	0.304	0.394	-4.079	0.000	
Koria et al. 2018	0.294	0.338	0.386	-6.414	0.000	
Madeghe et al. 2016	0.090	0.130	0.184	-9.041	0.000	
Mahenge et al. 2018	0.109	0.136	0.169	-14.172	0.000	
Mbarak et al. 2019	0.167	0.205	0.248	-10.760	0.000	
Nakku, 2006	0.044	0.062	0.086	-14.931	0.000	
Nampijja et al. 2019	0.263	0.313	0.368	-6.303	0.000	
Necho et al. 2020	0.227	0.270	0.317	-8.591	0.000	
Obindo et al. 2013	0.395	0.444	0.493	-2.218	0.027	
Odinka, 2019	0.290	0.333	0.379	-6.741	0.000	
Olaronipa et al. 2019	0.071	0.093	0.123	-14.669	0.000	
Ongeri et al. 2018	0.136	0.187	0.253	-7.491	0.000	
Owoeye et al. 2006	0.182	0.230	0.286	-8.068	0.000	
Pallowak et al. 2019	0.120	0.145	0.175	-15.533	0.000	
Peltzer et al. 2018	0.320	0.355	0.392	-7.439	0.000	
Pingo et al. 2017	0.332	0.456	0.585	-0.661	0.508	
Ramchan et al. 2009	0.143	0.164	0.188	-19.302	0.000	
Rogathi et al. 2017	0.100	0.118	0.140	-20.829	0.000	
Rotheram-Fuller et al. 2018	0.219	0.243	0.269	-16.299	0.000	
Sofagah et al. 2018	0.279	0.326	0.377	-6.380	0.000	
Shamu et al. 2016	0.187	0.214	0.243	-15.403	0.000	
Shitu et al. 2019	0.204	0.237	0.272	-12.155	0.000	
Stellenberg & Abraham, 2015	0.426	0.503	0.580	0.079	0.937	
Stewart et al. 2019	0.265	0.303	0.345	-8.553	0.000	
Tomlinson et al. 2006	0.169	0.236	0.319	-5.536	0.000	
Toru et al. 2018	0.051	0.072	0.099	-14.170	0.000	
Tungchama et al. 2018	0.185	0.218	0.256	-12.131	0.000	
Turan et al. 2014	0.186	0.252	0.332	-5.491	0.000	
Wemakor et al. 2016	0.236	0.279	0.326	-8.357	0.000	
Weobong et al. 2015	0.035	0.038	0.041	-71.420	0.000	
Wubetu et al. 2015	0.119	0.156	0.201	-10.754	0.000	
Yator et al. 2016	0.393	0.480	0.568	-0.451	0.652	
	0.185	0.221	0.262	-11.129	0.000	

Figure 2: Forest plot showing the prevalence estimate of postpartum depression in sub-Saharan Africa.

the early postpartum period.[82] Similar studies have shown that the optimum period for diagnosing PPD is 6–12 weeks postpartum.[83] Our review revealed an epidemiologically important higher rate of PPD using the EPDS compared to the PHQ and DSM-IV. This finding is consistent with that of Ayoub et al.[83], who found a higher rate of PPD using EPDS, especially if the assessment was made within 6–12 weeks postpartum. The median class for PPD assessment was within 12 weeks. Our study contradicts a previous African-based review[29] which found no variation in the magnitude of PPD with differences in assessment measures. While PHQ has been shown to be a sensitive measure of PPD in the sub-Saharan African setting[28], early detection and PPD optimisation may be better with EPDS. Pregnancy-related depression could be reliably

and validly diagnosed with both the PHQ-9 and EPDS.[83] The PHQ-9 measures somatic symptoms, while the EPDS covers simultaneous anxiety and depressive symptoms in early pregnancy. Administering both scales concurrently may further aid clinicians in identifying antepartum depressive disorders in the clinical setting.[84] However, when concurrent use of the duo is not feasible, we recommend using the EPDS for women in sub-Saharan Africa.

Generally, the burden of PPD has increased compared to values reported about 10 years ago.[11] Similarly, we observed a shift in the country-based distribution of PPD burden. Previously, Zimbabwe had the highest rate of PPD.[11] In contrast, our review found South Africa to have the highest

Funnel Plot of Standard Error by Logit event rate

Figure 3: Funnel plot displaying degree of publication bias.

Distribution of True Effects

The mean effect size is 0.22 with a 95% confidence interval of 0.18 to 0.26.
The true effect size in 95% of all comparable populations falls in the interval 0.05 to 0.61.

Figure 4: The prediction interval explaining the extent of variation in true prevalence.

PPD prevalence, even with a larger pool of studies than Zimbabwe, which had the second highest prevalence. Recall that sample size is negatively associated with the PPD burden.[14] Similarly, Tanzania recorded the lowest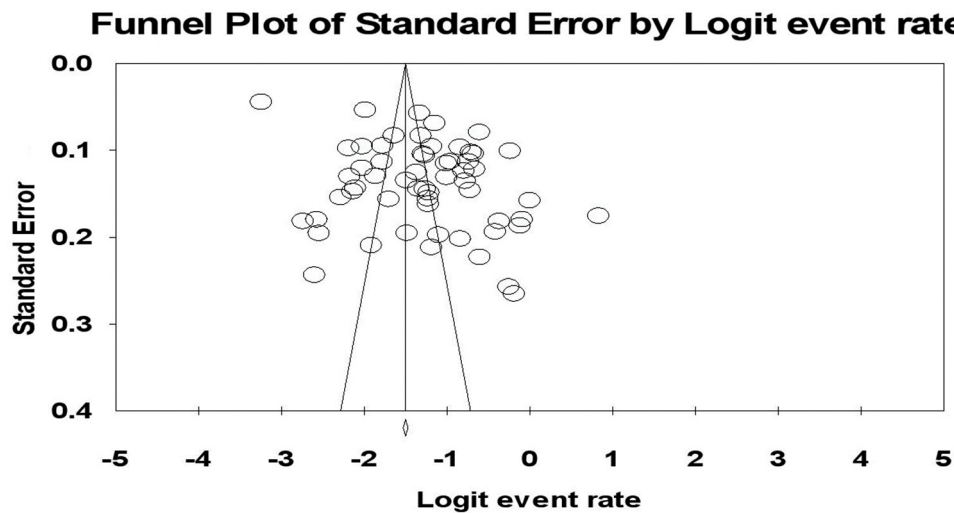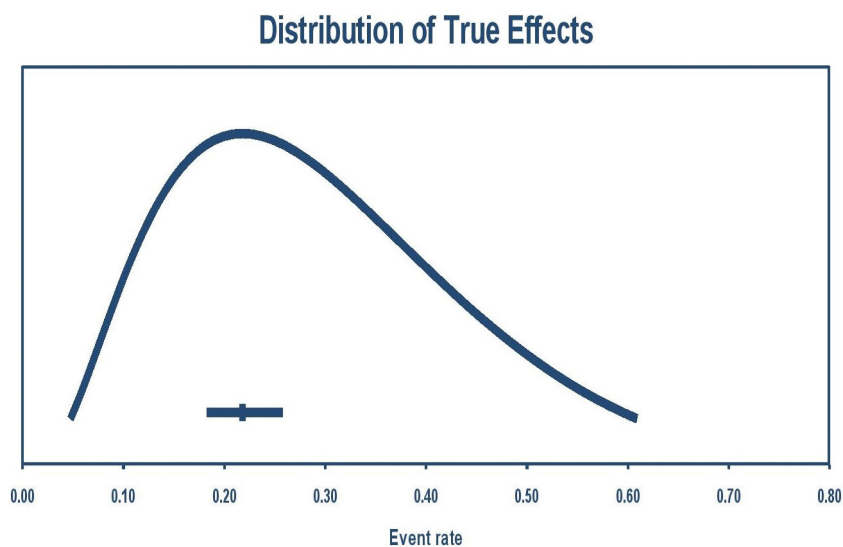 burden, followed by Uganda. This may be connected to the variation in HIV prevalence in sub-Saharan, with Tanzania having one of the lowest HIV prevalence rates.[85] Nonetheless, the increased burden of PPD in the region is not unconnected to the lack of mental health services in the region.[86] There are no mental health services in sub-Saharan communities where over 60% of the African populace lives.[87] In these communities, traditional healers experience increasing patronage from perinatal woman who may have the wrong perception of the cause of PPD.[88] According to Nakku et al.[88], factors such as the unavailability of mental health services within communities, poor attitudes and stigma towards perinatal mental health challenges, insufficient level of mental health literacy and gaps in the health system, poor social support as well as low income constitute putative barriers to utilisation of mental health services in a rural African district.

It is not surprising that a cross-sectional design is better for diagnosing PPD than a longitudinal approach, as PPD is a treatable condition. A fundamental problem of cross-sectional studies is overestimation.[14] Nonetheless, it is ethical to assume that individuals diagnosed with PPD early on will obtain treatment and as a result, the prevalence is expected to be lower over time. Hence, when using the longitudinal approach in estimating the prevalence of PPD, it is necessary to compute prevalence for different postpartum periods and report them separately to avoid underestimating the burden, or a correction/ weighting factor should be modelled to aid an unbiased aggregation of the prevalence points from the different study designs in a meta-analysis. The result is that we know we will need a statistical approach that considers differences in study design, the timing of diagnosis, and setting/region when conducting further research on PPD burden and when estimating pooled PPD prevalence for policy purposes in sub-Saharan Africa.

Implications of findings for practice and/or policy

The high burden of depressive symptoms amongst postpartum women portends a danger for women, children, family and the sub-Saharan African society at large, where poverty limits access to early medical care. By implication, to tackle the upsurge in postpartum depressive symptoms in sub-Saharan Africa, policy should be promulgated in favour of the availability and accessibility of mental health care amongst perinatal women in the region. Currently, in sub-Saharan Africa, mental health services are scarcely available.[86] For example, in the Nigerian and Ugandan health institutions, limited mental health services are

Table 2: Subgroup analyses displaying prevalence per study characteristics

Characteristics	n	Prevalence (CI) (%)	PI (%)	I²	Egger's t-value	p
Timing of diagnosis						
< 3 months	36	25.6 (21.5–30.1)[a]	8.0–58.0	96.187	1.7288	0.093
3–6 months	7	16.3 (11.8–22.2)[b]	5.0–43.0	93.245	0.0678	0.947
6–9 months	10	16.5 (11.5–23.0)[b]	4.0–49.0	97.747	0.3448	0.739
>12 months	5	16.8 (5.9–39.6)[b]	0.0–96.0	99.57	1.7028	0.187
Study design						
Cross-sectional	37	24.1 (20.5–28.0)[a]	8.0–53.0	96.383	1.313	0.198
Longitudinal	21	18.4 (12.8–25.7)[b]	5.0–49.0	98.000	2.854	0.010
Sampling technique						
Probability	12	24.9 (21.2–29.0)[a]	13.0–43.0	91.868	1.565	0.149
Non-probability	45	20.8 (16.5–25.8)[a]	4.0–65.0	98.452	3.249	0.002
Region						
East Africa	23	18.5 (15.2–22.3)[a]	7.0–42.0	95.986	0.280	0.782
Southern Africa	16	29.5 (24.1–35.6)[b]	11.0–58.0	95.168	1.583	0.136
West Africa	19	20.4 (12.9–30.7)[a]		99.039	3.448	0.003
Income/economy						
Low-income	29	19.7 (16.5–23.4)[a]	7.0–45.0	95.822	0.655	0.518
Low-middle	19	21.1 (13.6–31.2)[a]	2.0–77.0	99.084	3.586	0.002
Upper-middle	10	30.6 (23.6–38.7)[b]	10.0–63.0	95.635	1.313	0.226
Instrument						
EPDS	37	23.3 (20.1–26.8)[a]	9.0–49.0	95.510	0.290	0.773
PHQ	9	17.2 (9.4–29.6)[b]	1.0–75.0	99.241	2.615	0.035
DSM-IV	5	22.6 (14.6–33.1)[b,T]	5.0–71.0	94.114	1.431	0.248
Others						
Country						
Ghana	7	17.3 (6.5–38.8)[a]	0.0–93.0	99.436	2.514	0.054
Ethiopia	11	20.3 (15.6–26.0)[a]	7.0–47.0	95.675	1.387	0.199
Kenya	4	24.4 (12.9–41.4)[b]	1.0–93.0	93.972	4.625	0.044
Nigeria	10	21.8 (15.6–29.7)[a]	1.0–84.0	96.446	0.426	0.682
South Africa	10	30.6 (23.6–38.7)[c]	10.0–63.0	95.635	1.313	0.226
Tanzania	4	13.5 (10.1–17.9)[ad]	1.0–45.0	89.185	1.678	0.235
Uganda	3	16.9 (8.3–31.2)[ad]	0.0–100.0	97.423	0.455	0.728
Zimbabwe	4	29.3 (22.2–37.5)[cb]	7.0–70.0	88.288	1.574	0.256
Period						
2005–2010	10	16.4 (11.9–22.1)[a]	1.0–100.0	94.335	8.928	0.994
2010–2015	11	26.9 (12.6–48.4)[b]	1.0–94.0	99.447	4.603	0.001
2015–2021	37	22.1 (19.1–25.3)[a,T]	9.0–46.0	95.766	0.839	0.407

T = significant difference in prevalence by approximately 6 points

PI, prediction interval; EPDS, Edinburgh Postnatal Depression Scale; PHQ, Patient Health Questionnaire; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders 4th ed

[a,b,c,d]Bonferroni post-hoc test; the same superscript indicates no significant difference, while different superscripts indicate a significant difference

Table 3: Meta-regression showing the degree of heterogeneity explained by the difference in study characteristics

Characteristics	OR	*p*	Q	R²	Summary *p*	% Total variance explained by the model
Timing of diagnosis						
< 12 weeks (ref)						
3–6 months	-0.536	0.158	7.56	0.16	0.056	0.117
6–9 months	-0.555	0.093				
>12 months	-0.671	0.025*				
Study design						
Cross-sectional (ref)						
Longitudinal	-0.328	0.143	2.15	0.13	0.143	0.092
Sampling technique						
Probability (ref)						
Non-probability	0.236	0.396	0.72	0.02	0.396	0.0141
Region						
East Africa (ref)						
Southern Africa	0.67	0.015*	6.16	0.08	0.046	0.057
West Africa	0.169	0.518				
Income/economy						
Low-income (ref)						
Low-middle	0.113	0.655	4.03	0.04	0.134	0.029
Upper-middle	0.628	0.046*				
Instrument						
EPDS (ref)	0.153	0.696	2.63	0.29	0.621	0.208
PHQ	0.041	0.932				
DSM-IV	-0.233	0.601				
Others	-0.505	0.54				
Country						
South Africa (ref)						
Ghana	-0.755	0.053	16.49	0.16	0.124	0.113
Ethiopia	-0.562	0.106				
Kenya	-0.322	0.499				
Nigeria	-0.466	0.191				
Tanzania	-1.214	0.027*				
Uganda	-0.788	0.13				
Zimbabwe	-0.066	0.889				
Period						
2005–2010 (ref)						
2010–2015	0.65	0.082	3.05	0.04	0.218	0.028
2015–2021	0.384	0.211				
Sample size	0.0002	<0.001*	17.85	0.55	<0.0001*	0.402

*significant at α < 0.01

ref, reference category; EPDS, Edinburgh Postnatal Depression Scale; PHQ, Patient Health Questionnaire; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders 4th ed

available only at tertiary institutions, whereas most people live in rural areas where no mental health services are available.[88] Furthermore, no guidelines exist regarding the diagnosis of postpartum depressive symptoms in sub-Saharan Africa, although postpartum depressive symptoms may best be diagnosed within 12 weeks postpartum using EPDS. By implication, many trained personnel will be required to ensure that every postpartum woman is tested within 12 weeks. The region must be prepared for large-scale pre-emptive intervention, which must be guided by risk stratification and targeted screening to curb a growing PPD burden and the consequences thereof.[89,90] Overall, sub-Saharan Africa must adopt an holistic approach to curb the prevalence of PPD. Such approaches must comprise curative and preventive arms, providing mental healthcare services at the community level, driving health campaigns on awareness of PPD and its risk factors for pregnant women and communities, and addressing clinically relevant risk factors.

Limitations

The presence of publication bias constitutes a limitation of this study. Also, the aggregation of the prevalence points from longitudinal and cross-sectional studies without a weighting factor constitutes a limitation to the synthesised estimate.

Conclusions

The prevalence of PPD in sub-Saharan Africa is significant, with South Africa and Uganda reporting the highest and lowest burdens, respectively. Detection of PPD can best be optimised within 12 weeks postpartum with the use of the EPDS. In order to improve precision when estimating the regional burden of PPD, the correction factor in sample size estimation should be based on the difference in the prevalence of PPD between the study country and region. Early detection and intervention may help reduce the burden of depressive symptoms amongst postpartum women. The high prevalence of PPD in sub-Saharan Africa should stimulate further research on risk stratification and advocacy for incorporating mental health services across the different tiers of health care. Risk stratification may help identify individuals with the greatest need for pre-emptive interventional care.

Competing interests

We have no competing interests to declare.

Authors' contributions

M.N.: Conceptualisation; methodology; data collection; statistical analysis; writing – the initial draft; writing – revisions. M.U.: Data collection; statistical analysis; writing – the initial draft ; writing – revisions. A.C.A-B.: Methodology; writing – revisions. A.J.O.: Data collection; writing – revisions. P.I.U.: Data collection; writing – revisions. E.N.: Conceptualisation; data collection; writing – revisions. All authors read and approved the final manuscript.

References

1. Rothman R. The health visitor's role and postnatal depression: An overview. Br J Midwifery. 2006;14(11):658–660. https://doi.org/10.12968/bjom.2006.14.11.22253

2. United Nations. Political definition of Major Regions. Archived from the original on 20 April 2010. Retrieved 15 December 2010.

3. Parry Smith WR, Papadopoulou A, Thomas E, Tobias A, Price MJ, Meher S, et al. Uterotonic agents for first-line treatment of postpartum haemorrhage: A network meta-analysis. Cochrane Database Syst Rev. 2020;11(11), CD012754. https://doi.org/10.1002/14651858.CD012754.pub2

4. American College of Obstetricians and Gynecologists. Postpartum depression. Retrieved 1 April 2021.

5. Barber JS, East PL. Home and parenting resources available to siblings depending on their birth intention status. Child Dev. 2009;80(3):921–939. https://doi.org/10.1111/j.1467-8624.2009.01306.x

6. Bartell SS. The birth of a second child: A unique impact on the family system. Int J Childbirth Educ. 2004;19:4–7.

7. Hatters Friedman S, Resnick PJ. Child murder by mothers: Patterns and prevention. World Psychiatry. 2007;6(3):137–141.

8. Sulyman D, Ayanda KA, Dattijo LM, Aminu BM. Postnatal depression and its associated factors among northeastern sub-Saharan African women. Ann Trop Med Public Health. 2016;9(3):184–190. https://doi.org/10.4103/1755-6783.179099

9. Verbeek T, Bockting CL, Van Pampus MG, Ormel J, Meijer JL, Hartman CA, et al. Postpartum depression predicts offspring mental health problems in adolescence independently of parental lifetime psychopathology. J Affect Disord. 2012;136(3):948–954. https://doi.org/10.1016/j.jad.2011.08.035

10. Beck CT, Records K, Rice M. Further development of the postpartum depression predictors inventory-revised. J Obstet Gynecol Neonatal Nurs. 2006;35(6):735–745. https://doi.org/10.1111/j.1552-6909.2006.00094.x

11. Parsons CE, Young KS, Rochat TJ, Kringelbach ML, Stein A. Postnatal depression and its effects on child development: A review of evidence from low- and middle-income countries. Br Med Bull. 2012;101:57–79. https://doi.org/10.1093/bmb/ldr047

12. Obindo JT, Ekwempu CC, Ocheke AN, Piwuna CG, Adegbe EO, Omigbodun OO. Prevalence of postpartum depression in a teaching hospital in north central Nigeria and the associated correlates. Highland Med Res J. 2003;13(2):71–75.

13. Adewuya AO, Fatoye FO, Ola BA, Ijaodola OR, Ibigbami SM. Sociodemographic and obstetric risk factors for postpartum depressive symptoms in Nigerian women. J Psychiatr Pract. 2005;11(5):353–358. https://doi.org/10.1097/00131746-200509000-00009

14. Dadi AF, Miller ER, Mwanri L. Postnatal depression and its association with adverse infant health outcomes in low- and middle-income countries: A systematic review and meta-analysis. BMC Pregnancy Childbirth. 2020;20(1):416. https://doi.org/10.1186/s12884-020-03092-7

15. Atuhaire C, Rukundo GZ, Nambozi G, Ngonzi J, Atwine D, Cumber SN, et al. Prevalence of postpartum depression and associated factors among women in Mbarara and Rwampara districts of south-western Uganda. BMC Pregnancy Childbirth. 2021;21(1):503. https://doi.org/10.1186/s12884-021-03967-3

16. Shamseer L, Moher D, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015: Elaboration and explanation. BMJ. 2015;350, g7647. https://doi.org/10.1136/bmj.g7647

17. Hoy D, Brooks P, Woolf A, Blyth F, March L, Bain C, et al. Assessing risk of bias in prevalence studies: Modification of an existing tool and evidence of interrater agreement. J Clin Epidemiol. 2012;65(9):934–939. https://doi.org/10.1016/j.jclinepi.2011.11.014

18. Wang KS, Liu X. Statistical methods in the meta-analysis of prevalence of human diseases. Biostat Epidemiol. 2016;2:20–24.

19. DerSimonian R, Laird N. Meta-analysis in clinical trials. Control Clin Trials. 1986;7(3):177–188. https://doi.org/10.1016/0197-2456(86)90046-2

20. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editors. Cochrane handbook for systematic reviews of interventions version 6.4 [document on the Internet]. c2023 [cited 2023 Aug]. Available from: www.training.cochrane.org/handbook

21. Abadiga M. Magnitude and associated factors of postpartum depression among women in Nekemte town, East Wollega zone, west Ethiopia, 2019: A community-based study. PLoS ONE. 2019;14(11), e0224792. https://doi.org/10.1371/journal.pone.0224792

22. Abebe A, Tesfaw G, Mulat H, Hibdye G, Yohannes K. Postpartum depression and associated factors among mothers in Bahir Dar Town, Northwest Ethiopia. Ann Gen Psychiatry. 2019;18:19. https://doi.org/10.1186/s12991-019-0244-4

23. Abiodun OA. Postnatal depression in primary care populations in Nigeria. Gen Hosp Psychiatry. 2006;28(2):133–136. https://doi.org/10.1016/j.genhosppsych.2005.11.002

24. Adewuya AO, Afolabi OT. The course of anxiety and depressive symptoms in Nigerian postpartum women. Arch Womens Ment Health. 2005;8(4):257–259. https://doi.org/10.1007/s00737-005-0089-5

25. Adewuya AO. Early postpartum mood as a risk factor for postnatal depression in Nigerian women. Am J Psychiatry. 2006;163(8):1435–1437. https://doi.org/10.1176/ajp.2006.163.8.1435

26. Agbaje OS, Anyanwu JI, Umoke PI, Iwuagwu TE, Iweama CN, Ozoemena EL, et al. Depressive and anxiety symptoms and associated factors among postnatal women in Enugu-North Senatorial District, South-East Nigeria: A cross-sectional study. Arch Public Health. 2019;77:1. https://doi.org/10.1186/s13690-018-0329-6

27. Anato A, Baye K, Tafese Z, Stoecker BJ. Maternal depression is associated with child undernutrition: A cross-sectional study in Ethiopia. Matern Child Nutr. 2020;16(3), e12934. https://doi.org/10.1111/mcn.12934

28. Anokye R, Acheampong E, Budu-Ainooson A, Obeng EI, Akwasi AG. Prevalence of postpartum depression and interventions utilized for its management. Ann Gen Psychiatry. 2018;17:18. https://doi.org/10.1186/s12991-018-0188-0

29. Arach AAO, Nakasujja N, Nankabirwa V, Ndeezi G, Kiguli J, Mukunya D, et al. Perinatal death triples the prevalence of postpartum depression among women in Northern Uganda: A community-based cross-sectional study. PLoS ONE. 2020;15(10), e0240409. https://doi.org/10.1371/journal.pone.0240409

30. Azale T, Fekadu A, Hanlon C. Postpartum depressive symptoms in the context of high social adversity and reproductive health threats: A population-based study. Int J Ment Health Syst. 2018;12:42. https://doi.org/10.1186/s13033-018-0219-x

31. Bitew T, Hanlon C, Medhin G, Fekadu A. Antenatal predictors of incident and persistent postnatal depressive symptoms in rural Ethiopia: A population-based prospective study. Reprod Health. 2019;16(1):28. https://doi.org/10.1186/s12978-019-0690-0

32. Baggaley RF, Ganaba R, Filippi V, Kere M, Marshall T, Sombié I, et al. Detecting depression after pregnancy: The validity of the K10 and K6 in Burkina Faso. Trop Med Int Health. 2007;12(10):1225–1229. https://doi.org/10.1111/j.1365-3156.2007.01906.x

33. Chibanda D, Mangezi W, Tshimanga M, Woelk G, Rusakaniko P, Stranix-Chibanda L, et al. Validation of the Edinburgh Postnatal Depression Scale among women in a high HIV prevalence area in urban Zimbabwe. Arch Womens Ment Health. 2010;13(3):201–206. https://doi.org/10.1007/s00737-009-0073-6

34. Chinawa JM, Odetunde OI, Ndu IK, Ezugwu EC, Aniwada EC, Chinawa AT, et al. Postpartum depression among mothers as seen in hospitals in Enugu, South-East Nigeria: An undocumented issue. Pan Afr Med J. 2016;23, Art. #180. https://doi.org/10.11604/pamj.2016.23.180.8244

35. Choi KW, Sikkema KJ, Vythilingum B, Geerts L, Faure SC, Watt MH, et al. Maternal childhood trauma, postpartum depression, and infant outcomes: Avoidant affective processing as a potential mechanism. J Affect Disord. 2017;211:107–115. https://doi.org/10.1016/j.jad.2017.01.004

36. Dlamini LP, Mahanya S, Dlamini SD, Shongwe MC. Prevalence and factors associated with postpartum depression at a primary healthcare facility in Eswatini. S Afr J Psychiatr. 2019;25, Art. #1404. https://doi.org/10.4102/sajpsychiatry.v25i0.1404

37. Dow A, Dube Q, Pence BW, Van Rie A. Postpartum depression and HIV infection among women in Malawi. J Acquir Immune Defic Syndr. 2014;65(3):359–365. https://doi.org/10.1097/QAI.0000000000000050

38. Duma N, Madiba T. The prevalence of peripartum depression and its relationship to mode of delivery and other factors among mothers in Ixopo, KwaZulu-Natal, South Africa. S Afr J Psychol. 2020;50(4):530–539. https://doi.org/10.1177/0081246320931355

39. Gebregziabher NK, Netsereab TB, Fessaha YG, Alaza FA, Ghebrehiwet NK, Sium AH. Prevalence and associated factors of postpartum depression among postpartum mothers in central region, Eritrea: A health facility based survey. BMC Public Health. 2020;20(1):1614. https://doi.org/10.1186/s12889-020-09676-4

40. Gold KJ, Spangenberg K, Wobil P, Schwenk TL. Depression and risk factors for depression among mothers of sick infants in Kumasi, Ghana. Int J Gynaecol Obstet. 2013;120(3):228–231. https://doi.org/10.1016/j.ijgo.2012.09.016

41. Guo N, Bindt C, Te Bonle M, Appiah-Poku J, Hinz R, Barthel D, et al. Association of antepartum and postpartum depression in Ghanaian and Ivorian women with febrile illness in their offspring: A prospective birth cohort study. Am J Epidemiol. 2013;178(9):1394–1402. https://doi.org/10.1093/aje/kwt142

42. Holm-Larsen CE, Madsen FK, Rogathi JJ, Manongi R, Mushi D, Meyrowitsch DW, et al. Postpartum depression and child growth in Tanzania: A cohort study. BJOG: Int J Obstet Gynaecol. 2019;126(5):590–598. https://doi.org/10.1111/1471-0528.15495

43. Hung KJ, Tomlinson M, Le Roux IM, Dewing S, Chopra M, Tsai AC. Community-based prenatal screening for postpartum depression in a South African township. Int J Gynaecol Obstet. 2014;126(1):74–77. https://doi.org/10.1016/j.ijgo.2014.01.011

44. January J, Chivanhu H, Chiwara J, Denga T, Dera K, Dube T, et al. Prevalence and the correlates of postnatal depression in an urban high density suburb of Harare. Cent Afr J Med. 2015;61:1–4.

45. January J, Mutamba N, Maradzika J. Correlates of postnatal depression among women in Zimbabwean semi-urban and rural settings. J Psychol Afr. 2017;27(1):93–96. https://doi.org/10.1080/14330237.2016.1268299

46. Kerie S, Menberu M, Niguse W. Prevalence and associated factors of postpartum depression in Southwest, Ethiopia, 2017: A cross-sectional study. BMC Res Notes. 2018;11(1):623. https://doi.org/10.1186/s13104-018-3730-x

47. Madeghe BA, Kimani VN, Vander Stoep A, Nicodimos S, Kumar M. Postpartum depression and infant feeding practices in a low income urban settlement in Nairobi-Kenya. BMC Res Notes. 2016;9(1):506. https://doi.org/10.1186/s13104-016-2307-9

48. Mahenge B, Stöckl H, Abubakari A, Mbwambo J, Jahn A. Physical, sexual, emotional and economic intimate partner violence and controlling behaviors during pregnancy and postpartum among women in Dar es Salaam, Tanzania. PLoS ONE. 2016;11(10), e0164376. https://doi.org/10.1371/journal.pone.0164376

49. Mbarak B, Kilewo C, Kuganda S, Sunguya BF. Postpartum depression among women with pre-eclampsia and eclampsia in Tanzania; a call for integrative intervention. BMC Pregnancy Childbirth. 2019;19(1):270. https://doi.org/10.1186/s12884-019-2395-3

50. Nakku JE, Nakasi G, Mirembe F. Postpartum major depression at six weeks in primary health care: Prevalence and associated factors. Afr Health Sci. 2006;6(4):207–214.

51. Nampijja M, Natamba B, Mpango R, Kinyanda E.The burden and risk factors for postnatal depression and depressive symptomatology among women in Kampala. Trop Doct. 2019;49(3):170–177. https://doi.org/10.1177/0049475519837107

52. Necho M, Belete A, Zenebe Y. The association of intimate partner violence with postpartum depression in women during their first month period of giving delivery in health centers at Dessie town, 2019. Ann Gen Psychiatry. 2020;19:59. https://doi.org/10.1186/s12991-020-00310-6

53. Odinka JI, Nwoke M, Chukwuorji JC, Egbuagu K, Mefoh P, Odinka PC, et al. Post-partum depression, anxiety and marital satisfaction: A perspective from Southeastern Nigeria. S Afr J Psychiatr. 2018;24:1109. https://doi.org/10.4102/sajpsychiatry.v24i0.1109

54. Okronipa HE, Marquis GS, Lartey A, Brakohiapa L, Perez-Escamilla R, Mazur RE. Postnatal depression symptoms are associated with increased diarrhea among infants of HIV-positive Ghanaian mothers. AIDS Behav. 2012;16(8):2216–2225. https://doi.org/10.1007/s10461-012-0153-x

55. Ongeri L, Wanga V, Otieno P, Mbui J, Juma E, Stoep AV, et al. Demographic, psychosocial and clinical factors associated with postpartum depression in Kenyan women. BMC Psychiatry. 2018;18(1):318. https://doi.org/10.1186/s12888-018-1904-7

56. Owoeye AO, Aina OF, Morakinyo O. Risk factors of postpartum depression and EPDS scores in a group of Nigerian women. Trop Doct. 2006;36(2):100–103. https://doi.org/10.1258/004947506776593341

57. Pellowski JA, Bengtson AM, Barnett W, DiClemente K, Koen N, Zar HJ, et al. Perinatal depression among mothers in a South African birth cohort study: Trajectories from pregnancy to 18 months postpartum. J Affect Disord. 2019;259:279–287. https://doi.org/10.1016/j.jad.2019.08.052

58. Peltzer K, Rodriguez VJ, Lee TK, Jones D. Prevalence of prenatal and postpartum depression and associated factors among HIV-infected women in public primary care in rural South Africa: A longitudinal study. AIDS Care. 2018;30(11):1372–1379. https://doi.org/10.1080/09540121.2018.1455960

59. Pingo J, Van den Heuvel LL, Vythilingum B, Seedat S. Probable postpartum hypomania and depression in a South African cohort. Arch Womens Ment Health. 2017;20(3):427–437. https://doi.org/10.1007/s00737-017-0719-8

60. Ramchandani PG, Richter LM, Stein A, Norris SA. Predictors of postnatal depression in an urban South African cohort. J Affect Disord. 2009;113(3):279–284. https://doi.org/10.1016/j.jad.2008.05.007

61. Rogathi JJ, Manongi R, Mushi D, Rasch V, Sigalla GN, Gammeltoft T, et al. Postpartum depression among women who have experienced intimate partner violence: A prospective cohort study at Moshi, Tanzania. J Affect Disord. 2017;218:238–245. https://doi.org/10.1016/j.jad.2017.04.063

62. Rotheram-Fuller EJ, Tomlinson M, Scheffler A, Weichle TW, Hayati Rezvan P, Comulada WS, et al. Maternal patterns of antenatal and postnatal depressed mood and the impact on child health at 3-years postpartum. J Consult Clin Psychol. 2018;86(3):218–230. https://doi.org/10.1037/ccp0000281

63. Sefogah PE, Samba A, Mumuni K, Kudzi W. Prevalence and key predictors of perinatal depression among postpartum women in Ghana. Int J Gynaecol Obstet. 2020;149(2):203–210. https://doi.org/10.1002/ijgo.13124

64. Shamu S, Zarowsky C, Roelens K, Temmerman M, Abrahams N. High-frequency intimate partner violence during pregnancy, postnatal depression and suicidal tendencies in Harare, Zimbabwe. Gen Hosp Psychiatry. 2016;38:109–114. https://doi.org/10.1016/j.genhosppsych.2015.10.005

65. Shitu S, Geda B, Dheresa M. Postpartum depression and associated factors among mothers who gave birth in the last twelve months in Ankesha district, Awi zone, North West Ethiopia. BMC Pregnancy Childbirth. 2019;19(1):435. https://doi.org/10.1186/s12884-019-2594-y

66. Stellenberg EL, Abrahams JM. Prevalence of and factors influencing postnatal depression in a rural community in South Africa. Afr J Prim Health Care Fam Med. 2015;7(1), Art. #874. https://doi.org/10.4102/phcfm.v7i1.874

67. Stewart DE, Vigod SN. Postpartum depression: Pathophysiology, treatment, and emerging therapeutics. Annu Rev Med. 2019;70:183–196. https://doi.org/10.1146/annurev-med-041217-011106

68. Tomlinson M, Cooper PJ, Stein A, Swartz L, Molteno C. Post-partum depression and infant growth in a South African peri-urban settlement. Child Care Health Dev. 2006;32(1):81–86. https://doi.org/10.1111/j.1365-2214.2006.00598.x

69. Toru T, Chemir F, Anand S. Magnitude of postpartum depression and associated factors among women in Mizan Aman town, Bench Maji zone, Southwest Ethiopia. BMC Pregnancy Childbirth. 2018;18(1):442. https://doi.org/10.1186/s12884-018-2072-y

70. Tungchama FP, Obindo JT, Armiya'u AY, Maigari YT, Davou FJ, Goar SG, et al. Prevalence and sociodemographic correlates of postpartum depression among women attending postnatal and/or children's welfare clinics in a tertiary hospital, Jos, Nigeria. Sahel Med J. 2018;21(1):23–30. https://doi.org/10.4103/smj.smj_39_16

71. Turan B, Stringer KL, Onono M, Bukusi EA, Weiser SD, Cohen CR, et al. Linkage to HIV care, postpartum depression, and HIV-related stigma in newly diagnosed pregnant women living with HIV in Kenya: A longitudinal observational study. BMC Pregnancy Childbirth. 2014;14:400. https://doi.org/10.1186/s12884-014-0400-4

72. Wemakor A, Mensah KA. Association between maternal depression and child stunting in Northern Ghana: A cross-sectional study. BMC Public Health. 2016;16(1):869. https://doi.org/10.1186/s12889-016-3558-z

73. Weobong B, Ten Asbroek AH, Soremekun S, Danso S, Owusu-Agyei S, Prince M, et al. Determinants of postnatal depression in rural Ghana: Findings from the DON population based cohort study. Depress Anxiety. 2015;32(2):108–119. https://doi.org/10.1002/da.22218

74. Wubetu AD, Engidaw NA, Gizachew KD. Prevalence of postpartum depression and associated factors among postnatal care attendees in Debre Berhan, Ethiopia, 2018. BMC Pregnancy Childbirth. 2020;20(1):189. https://doi.org/10.1186/s12884-020-02873-4

75. Yator O, Mathai M, Vander Stoep A, Rao D, Kumar M. Risk factors for postpartum depression in women living with HIV attending prevention of mother-to-child transmission clinic at Kenyatta National Hospital, Nairobi. AIDS Care. 2016;28(7):884–889. https://doi.org/10.1080/09540121.2016.1160026

76. Almond P. Postnatal depression: A global public health perspective. Perspect Public Health. 2009;129(5):221–227. https://doi.org/10.1177/1757913909343882

77. Bauman BL, Ko JY, Cox S, D'angelo Mph DV, Warner L, Folger S, et al. Vital signs: Postpartum depressive symptoms and provider discussions about perinatal depression - United States, 2018. Morbidity and Mortality Weekly Report. 2020;69(19):575–581. https://doi.org/10.15585/mmwr.mm6919a2

78. Gona PN, Gona CM, Ballout S, Rao SR, Kimokoti R, Mapoma CC, et al. Burden and changes in HIV/AIDS morbidity and mortality in Southern Africa Development Community countries, 1990-2017. BMC Public Health. 2020;20(1):867. https://doi.org/10.1186/s12889-020-08988-9

79. Li Q, Yang S, Xie M, Wu X, Huang L, Ruan W, et al. Impact of some social and clinical factors on the development of postpartum depression in Chinese women. BMC Pregnancy Childbirth. 2020;20(1):226. https://doi.org/10.1186/s12884-020-02906-y

80. Andrews-Fike C. A review of postpartum depression. Prim Care Companion J Clin Psychiatry. 1999;1(1):9–14. https://doi.org/10.4088/PCC.v01n0103

81. Suresh K, Chandrashekara S. Sample size estimation and power analysis for clinical research studies. J Hum Reprod Sci. 2012;5(1):7–13. https://doi.org/10.4103/0974-1208.97779

82. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 5th ed. Washington DC: American Psychiatric Association; 2013. https://doi.org/10.1176/appi.books.9780890425596

83. Ayoub K. Prevalence of postpartum depression among recently delivering mothers in Nablus District and its associated factors [MPH thesis]. Nablus: An-Najah National University; 2014.

84. Zhong Q, Gelaye B, Rondon M, Sánchez SE, García PJ, Sánchez E, et al. Comparative performance of Patient Health Questionnaire-9 and Edinburgh Postnatal Depression Scale for screening antepartum depression. J Affect Disord. 2014;162:1–7. https://doi.org/10.1016/j.jad.2014.03.028

85. Dwyer-Lindgren L, Cork MA, Sligar A, Steuben KM, Wilson KF, Provost NR, et al. Mapping HIV prevalence in sub-Saharan Africa between 2000 and 2017. Nature. 2019;570(7760):189–193. https://doi.org/10.1038/s41586-019-1200-9

86. Baron E, Bass J, Murray SM, Schneider M, Lund C. A systematic review of growth curve mixture modelling literature investigating trajectories of perinatal depressive symptoms and associated risk factors. J Affect Disord. 2017;223:194–208. https://doi.org/10.1016/j.jad.2017.07.046

87. United Nations. Mental health matters: Social inclusion of youth with mental health conditions. New York: UN Division for Social Policy and Development; 2014.

88. Nakku JEM, Okello ES, Kizza D, Honikman S, Ssebunnya J, Ndyanabangi S, et al. Perinatal mental health care in a rural African district, Uganda: A qualitative study of barriers, facilitators and needs. BMC Health Serv Res. 2016;16:295. https://doi.org/10.1186/s12913-016-1547-7

89. Nweke M, Ukwuoma M, Adiuku-Brown AC, Ugwu P, Nseka E. Characterization and stratification of the correlates of postpartum depression in sub-Saharan Africa: A systematic review with meta-analysis. Womens Health (Lond). 2022;18. https://doi.org/10.1177/17455057221118773

90. Nwagha T, Nweke M. Stratification of risk factors of lung cancer-associated venous thromboembolism and determining the critical point for preemptive intervention: A systematic review with meta-analysis. Clin Med Insights Oncol. 2023;17. https://doi.org/10.1177/11795549231175221

Appendix 1: Pilot search strategies

Search terms	Date of search	Databases	Articles retrieved
(Depression) [All Fields] AND (postnatal or postpartum) [All Fields] AND (Africa [MeSH Terms])	27/10/2020	PubMed	326

AUTHORS:
Sello P. Moloi[1,2] (iD)
Jonathan O. Okonkwo[1] (iD)
Raymond Jansen[1] (iD)

AFFILIATIONS:
[1]Department of Environmental, Water and Earth Sciences, Faculty of Science, Tshwane University of Technology, Pretoria, South Africa
[2]National Research Foundation, Pretoria, South Africa

CORRESPONDENCE TO:
Jonathan Okonkwo

EMAIL:
OkonkwoOJ@tut.ac.za

Defluoridation of water through the application of carbonised bone as a green adsorbent: A review

Flouride contamination of water is recognised as a serious challenge facing humanity. Consumption of water that contains excessive amounts of fluoride can result in fluorosis. Consequently, concerted efforts have been made to develop cheap, effective and green techniques/materials to remove fluoride from water, particularly potable water. Bone char prepared from bovine, swine, and equine bones has been used extensively in this regard, and is the most promising, cheap and green material for treating drinking water with high fluoride concentration, particularly in developing countries. However, research on bone char prepared from bones of animals in the wild, as well as those from domestic and semi-wild animals treated with antibiotics to enhance growth, has been scanty. Such research is important as the use of antibiotics may alter the composition of bones, and thus their potential as a green adsorbent to remove fluoride may be impaired. Furthermore, little attempt has been made so far to package char bones for easy application domestically, particularly in rural communities.

Significance:
- Contamination of water by fluoride is a major problem globally.
- Various techniques and materials have been employed for water defluoridation, including the use of bone char, which has several advantages.
- Bone char prepared from bones of animals in the wild and those from domestic and semi-wild animals treated with antibiotics to enhance growth should be further investigated.
- Cheaper and less elaborate processes and packaging are required to scale down the use of bone char at domestic level.

Introduction

Covering approximately 75% of the surface of the earth, water is the most common substance on earth and one of the major elements that are essential for sustaining all life forms on earth.[1] Water has the ability to dissolve almost all substances with which it comes into contact, hence it is frequently referred to as a universal solvent.[2] About 60% of the available fresh water is found in just nine countries and the distribution of water is still uneven within these countries.[3] The shortage of water is observed as an arduous challenge that the modern world is facing. Fluoride (F) in water affects millions of people worldwide and hence is a major contributor to the world's water crisis.[4] Arid and semi-arid regions are ideal sites for contamination of drinking water by fluoride.[5] Groundwater is one of the primary sources of water for daily needs in many regions of the world.[6] Around 200 million people from 25 nations in the world, including highly populated countries like China and India, and a significant population from East Africa, are worst affected by the presence of excess fluoride in their drinking water.[7-10]

South Africa is a water-stressed country. However, like other areas of the globe facing severe water stress have demonstrated, there are solutions available to cultivate a more secure water future for South Africa.[11] These solutions include using groundwater more frequently in areas where the supply is sustainable.[12] Most rural communities in Africa, including South Africa, depend on groundwater as the foremost water reservoir. However, groundwater is prone to contamination by chemicals that occur in nature, including fluoride.[13-15] Fluorine is a highly reactive element of fluoride which is found naturally as calcium fluoride (CaF_2).[16] Due to geogenic causes, the distribution of fluoride in the environment is unbalanced.[17] Fluoride has a strong liking for the acquisition of electrons, hence the formation of negative fluoride ions (F^-). As a result, fluoride forms composites with numerous positively charged ions, which constitute about 0.75% of the earth's crust. Seawater contains fluoride that is about 1 mg/L in concentration, while lakes and rivers, and groundwater have fluoride concentrations of 0.5 mg/L and 1–35 mg/L, respectively.[18,19]

Origins of fluoride

Fluoride originates from both natural sources, such as volcanic activities, as well as anthropogenic sources, such as pesticides and industrial waste. The origins of fluoride are discussed below.[20]

Natural sources

Typical natural sources of fluoride are soil, water, forage and grasses, and volcanic activity. Soil normally has a fluoride content that ranges from 150 mg/kg to 400 mg/kg[21,22], and this content varies in other natural sources based on alkalinity and temperature[23,24]. The level of fluoride in clay soil is 1000 mg/kg.[25] The contamination of soil with fluoride is a result of using phosphorus fertilisers which have 1–1.5% fluorine.[25] The toxicity of fluoride-contaminated soil comes after the inhalation of soil contaminants which have vaporised or through contaminated groundwater after the leaching of fluoride from adjacent fluoride-contaminated soil.[21] A concentration above 2.6 mg/L F is considered to be highly contaminated.[26] It was found that the level of fluoride in groundwater is higher than that in surface water due to percolation of fluoride from the soil to groundwater through a leaching process.[21] Some studies have also found grasses and forage that have higher levels of fluoride than those in industrialised areas.[21,23] Volcanic ash contains

a high level of fluoride and contamination of the geochemical cycle with fluoride occurs frequently.[21,27] Fluoride from a volcanic eruption may cover a wide area and remain for many years. After decaying and leaching, fluoride wreaks havoc on domestic and wild animals.[21]

Anthropogenic sources

Human activities that bring about anthropogenic fluoride contamination include the development of industries, the introduction of motor vehicles, the use of fluoride-containing pesticides, and the deliberate addition of fluoride to drinking water supplies, toothpaste and mouthwashes, refrigerants, and fire extinguishers. The average concentration of fluoride in normal areas (unpolluted/non-industrialised) is generally less than 0.1 $\mu g/m^3$.[21]

Fluoride levels at global scale

Around the world, 23 nations, including South Africa, are situated in the condemnatory region with regard to their fluoride levels.[21] A global indication of maximum fluoride levels in drinking water is shown in Supplementary table 1.

Human exposure

Water scarcity, rapid population growth, and unfavourable climate changes have led to more communities relying on drinking water that has excessive fluoride content.[28] Water, food and oral hygiene products are the main sources of fluoride exposure for human beings.[29,30]

Toxicity of fluoride

The World Health Organization (WHO) guidelines indicate 1.5 mg/L as the highest fluoride content that is permissible in drinking water.[31] Consumption of fluoride above the permitted limit of 1.5 mg/L is considered harmful to health and could result in dental fluorosis among other effects.[32-34] As stated by WHO[32], consumption of water containing higher fluoride levels, in the range of 3 mg/L to 6 mg/L, could account for skeletal fluorosis.[32,35,36]

Defluoridation techniques

This section focuses on various techniques and materials employed globally for water defluoridation. Defluoridating methods may loosely be classified into two categories: 1) additive methods and 2) adsorptive methods.[37] Various methods and materials, as listed in Table 1, have been used to defluoridate drinking water.

Adsorption

This technique involves the adsorption of fluoride ions onto the surface of an active agent. In the adsorption method, a bed of greater surface activity is chosen, and water is passed through the bed. Due to surface activity, fluoride ions get preferentially adsorbed onto the bed surface, thereby causing a reduction in fluoride ions in the exit stream.[39]

Ion exchange

This technique utilises synthetic chemicals. Anion and cation exchange resins are used to remove fluoride. These resins are commercially produced and hence they are expensive and not cost-effective in most circumstances.[37,42]

Precipitation and coagulation

Precipitation methods rely on the addition of a chemical precipitant or coagulant to transform dissolved, moderately soluble fluoride salts into insoluble fluorapatite. Sedimentation or filtration is then required for separation of the solids from the liquid, thereby removing the fluoride.[38]

Reverse osmosis

Reverse osmosis has emerged as the ideal method for water defluoridation, thus providing safe drinking water without presenting the challenges that are normally associated with other methods such as ion exchange resins, addition of chemicals to achieve coagulation and total reliance on electricity to sustain electrocoagulation.[43] Reverse osmosis uses a semi-permeable membrane that traps a significant array of contaminants but allows water molecules to pass through when water is pushed under pressure through the membrane.[39]

Nanofiltration membrane

Nanofiltration is reported to be the most recent improvement among all the membrane processes that are used for defluoridation of drinking water and wastewater. Properties of nanofiltration membrane cut across reverse osmosis and ultrafiltration.[44,45]

Electrocoagulation

Electrocoagulation is a process that uses an electrical charge to destabilise and aggregate contaminant particles, ions, and colloids to hold them in solution. Instead of expensive chemical reagents, the process removes heavy metals, suspended solids, emulsified organics and many other contaminants from water using electricity. It is a complex process that takes place through serial steps.[46-49]

The fluoride removal performance of these different defluoridation techniques is compared in Supplementary table 2.

Water defluoridation using bone char

Mechanism of defluoridation

Bone char is one of the most promising methods for the treatment of drinking water with excess fluoride concentration in developing countries.[25] This is mainly because it is rather inexpensive as it is produced from animal bones, can be synthesised in adequate amounts and to the desired quality[50], and can also be regenerated by simply reheating[51]. The ability of bone to defluoridate water was reported by Smith and Smith in 1937.[52] Several scholars and organisations have investigated defluoridation by bone char and the outcomes show

Table 1: Different defluoridation techniques and the materials employed[37-41]

Adsorption	Ion exchange	Precipitation and coagulation	Membrane	Electrocoagulation
Activated alumina	Anion exchange resins: NCL poly anion resin, Tulsion A27, Lewatit–MIH–59, Amberlite IRA–400, Deacedodite	Contact precipitation	Reverse osmosis	Electrolysis: aluminium electrode
Bone char	Cations	IISc Method	Electrodialysis	
Calcined clay		Nalgonda technique: lime, alum, lime and alum	Direct contact membrane distillation (DCMD)	
Mud pots				
Bio-adsorbents: tea ash, eggshell powder				

impressive efficiency.[53-55] Bone char consists largely of hydroxyapatite and a significant amount of calcium carbonate while carbon accounts only for 10% (w/w).[56] Scanning electron microscopy (SEM) and transmission electron microscopy (TEM) are used to assess, respectively, the outside and internal morphology of the prepared bone char, whereas the chemical composition is often determined by the application of energy dispersive X-ray spectroscopy (EDX).[57]

There are two basic ways to char bones for the production of bone char: (1) calcination, in which atmospheric air provides oxygen during the heating process, and (2) pyrolysis, in which the heating process takes place in an oxygen-deprived environment. According to previous investigations, bone char that is produced by calcination at temperatures higher than 600 °C inhibits the adsorption process[58,59], whereas bone char produced at temperatures below 400 °C may influence the taste and odour of the treated water.[60]

Water defluoridation by charred bone is achieved through ion exchange and adsorption that result in the replacement of the carbonate of the apatite in the bone char with the fluoride in the water.[38,58] This defluoridation technique involves more than one reaction, as shown in Equations 1 and 2, and the efficiency of bone char to remove fluoride from water depends on the level of fluoride in the raw water, pH, contact time and amount of char used (available surface area).[61]
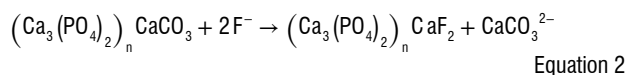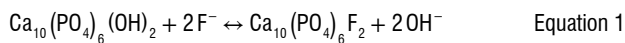
$$Ca_{10}(PO_4)_6(OH)_2 + 2F^- \leftrightarrow Ca_{10}(PO_4)_6F_2 + 2OH^- \qquad \text{Equation 1}$$

$$\left(Ca_3(PO_4)_2\right)_n CaCO_3 + 2F^- \rightarrow \left(Ca_3(PO_4)_2\right)_n CaF_2 + CaCO_3^{2-}$$
$$\text{Equation 2}$$

There is no universal water defluoridation method that perfectly meets all social, financial, economic, environmental and technical requirements. However, using charred bone as a medium for water defluoridation has a few advantages.[62] With this defluoridation approach, there is no daily addition of chemicals, working load and continuous power supply. High removal efficiency can be achieved. There is no demand for skilled operation as the defluoridation set-up is simple to construct. Construction materials are cheap and widely available. The method comes at a controlled risk of declination of the inherent quality of water, and it is a highly profitable technique that could remove a maximum of 66% of fluoride.[38,63]

Bones from wild animals

While extensive research has been done on defluoridation using bone char from bones obtained from domestic bovines, porcines, ovines, caprines, chickens, fish, camels, and other undeclared sources (Supplementary table 3), there is no mention of the use of bones of animals in the wild. The bones of wild animals, including birds, even those in captivity, are believed to have higher calcium content than the bones of domestic animals because their main diet is based on natural materials. Studies have found a relationship between meat protein and elevated levels of growth hormone, which in turn is connected to increased laying down of calcium and phosphorus, as well as other minerals, on the matrix of bone, thus minimising bone fractures.[64] Therefore, the use of such bones is envisaged to enhance defluoridation of water. On the other hand, the use of antibiotics in domestic and semi-wild animals to fight and prevent infections that lead to diseases and to promote growth in animals is likely to have some impact on bone composition, thus establishing yet another difference in the composition of bones from wild, semi-wild and production animals.[65] Rosol et al.[66] found that mithramycin, an antibiotic that has anti-tumour properties, inhibited the stimulation of bone resorption in dogs. Phenylbutazone, a non-steroidal compound that is commonly used for temporary treatment of pain, fever and stiffness in animals, was found to reduce the rate at which some minerals align in order to respond to defects in cortical bone in horses.[67]

Defluoridation in Africa

There is a relationship between volcanic activities, hot springs (particularly those that have higher pH), gases that are released from earth's crust, and igneous and metamorphic rocks.[68-71]

The world map in Amini et al.[72] illustrates the likelihood of fluoride concentrations that surpass 1.5 mg/L in groundwater. Figure 1 shows the African countries with elevated fluoride concentrations in water.

Various studies have demonstrated a relationship between most cases of dental fluorosis that occur in South Africa and the concentration of fluoride in groundwater that is intended for drinking[74-76], therefore rural communities are those mainly affected by dental fluorosis. Fluoride concentrations that exceed the WHO threshold of 1.5 mg/L have been reported in many areas and provinces in South Africa, such as the Free State, Limpopo, North-West and KwaZulu-Natal Provinces.[15,31,71]

Column and batch studies for defluoridation

Removal of fluoride from the solution is normally managed through a column[77,78] or batch set-up. In a column set-up, a vertical column that contains the prepared adsorbent is used, and the solution is run from the top through the adsorbent as pulled and guided by gravity. The treated solution is collected at the bottom of a column. In a batch set-up, both the prepared adsorbent and the solution are mixed and adsorption is allowed to take place over a controlled period of time with continuous stirring throughout the process. As a result of simultaneous mixing of the adsorbent and the solution, and the subsequent agitation of the mixture, the resultant solution requires separation that could be achieved by filtration, centrifugation or decantation.[79]

Adsorption isotherms are needed to better describe the interaction of solutes with the adsorbents. In line with the optimised use of the adsorbent, it is just as important to measure the rate of adsorption at constant concentration. Adsorption kinetics are important to assess the dispersal of the adsorbate at the pore level of the adsorbent. Batch adsorption experiments are required to carry out adsorption isotherms and kinetics tests.[25] The maximum capacity of the adsorbent under the set operating conditions is indicated by the quantity of fluoride ions adsorbed at the equilibrium time. Adsorption isotherm and kinetics experiments are key in optimising the use of adsorbents.[80] Ion chromatography is used to analyse the total fluoride in water.

Research gaps on carbonised bone in Africa

Sorbent prepared from charred bone is an affordable method of removing fluoride from groundwater, and bones from production animals are most extensively used in this regard. However, the issue of optimum particle size required for maximum fluoride removal has not been given adequate attention, particularly for bones from wild animals, including birds. It is generally accepted that the finer the particles the higher the uptake, because of the increase in surface area. However, very fine particles also have the problem of aggregation, which impedes adsorption. Therefore, determination of the optimum particle size to achieve maximum removal is essential.

It is also generally believed that the removal of fluoride is the direct reaction of calcium compound(s) in the bone with fluoride in the water to form calcium fluoride compounds as indicated in the equations above. However, it is still not clear whether the driving mechanism is that of mesoporosity or microporosity. It is important to know the driving mechanism as this will assist when considering the recovery of fluoride from the adsorbent. A detailed morphology of bone char will heighten our understanding of the structural arrangement of the bone(s) under study, and hence of the driving process(es). Furthermore, information on the defluoridation of water using various types of bone char prepared from the bones of wild animals is still scarce. The use of antibiotics to treat and control clinical diseases in domestic and semi-wild animals and to enhance growth in production animals is likely to have an impact on bone composition, thus establishing yet another difference in the composition of bones from wild, semi-wild and production animals.

Studies on defluoridation of water using bone char have so far used batch and column models. These models have been tested and are reliable. However, the scaling down of the use of bone char to a domestic level requires a cheaper and less elaborate process and packaging. Packaging of bone char in the form of a tea bag can be very useful at domestic levels, particularly for low-resourced rural communities that are commonly found in Africa and some parts of Asia and South America.
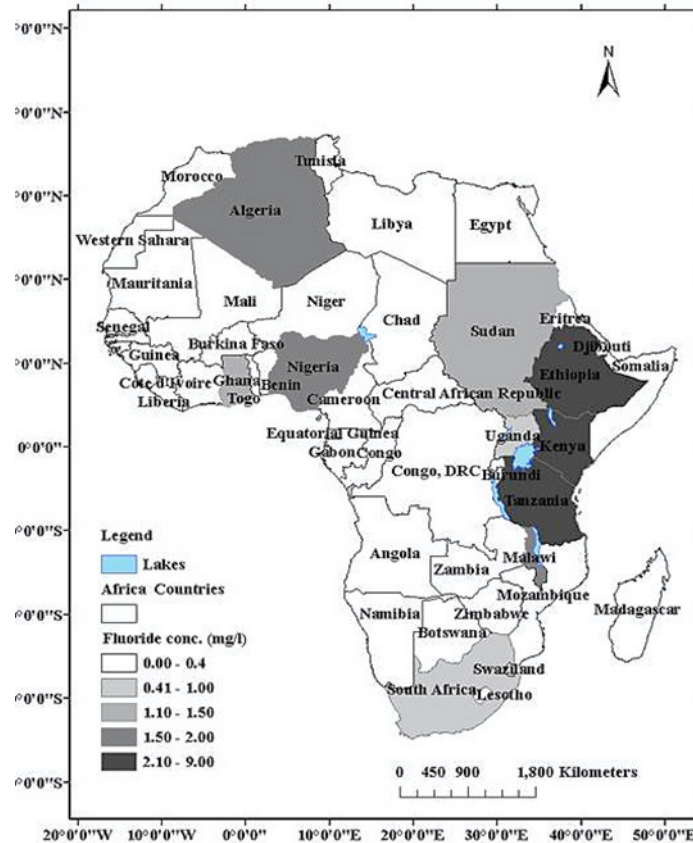
Source: Malago et al.[73] (under CC-BY-4.0 licence)

Figure 1: Map of Africa illustrating the countries with high fluoride concentrations in their water.

Conclusions

This review highlights the scarcity of water, the subsequent reliance on groundwater as an alternative source of potable water in some nations and communities, the contamination of groundwater with fluoride, fluoride toxicity, and different water defluoridation techniques. When all methods are compared, there is no universal water defluoridation method that perfectly meets all social, financial, economic, environmental, and technical requirements. However, using bone char as an adsorbent in fluoride removal has a few advantages. With this defluoridation approach, there is no daily addition of chemicals, working load and continuous power supply. There is no demand for skilled operation as the defluoridation set-up requires minimal effort to construct. The materials required for construction are affordable and abundantly available. There is a low risk of the original water declining with this method. Adsorption using bone char is a useful technique that has the capacity to remove up to 66% of fluoride. Ingestion of fluoride affects human health in various significant ways, and can negatively impact quality of life. Whilst fluorosis is not deadly, moderate fluorosis causes the deterioration of dental aesthetics and severe fluorosis can lead to disorders of the skeletal system. In the absence of any practical cure, prevention of fluorosis remains the only solution. The key and immediate preventive measure is to consume water that has an acceptable fluoride level. This can be achieved by treating drinking water that is contaminated with fluoride to significantly lower the fluoride level. That being the case, water purification techniques should be further investigated, and new ones developed to ultimately achieve a method that is both safe and inexpensive.

Acknowledgements

Competing interests

We have no competing interests to declare.

Authors' contributions

S.P.M.: Writing – initial draft; project management; data collection. J.O.O.: Conceptualisation; student supervision; writing – revisions. R.J.: Student supervision.

References

1. Nagendra Rao C, editor. Fluoride and environmental review. In: Bunch MJ, Suresh VM, Kumaran TV, editors. Proceedings of the Third International Conference on Environment and Health; 2003 December 15–17; Chennai, India. Chennai: Department of Geography, University of Madras and Faculty of Environmental Studies, York University; 2003. p. 386–399.

2. Tewari A, Dubey A. Defluoridation of drinking water: Efficacy and need. J Chem Pharm Res. 2009;1(1):31–37.

3. Baker B, Aldridge C, Omer A. Water: Availability and use. Publication 3011 (POD-04-23). Mississippi State University Extension; 2016.

4. Kumari S, Khan S. Defluoridation technology for drinking water and tea by green synthesized Fe(3)O(4)/Al(2)O(3) nanoparticles coated polyurethane foams for rural communities. Sci Rep. 2017;7(1):8070. https://doi.org/10.1038/s41598017085947

5. Yihunu EW, Yu H, Junhe W, Kai Z, Teffera ZL, Weldegebrial B, et al. A comparative study on defluoridation capabilities of biosorbents: Isotherm, kinetics, thermodynamics, cost estimation and regeneration study. Environ Eng Res. 2020;25(3):384–392. https://doi.org/10.4491/eer.2019.097

6. Schmoll O, Howard G, Chilton J, Chorus I. Protecting groundwater for health: Managing the quality of drinking water sources. London: IWA Publishing; 2006. p. 697.

7. Ayoob S, Gupta A, Bhat VT. A conceptual overview on sustainable technologies for the defluoridation of drinking water. Crit Rev Environ Sci Technol. 2008;38(6):401–470. https://doi.org/10.1080/10643380701413310

8. Sankannavar R, Chaudhari S. An imperative approach for fluorosis mitigation: Amending aqueous calcium to suppress hydroxyapatite dissolution in defluoridation. J Environ Manage. 2019;245:230–237. https://doi.org/10.1016/j.jenvman.2019.05.088

9. Susheela A. Fluorosis in developing countries: Remedial measures and approaches. Proc Indian Natl Sci Acad. 2002;68(5):389–400. https://doi.org/10.1007/978-1-4615-0893-9_20

10. Mumtaz N, Pandey G, Labhasetwar PK. Global fluoride occurrence, available technologies for fluoride removal, and electrolytic defluoridation: A review. Crit Rev Environ Sci Technol. 2015;45(21):2357–2389. https://doi.org/10.1080/10643389.2015.1025638

11. Blignaut J, Van Heerden J. The impact of water scarcity on economic development initiatives. Water SA. 2009;35(4):415–420. https://doi.org/10.4314/wsa.v35i4.76800

12. Donnenfeld Z, Hedden S, Crookes C. A delicate balance: Water scarcity in South Africa. ISS South Afr Rep. 2018;2018(13):1–24

13. Brindha K, Elango L. Fluoride in groundwater: Causes, implications and mitigation measures. In: Fluoride properties, applications and environmental management. New York: Nova Publishers; 2011; 1. p. 111–136.

14. Rafique T, Naseem S, Ozsvath D, Hussain R, Bhanger MI, Usmani TH. Geochemical controls of high fluoride groundwater in Umarkot sub-district, Thar Desert, Pakistan. Sci Total Environ. 2015;530:271–278. https://doi.org/10.1016/j.scitotenv.2015.05.038

15. Ncube E, Schutte C. The occurrence of fluoride in South African groundwater: A water quality and health problem. Water SA. 2005;31(1):35–40. https://doi.org/10.4314/wsa.v31i1.5118

16. Keri RS, Hosamani KM, Reddy HRS, Nataraj SK, Aminabhavi TM. Application of the electrodialytic pilot plant for fluoride removal. J Water Chem Technol. 2011;33(5):293–300. https://doi.org/10.3103/S1063455X11050043

17. Linhares DPS, Garcia PV, dos Santos Rodrigues A. Fluoride in volcanic areas: A case study in medical geology. In: Makan A, editor. Environmental health-management and prevention practices. London: IntechOpen; 2019. https://doi.org/10.5772/intechopen.86058

18. Alkurdi SSA, Al-Juboori RA, Bundschuh J, Hamawand I. Bone char as a green sorbent for removing health threatening fluoride from drinking water. Environ Int. 2019;127:704–719. https://doi.org/10.1016/j.envint.2019.03.065

19. Tripathy SS, Bersillon J-L, Gopal K. Removal of fluoride from drinking water by adsorption onto alum-impregnated activated alumina. Sep Purif Technol. 2006;50(3):310–317.

20. Ali S, Thakur SK, Sarkar A, Shekhar S. Worldwide contamination of water by fluoride. Environ Chem Lett. 2016;14(3):291–315. https://doi.org/10.1007/s10311016-0563-5

21. Bharti VK, Giri A, Kumar K. Fluoride sources, toxicity and its amelioration: A review. Ann Environ Sci Toxicol. 2017;2(1):021–032. https://doi.org/10.17352/aest.000009

22. Brindha K, Rajesh R, Murugan R, Elango L. Fluoride contamination in groundwater in parts of Nalgonda District, Andhra Pradesh, India. Environ Monit Assess. 2011;172(1):481–492. https://doi.org/10.1007/s10661-010-1348-0

23. Zhu H, Wang H, Wang G, Zhang K, editors.Removal of fluorine from water by the aluminum-modified bone char. International Conference on Biology, Environment and Chemistry; Hangzhou, China. IEEE; 2011. p. 6103–6105. http://dx.doi.org/10.1109/ICISE.2010.5690593

24. Radostits O, Gay C, Blood D, Hinchcliff K. A textbook of the diseases of cattle, sheep, pigs, goats and horses. Vet Sci. 2000;9:603–700

25. Rojas-Mayorga C, Bonilla-Petriciolet A, Aguayo-Villarreal IA, Hernandez-Montoya V, Moreno-Virgen M, Tovar-Gómez R, et al. Optimization of pyrolysis conditions and adsorption properties of bone char for fluoride removal from water. J Anal Appl Pyrolysis. 2013;104:10–18. https://doi.org/10.1016/j.jaap.2013.09.018

26. Rugayah N, Muladno, Nuraini H, Salundik. Chicken bone charcoal for defluoridation of groundwater in Indonesia. Int J Poult Sci. 2014;13:591–596. https://doi.org/10.3923/ijps.2014.591.596

27. D'Alessandro W. Human fluorosis related to volcanic activity: A review. Environ Toxicol. 2006;1:21–30. https://doi.org/10.2495/ETOX060031

28. Wambu EW, Agong SG, Anyango B, Akuno W, Akenga T. High fluoride water in Bondo-Rarieda area of Siaya County, Kenya: A hydro-geological implication on public health in the Lake Victoria Basin. BMC Public Health. 2014;14(1):462. https://doi.org/10.1186/1471-2458-14-462

29. Guth S, Hüser S, Roth A, Degen G, Diel P, Edlund K, et al. Toxicity of fluoride: Critical evaluation of evidence for human developmental neurotoxicity in epidemiological studies, animal experiments and in vitro analyses. Arch Toxicol. 2020;94(5):1375–1415. https://doi.org/10.1007/s00204-020-02725-2

30. EFSA Panel on Dietetic Products, Nutrition and Allergies. Scientific opinion on dietary reference values for fluoride. EFSA J. 2013;11(8):3332. https://doi.org/10.2903/j.efsa.2013.3332

31. World Health Organization (WHO). Guidelines for drinking-water quality. WHO Chron. 2011;38(4):104–108

32. World Health Organization (WHO). Fluoride in drinking water. Background document for development of WHO guidelines for drinking-water quality. Geneva: World Health Organization; 2004.

33. South African Department of Water Affairs and Forestry (DWAF). South African water quality guidelines. Pretoria: DWAF; 1996.

34. Durowoju O, Odiyo J, Ekosse G-I. Hydrogeochemical setting of geothermal springs in Limpopo Province, South Africa. Res J Chem Environ. 2015;19(1):77–88. https://doi.org/10.3390/app9081688

35. Ologundudu TO, Durowoju OS, Odiyo JO, Ekosse G-IE. Defluoridation of groundwater from Siloam Village, Limpopo Province, South Africa, using vermiculite modified with hexadecyltrimethylammonium. Cogent Eng. 2020;7(1):1795050. https://doi.org/10.1080/23311916.2020.1795050

36. Mohammadi AA, Yousefi M, Yaseri M, Jalilzadeh M, Mahvi AH. Skeletal fluorosis in relation to drinking water in rural areas of West Azerbaijan, Iran. Sci Rep. 2017;7(1):17300. https://doi.org/10.1038/s41598-017-17328-8

37. Ingle NA, Dubey HV, Kaur N, Sharma I. Defluoridation techniques: Which one to choose. J Health Res Rev. 2014;1(1):1–4. https://doi.org/10.4103/2394-2010.143315

38. Piddennavar R, Krishnappa P. Review on defluoridation techniques of water. Int J Eng Sci. 2013;2(3):86–94.

39. Khairnar MR, Dodamani AS, Jadhav HC, Naik RG, Deshmukh MA. Mitigation of fluorosis – a review. J Clin Diagn Res. 2015;9(6):Ze05–Ze09. https://doi.org/10.7860/JCDR/2015/13261.6085

40. Boubakri A, Bouchrit R, Hafiane A, Bouguecha SA-T. Fluoride removal from aqueous solution by direct contact membrane distillation: Theoretical and experimental studies. Environ Sci Pollut Res. 2014;21(17):10493–10501. https://doi.org/10.1007/s11356-014-2858-z

41. Karunanithi M, Agarwal R, Qanungo K. A review of fluoride removal from groundwater. Period Polytech Chem Eng. 2019;63(3):425–437. https://doi.org/10.3311/PPch.12076

42. Chibi C, Haarhoff J. A promising approach to fluoride removal in rural drinking water supplies. In: Proceedings of WISA Biennial Conference; 2000 May 31–June 01; Sun City, South Africa. Johannesburg: WISA; 2000.

43. Hu K, Dickson JM. Nanofiltration membrane performance on fluoride removal from water. J Membr Sci. 2006;279(1–2):529–538. https://doi.org/10.1016/j.memsci.2005.12.047

44. Waghmare SS, Arfin T. Fluoride removal from water by various techniques. Int J Innov Sci Eng Technol. 2015;2:560–571.

45. Tahaikt M, El Habbani R, Haddou AA, Achary I, Amor Z, Taky M, et al. Fluoride removal from groundwater by nanofiltration. Desalination. 2007;212(1–3):46–53. https://doi.org/10.1016/j.desal.2006.10.003

46. Bayramoglu M, Kobya M, Can OT, Sozbir M. Operating cost analysis of electrocoagulation of textile dye wastewater. Sep Purif Technol. 2004;37(2):117–125. https://doi.org/10.1016/j.seppur.2003.09.002

47. Takdastan A, Emami Tabar S, Neisi A, Eslami A. Fluoride removal from drinking water by electrocoagulation using iron and aluminum electrodes. Jundishapur J Health Sci. 2014;6(3):e21718. https://doi.org/10.5812/jjhs.21718

48. Mumtaz N, Pandey G, Labhasetwar PK. Assessment of electrolytic process for water defluoridation. J Issues. 2014;1(9):175–182.

49. Mumtaz N, Pandey G, Labhasetwar P, Andey S. Evaluation of operational parameters involved in electrolytic defluoridation process. Int J Civil Struct Environ Infrastruct Eng Res Dev. 2012;2(4):23–32.

50. Kariuki S, Ngari M, Mavura W, Ollengo M, Ongoma P. Effect of essential mineral ions from aqueous media on adsorption of fluoride by bone char. J Environ Sci Toxicol Food Technol. 2015;9:9–17.

51. Rezayee A, Ramin M, Gh G, Valipour F. Designing of bioaerosol production system for removing *Escherichia coli* from contaminated air using bone char. J Mil Med. 2011;13(2):89–95.

52. Dahi E. Africa's U-Turn in defluoridation policy: From the Nalgonda technique to bone char. Fluoride. 2016;49(4):401–416.

53. Jacobsen P, Dahi E, editors. Bone char based bucket defluoridator in Tanzanian households. In: 2nd International Workshop on Fluorosis and Defluoridation of Water; 1997 November 19–25; Ethiopia. Nazreth:; The International Society for Fluoride Research; 1997. p. 156–159.

54. Mavura W, Mwanyika F, Wrensford G. Construction and optimisation of a cartridge filter for removing fluoride in drinking water. Bull Chem Soc Ethiop. 2004;18(1). https://doi.org/10.4314/bcse.v18i1.61630

55. Watanesk R, Watanesk S, editors. Comparative study of fluoride sorption behaviour on activated carbon and bone char. In: 3rd International Workshop on Fluorosis Prevention and Defluoridation of Water; 2000 November 20–24; Thailand. Chiang Mai: ISFT, EnDeCo, & ICOH.; 2000. p. 80.

56. Kanyora A, Kinyanjui T, Kariuki S, Chepkwony C. Efficiency of various sodium solutions in regeneration of fluoride saturated bone char for de-fluoridation. IOSR J Environ Sci Toxicol Food Technol. 2014;8:10–16. https://doi.org/10.9790/2402-081031016

57. Terasaka S, Kamitakahara M, Yokoi T, Ioku K. Effect of preparation temperature on the ability of bone char to remove fluoride ion and organic contaminants. J Ceram Soc Jpn. 2014;122(1432):995–999. https://doi.org/10.2109/jcersj2.122.995

58. Kaseva ME. Optimization of regenerated bone char for fluoride removal in drinking water: A case study in Tanzania. J Water Health. 2006;4(1):139–147. https://doi.org/10.2166/wh.2006.0011

59. Albertus J, Bregnhøj H, Kongpun M. Bone char quality and defluoridation capacity in contact precipitation. In: 3rd International Workshop on Fluorosis Prevention and Defluoridation of Water; 2000 November 20; Thailand. Chiang Mai: ISFR, EnDeCo & ICOH; 2002. p. 61–72.

60. Mutchimadilok Y, Smittakorn S, Mongkolnchai-arunya S, Durnford D. Defluoridation with locally produced Thai bone char. Adv Environ Chem. 2014;2014:483609. https://doi.org/10.1155/2014/483609

61. Bregnhøj H, Dahi E, Jensen M, editors. Modelling defluoridation of water in bone char columns. In: Proceedings of the 1st International Workshop on Fluorosis and Defluoridation of Water; 1995 October 18–21; Ngurdoto, Tanzania. The International Society for Fluoride Research; 1995. p. 18–22.

62. Feenstra L, Vasak L, Griffioen J. Fluoride in groundwater: Overview and evaluation of removal methods. Report No.: SP 2007–1. Utrecht: International Groundwater Resources Assessment Centre; 2007.

63. Mobeen N, Kumar P. Defluoridation techniques – a critical review. Asian J Pharm Clin Res. 2017;10(6):64–71. https://doi.org/10.22159/ajpcr.2017.v10i6.13942

64. Heaney RP, Layman DK. Amount and type of protein influences bone health. Am J Clin Nutr. 2008;87(5):1567S–1570S. https://doi.org/10.1093/ajcn/87.5.1567S

65. Landers TF, Cohen B, Wittum TE, Larson EL. A review of antibiotic use in food animals: Perspective, policy, and potential. Pub Health Rep. 2012;127(1):4–22. https://doi.org/10.1177/003335491212700103

66. Rosol T, Chew D, Couto C, Ayl R, Nagode L, Capen C. Effects of mithramycin on calcium metabolism and bone in dogs. Vet Pathol. 1992;29(3):223–229. https://doi.org/10.1177/030098589202900306

67. Rohde C, Anderson DE, Bertone AL, Weisbrode SE. Effects of phenylbutazone on bone activity and formation in horses. Am J Vet Res. 2000;61(5):537–543. https://doi.org/10.2460/ajvr.2000.61.537

68. AbuZeid K, Elhatow L, editors. Impact of fluoride content in drinking water. Cairo: Arab Water Healthy Conference Egypt; 2007.

69. Brunt R, Vasak L, Griffioen J. Fluoride in groundwater: Probability of occurrence of excessive concentration on global scale. The Netherlands: Igrac; 2004.

70. Fawell J, Bailey K, Chilton J, Dahi E, Magara Y. Fluoride in drinking-water. London: IWA Publishing; 2006. p. 144

71. Thole B. Ground water contamination with fluoride and potential fluoride removal technologies for East and Southern Africa. Perspectives in Water Pollution. Blantyre: INTECH Open Science; 2013. p. 66–90. https://doi.org/10.5772/54985

72. Amini M, Mueller K, Abbaspour KC, Rosenberg T, Afyuni M, Møller KN, et al. Statistical modeling of global geogenic fluoride contamination in groundwaters. Environ Sci Technol. 2008;42(10):3662–3668. https://doi.org/10.1021/es071958y

73. Malago J, Makoba E, Muzuka AN. Fluoride levels in surface and groundwater in Africa: A review. Am J Water Sci Eng. 2017;3(1):1–17. https://doi.org/10.11648/j.ajwse.20170301.11

74. McCaffrey L, Willis J. Distribution of fluoride-rich groundwater in the eastern and Mogwase regions of the Northern and North-West Provinces. Pretoria: Water Research Commission; 2001.

75. Fayazi M. Regional groundwater investigation on the Northern Springbok flats. GH Report No.: 3684. Pretoria: Department of Water Affairs and Forestry; 1994.

76. Du Plessis JB. What would be the maximum concentration of fluoride in water that would not cause dental fluorosis? Fluoride and fluorosis. The status of South African Research, North West Province. 1995;4.

77. Gogoi S, Dutta RK. Column evaluation of a water defluoridation technique based on phosphoric acid-enhanced limestone adsorption. Indian J Chem Technol. 2017;24(4):374–382. http://op.niscpr.res.in/index.php/IJCT/article/view/11881/0

78. Teusner A, Butler R, Le Clech P. Impact of pretreatment on defluoridation of drinking water by bone char adsorption. J Humanit Eng. 2016;4(1). https://doi.org/10.36479/jhe.v4i1.48

79. Loebenstein WV. Batch adsorption from solution. J Res Natl Bur Stand A Phys Chem A. 1962;66:503–515. https://doi.org/10.6028/jres.066A.052

80. Alagumuthu G, Veeraputhiran V, Venkataraman R. Adsorption isotherms on fluoride removal: Batch techniques. Arch Appl Sci Res. 2010;2(4):170–185.

AUTHORS:
Kojo S. Acquah[1,2]
David W. Gammon[1]
Denzil R. Beukes[3]

AFFILIATIONS:
[1]Department of Chemistry, University of Cape Town, Cape Town, South Africa
[2]Division of Medicinal Chemistry, Department of Pharmaceutical Sciences, University of Connecticut, Storrs, Connecticut, USA
[3]School of Pharmacy, University of the Western Cape, Cape Town, South Africa

CORRESPONDENCE TO:
Denzil Beukes

EMAIL:
dbeukes@uwc.ac.za

South African actinobacteria: A treasure trove of novel bioactive metabolites for drug discovery

Although South Africa is known as one of the most biodiverse countries in the world, based on its unique plants and animals, microorganisms have received much less attention. Microorganisms in general and actinobacteria in particular are an underexplored source of new medicines. Recent studies have demonstrated the presence of diverse cultivable actinobacteria from various biomes. However, investigations of the natural product diversity associated with these microorganisms are lacking. We hereby present a review of natural products isolated from South African actinobacteria together with their biological activities. Many of these natural products are structurally novel and include compounds belonging to the following classes: anthraquinones, isoflavonoids, ketolides, macrolides, macrolactams, tripeptides and depsipeptides. They show a wide range of biological activities including antibacterial, antifungal, cytotoxic and antitumour activities.

Significance:

- This review highlights the importance of actinobacteria in the discovery of new medicines and summarises the state-of-the-art on their research in South Africa.

- We reveal a gap in the exploitation of this resource and emphasise the opportunities for multidisciplinary research.

Introduction

Natural products from plants, invertebrates and microorganisms have played an important role in the development of new medicines and agrochemicals.[1] Microbial natural products, in particular, offer significant advantages over natural products produced by macroorganisms. These advantages include a reduced impact on the environment (ecosystems), and hence reduced competition with food crops for arable land, as well as relative ease of production and manipulation of biosynthetic pathways to produce novel compounds for commercial exploitation.

As one of the most biodiverse countries in the world[2], South Africa has a rich tradition of natural products research.[3] However, the main focus of these endeavours has been on plant natural products, and more recently marine natural products[4,5], with much less attention on microbial natural products.

One of the most important sources of medicinally important natural products are the actinobacteria. Actinobacteria, also known as actinomycetes, are filamentous Gram-positive bacteria with high guanine and cytosine (G+C) content in their DNA. The phylum Actinobacteria comprises 20 orders and more than 50 families.[6] This bacterial phylum is widely distributed in terrestrial and both fresh and marine aquatic environments. Some can thrive under extreme conditions such as hyperaridity, high salinity, cold, high pressure, low pH (acidic) and heavy metal contaminated ecosystems.[7] Actinobacteria are either free living, such as soil-dwelling bacteria, or living in association with other organisms, like the plant commensals or those living in and/or on the surfaces of animals like ants, termites and marine invertebrates.[8-10] Some actinobacteria are also plant and animal pathogens[11], while others find use in agriculture, biotechnology, and medicine. In agriculture, they are saprophytic and break down dead plant and animal remains, hastening decomposition.[8] They also aid in nitrogen fixation and are known to produce plant growth promoters, insecticides, herbicides and fungicides.[9] Natural products produced by actinobacteria are structurally diverse and have shown diverse biological activities, including antioxidant, antimalarial, anthelmintic, antifungal, enzyme inhibitory, antibacterial, anticancer, immunosuppressive and cardiovascular properties.[9,12,13] Antibiotics are the largest class of drugs discovered from actinomycetes, as they produce about 70% of all naturally derived antibiotics currently in clinical use.[9] Most of these antibiotics were discovered during the 'golden era' of antibiotic drug discovery and include the aminoglycosides, β-lactams, glycopeptides, macrolides, rifamycins and tetracyclines.[12]

In this review, we discuss the natural products produced by actinobacterial strains isolated from South African environments. The many excellent studies focusing only on the distribution and description of new actinobacteria species, as well as those only reporting on the biological activity of crude extracts and the discovery of enzymes fall outside the scope of this review.

Natural products from South African actinobacteria

Several novel actinomycete strains have been isolated from South African soils, flora and fauna, in both terrestrial and marine environments, and have been shown to contain bio-active secondary metabolites (Table 1). These strains include species of the ubiquitous *Streptomyces* genus and the less isolated rare genera *Actinomadura*, *Actinosynnema*, *Amycolatopsis*, *Gordonia*, *Kribbella*, *Nocardia*, *Nonomuraea*, *Rhodococcus*, *Streptosporangium*, *Saccharopolyspora* and *Tsukamurella*.[14-22]

Streptomyces

The first report of a South African actinomycete-derived secondary metabolite was a tetraene macrolide natamycin **1** (Figure 1) (also known as pimaricin, natacyn, tennecetin and E235) which was patented in 1955 for its antifungal activity.[23,24] This antibiotic was first purified in 1955 from the extract of the culture broth of *Streptomyces natalensis*,

Table 1: Secondary metabolites from South African actinomycetes and their biological activities. Structures of compounds **1–122** are shown in the text.

Compounds	Name	Structural class	Bioactivity	Producing species	Discovery tool	Reference
1	Natamycin	Tetraene macrolide	Antifungal	*Streptomyces natalensis*		23-32
2–10	Altromycins	New anthraquinone pluramycin type	Antibacterial, antitumour and cytotoxicity	Strain AB 1246E-26	Antibiotic screening	33-37
11–51	Platensimycin, platencin, their analogues and unrelated compounds	New ketolides	Antibacterial	*Streptomyces platensis* strain MA7327	Antisense screening	38-48
52–60	Natalamycin A, reblastatin, geldanamycin and its derivatives	Ansamycin macrolides	Antifungal	*Streptomyces* strain M56		49
61	17-Hydroxycyclooctatin		Cytotoxicity	*Streptomyces* sp. M56	Liquid chromatography–mass spectrometry (LCMS) based	50
62–75	Termisoflavones A–D and other isoflavonoids	Isoflavonoids		*Streptomyces* sp. RB1	Antibiotic screening	51,52
76	1-*O*-(2-aminobenzoyl)-α-L-rhamnopyranoside (ABR)		Cytotoxicity	*Streptomyces* sp. RB1	Bioactivity	53
77–80	Dentigerumycin, Dentigerumycin C–D	Cyclic depsipeptides	Antifungal	*Streptomyces* sp. M41	Bioactivity and LCMS based	54
81–83	Krisynomycins	Cyclic nonribosomal peptides	Antibacterial	*Streptomyces canus* strain CA-091830	Bioactivity and LCMS based	55,56
84–89	Rubterolones A–F	Tropolone alkaloids	Antifungal	*Actinomadura* sp. 5-2	Bioactivity and LCMS based	57
90–92	Natalenamides A–C	Cyclic tripeptides	Cytotoxicity	*Actinomadura* sp. RB99	LCMS based	58
93–107	Polychlorinated and polybrominated analogues of daidzein and genistein	Isoflavonoids	Antibacterial and antifungal	*Actinomadura* sp. RB99	LCMS based	59
108	Fridamycin A	Type II polyketide synthase-derived	Antibiotic and antitumour	*Actinomadura* sp. RB99	LCMS based	60
109–112	Macrotermycins A–D	Glycosylated polyketide macrolactams	Antifungal	*Amycolatopsis* sp. M39	Bioactivity and LCMS based	61
113–122	Speibonoxamine, dehydroxylated desferrioxamine analogues and diketopiperazines	Desferrioxamine		*Kribbella speibonae* strain SK5	Bioactivity and LCMS based	62,63
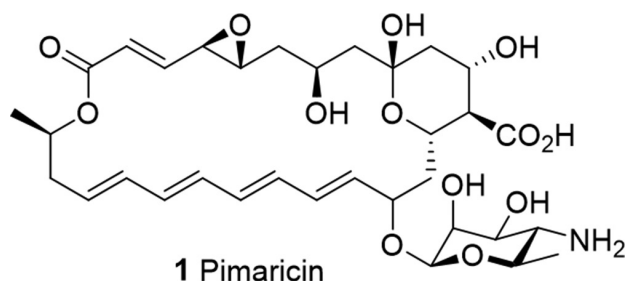
**Figure 1:** Natamycin **1** from *Streptomyces natalensis*.

isolated from a soil sample collected from Pietermaritzburg (giving rise to its original name, pimaricin) in the KwaZulu-Natal Province, South Africa.[24] Natamycin **1** was later also isolated from *Streptomyces gilvosporeus, Streptomyces chattanogenesis* and *Streptomyces lydicus*.[25-27] It is active against a variety of saprophytic and parasitic fungi and is therefore used commercially as a preservative.[28] It acts by binding to ergosterol in the fungal cell wall, thereby inhibiting fungal growth.[29,30] Natamycin **1** therefore has a wide spectrum of antifungal activity and minimal toxicity to mammalian cells. It also displays in vitro activity against numerous protozoa including *Trypanosoma* and *Acanthamoeba*.[31,32] Clinically, natamycin **1** is used to treat keratitis, and especially that caused by *Aspergillus fumigatus, Candida albicans* and *Acanthamoeba* sp.[28,32] It is also used to treat fungal infections caused by *Cephalosporium, Fusarium* and *Penicillium*, and has shown activity against *Alternaria, Colletotrichum, Curvularia, Lasiodiplodia, Scedosporium* and *Trichophyton*.[28] It is a commercial food additive which has been used for about half a century to prevent fungal growth on foods such as cheese, sausages, yoghurt, fruits, meats, baked confectioneries and beverages.[28]

A soil microbe antibiotic screening programme led to the isolation of actinomycete strain AB 1246E-26 from South African bushveld soil.[33] Although the genus of this actinomycete strain was not determined, preliminary characterisation narrowed the taxonomic assignment to either *Nocardia* or the defunct *Micropolyspora*.[33] Strain AB 1246E-26 showed activity against the antibiotic-sensitive strain of *Pseudomonas aeruginosa* K799/61 among other *P. aeruginosa* strains.[33,34] The organic extract of the whole fermentation broth of strain AB 1246E-26 was

subjected to a bioactivity guided isolation protocol to yield the novel anthraquinone-derived class of antibiotics called altromycins.[34,35] The altromycins A-I **2–10** (Figure 2) are nine closely related members of the pluramycin class of compounds with a single epoxide substituent, an amino-disaccharide and/or a 6-deoxy-3-O-methylaltrose attached to the conjugated ring systems of an anthraquinone-γ-pyrone core.[34-36] Altromycin B **3** was screened against 30 bacterial strains including *Staphylococcus aureus, Staphylococcus epidermidis, Micrococcus luteus, Enterococcus hirae, Streptococcus bovis, Streptococcus agalactiae, Streptococcus pyogenes, Escherichia coli, Enterobacter aerogenes, Klebsiella pneumoniae, Providencia stuartii, Pseudomonas aeruginosa, Pseudomonas cepacia* and *Acinetobacter* sp. clinical isolates.[33] Altromycin B **3** exhibited potent antibacterial activity against the clinical isolates of *Staphylococcus* and *Streptococcus* with a minimum inhibitory concentration (MIC) range of 0.39–3.12 μg/mL and 0.2–3.12 μg/mL, respectively, but displayed moderate to weak antibacterial activity against Gram-negative bacteria with an MIC range of 25 to >100 μg/mL.[33] The altromycins also showed cytotoxic activity against various cancer and tumour cell lines including cervical cancer (HeLa), human lung cancer (A549), colon tumour (HCT-8), murine leukemia cell (P388) lines and ovarian sarcoma (M5076).[37]

In their search for antibiotics that inhibit fatty acid biosynthesis in bacteria, researchers at Merck discovered the novel broad-spectrum antibiotic platensimycin **11** (Figure 3) from *Streptomyces platensis* strain MA7327, which was originally isolated from a soil sample collected in the Eastern Cape Province of South Africa.[38] An organic extract of the fermentation broth of *S. platensis* strain MA7327 was subjected to a unique antisense differential sensitivity whole-cell two-plate agar diffusion bioassay-guided fractionation process to yield platensimycin **11**.[38] Platensimycin **11** consists of a 3-amino-2,4-dihydroxybenzoic acid tethered via an amide bond to a C-17 tetracyclic enone which includes a bridge-head oxygen.[38,39] Another closely related compound, platencin **12** (Figure 3), was produced by *Streptomyces platensis* strain MA7339 using the same bioassay-guided fractionation procedure, although its biosynthetic gene cluster was also identified in strain MA7327.[40] Several other analogues **13–49** (Figure 3) – with modifications on or loss of the aromatic ring, modifications on the terpenoid and anilide moieties and a change in the length of the enone acid portion of platensimycin and platencin – have also been isolated from strain MA7327.[41-48] Compounds **50** and **51** (Figure 3), which are structurally different from the platensimycin group of compounds, were also isolated from strain MA7327.[41,42] Furthermore, other glycosylated analogues of platensimycin **11** and platencin **12** have
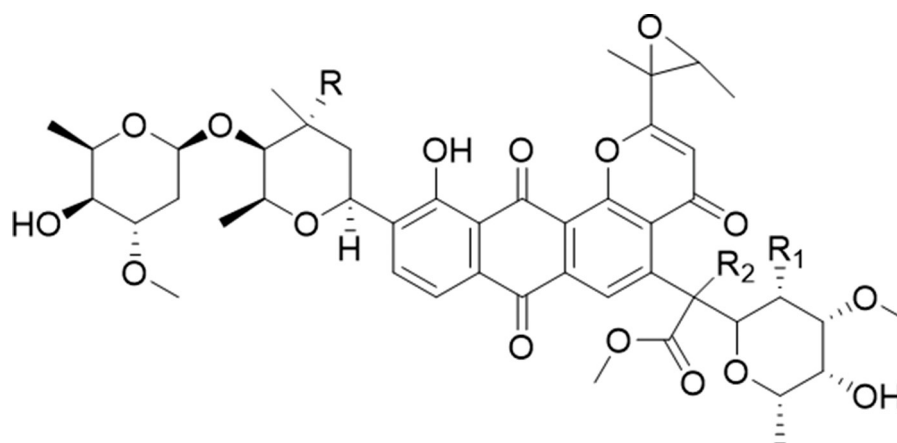
2 R = NHMe, R₁ = R₂ = OH, Altromycin A
3 R = N(Me)₂, R₁ = R₂ = OH, Altromycin B
4 R = NHMe, R₁ = H, R₂ = OH, Altromycin C
5 R = N(Me)₂, R₁ = H, R₂ = OH, Altromycin D
6 R = NHMe, R₁ = OH, R₂ = H, Altromycin E
7 R = N(Me)₂, R₁ = OH, R₂ = H, Altromycin F
8 R = NH₂, R₁ = OH, R₂ = OH, Altromycin G
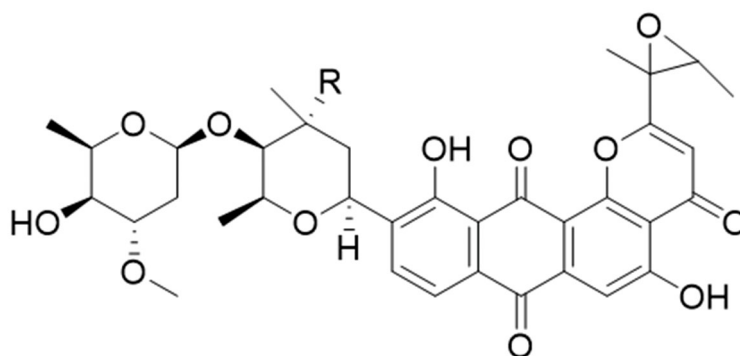
9 R = NHMe, Altromycin H
10 R = N(Me)₂, Altromycin I

Figure 2:   Altromycins A-I **2–10** from South African actinomycete strain AB 1246E-26.
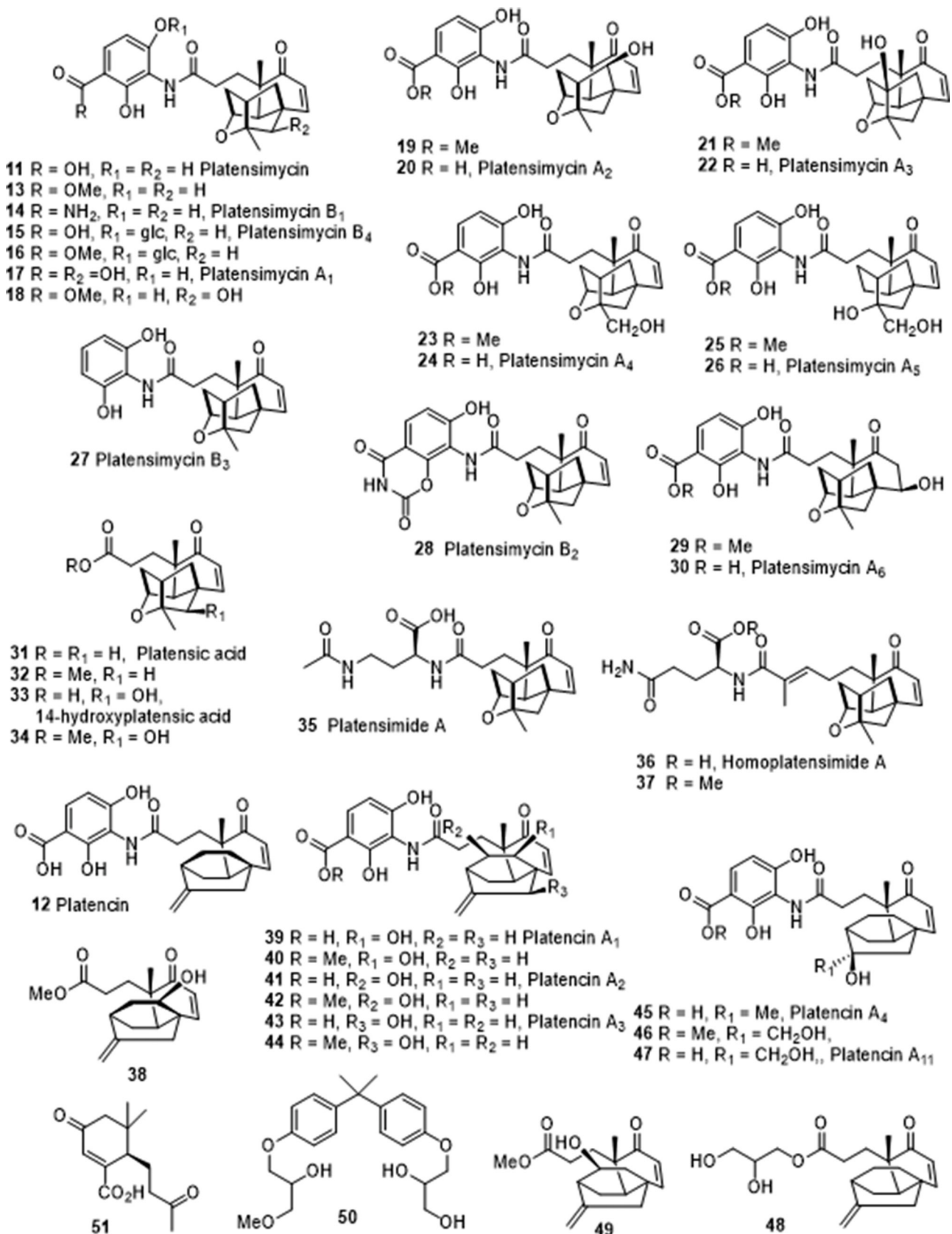
Figure 3: Secondary metabolites **11**–**51** from South African actinomycetes *Streptomyces platensis*.

been produced by an engineered mutant strain *S. platensis* SB12600.[40] Platensimycin **11** selectively inhibits the elongation-condensing enzyme FabF of the bacterial fatty acid synthesis pathway, while platencin **12** equally inhibits both the initiation condensing (FabH) and elongation

(FabF) enzymes.[39] Platensimycin, platencin and their analogues have shown potent in vitro activity against both cell-free and whole-cell systems including methicillin-resistant *Staphylococcus aureus*, vancomycin-resistant *Enterococcus* and *Mycobacterium tuberculosis*.[39]

Although the isolated analogues and synthesised ones did not exhibit improved activity compared with compounds **11** and **12**, they provide important structure–activity relationship information to determine the pharmacophore of **11** and **12**.

In their quest to isolate biologically active metabolites from termite-associated actinobacteria, Kim et al. discovered that *Streptomyces* strain M56, isolated from the fungal comb of a South African *Macrotermes natalensis* Mn802 colony, exhibited potent broad-spectrum antifungal activity.[49] Bioassay-guided isolation resulted in the purification of the novel fused bicyclic ansa macrolide natalamycin A **52** (Figure 4) alongside other ansa macrolides including reblastatin **53**, geldanamycin **54** and its derivates, 17-O-demethyl-geldanamycin **55**, 19-S-methylgeldanamycin **56**, 17-amino-17-demethoxy-geldanamycin **57**, methyl geldanamycinate derivative **58**, 17-amino-17-demethoxy-methyl geldanamycinate **59** and 19-[(1′S,4′R)-4′-hydroxy-1′-methoxy-2′-oxopentyl]geldanamycin **60**.[49] Although the fractions that yielded the isolated compounds showed strong antifungal activity against some strains, the natalamycins exhibited weak or no activity against *Saccharomyces cerevisiae* and other fungal isolates.[49] An untargeted dereplication of the liquid chromatography–mass spectrometry (LCMS) data of the methanol extract of *Streptomyces* strain M56 signified the presence of new metabolites.[50] The methanol extract of a large-scale culture of strain M56 was subjected to several chromatographic methods to yield the fused 5-8-5 tricyclic diterpene 17-hydroxycyclooctatin **61** (Figure 4).[50] Compound **61** is a potential ERα antagonist and exhibited weak cytotoxicity activity against MCF-7 human breast cancer cell lines with an IC_{50} value of 566.95 ± 0.48 μM.[50]

Further studies on the metabolites of the actinobacteria associated with the fungus-growing South African termite *M. natalensis* led to the isolation of *Streptomyces* strain RB1 which exhibited antibacterial activity against *Staphylococcus aureus* and *Candida albicans*.[51] Fractionation of the methanol (MeOH) extract of strain RB1 yielded the new isoflavonoid glycosides, termisoflavone A-C **62**–**64**. and other isoflavonoids **66**–**70**, **72**–**74** (Figure 5).[51] The isolated compounds showed no antifungal or antibacterial activity when screened against *C. albicans, C. neoformans, S. aureus*, and *E. coli,* but compounds **69** and **73** ameliorated cisplatin-induced kidney cell damage.[51] Further investigation of the MeOH extract of strain RB1 using LCMS- and NMR-based dereplication strategies led to the identification and subsequent isolation of another new isoflavonoid glycoside, termisoflavone D **65**, together with the known isoflavonoids **66**, **67**, **69**, **71**–**73** and **75** (Figure 5).[52] Isoflavonoid **69** displayed activity against glutamate-induced HT22 cells by preventing accumulation of intracellular reactive oxygen species.[52] Another study exploring the termite associated actinobacteria for reno- and kidney-protective drug discovery found that the MeOH extract of *Streptomyces* sp. RB1 exhibited a protective effect against cisplatin-induced cytotoxicity.[53] A bioassay (LLC-PK1 cells)-guided isolation process yielded the renoprotective 1-O-(2-aminobenzoyl)-α-L-rhamnopyranoside (ABR) **76** (Figure 5).[53]

Analysing the chemical and metabolomic profiles of actinobacteria derived from termite nests with an unbiased high-throughput high-performance liquid chromatography–high-resolution mass spectrometry based dereplication strategy revealed that *Streptomyces* sp. M41 isolated from the South African termite *M. natalensis* produces new complex nonribosomal peptide polyketide synthase (NRPS/PKS) hybrid compounds.[54] Chromatographic purification of a large-scale culture of strain M41 interestingly yielded new analogues (two linear **77**, **78** and one cyclic **79**) of the cyclic depsipeptide dentigerumycin **80** (Figure 6).[54]

The South African soil actinomycete, *Streptomyces canus* strain CA-091830, is a producer of the cyclic depsipeptides krisynomycins A–C **81**–**83** (Figure 7).[55,56] Krisynomycin A was initially isolated based on a screening project with the aim of identifying and isolating imipenem potentiators against methicillin-resistant *Staphylococcus aureus* (MRSA).[55] Further investigation led to the isolation of Krisynomycin B and C, which are chlorinated analogues of Krisynomycin A.[56] Although compounds **81**–**83** showed weak activity against MRSA, the activity was improved when they were tested in combination with sub-lethal concentrations of imipenem.[56]

52 Natalamycin A

53 Reblastatin

54 R = OMe, R₁ = H Geldanamycin
55 R = OH, R₁ = H
56 R = OMe, R₁ = SMe
57 R = NH₂, R₁ = H

58 R = OMe
59 R = NH₂

60

61 17-hydroxycyclooctatin
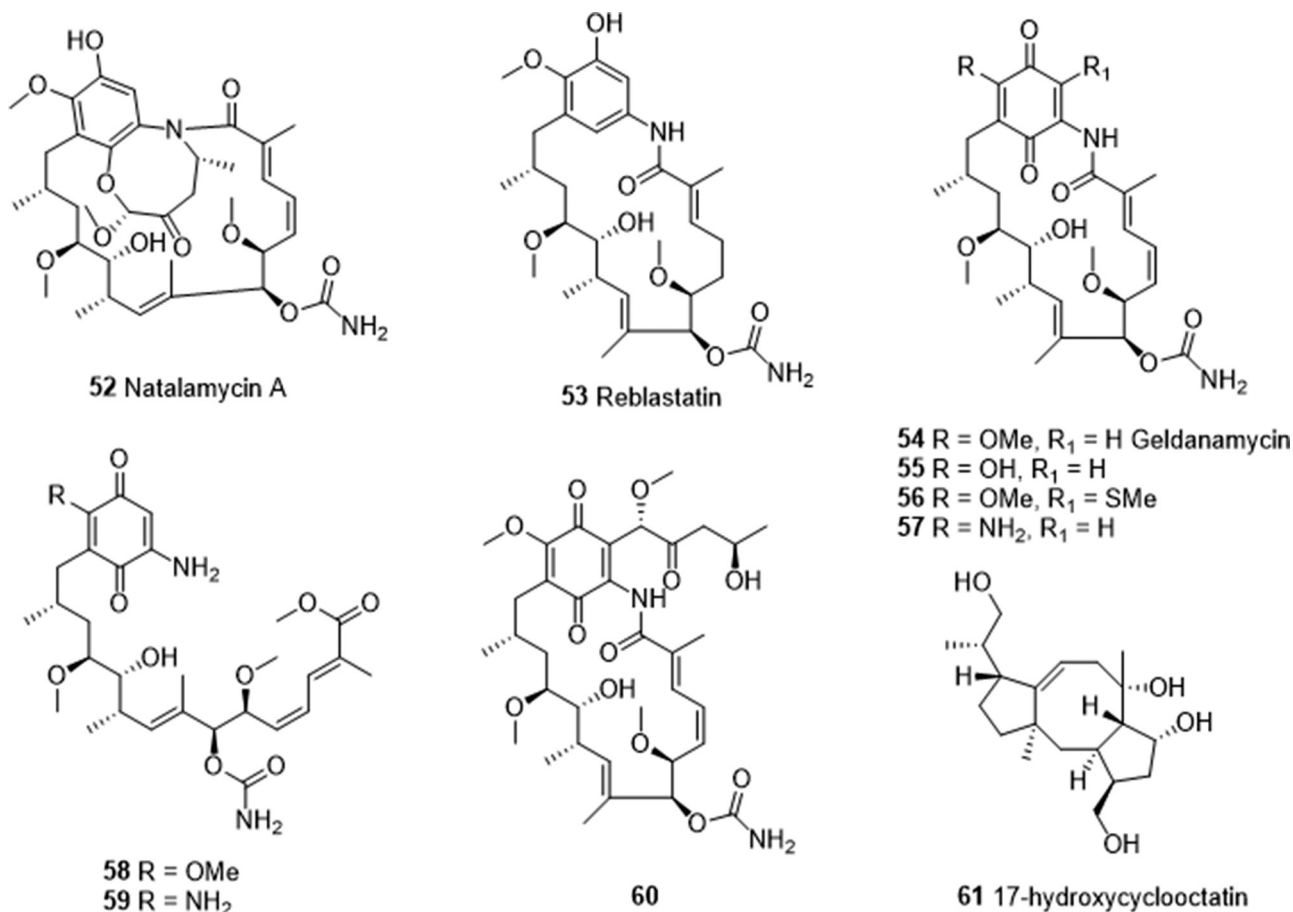
Figure 4: Secondary metabolites **52**–**61** from South African actinomycetes *Streptomyces* strain M56.

62 R = OH, R₁ = H, R₂ = Me, Termisoflavone A
63 R = OH, R₁ = R₂ = Me, Termisoflavone B
64 R = H, R₁ = R₂ = Me, Termisoflavone C
65 R = R₁ = H, R₂ = Me, Termisoflavone D
66 R = OH, R₁ = R₂ = H
67 R = R₁ = R₂ = H

68 R = OH, R₁ = R₂ = H
69 R = R₁ = R₂ = H
70 R = R₂ = H, R₁ = OMe
71 R = OH, R₁ = H, R₂ = Me

72 R = OH
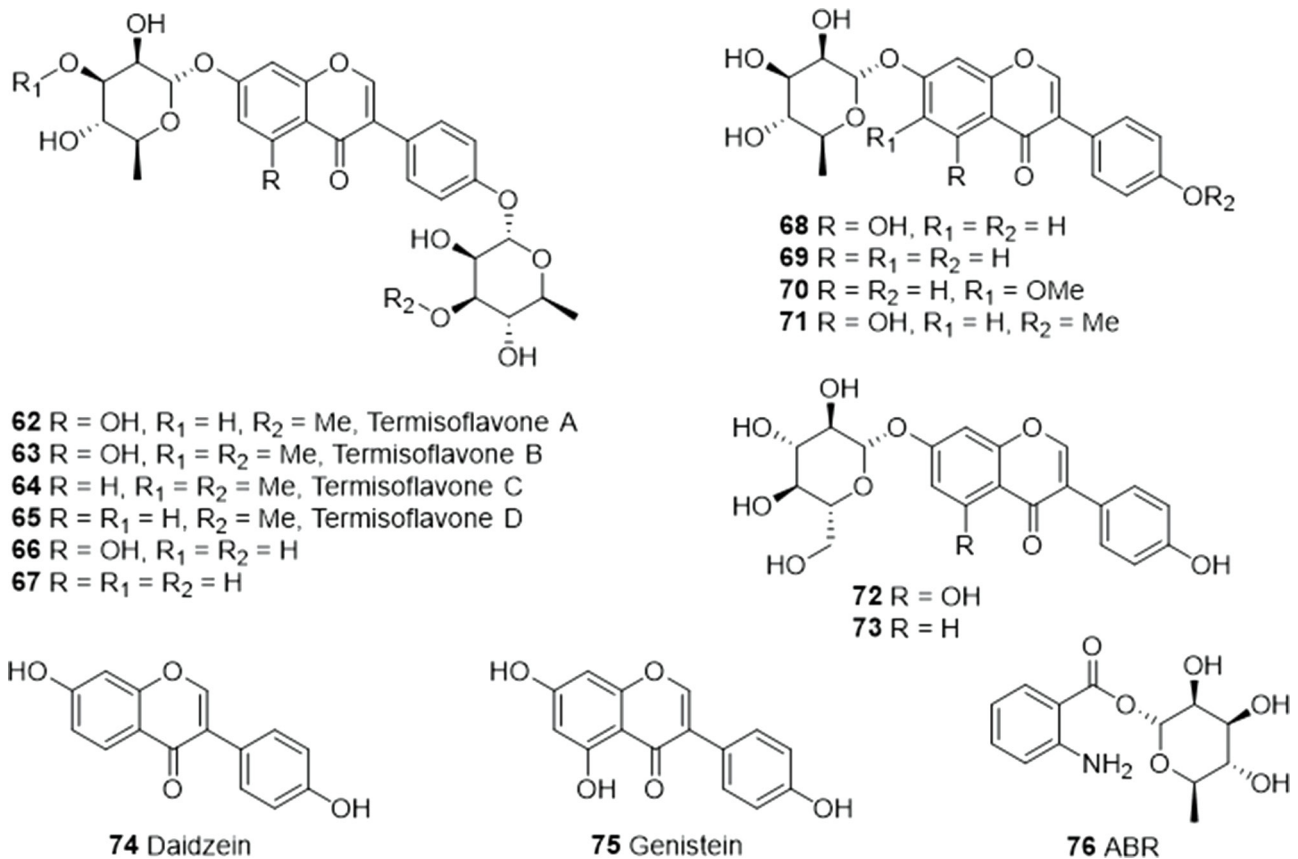**73** R = H

74 Daidzein

75 Genistein

76 ABR

Figure 5: Secondary metabolites **62–76** from South African actinomycetes *Streptomyces* strain RB1.

77 R = OH, Dentigerumycin C
78 R = NH₂, Dentigerumycin D
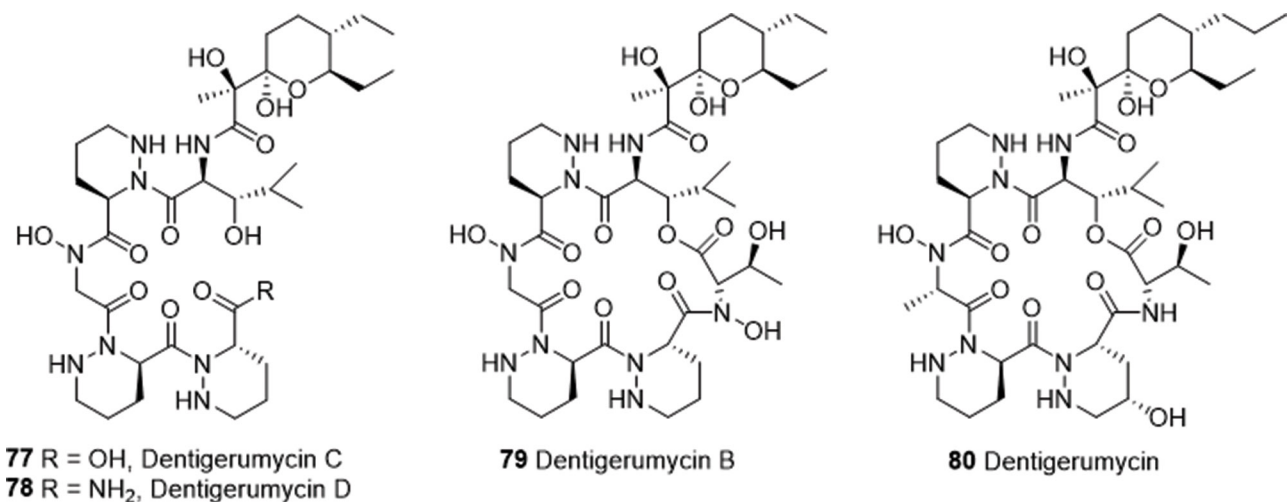
79 Dentigerumycin B

80 Dentigerumycin

Figure 6: Dentigerumycins A–D **77–80** from South African actinomycetes *Streptomyces* sp. M41.
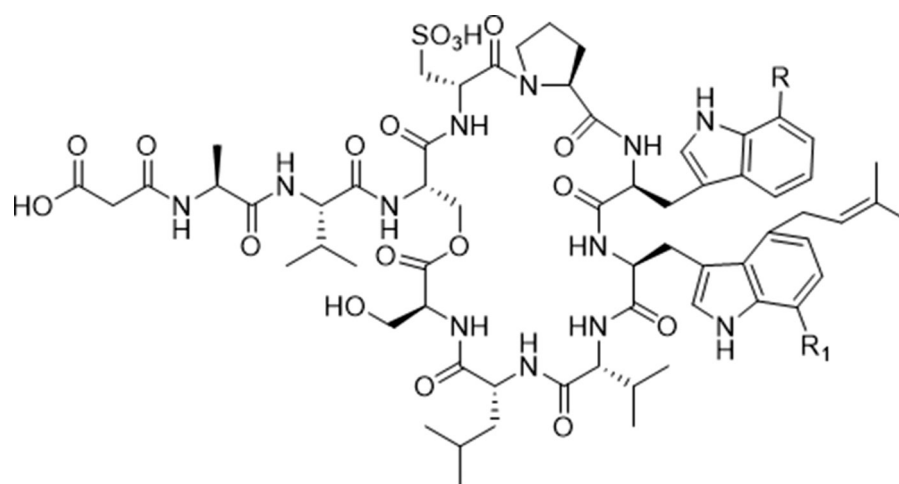
Rare actinomycete strains

The rare actinomycete *Actinomadura* sp. 5-2, which was recovered from the gut of the fungus-growing termite *M. natalensis*, produced novel, highly substituted tropolone alkaloids, rubterolones A–F **84–89** (Figure 8).[57] These compounds were detected by both bioactivity and high-resolution mass spectrometry based dereplication techniques, and subsequently isolated from the organic extract of a culture of strain 5-2.[57] Curiously, compounds **84–89** did not show any significant antifungal activity.[57]

The crude extract of another *Actinomadura* isolate, strain RB99, isolated from the surface of the termite *M. natalensis*, was analysed by liquid chromatography (LC)/ultraviolet (UV)/mass spectrometry (MS) and shown to produce new compounds.[58] A spectrometry guided isolation

led to the discovery of three new cyclic tripeptides named natalenamides A–C **90–92** (Figure 9).[58] The isolated compounds exhibited weak cytotoxicity when screened against HepG2 and HeLa/A549 cells.[58] Compound **92** showed significant activity against IBMX-mediated melanin synthesis in a dose-dependent manner.[58]

Analyses of the high-resolution tandem mass spectrometry (HR-MS²) data of the MeOH extract of strain RB99 and further exploration of the HR-MS² data on the Global Natural Product Social (GNPS) molecular networking platform showed that strain RB99 produces polyhalogenated isoflavonoids.[59] Seoung et al. proved that *Actinomadura* sp. RB99 can bio-transform the plant-based daidzein **74** and genistein **75** isoflavonoids contained in the ISP2 growth medium to polyhalogenated derivatives.[59] Optimisation of the growth

81 R = R₁ =H, Krisynomycin A
82 R = Cl, R₁ = H, Krisynomycin B
83 R = R₁ = Cl, Krisynomycin C

Figure 7: Krisynomycins **81–83** from South African actinomycetes *Streptomyces canus* strain CA-091830.

84 Rubterolone A

86 Rubterolone C

88 Rubterolone E

85 Rubterolone B

87 Rubterolone D
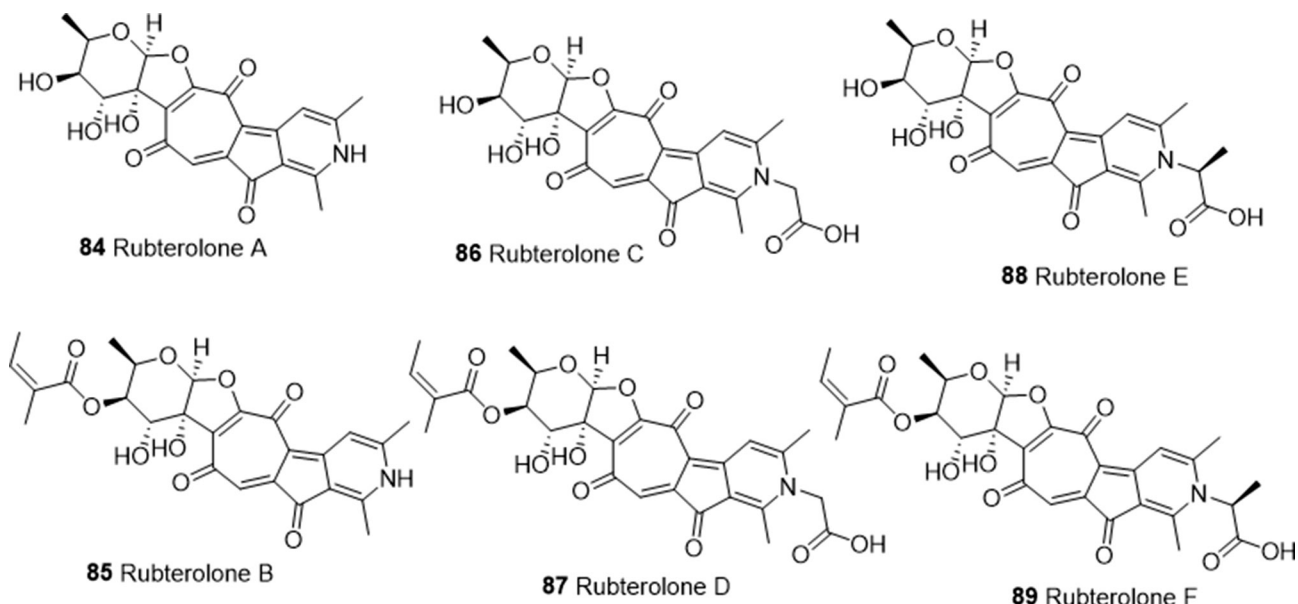
89 Rubterolone F

Figure 8: Rubterolones A–F (**84–89**) from South African actinomycetes *Actinomadura* sp. 5-2.

medium (ISP2 augmented with NaCl or KBr) led to the production and subsequent MS-guided purification of eight polychlorinated analogues **93–100** (Figure 9), of which six were new, and seven novel polybrominated analogues **101–107** (Figure 9), of daidzein and genistein.[59] The isolated chlorinated analogues did not exhibit any antibacterial or antifungal activities against *E. coli, S. aureus, S. epidermidis* and *C. albicans,* but the brominated analogues **101** and **105** were active against *Helicobacter pylori*.[59] Additional analysis of the LCMS data of the MeOH extract of strain RB99 led to the detection and isolation of the antibiotic and antitumour agent fridamycin A **108** (Figure 9), which is a type II polyketide.[60] Fridamycin A **108** showed good antidiabetic properties in 3T3-L1 adipocytes and could serve as a promising lead for type 2 diabetes drug discovery.[60]

Metabolomic and bioactivity profiling of termite-associated actinomycetes led to the detection and subsequent isolation of four new 20-membered glycosylated polyketide macrolactams, named macrotermycins A–D **109–112** (Figure 10), from the organic extract of the rare actinomycete *Amycolatopsis* sp. M39.[61] Strain M39 was also isolated from the termite

M. natalensis and its organic crude extract exhibited a unique metabolomic profile and was active against the termite fungal garden competitor *Pseudoxylaria* spp.[61] Only compounds **109** and **111** were active against *Pseudoxylaria* sp.[61]

A rare actinomycete, *Kribbella speibonae* strain SK5, isolated from a soil sample collected from Stellenbosch in the Western Cape Province of South Africa, displayed strong antimycobacterial activity against *Mycobacterium aurum* strain A+.[62] Chemical and metabolomic profiling of an organic extract of a liquid culture of this strain showed that it is a prolific producer of hydroxamate siderophores, including new dehydroxylated desferrioxamine analogues and diketopiperazines (DKP).[63] Two new dehydroxylated desferrioxamines, speibonoxamine **113** and desoxy-desferrioxamine D1 **114** (Figure 11), alongside already reported desferrioxamines **115–118** and a DKP **119** (Figure 11), were subsequently isolated from an organic extract of a liquid culture of strain SK5.[63] The plausible structures of three new dehydroxylated analogues **120–122** were determined by the GNPS molecular network and MS/MS fragmentation analyses.[63]
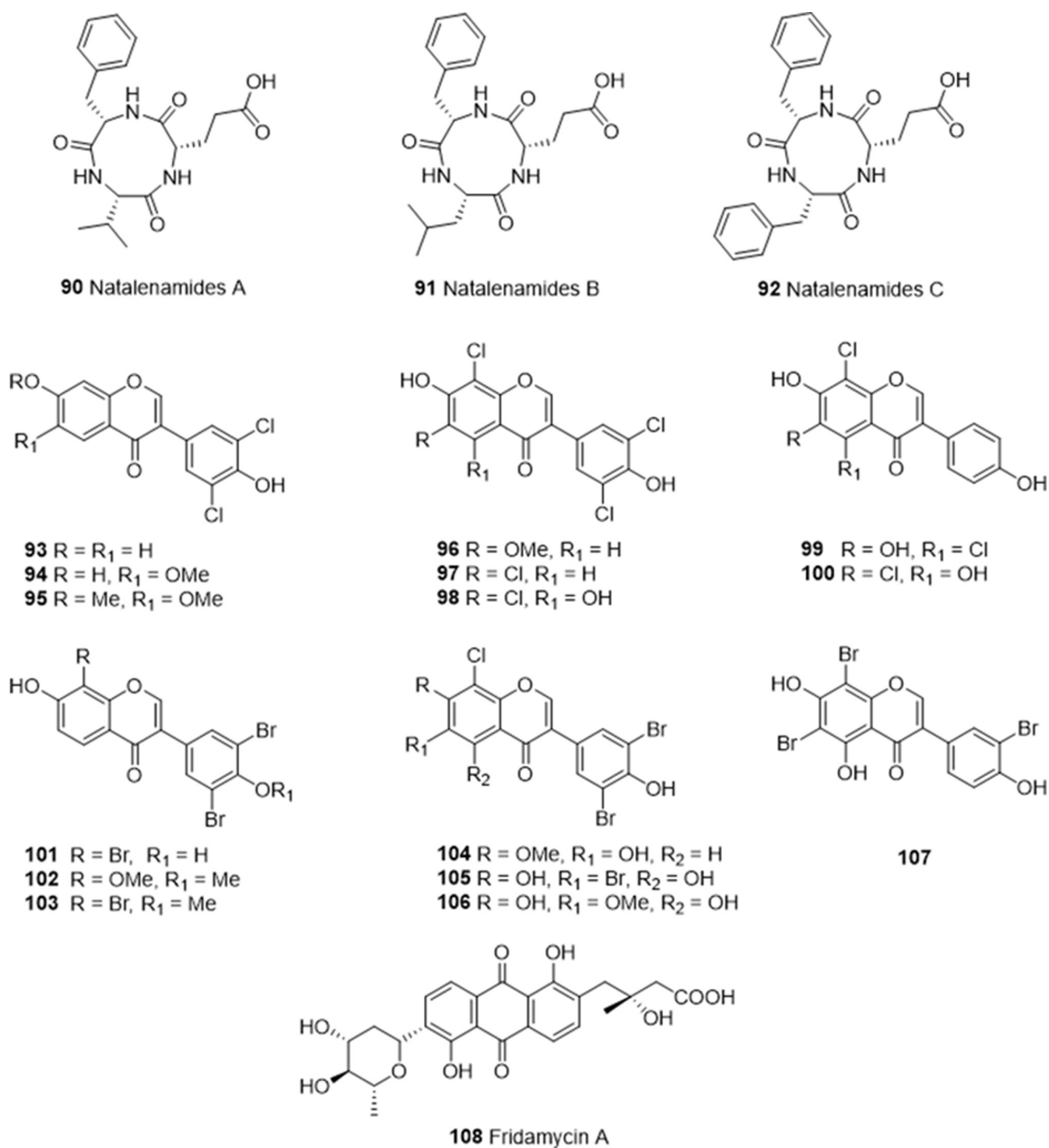
90 Natalenamides A

91 Natalenamides B

92 Natalenamides C

93 R = R₁ = H
94 R = H, R₁ = OMe
95 R = Me, R₁ = OMe

96 R = OMe, R₁ = H
97 R = Cl, R₁ = H
98 R = Cl, R₁ = OH

99 R = OH, R₁ = Cl
100 R = Cl, R₁ = OH

101 R = Br, R₁ = H
102 R = OMe, R₁ = Me
103 R = Br, R₁ = Me

104 R = OMe, R₁ = OH, R₂ = H
105 R = OH, R₁ = Br, R₂ = OH
106 R = OH, R₁ = OMe, R₂ = OH

107

108 Fridamycin A

Figure 9: Secondary metabolites from South African actinomycetes **90–108**.

Conclusions and future prospects

Biodiversity is "more than just legs and leaves"[64] and South Africa's microbial biodiversity presents a tremendous opportunity for the natural products chemist and those interested in drug discovery. In this review, we have described 122 compounds and shown that South African actinobacteria are prolific producers of novel, bioactive metabolites. Interestingly, the first compound described from a South African actinomycete, natamycin, is also the only one that has made it to the clinic. Other compounds, such as platensimycin and geldanamycin, have shown promise but either lack efficacy in humans or showed toxic side effects which prevented their development as drugs. South African researchers interested in natural products based drug discovery face the same challenges as elsewhere in the world.

These challenges include the significant cost of drug development, re-isolation of previously reported compounds, and lack of interest in natural products for drug development. Nevertheless, the compounds reviewed here present only the tip of the iceberg and many more species remain to be discovered and studied for natural product production. Furthermore, with innovations and technological advancement in purification, structure elucidation, chemical biology, genome sequencing and mining, dereplication, and bioinformatic, cheminformatic and metabolomic tools like the GNPS molecular networking, microbial natural product drug discovery in South Africa shows great potential. It is worth mentioning that, apart from our research on the chemistry of the metabolites of the South African rare actinomycete *Kribbella speibonae* strain SK5, all the research on South African actinomycetes reported here was done by research
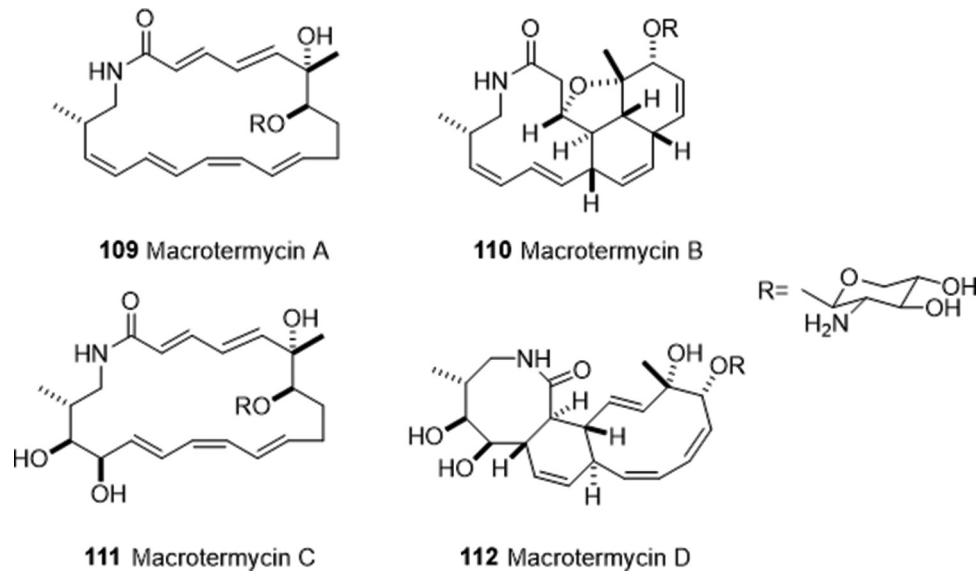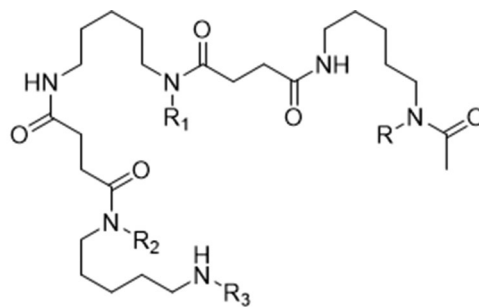
Figure 10: Macrotermycins A–D **109**–**112** from South African actinomycetes *Amycolatopsis* sp. M39.

113 R = R$_1$ = R$_2$ = H, R$_3$ = COCH$_3$, Speibonoxamine
114 R = H, R$_1$ = R$_2$ = OH, R$_3$ = COCH$_3$, desoxy-desferrioxamine D$_1$
115 R = R$_1$ = R$_2$ = OH, R$_3$ = COCH$_3$, Desferrioxamine D$_1$
116 R = R$_1$ = R$_2$ = OH, R$_3$ = H, desferrioxamine B
120 R = R$_1$ = H, R$_2$ = OH, R$_3$ = COCH$_3$, *di*desoxy-desferrioxamine D$_1$
121 R = H, R$_1$ = R$_2$ = OH, R$_3$ = H, desoxy-desferrioxamine B
122 R = R$_2$ = R$_3$ = H, R$_1$ = OH, *di*desoxy-desferrioxamine B
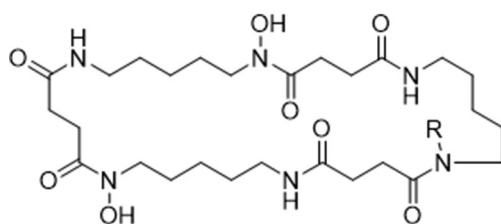
117 R = H, Desoxynocardamine
118 R = OH, Desferrioxamine E
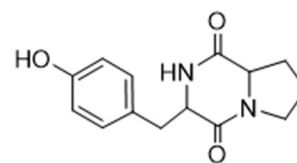
119 R =hexahydro-3-((4-hydroxyphenayl)methyl)-pyrrolo[1,2-a]pyrazine-1,4-dione

Figure 11: Secondary metabolites **113**–**122** from the South African *Kribbella speibonae* strain SK5.

groups based outside of South Africa. This represents a challenge and an opportunity for closer collaboration between South African researchers (microbiologists, pharmacologists and chemists) in order to fully explore the opportunities presented by South African microbial biodiversity.

Acknowledgements

Competing interests

We have no competing interests to declare.

Authors' contributions

K.S.A: Conceptualisation; data collection; writing – the initial draft.
D.W.G: Conceptualisation; student supervision; writing – revisions.
D.R.B: Conceptualisation; data collection; student supervision; writing – revisions.

References

1. Newman DJ, Cragg GM. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. J Nat Prod. 2020;83(3):770–803. https://doi.org/10.1021/acs.jnatprod.9b01285

2. UN Environment Programme World Conservation Monitoring Centre (UNEP-WCMC). Biodiversity A-Z. Cambridge, UK: UNEP-WCMC; 2014. Available from: https://www.biodiversitya-z.org/content/megadiverse-countries

3. Drewes SE. Natural products research in South Africa: 1890–2010. S Afr J Sci. 2012;108(5/6), Art. #765. https://doi.org/10.4102/sajs.v108i5/6.765

4. Davies-Coleman MT. Natural products research in South Africa: End of an era on land or the beginning of an endless opportunity in the sea? S Afr J Chem. 2010;63:105–113.

5. Davies-Coleman MT, Veale CG. Recent advances in drug discovery from South African marine invertebrates. Mar Drugs. 2015;13(10):6366–6383. https://doi.org/10.3390/md13106366

6. Parte AC, Sardá Carbasse J, Meier-Kolthoff JP, Reimer LC, Göker M. List of prokaryotic names with standing in nomenclature (LPSN) moves to the DSMZ. Int J Syst Evol Microbiol. 2020;70:5607–5612. https://doi.org/10.1099/ijsem.0.004332

7. Merino N, Aronson HS, Bojanova DP, Feyhl-Buska J, Wong ML, Zhang S, et al. Living at the extremes: Extremophiles and the limits of life in a planetary context. Front Microbiol. 2019;10, Art. #780. https://doi.org/10.3389/fmicb.2019.00780

8. Mayfield CI, Williams ST, Ruddick SM, Hatfield HL. Studies on the ecology of actinomycetes in soil IV. Observations on the form and growth of Streptomycetes in soil. Soil Biol Biochem. 1972;4(1):79–91. https://doi.org/10.1016/0038-0717(72)90045-4

9. Barka EA, Vatsa P, Sanchez L, Gaveau-Vaillant N, Jacquard C, Klenk HP, et al. Taxonomy, physiology, and natural products of Actinobacteria. Microbiol Mol Biol Rev. 2016;80(1):143. https://doi.org/10.1128/MMBR.00019-15

10. Benndorf R, Martin K, Küfner M, de Beer ZW, Vollmers J, Kaster AK, et al. Actinomadura rubteroloni sp. nov. and Actinomadura macrotermitis sp. nov., isolated from the gut of the fungus growing-termite Macrotermes natalensis. Int J Syst Evol. 2020;70(10):5255–5262. https://doi.org/10.1099/ijsem.0.004403

11. Salwan R, Sharma V. The role of actinobacteria in the production of industrial enzymes. In: Singh BP, Gupta VK, Passari AK, editors. New and future developments in microbial biotechnology and Bioengineering. Amsterdam: Elsevier; 2018. p. 165–177. https://doi.org/10.1016/B978-0-444-63994-3.00011-4

12. Genilloud O. Actinomycetes: Still a source of novel antibiotics. Nat Prod Rep. 2017;34(10):1203–1232. https://doi.org/10.1039/C7NP00026J

13. Taechowisan T, Chaisaeng S, Phutdhawong WS. Antibacterial, antioxidant and anticancer activities of biphenyls from Streptomyces sp. BO-07: An endophyte in Boesenbergia rotunda (L.) Mansf A. Food Agric Immunol. 2017;28(6):1330–1346. https://doi.org/10.1080/09540105.2017.1339669

14. Rohland J, Meyers PR. Streptomyces fractus sp. nov., a novel streptomycete isolated from the gut of a South African termite. Antonie van Leeuwenhoek. 2015;107(5):1127–1134. https://doi.org/10.1007/s10482-015-0404-8

15. Cook AE, Le Roes M, Meyers PR. Actinomadura napierensis sp. nov., isolated from soil in South Africa. Int J Syst Evol. 2005;55(2):703–706. https://doi.org/10.1099/ijs.0.63359-0

16. Everest GJ, Le Roes-Hill M, Rohland J, Enslin S, Meyers PR. Amycolatopsis roodepoortensis sp. nov. and Amycolatopsis speibonae sp. nov.: Antibiotic-producing actinobacteria isolated from South African soils. J Antibiot. 2014;67(12):813–818. https://doi.org/10.1038/ja.2014.79

17. Everest GJ, Meyers PR. Kribbella hippodromi sp. nov., isolated from soil from a racecourse in South Africa. Int J Syst Evol Microbiol. 2008;58(2):443–446. https://doi.org/10.1099/ijs.0.65278-0

18. Le Roes M, Goodwin CM, Meyers PR. Gordonia lacunae sp. nov., isolated from an estuary. Syst Appl Microbiol. 2008;31(1):17–23. https://doi.org/10.1016/j.syapm.2007.10.001

19. Benndorf R, Schwitalla JW, Martin K, De Beer ZW, Vollmers J, Kaster AK, et al. Nocardia macrotermitis sp. nov. and Nocardia aurantia sp. nov., isolated from the gut of the fungus-growing termite Macrotermes natalensis. Int J Syst Evol Microbiol. 2020;70(10):5226–5234. https://doi.org/10.1099/ijsem.0.004398

20. Le Roes M, Meyers PR. Nonomuraea candida sp. nov., a new species from South African soil. Antonie van Leeuwenhoek. 2008;93(1):133–139. https://doi.org/10.1007/s10482-007-9187-x

21. Sibanda T, Mabinya LV, Mazomba N, Akinpelu DA, Bernard K, Olaniran AO, et al. Antibiotic producing potentials of three freshwater actinomycetes isolated from the Eastern Cape Province of South Africa. Int J Mol Sci. 2010;11(7):2612–2623. https://doi.org/10.3390/ijms11072612

22. Le Roes-Hill M, Durrell K, Prins A, Meyers PR. Streptosporangium minutum sp. nov., isolated from garden soil exposed to microwave radiation. J Antibiot. 2018;71(6):564–574. https://doi.org/10.1038/s41429-018-0036-0

23. Struyk AP, Waisvisz JM. (Gist-Brocades, Delft): Pimaricin and process of producing same. US-Patent. 1955;3:892.

24. Struyk AP, Drost G, Haisvisz JM, Van Eek T, Hoogerheide JC. Pimaricin, a new antifungal antibiotic. Antibiot Annu. 1957/1958;5:878–885.

25. Chen GQ, Lu FP, Du LX. Natamycin production by Streptomyces gilvosporeus based on statistical optimization. J Agric Food Chem. 2008;56(13):5057–5061. https://doi.org/10.1021/jf800479u

26. Du YL, Chen SF, Cheng LY, Shen XL, Tian Y, Li YQ. Identification of a novel Streptomyces chattanoogensis L10 and enhancing its natamycin production by overexpressing positive regulator ScnRII. J Microbiol. 2009;47(4):506–513. https://doi.org/10.1007/s12275-009-0014-0

27. Lu CG, Liu WC, Qiu JY, Wang HM, Liu T, Liu DW. Identification of an antifungal metabolite produced by a potential biocontrol actinomyces strain A01. Braz J Microbiol. 2008;39(4):701–707. https://doi.org/10.1590/S1517-838220080004000020

28. Aparicio JF, Barreales EG, Payero TD, Vicente CM, De Pedro A, Santos-Aberturas J. Biotechnological production and application of the antibiotic pimaricin: Biosynthesis and its regulation. Appl Microbiol Biotechnol. 2016;100(1):61–78. https://doi.org/10.1007/s00253-015-7077-0

29. Te Welscher YM, Jones L, Van Leeuwen MR, Dijksterhuis J, De Kruijff B, Eitzen G, et al. Natamycin inhibits vacuole fusion at the priming phase via a specific interaction with ergosterol. Antimicrob Agents Chemother. 2010;54(6):2618–2625. https://doi.org/10.1128/AAC.01794-09

30. Te Welscher YM, Hendrik H, Balagué MM, Souza CM, Riezman H, De Kruijff B, et al. Natamycin blocks fungal growth by binding specifically to ergosterol without permeabilizing the membrane. J Biol Chem. 2008;283(10):6393–6401. https://doi.org/10.1074/jbc.M707821200

31. Rolón M, Seco EM, Vega C, Nogal JJ, Escario JA, Gómez-Barrio A, et al. Selective activity of polyene macrolides produced by genetically modified Streptomyces on Trypanosoma cruzi. Int J Antimicrob Agents. 2006;28(2):104–109. https://doi.org/10.1016/j.ijantimicag.2006.02.025

32. Sunada A, Kimura K, Nishi I, Toyokawa M, Ueda A, Sakata T, et al. In vitro evaluations of topical agents to treat Acanthamoeba keratitis. Ophthalmology. 2014;121(10):2059–2065. https://doi.org/10.1016/j.ophtha.2014.04.013

33. Jackson M, Karwowski JP, Theriault RJ, Hardy DJ, Swanson SJ, Barlow GJ, et al. Altromycins, novel pluramycin-like antibiotics I. Taxonomy of the producing organism, fermentation and antibacterial activity. J Antibiot. 1990;43(3):223–228. https://doi.org/10.7164/antibiotics.43.223

34. Brill GM, Mcalpine JB, Whittern DN, Buko AM. Altromycins, novel pluramycin-like antibiotics II. Isolation and elucidation of structure. J Antibiot. 1990;43(3):229–237. https://doi.org/10.7164/antibiotics.43.229

35. Brill GM, Jackson M, Whittern DN, Buko AM, Hill P, Chen RH, et al. Altromycins E, F, G, H and I; additional novel components of the altromycin complex. J Antibiot. 1994;47(10):1160–1164. https://doi.org/10.7164/antibiotics.47.1160

36. Hansen M, Hurley L. Altromycin B threads the DNA helix interacting with both the major and the minor grooves to position itself for site-directed alkylation and guanine N7. J Am Chem Soc. 1995;117(9):2421–2429. https://doi.org/10.1021/ja00114a006

37. McAlpine JB, Karwowski JP, Jackson M, Brill GM, Kadam S, Shen L, et al. Altromycins: A new family of antitumor antibiotics-discovery and biological evaluation. In: Valeriote FA, Corbett TH, Baker LH, editors. Anticancer drug discovery and development: Natural products and new molecular models. Boston, MA: Springer; 1994. p. 95–117. https://doi.org/10.1007/978-1-4615-2610-0_6

38. Singh SB, Jayasuriya H, Ondeyka JG, Herath KB, Zhang C, Zink DL, et al. Isolation, structure, and absolute stereochemistry of platensimycin, a broad spectrum antibiotic discovered using an antisense differential sensitivity strategy. J Am Chem Soc. 2006;128(36):11916–11920. https://doi.org/10.1021/ja062232p

39. Martens E, Demain AL. Platensimycin and platencin: Promising antibiotics for future application in human medicine. J Antibiot. 2011;64(11):705–710. https://doi.org/10.1038/ja.2011.80

40. Yu Z, Rateb ME, Smanski MJ, Peterson RM, Shen B. Isolation and structural elucidation of glucoside congeners of platencin from *Streptomyces platensis* SB12600. J Antibiot. 2013;66(5):291–294. https://doi.org/10.1038/ja.2013.1

41. Zhang C, Ondeyka J, Herath K, Jayasuriya H, Guan Z, Zink DL, et al. Platensimycin and platencin congeners from *Streptomyces platensis*. J Nat Prod. 2011;74(3):329–340. https://doi.org/10.1021/np100635f

42. Zhang C, Ondeyka J, Dietrich L, Gailliot FP, Hesse M, Lester M, et al. Isolation, structure and biological activities of platencin A2-A4 from *Streptomyces platensis*. Bioorg Med Chem. 2010;18(7):2602–2610. https://doi.org/10.1016/j.bmc.2010.02.030

43. Herath KB, Zhang C, Jayasuriya H, Ondeyka JG, Zink DL, Burgess B, et al. Structure and semisynthesis of platensimide A, produced by *Streptomyces platensis*. Org Lett. 2008;10(9):1699–1702. https://doi.org/10.1021/ol900695v

44. Singh SB, Jayasuriya H, Herath KB, Zhang C, Ondeyka JG, Zink DL, et al. Isolation, enzyme-bound structure, and activity of platensimycin A1 from *Streptomyces platensis*. Tetrahedron Lett. 2009;50(37):5182–5185. https://doi.org/10.1016/j.tetlet.2009.06.118

45. Zhang C, Ondeyka J, Zink DL, Burgess B, Wang J, Singh SB. Isolation, structure and fatty acid synthesis inhibitory activities of platensimycin B1-B3 from *Streptomyces platensis*. Chem Comm. 2008(40):5034–5036. https://doi.org/10.1039/b810113b

46. Zhang C, Ondeyka J, Guan Z, Dietrich L, Burgess B, Wang J, et al. Isolation, structure and biological activities of platensimycin B4 from *Streptomyces platensis*. J Antibiot. 2009;62(12):699–702. https://doi.org/10.1038/ja.2009.106

47. Singh SB, Ondeyka JG, Herath KB, Zhang C, Jayasuriya H, Zink DL, et al. Isolation, enzyme-bound structure and antibacterial activity of platencin A1 from *Streptomyces platensis*. Bioorg Med Chem Lett. 2009;19(16):4756–4759. https://doi.org/10.1016/j.bmcl.2009.06.061

48. Jayasuriya H, Herath KB, Ondeyka JG, Zink DL, Burgess B, Wang J, et al. Structure of homoplatensimide A: A potential key biosynthetic intermediate of platensimycin isolated from *Streptomyces platensis*. Tetrahedron Lett. 2008;49(22):3648–3651. https://doi.org/10.1016/j.tetlet.2008.03.155

49. Kim KH, Ramadhar TR, Beemelmanns C, Cao S, Poulsen M, Currie CR, et al. Natalamycin A, an ansamycin from a termite-associated *Streptomyces* sp. Chem Sci. 2014;5(11):4333-–4338. https://doi.org/10.1039/C4SC01136H

50. Lee SR, Lee D, Park M, Lee JC, Park HJ, Kang KS, et al. Absolute configuration and corrected NMR assignment of 17-hydroxycyclooctatin, a fused 5-8-5 tricyclic diterpene. J Nat Prod. 2020;83(2):354–361. https://doi.org/10.1021/acs.jnatprod.9b00837

51. Kang HR, Lee D, Benndorf R, Jung WH, Beemelmanns C, Kang KS, et al. Termisoflavones A-C, isoflavonoid glycosides from termite-associated *Streptomyces* sp. RB1. J Nat Prod. 2016;79(12):3072–3078. https://doi.org/10.1021/acs.jnatprod.6b00738

52. Lee SR, Song JH, Song JH, Ko HJ, Baek JY, Trinh TA, et al. Chemical identification of isoflavonoids from a termite-associated *Streptomyces* sp. RB1 and their neuroprotective effects in murine hippocampal HT22 cell line. Int J Mol Sci. 2018;19(9), Art. #2640. https://doi.org/10.3390/ijms19092640

53. Lee D, Kang KS, Lee HJ, Kim KH. Chemical characterization of a renoprotective metabolite from termite-associated *Streptomyces* sp. RB1 against cisplatin-induced cytotoxicity. Int J Mol Sci. 2018;19(1), Art. #174. https://doi.org/10.3390/ijms19010174

54. Wyche TP, Ruzzini AC, Beemelmanns C, Kim KH, Klassen JL, Cao S, et al. Linear peptides are the major products of a biosynthetic pathway that encodes for cyclic depsipeptides. Org Lett. 2017;19(7):1772–1775. https://doi.org/10.1021/acs.orglett.7b00545

55. Therien AG, Huber JL, Wilson KE, Beaulieu P, Caron A, Claveau D, et al. Broadening the spectrum of β-lactam antibiotics through inhibition of signal peptidase type I. Antimicrob Agents Chemother. 2012;56(9):4662–4670. https://doi.org/10.1128/AAC.00726-12

56. Pérez-Bonilla M, Oves-Costales D, González I, De la Cruz M, Martin J, Vicente F, et al. Krisynomycins, imipenem potentiators against methicillin-resistant *Staphylococcus aureus*, produced by *Streptomyces canus*. J Nat Prod. 2020;83(9):2597–2606. https://doi.org/10.1021/acs.jnatprod.0c00294

57. Guo H, Benndorf R, Leichnitz D, Klassen JL, Vollmers J, Görls H, et al. Isolation, biosynthesis and chemical modifications of rubterolones A-F: Rare tropolone alkaloids from *Actinomadura* sp. 5-2. Chem Eur J 2017;23:9338–9345. https://doi.org/10.1002/chem.201701005

58. Lee SR, Lee D, Yu JS, Benndorf R, Lee S, Lee DS, et al. Natalenamides A-C, cyclic tripeptides from the termite-associated *Actinomadura* sp. RB99. Molecules. 2018;23(11), Art. #3003. https://doi.org/10.3390/molecules23113003

59. Rak Lee S, Schalk F, Schwitalla JW, Benndorf R, Vollmers J, Kaster AK, et al. Polyhalogenation of isoflavonoids by the termite-associated *Actinomadura* sp. RB99. J Nat Prod. 2020;83(10):3102–3110. https://doi.org/10.1021/acs.jnatprod.0c00676

60. Yoon SY, Lee SR, Hwang JY, Benndorf R, Beemelmanns C, Chung SJ, et al. Fridamycin A, a microbial natural product, stimulates glucose uptake without inducing *adipogenesis*. Nutrients. 2019;11(4), Art. #765. https://doi.org/10.3390/nu11040765

61. Beemelmanns C, Ramadhar TR, Kim KH, Klassen JL, Cao S, Wyche TP, et al. Macrotermycins A-D, glycosylated macrolactams from a termite-associated *Amycolatopsis* sp. M39. Org Lett. 2017;19(5):1000–1003. https://doi.org/10.1021/acs.orglett.6b03831

62. Curtis SM, Norton I, Everest GJ, Pelser JG, De Kock MC, Meyers PR. Development of a *Kribbella*-specific isolation medium and description of *Kribbella capetownensis* sp. nov. and *Kribbella speibonae* sp. nov., isolated from soil. Antonie van Leeuwenhoek. 2020;113(5):617–628. https://doi.org/10.1007/s10482-019-01365-6

63. Acquah KS, Beukes DR, Warner DF, Meyers PR, Sunassee SN, Maglangit F, et al. Novel South African rare actinomycete *Kribbella speibonae* strain SK5: A prolific producer of hydroxamate siderophores including new dehydroxylated congeners. Molecules. 2020;25(13), Art. #2979. https://doi.org/10.3390/molecules25132979

64. Cowan DA, Valverde A, Wingfield MJ, Rybicki EP, Tuffin MI. Biodiversity: So much more than legs and leaves. S Afr J Sci. 2013;109(11/12), Art. #a0037. https://doi.org/10.1590/sajs.2013/a0037

AUTHORS:
Kubrat A. Oyinlola[1] iD
Gbemisola E. Ogunleye[2] iD
Augustina I. Balogun[2] iD
Oluwarotimi Joseph[2] iD

AFFILIATIONS:
[1]Department of Microbiology, University of Ibadan, Ibadan, Nigeria
[2]Department of Biological Sciences, Koladaisi University, Ibadan, Nigeria

CORRESPONDENCE TO:
Kubrat Oyinlola

EMAIL:
bibimola.abiola@gmail.com

Comparative study: Garlic, ginger and turmeric as natural antimicrobials and bioactives

Biologically active compounds in most spices possess antimicrobial and other important biomedical properties. There have been huge demands for natural immunity boosters (spices and herbs), considering the recent global pandemic and challenges relating to drug-resistant pathogens. This study was designed to compare the efficacy of ginger, garlic and turmeric spices against some pathogenic microorganisms. Aqueous extraction of spices, antimicrobial sensitivity and minimum inhibitory concentration tests were done using standard microbiological methods. Bioactive compounds were estimated using the gas chromatography–mass spectrometry (GC-MS) method. Aqueous extracts of ginger inhibited the growth of all test isolates except *Streptococcus pneumoniae*, with inhibition zones ranging between 0.9 mm and 13.5 mm. *Escherichia coli, S. pneumoniae* and *Haemophilus influenzae* were resistant to turmeric extracts, while the extract of garlic inhibited only four of the test pathogens. Inhibition zones for turmeric ranged between 4.4 mm and 10.9 mm, while those for garlic were between 4.7 mm and 11.5 mm. All the spice extracts did not inhibit microbial growth at 10–40%. An antibiotic spectrum indicated that *Bacillus* sp. was resistant to all but one, nitrofurantoin, which also inhibited the growth of almost all pathogens, except *H. influenzae*, with zones ranging between 10.5 mm and 11.6 mm. All test pathogens were resistant to cloxacillin except *E. coli* (10.6 mm). The major phyto-active compounds present in ginger are 2-Butanone,4-(4-hydroxy-3-methoxyphenyl), 1,3-Cyclohexadiene and 1-(4-Hydroxy-3-methoxyphenyl).

Significance:

Conclusively, ginger, turmeric and garlic have varied inhibitory activities against diverse organisms, indicating their antimicrobial properties; however, ginger showed a higher inhibitory effect and more diverse antimicrobial property amongst selected isolates. Furthermore, certain bioactive compounds of biomedical importance were present. We therefore recommend the use of these spices as alternative natural food preservatives against spoilage organisms, as well as potential natural sources for bioactive compounds in drug development against pathogens.

Introduction

Spice is an "aromatic vegetable substance in the whole, broken, or ground form, the significant function of which in food is seasoning rather than nutrition", according to the US Food and Drug Administration[1]. Globally, spice is used to enhance both the flavour and aroma of foods; however, their usage in preserving food quality and in the treatment of infections and diseases has been widely recognised[2], thus indicating their antimicrobial properties, as a result of certain naturally derived bioactive components[3]. Being organic and naturally plant based, spices are more widely accepted than synthetic additives like butylated hydroxytoluene (BHT) and butylated hydroxyanisole (BHA), and, as such, their safety has not been questioned when compared. They have been widely employed in food processing and preparation, due to their antimicrobial properties that ensure improved food quality as well as prevent food spoilage.[4] Numerous phytochemical compounds in spices, such as isoflavones, anthocyanins, flavonoids, phenolic compounds, sulfur-containing compounds, tannins and alkaloids, acts as antimicrobial compounds and photoprotectants.[5] Early in the 20th century, the majority of the world's population depended on traditional preparations to address most medical conditions. Spices are regarded as safe and their efficacy against certain ailments has been recorded.[6] Various studies have revealed that oils and alkaloids in most spices possess antimicrobial, antiparasitic, immune booster, antioxidant and other important biological properties. Most recently, there has been huge demand for natural immunity boosters like spices and herbs, in response to the COVID-19 global pandemic.

However, the antimicrobial property spectrum of different spices varies[7], as it is difficult to predict the mechanism of microbial susceptibility. Different constituents may affect several targets, such as a microorganism's cell membrane, enzymes and/or their genetic material.[4,7] Research studies to tackle antibiotic drug resistance in microorganisms, through the quest for an alternative, have been widely prompted as some of these spices have shown great antimicrobial activities. The aim of this study was to evaluate the antimicrobial effect of the aqueous extracts of garlic (*Allium sativum*), ginger (*Zingiber officinale*) and turmeric (*Curcuma longa*) against some important foodborne and pathogenic microorganisms, as well as compare their activity with that of selected antibiotics and determine the presence of bioactive compounds.

Materials and methods

Sample collection

Fresh garlic (*Allium sativum*), turmeric (*Curcuma longa*) and ginger (*Zingiber officinale*) roots were locally sourced from a produce market in Ibadan, Oyo State, Nigeria. They were labelled and transported to the laboratory for immediate microbiological analysis.

Culture collection

Microorganisms used were obtained from the culture collection centre of the University College Hospital, Ibadan, and included *Escherichia coli, Candida albicans, Staphylococcus epidermidis, Staphylococcus aureus,*

Figure 1: Whole ginger, garlic and turmeric spices.

Pseudomonas aeruginosa, Haemophilus influenzae, Bacillus sp. *and Streptococcus pneumoniae.* The bacterial strains were resuscitated on nutrient agar (LabM, UK) before use.

Preparation of spice extracts

The modified method of Joe et al.[8] was used for the preparation of spices for extraction. The fresh spices were cleaned, peeled and subsequently washed in sterile distilled water (Figure 1). Samples were surface sterilised with 70% ethanol which was allowed to evaporate, after which the samples were rinsed with sterile distilled water. Samples (250 g of each) were cut into small sizes, crushed and blended, using a laboratory blender, to get a fine paste. The resulting mixture (of spices and sterile distilled water) was filtered through Whatman filter paper (No. 1) and sterilised using a membrane filter (0.45-μm filter unit, Merck). The filtrate was used as the 100% extract concentration. Appropriate volumes of sterile distilled water were mixed with the concentrate to obtain different concentrations (10%, 20%, 30% and 40%).

Antimicrobial sensitivity test

The agar well diffusion method was used[9] to test for the antimicrobial activity of the extracts. Test cultures were maintained on sterile nutrient broth for 18 h. They were further diluted out to 0.5 McFarland standard of approximately 1.5×10^8 CFU/mL. Cultures (0.1 mL) were aseptically inoculated on sterile Mueller–Hinton Agar (MHA; LabM, UK) plates and spread evenly, using a sterile cotton swab. Wells were made using a sterilised cork borer, and equal amounts (0.1 mL) of the different extracts (concentrated) were introduced into respective wells on the plates. Incubation was done in an upright position at 37 °C for 24 h, and the diameter of inhibition zones was measured (Figure 2).

Antibiotic sensitivity test

Cultures were analysed using a modified Kirby-Bauer Disc Diffusion method for antibiotic sensitivity. Augmentin (amoxicillin/clavulanate, 30 μg), ofloxacin (5 μg), cloxacillin (5 μg), erythromycin (5 μg), ceftriaxone (30 μg), cefuroxime (5 μg), ceftazidime (30 μg) and nitrofurantoin (5 μg) were obtained from Hi-media, India. Aliquots (0.1 mL) of 18-hour-old cultures (0.5 McFarland) were aseptically swabbed on sterile MHA plates. Forceps, sterilised by flaming, were used to aseptically place the antibiotic discs over the seeded MHA plates. The discs were placed accordingly to prevent overlapping of expected inhibition zones. Incubation was done in an upright position at 37 °C for 24 h, and the diameter of inhibition zones was measured.

Minimum inhibitory concentration

The broth dilution method[10] was used to determine the minimum inhibitory concentrations (MICs). Aliquots (0.5 mL) of each test organism (18-hour-old culture) were added to various concentrations (10%, 20%, 30% and 40% v/v) of the extracts, prepared by diluting with sterile distilled water, with a final
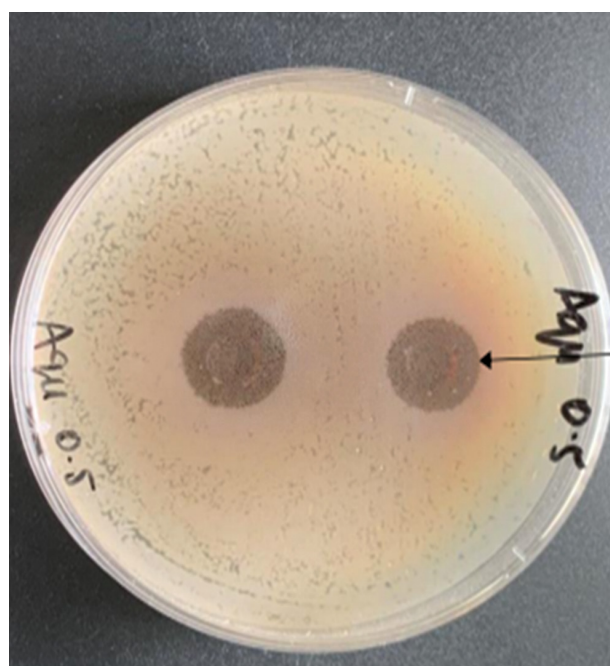

Figure 2: Representative plate showing zones of inhibition of the ginger aqueous extract against *Bacillus* sp.

volume of 5 mL. The cultures were incubated at 37 °C for 24 h. Numbers of cells were determined using a spectrophotometer (Cecil CE 1011, Cambridge, UK) at 540 nm and were compared to the initial cell numbers.

Analysis of bioactive compounds

Identification and quantification of organic compounds in the extract with the most antimicrobial activity was done using gas chromatography–mass spectroscopy (GC-MS).[11]

Statistical analysis

Data obtained were subjected to statistical analysis (SPSS) at a 5% level of significance and are presented as the mean of replicates $\pm$ standard deviation.

Results

The results of the agar well diffusion test indicate that extracts of garlic, ginger and turmeric showed different degrees of growth inhibition, depending on the bacterial strain (Table 1). The ginger extract showed the broadest antibacterial activity by inhibiting the growth of all bacterial strains except one, with the diameters of inhibition zones ranging between 0.9 mm and 13.5 mm. The highest inhibition zone recorded (13.5 mm) was

Table 1: Antimicrobial activities of spice extracts (100%) against selected microorganisms

Isolate	Spices / diameter of inhibition zone (mean values in mm)		
	Turmeric	Ginger	Garlic
Escherichia coli	–	12.5	11.5
Bacillus sp.	4.4	12.1	–
Staphylococcus aureus	6.1	5.8	10.6
Staphylococcus epidermidis	2.3	5.1	–
Streptococcus pneumoniae	–	–	–
Candida sp.	10.3	12.6	10.2
Haemophilus influenzae	–	0.9	–
Pseudomonas aeruginosa	10.9	13.5	4.7

– *No inhibition zone / resistant*

against *P. aeruginosa,* while *S. aureus, S. epidermidis, S. pneumoniae, H. influenzae* and *Salmonella* sp. had inhibition zones that were less than 6 mm, indicating sensitivity of the ginger to the growth of these organisms.

Only three microbial isolates (*S. aureus, Candida* sp. and *P. aeruginosa*) showed inhibition sensitivity, indicated by inhibition zones greater than 6 mm, with the turmeric extract, with inhibition zones of 6.1 mm, 10.3 mm and 10.9 mm, respectively. However, the inhibition zones observed for the garlic extract ranged between 4.7 mm and 11.5 mm, the highest recorded against *E. coli. Bacillus* sp.*, S. epidermidis, S. pneumonia* and *H. influenzae* were all resistant to the garlic extract. Overall, amongst the three spice extracts, ginger showed the highest inhibitory effect (13.5 mm), as well as a wide target range.

The sensitivity of different concentrations of the turmeric extract on the selected isolates is shown in Table 2. Against *E. coli*, the MIC of the extract was recorded at 10% with an optical density (OD) of 1.581 from an initial concentration of 1.642 OD at 540 nm. There was an increase in the OD of all other isolates, even when the concentration was increased from 10% to 40%.

The ginger extract, at 40%, was the MIC against *Bacillus* sp. with an OD of 1.499, while 30% was the MIC against *S. epidermidis, H. influenzae* and *P. aeruginosa,* with ODs at 540 nm of 1.490, 1.486 and 1.481, respectively. However, 10% was the MIC for the inhibition of *E. coli* (1.677). The concentrations used (10–40%) did not inhibit the growth of *S. aureus, S. pneumonia* and *C. albicans*. (Table 3).

Table 2: Effect of different concentrations of the turmeric extract on microbial growth

Test isolate	24-hour-old culture concentration	Extract concentration (%) / OD at 540 nm after incubation for 24 h			
		10	20	30	40
Escherichia coli	1.725	1.642 ± 0.00[c]	1.933 ± 0.00[a]	1.643 ± 0.00[c]	1.681 ± 0.00[b]
Bacillus sp.	1.525	1.915 ± 0.00[a]	1.682 ± 0.00[b]	1.670 ± 0.00[c]	1.564 ± 0.00[d]
Staphylococcus aureus	1.514	1.545 ± 0.00[a]	1.540 ± 0.00[b]	1.529 ± 0.00[c]	1.525 ± 0.00[c]
Staphylococcus epidermidis	1.501	1.550 ± 0.01[a]	1.541 ± 0.00[ab]	1.541 ± 0.01[ab]	1.533 ± 0.00[b]
Streptococcus pneumoniae	1.477	1.575 ± 0.00[a]	1.573 ± 0.00[a]	1.571 ± 0.01[a]	1.505 ± 0.00[b]
Candida sp.	1.440	1.732 ± 0.00[a]	1.707 ± 0.00[b]	1.619 ± 0.00[d]	1.660 ± 0.00[c]
Haemophilus influenzae	1.511	1.708 ± 0.00[a]	1.670 ± 0.00[c]	1.686 ± 0.00[b]	1.648 ± 0.00[d]
Pseudomonas aeruginosa	1.497	1.732 ± 0.02[a]	1.699 ± 0.00[b]	1.697 ± 0.01[b]	1.689 ± 0.00[c]

Values are mean ± standard deviation of three replicates

Means reported with the same superscript in each row indicate no significant difference ($p \leq 0.05$)

Table 3: Effect of different concentrations of the ginger extract on microbial growth

Test isolate	24-hour-old culture concentration	Extract concentration (%) / OD at 540 nm after incubation for 24 h			
		10	20	30	40
Escherichia coli	1.725	1.677 ± 0.00[a]	1.645 ± 0.00[b]	1.645 ± 0.01[b]	1.488 ± 0.00[c]
Bacillus sp.	1.525	1.661 ± 0.00[b]	1.680 ± 0.00[a]	1.661 ± 0.00[b]	1.499 ± 0.00[c]
Staphylococcus aureus	1.514	1.667 ± 0.00[a]	1.561 ± 0.00[b]	1.549 ± 0.00[c]	1.533 ± 0.00[d]
Staphylococcus epidermidis	1.501	1.575 ± 0.01[a]	1.575 ± 0.01[a]	1.490 ± 0.02[b]	1.477 ± 0.00[c]
Streptococcus pneumoniae	1.477	1.549 ± 0.00[c]	1.627 ± 0.00[a]	1.578 ± 0.00[b]	1.481 ± 0.00[d]
Candida sp.	1.440	1.670 ± 0.01[a]	1.573 ± 0.00[b]	1.573 ± 0.00[b]	1.490 ± 0.00[c]
Haemophilus influenzae	1.511	1.664 ± 0.01[a]	1.516 ± 0.00[b]	1.486 ± 0.00[c]	1.479 ± 0.00[c]
Pseudomonas aeruginosa	1.497	1.674 ± 0.00[a]	1.510 ± 0.00[b]	1.481 ± 0.00[c]	1.360 ± 0.01[d]

Values are mean ± standard deviation of three replicates

Means reported with the same superscript in each row indicate no significant difference ($p \leq 0.05$)

Table 4: Effect of different concentrations of the garlic extract on microbial growth

Test isolate	24-hour-old culture concentration	MIC (%) at 540 nm after incubation for 24 h			
		10	**20**	**30**	**40**
Escherichia coli	1.725	1.666 ± 0.00[c]	1.782 ± 0.00[a]	1.796 ± 0.00[a]	1.747 ± 0.01[b]
Bacillus sp.	1.525	1.661 ± 0.01[d]	1.670 ± 0.00[c]	1.816 ± 0.00[a]	1.725 ± 0.00[b]
Staphylococcus aureus	1.514	1.697 ± 0.00[c]	1.854 ± 0.00[a]	1.706 ± 0.00[b]	1.654 ± 0.00[d]
Staphylococcus epidermidis	1.501	1.781 ± 0.00[b]	1.835 ± 0.00[a]	1.728 ± 0.00[c]	1.649 ± 0.01[d]
Streptococcus pneumoniae	1.477	1.703 ± 0.00[b]	1.778 ± 0.00[a]	1.699 ± 0.01[c]	1.660 ± 0.00[d]
Candida sp.	1.440	1.826 ± 0.01[a]	1.713 ± 0.00[b]	1.674 ± 0.00[c]	1.661 ± 0.00[d]
Haemophilus influenzae	1.511	1.778 ± 0.00[a]	1.747 ± 0.00[b]	1.781 ± 0.00[a]	1.646 ± 0.00[c]
Pseudomonas aeruginosa	1.497	1.648 ± 0.01[d]	1.848 ± 0.00[a]	1.745 ± 0.00[b]	1.682 ± 0.00[c]

Values are mean ± standard deviation of three replicates

Means reported with the same superscript in each row indicate no significant difference ($p \leq 0.05$)

Table 5: Minimum inhibitory concentrations (MIC) for antimicrobial activity of spice extracts

Spice extract	MIC (%)							
	E. coli	*Bacillus* sp.	*S. aureus*	*S. epidermidis*	*S. pneumoniae*	*Candida* sp.	*H. influenzae*	*P. aeruginosa*
Ginger	10	40	–	40	–	–	40	40
Turmeric	10	–	–	–	–	–	–	–
Garlic	10	–	–	–	–	–	–	–

Table 6: Antibiotic susceptibility spectrum of selected isolates

Antibiotic	Test isolates / zone of inhibition (mm)							
	E. coli	*Bacillus* sp.	*S. aureus*	*S. epidermidis*	*S. pneumoniae*	*Candida* sp.	*H. influenzae*	*P. aeruginosa*
Augmentin	11.7	–	10.8	12.4	11.6	10.6	–	12.1
Ofloxacin	–	–	10.6	10.6	–	10.8	–	10.9
Cloxacillin	10.6	–	–	–	–	–	–	–
Erythromycin	10.2	–	10.6	–	–	10.7	13.1	10.5
Ceftriaxone	12.1	–	11.8	11.6	–	–	–	10.4
Cefuroxime	–	–	10.9	10.6	–	11.6	–	–
Ceftazidime	–	–	–	–	11.6	10.6	10.9	–
Nitrofurantoin	11.6	10.5	10.5	10.7	11.2	10.6	–	10.8

– No inhibition zone / resistant

Similar to the turmeric extract, only *E. coli* was inhibited at a 10% concentration of the garlic extract, with an OD of 1.666 at 540 nm. All other isolates were not inhibited by the extract concentrations (10–40%) used (Tables 4 and 5).

The antibiotic sensitivity pattern of selected isolates (Table 6) indicates that *Bacillus* sp. was resistant to all but one antibiotic (nitrofurantoin). *S. aureus* and *Candida* sp. were resistant to two of the tested antibiotics, cloxacillin for both, then ceftazidime and ceftriaxone, respectively. Nitrofurantoin, Augmentin and erythromycin showed sensitivity against a wide range of isolates; however, the highest inhibition zone (13.1 mm) was recorded for erythromycin against *H. influenzae*.

The ginger aqueous extract that showed the highest inhibitory effect was analysed for the phyto-active compounds present using GC-MS, and the results are presented in Figures 3–6. This extract generated 27 constituents, with major ones at peaks 18 (peak area 17.70%), 11 (peak area 13.30%) and 23 (peak area 10.84%) comprising 2-Butanone,4-(4-hydroxy-3-methoxyphenyl), 1,3-Cyclohexadiene and 1-(4-Hydroxy-3-methoxyphenyl), respectively.
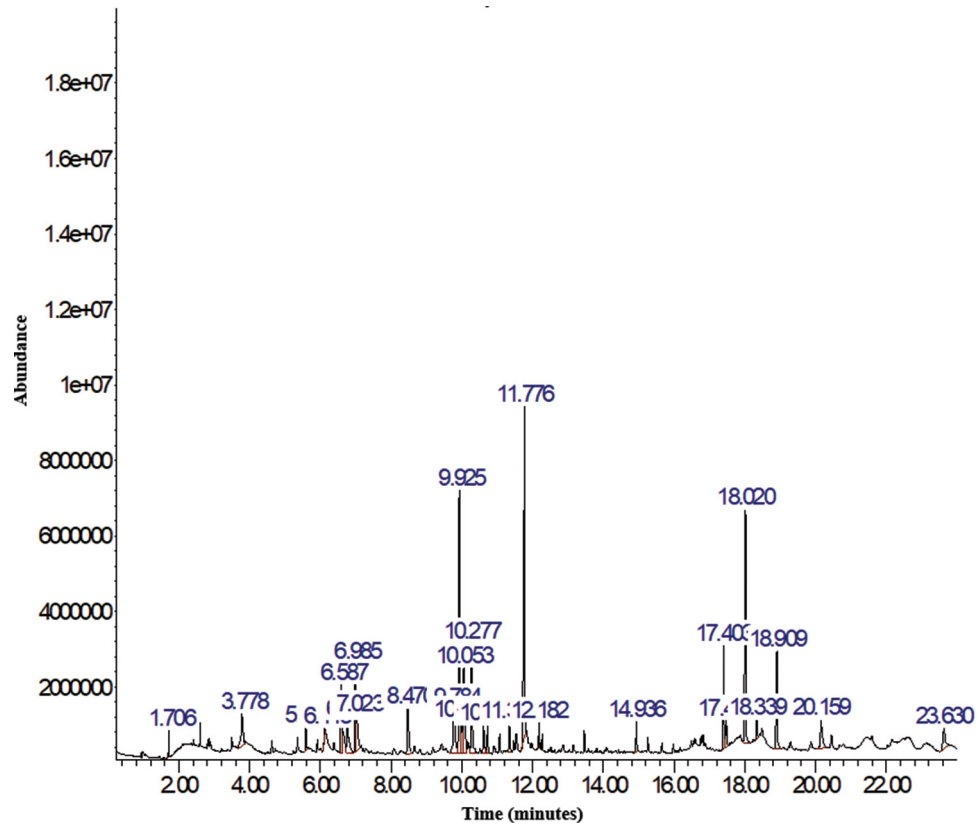
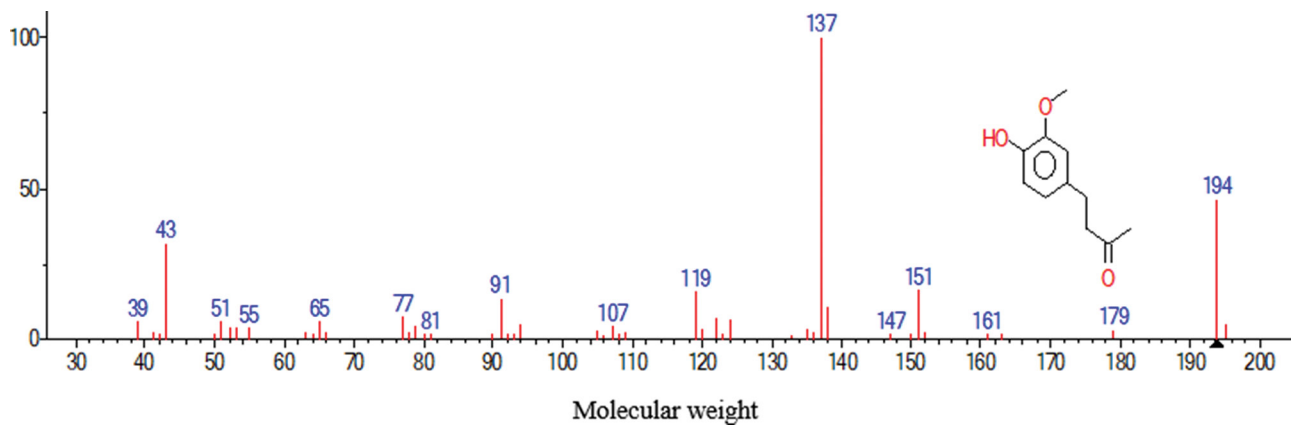**Figure 3:** Total ion chromatogram of the ginger extract.

Figure 4: Chromatogram showing 2-Butanone,4-(4-hydroxy-3-methoxyphenyl) (% composition by area = 17.70; molecular weight 194).
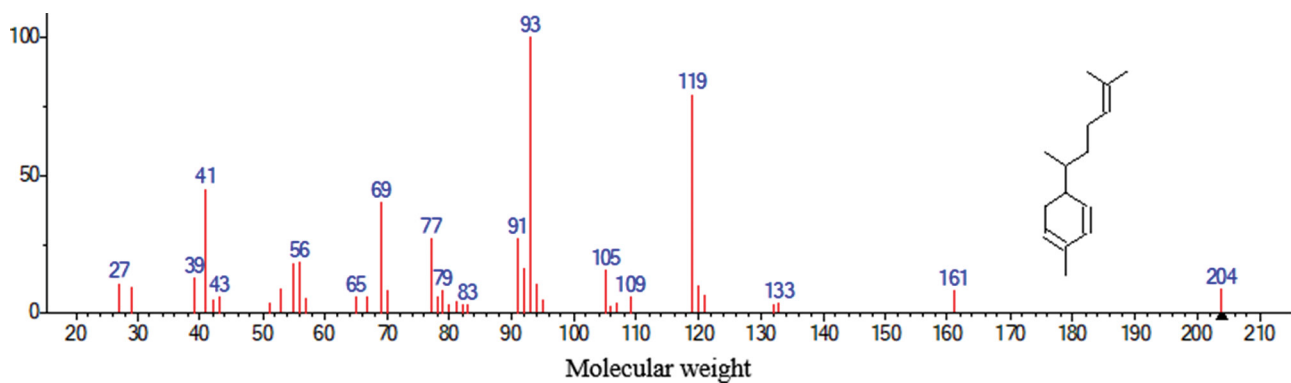

Figure 5: Chromatogram showing 1,3-Cyclohexadiene (% composition by area = 13.30; molecular weight 204).
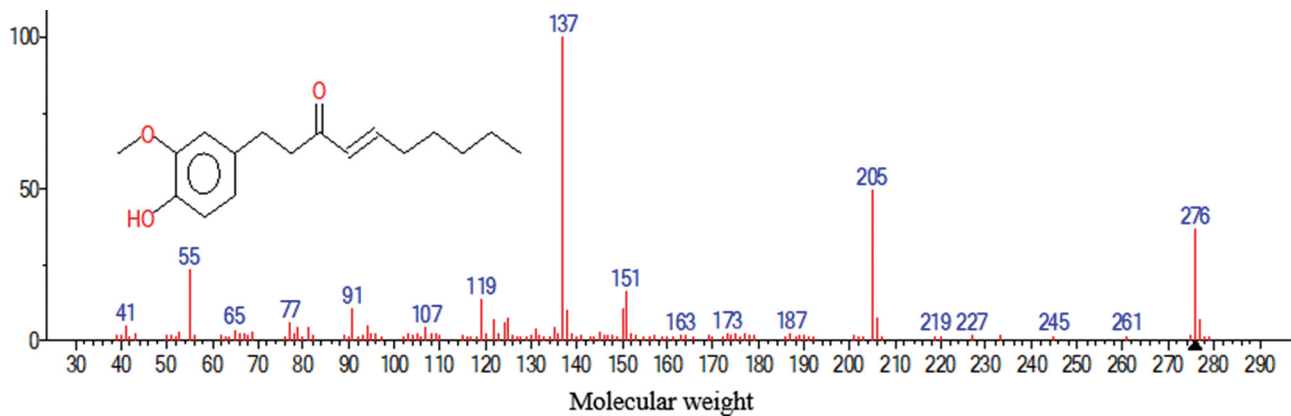
Figure 6: Chromatogram showing 1-(4-Hydroxy-3-methoxyphenyl) (% composition by area = 10.84; molecular weight 276).

Discussion

Use of spice and plant extracts has generally gained wide acceptability[11-14] because they are easy to use and highly efficient. Spices have beneficial biological functions such as bactericidal, bacteriostatic, fungistatic, anthelmintic, medicinal and flavouring properties.[15] In this study, not all the spice extracts showed inhibitory activity against all the test organisms. In contrast to our results which show that ginger exhibited the highest antimicrobial effect on test isolates, it was previously reported[13] that extracts of garlic had a strong effect against some pathogens, turmeric extracts exhibited a weak effect, and ginger extracts showed no inhibitory effect against any test bacteria during a study on the efficacy of spice extracts on bacterial isolates from meat products. It has been established that the antimicrobial activity of different spices may vary[7] towards different microorganisms, as the mechanism of microbial susceptibility is not predictable. The mechanism of action could be inhibition of cell wall synthesis, depolarisation of the cell membrane, inhibition of protein synthesis, inhibition of nuclei acid synthesis or inhibition of metabolic pathways.[16] The method followed may also result in such disparity, as reported in another study in which the antibacterial activity of spices was less evident when the paper disc method was used instead of the agar well assay method.[17]

Many studies have evaluated the antimicrobial activity of ginger, which has been shown to have a promising inhibitory effect against some pathogenic bacteria and fungi.[18-21] Ginger extracts were reported to have exerted an inhibitory effect against *Bacillus cereus, S. aureus and P. aeruginosa*.[14-21] In a study conducted using different spices (cinnamon, black pepper, cloves, turmeric and ajwain), antimicrobial zones of inhibition against *S. aureus* and *Klebsiella pneumoniae* ranged between 4 mm and 15 mm.[22] The maximum inhibition zone (13.5 mm) in the present study was within this reported range.

The mechanism of inhibitory action is dependent on the spices used. The major antimicrobial compound in garlic is allicin[23], while ginger has zingerone and gingerol. All have been reported to have strong inhibitory activities.[24] However, the general hypothesis is the integration of phenolic compounds to membrane proteins, thus leading to partitions in the lipid bilayer and effects on permeability, and subsequent membrane disruption.[25] The presence of lipophilic oils and hydrophilic antioxidants also contributes to good antibacterial activity.[26-28]

The minimum inhibitory concentration test against selected pathogens indicated a 10% MIC against *E. coli* for turmeric, garlic and ginger extracts. However, at 40%, ginger also inhibited the growth of a few other isolates, whereas most isolates were not inhibited by the garlic and turmeric concentrations used. Albaridi and Yehia[14] reported a 100 mg/mL concentration of ginger extract as the MIC against some selected pathogens. Different microorganisms responded differently to spice extracts at different concentrations.[8]

When comparing the inhibitory effect of these spices with those of different antibiotics, ginger particularly exhibited higher activity against *E. coli, Bacillus* sp*., Candida* sp. *and P. aeruginosa* than did the different antibiotics used. However, the inhibitory effect against *S. aureus, S. epidermidis, S. pneumoniae and H. influenzae* was more evident when antibiotics were used. Thus, ginger could be a potential source for drug development.

More than 60 bioactive constituents (volatile and nonvolatile compounds) are known to be present in ginger[29]; however, 2-Butanone,4-(4-hydroxy-3-methoxyphenyl), a phyto-active component of ginger, with a major peak area as revealed on the chromatogram, has been reported to have antioxidant, anti-inflammatory, anticancer and antimicrobial activities[30-33]. 1,3-Cyclohexadiene and 1-(4-Hydroxy-3-methoxyphenyl) have also been reported to have exhibited antimicrobial effects.[34] Ginger can thus be said to be of medicinal importance, rather than being highly and widely recognised only as a flavouring agent. It could be widely used in mitigating many human and other animal diseases.

Conclusion

Our findings demonstrate that different spices used have significant and varied activity against diverse organisms, indicating that natural products like ginger, turmeric and garlic have antimicrobial properties. Ginger showed the highest inhibitory effect against a wide range of isolates. A 10% minimum inhibitory concentration of all spices used inhibited the growth of *E. coli*, while concentrations up to 40% did not inhibit most isolates. Bioactive compounds of biomedical importance were present in the analysed ginger extract. Spices can be used as alternative natural food preservatives against spoilage organisms and as potential natural sources for drug development and use against pathogens.

Competing interests

We have no competing interests to declare.

Authors' contributions

K.A.O.: Conceptualisation; methodology; data collection; data analysis; writing the initial draft and revisions; project leadership. G.E.O.: Conceptualisation; methodology; data collection; validation; project management. A.I.B.: Sample analysis; methodology; student supervision. O.J.J.: Validation; data analysis; writing revisions.

References

1. Sung B, Prasad S, Yadav VR, Aggarwal BB. Cancer cell signaling pathways targeted by spice-derived nutraceuticals. Nutr Cancer. 2012;64:173197. https://doi.org/10.1080/01635581.2012.630551

2. Panpatil VV, Tattari S, Kota N, Nimgulkar C, Polasa K. In vitro evaluation of antioxidant and antimicrobial activity of spice extracts of ginger, turmeric and garlic. J Pharmacogn Phytochem. 2013;2(3):143–148.

3. Maharjan R, Thapa S, Acharya A. Evaluation of antimicrobial activity and synergistic effect of spices against few selected pathogens. Tribhuvan Univ J Microbiol. 2019;6(1):1018. https://doi.org/10.3126/tujm.v6i0.26573

4. Tajkarimi MM, Ibrahim SA, Cliver DO. Antimicrobial herb and spice compounds in food. Food Control. 2010;21(9):1199–1218. https://doi.org/10.1016/j.foodcont.2010.02.003

5. Feldberg RS, Chang SC, Kotik AN, Nadler M, Neuwirth Z, Sundstrom DC, et al. In vitro mechanism of inhibition of bacterial cell growth by allicin. Antimicrob Agents Chemother. 1989;32(12):1763–1768. https://doi.org/10.1128/AAC.32.12.1763

6. Hora SL, Nair KK. Pollution of streams and conservation of fisheries. Proc Natl Inst Sci India. 1994;10:147–166

7. Hyldgaard M, Myging T, Meyer RL. Essential oils in food preservation: Mode of action, synergies, and interactions with food matrix components. Front Microbiol. 2012;3:12. https://doi.org/10.3389/fmicb.2012.00012

8. Joe MM, Jayachitra J, Vijayapriya M. Antimicrobial activity of some common spices against certain human pathogens. J Med Plants Res. 2009;3(11):1134–1136. https://doi.org/10.1093/ajcp/45.4_ts.493

9. Baur AW, Kirby WMM, Sherris JC, Trunk M. Antibiotic susceptibility testing by standardized single disc method. Am J Clin Path. 1966;45:493–496.

10. Jorgensen JH, Turnidge JD. Susceptibility test methods: Dilution and disk diffusion methods. In: Jorgensen JH, Carroll KC, Funke G, Pfaller MA, Landry ML, Richter SS, et al., editors. Manual of clinical microbiology. 11th ed. Washington DC: ASM Press; 2015. p. 1253–1273. https://doi.org/10.1128/9781555817381.ch71

11. Gavamukulya Y, Abou-Ellela F, Wamunyokoli F, El-Shemy HA. GC-MS analysis of bioactive phytochemicals present in ethanolic extracts of leaves of *Annona muricata*: A further evidence for its medicinal diversity. Pharmacogn J. 2015;7(5):300–304. https://doi.org/10.5530/pj.2015.5.9

12. Shamsuddeen U, Ameh JB, Oyeyi TI, Dantata AA. Study on the phytochemical and in vitro antibacterial activity of some spice extracts on some bacteria isolated from meat products. J Pure Appl Sci. 2009;2(1):101–104. https://doi.org/10.4314/bajopas.v2i1.58479

13. Sah SN, Khanal H, Acharya DR. Antibacterial activity of common spices extracts on bacterial isolates found in Kachhila, a Newari cuisine. Tribhuvan Univ J Microbiol. 2020;7(1):8–18. https://doi.org/10.3126/tujm.v7i0.33785

14. Albaridi NA, Yehia H. The real role of select herb and spice extracts against *Bacillus cereus* ATCC 14579 growth in cooked rice. Food Sci Tech. 2021;42. https://doi.org/10.1590/fst.08521

15. Pruthi JS. Spices and condiments: Chemistry, microbiology, biotechnology. Adv Food Res Suppl. New York: Academic Press; 1980.

16. Reygaert WC. An overview of the antimicrobial resistance mechanisms of bacteria. AIMS Microbiol. 2018;4(3):482–501. https://doi.org/10.3934/microbiol.2018.3.482

17. Indu MM, Hatha AAM, Abirosh C, Harsha U, Vivekanandan G. Antimicrobial activity of some of the South-Indian spices against serotypes of *Escherichia coli*, *Salmonella*, *Listeria monocytogenes* and *Aeromonas hydrophila*. Braz J Microbiol. 2006;37:153–158. https://doi.org/10.1590/S1517-83822006000200011

18. Mascolo N, Jain R, Jain SC, Capasso F. Ethnopharmacologic investigation of ginger (*Zingiber officinale*). J Ethnopharm. 1989;27(1–2):129–140. https://doi.org/10.1016/0378-8741(89)90085-8

19. Thongson C, Davidson PM, Mahakarnchanakul W, Vibulsresth P. Antimicrobial effect of Thai spices against *Listeria monocytogenes* and *Salmonella typhimurium* DT104. J Food Prot. 2005;68(10):2054–2058. https://doi.org/10.4315/0362-028X-68.10.2054

20. Mahady GB, Pendland SL, Stoia A, Hamill FA, Fabricant D, Dietz BM, et al. In vitro susceptibility of helicobacter pylori to botanical extracts used traditionally for the treatment of gastrointestinal disorders. Phytother Res. 2005;19(11):988–991. https://doi.org/10.1002/ptr.1776

21. Kaushik P, Goyal P. Evaluation of various crude extracts of *Zingiber officinale* rhizome for potential antibacterial activity: A study in vitro. Adv Appl Microbiol. 2011;1(1):7–12. https://doi.org/10.4236/aim.2011.11002

22. Harmalkar MN, Desai P. Comparative assessment of antibacterial activity of ginger extracts with antimicrobial agents. Intl J Pharmacol Biosci. 2011;5:33–38.

23. Wendakoon CN, Sakaguchi M. Inhibition of amino acid decarboxylase activity of *Enterobacter aerogenes* by active components in spices. J Food Prot. 1995;58(3):280–283. https://doi.org/10.4315/0362-028X-58.3.280

24. Conner DE. Naturally occurring compounds. In: Davidson PM, Branen AL, editors. Antimicrobials in foods. New York: Marcel Dekker Inc.; 1993. p. 441–468.

25. Hirasa K, Takemasa M. Spice science and technology. New York: Marcel Dekker Inc.; 1998. https://doi.org/10.1201/9780367800451

26. Pundir RK, Jain P. Comparative studies on the antimicrobial activity of black pepper (*Piper nigrum*) and turmeric (*Curcuma longa*) extracts. Int J Appl Biol Pharm Technol. 2010;1(2):492–501.

27. Masuda H, Hironaka S, Matsui Y, Hirooka S, Hirai M, Hirata Y, et al. Comparative study of the antioxidative activity of culinary herbs and spices, and hepatoprotective effects of three selected Lamiaceae plants on carbon tetrachloride induced oxidative stress in rats. Food Sci Technol Res. 2015;21:407–418. https://doi.org/10.3136/fstr.21.407

28. Coccimigli, J, Alipour M, Jiang ZH, Gottardo C, Suntres Z. Antioxidant, antibacterial, and cytotoxic activities of the ethanolic *Origanum vulgare* extract and its major constituents. Oxid Med Cell Longev. 2016;2016:1404505. https://doi.org/10.1155/2016/1404505

29. Bilal A, Muneeb U, Rehman I, Ahmad A, Saiema R, et al. A review on pharmacological properties of zingerone (4-(4-Hydroxy-3-methoxyphenyl)-2-butanone). Sci World J. 2015;2015:816364. https://doi.org/10.1155/2015/816364

30. Kim MK, Chung SW, Kim JM, Kim DH, Lee EK, Kim JY, et al. Modulation of age-related NF-κB activation by dietary zingerone via MAPK pathway. Exp Gerontol. 2010;45(6):419–426. https://doi.org/10.1016/j.exger.2010.03.005

31. Kumar L, Chhibber S, Harjai K. Zingerone inhibit biofilm formation and improve antibiofilm efficacy of ciprofloxacin against *Pseudomonas aeruginosa* PAO1. Fitoterapia. 2013;90:73–78. https://doi.org/10.1016/j.fitote.2013.06.017

32. Rajan I, Narayanan N, Rabindran R, Jayasree PR, Manish Kumar PR. Zingerone protects against stannous chloride-induced and hydrogen peroxide-induced oxidative DNA damage in vitro. Biological Trace Element Res. 2013;155(3):455–459. https://doi.org/10.1007/s12011-013-9801-x

33. Vinothkumar R, Sudha M, Nalini N. Chemopreventive effect of zingerone against colon carcinogenesis induced by 1,2-dimethylhydrazine in rats. Eur J Cancer Prev. 2014;23(5):361–371. https://doi.org/10.1097/CEJ.0b013e32836473ac

34. Tobal IE, Bautista R, Diez D, Garrido NM, García-García P. 1,3-Cyclohexadien-1-als: Synthesis, reactivity and bioactivities. Molecules. 2021;26(6):1772. https://doi.org/10.3390/molecules26061772

AUTHORS:
Domonique C. Basson[1]
Sandy van Vuuren[2]
Ida M. Risenga[1]

AFFILIATIONS:
[1]School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, Johannesburg, South Africa

[2]Department of Pharmacy and Pharmacology, University of the Witwatersrand, Johannesburg, South Africa

CORRESPONDENCE TO:
Ida Risenga

EMAIL:
Ida.Risenga@wits.ac.za

The effect of elevated carbon dioxide on the medicinal properties of *Portulacaria afra*

There is a global concern that rising atmospheric CO_2 concentrations may impact the medicinal or nutritional properties of medicinal plants. *Portulacaria afra* is a South African medicinal plant that is used by traditional healers to treat various skin conditions. The aim of this study was to determine whether elevated CO_2 concentrations would affect the medicinal properties of the leaves of *P. afra*. This was achieved by comparing the phytochemical presence, antioxidant and antimicrobial activity of the leaves of *P. afra* which were exposed to ambient (420 ppm) and elevated (600 ppm) CO_2 concentrations of plants grown in greenhouse conditions. The results revealed that leaf samples that were exposed to elevated CO_2 concentrations exhibited a significant increase in flavonoid presence, compared to the control group. The antioxidant activity of the leaves of *P. afra* (DPPH activity) remained mostly unchanged in the samples that were exposed to elevated CO_2 concentrations. The antimicrobial activity efficacy against *Cutibacterium acnes* increased with increasing global atmospheric CO_2 concentration. These findings suggest that *P. afra* is a resilient medicinal plant and that its leaves may continue to provide relief against certain ailments, despite rising atmospheric CO_2 concentrations.

Significance:

- *Portulacaria afra* is a South African medicinally important species that shows great resilience against elevated CO_2 concentrations.

- It is important to anticipate how changing environmental factors, such as rising CO_2 concentrations, may affect natural resources.

- The phytochemical profile and antioxidant and antimicrobial activities of the various plant parts either remained the same or increased after exposure to an elevated CO_2 concentration of 600 ppm.

Introduction

Prior to the industrialisation era, the atmospheric carbon dioxide (CO_2) concentration remained relatively stable at 280 ± 10 parts per million (ppm).[1] In the year 2022, the atmospheric CO_2 concentration reached an average of 417.2 ppm, representing a more than 50% increase. Currently, the energy sector remains heavily reliant on fossil fuels, and as a result, this value is predicted to continue rising at an unprecedented rate, potentially reaching 600 ppm by 2050.[2]

Carbon dioxide is an essential resource for the growth, development and overall survival of plants, as it forms a major component of photosynthesis – which produces sugars, carbohydrates and other organic molecules necessary for plant function.[3] As CO_2 is a limiting factor in photosynthesis, the rise in atmospheric CO_2 may result in an increase in the photosynthetic pathway. Previous studies have shown that elevated CO_2 concentrations can lead to an alteration of primary and secondary metabolism, resulting in an increase of up to 44% in plant biomass.[4] However, the alteration to secondary metabolites may result in an unequal distribution of chemicals throughout the plant. The increase in CO_2 concentration results in an accumulation of carbon-based phytochemicals, such as phenolics, phlobatannins and flavonoids.[3]

This increase in carbon-based molecules can lead to a decrease in the partitioning and allocation of other molecules to plant organs such as leaves.[3] Scientists warn that the alteration in the chemical composition of the plant may affect the potency of the medicinal properties of the plant.[5]

In a previous study, *Asclepias curassavica* was subjected to elevated CO_2 concentrations and ultimately exhibited a decrease in medicinal properties. However, this is not a consistent trend in the literature.[6] The effect of elevated CO_2 concentrations varies with species exhibiting a negative impact, a positive impact or no impact at all.[7] Thus, more studies need to be conducted on the effect of rising CO_2 concentrations on various medicinal plant species.

Portulacaria afra is a medicinal plant species endemic to South Africa and is also found in neighbouring countries such as Eswatini and Mozambique.[8] Despite being primarily known for its carbon sequestration capabilities,[9] ethnobotanical interviews conducted in 2013 revealed that *P. afra* is used traditionally to treat various skin conditions.[10] The small succulent leaves are predominantly used topically to treat conditions such as rashes, ringworm and acne.[11]

Given that a large population still relies on medicinal plants as a primary source of health care, it is imperative to determine the effect that rising CO_2 concentrations may have on these resources. In this study, we aimed to assess the effects of an elevated atmospheric CO_2 concentration on the physiological properties, phytochemical profile and biological activity of the leaves of *P. afra*. In order to accomplish this, *P. afra* plants were exposed to both ambient CO_2 concentration (420 ppm) and elevated CO_2 concentration (600 ppm) to determine whether there was a significant difference in the medicinal properties of the plant. Indicators of medicinal properties used in this study include the presence of 10 phytochemical groups, antioxidant activity against DPPH and H_2O_2, and antimicrobial properties against pathogens commonly associated with skin infections.

Materials and methods

Plant material

Portulacaria afra cuttings that exhibited no signs of wilting or disease were collected from the University of the Witwatersrand (26.1929° S, 28.0305° E) in December 2021 (South African summer time). The plants were identified by a botanist (IMR). A voucher specimen was deposited in the institute's herbarium (IMR001). The 90 *P. afra* plants were maintained in the Oppenheimer Life Sciences greenhouse, at the University of the Witwatersrand.

Treatment

Of the 90 plants that were maintained in the greenhouse, 30 were harvested and directly subjected to analysis as control samples. The remaining 60 plants were transferred to a Conviron® climate simulator to expose the plants to varying CO_2 concentrations. Of these, 30 were exposed to a CO_2 concentration of 420 ppm and the remaining 30 plants were exposed to a CO_2 concentration of 600 ppm. The temperature, light and relative humidity remained ambient and constant between the treatments. The plants were watered every second or third day, with approximately 200 mL water.

Each month, for a period of 3 months, 10 plants were harvested from each chamber and subjected to analysis.

Chlorophyll content

The fresh leaves (3 g) were mixed with 10 mL of 80% methanol and incubated in the dark for 24 h. The supernatant was then collected, and absorbance was measured using a spectrophotometer (Genesys 10S UV-VIS) at 645 nm and 663 nm to obtain the absorbance values. The chlorophyll content was calculated using Equation 1:

$$20.21 \times A_{645} + 8,02 \times A_{663} \hspace{2cm} \text{Equation 1}$$

This method followed the procedure outlined by Liu et al.[12]

Extraction

The crude plant extract used to determine the phytochemical presence and antioxidant activity was prepared using distilled water at 40 °C, 80% methanol, dichloromethane or n-hexane. These solvents were chosen due to the varying polarities between the solvents. The fresh leaves were placed in an air dryer at 40 °C for the duration of a week, or until sufficiently dry. Once sufficiently dried, the plant material was placed into a mechanical grinder to create powdered plant material. A mixture of 3 g of the powdered plant material and 30 mL of the respective solvent was placed on a mechanical shaker for 48 h. The hot water extracts were placed on a hot plate set at 40 °C and stirred with a magnetic stirrer. The extracts were filtered through a Whatman No. 2 filter paper. The plant extracts were stored at 4 °C prior to phytochemical and antioxidant analysis. The crude plant extracts used to determine the antimicrobial activity of *P. afra* were created by mixing powdered plant with dichloromethane: methanol at a ratio of 1:1.[13]

Preliminary phytochemical screening

A qualitative analysis was conducted to determine the presence or absence of 10 phytochemical groups in the leaves of *P. afra* using standard test methods.[14,15] The 10 phytochemical groups considered were saponins, phenolics, flavonoids, glycosides, tannins, terpenoids, steroids, coumarins, phlobatannins and volatile oils.

Total flavonoid content

An aluminium chloride colourimetric assay was used to determine the total flavonoid content in the leaves, stems and roots of *P. afra*.[16]

In a test tube, 3 mL of the crude plant extracts were mixed with 4 mL of 5% sodium nitrate. The mixture was incubated for 5 min. Following the incubation period, 3 mL of 10% aluminium chloride was added to the sample and incubated for an additional 6 min. A sodium hydroxide solution with a volume of 2 mL was then added to the sample. The sample was topped with 0.7 mL of distilled water to create a mixture

of 10 mL. The test tubes were incubated at ambient temperature for 1 h. The absorbance of each mixture was measured at 510 nm using the Genesys 10S UV-VIS spectrophotometer.

The test was conducted in triplicate. A calibration curve was created using a quercetin standard. The total flavonoid content was calculated from the calibration curve using Equation 2:

$$Y = 0.2388x - 0.0019 \hspace{2cm} \text{Equation 2}$$

Antioxidant activity

DPPH scavenging activity

The stable radical DPPH was used to determine the scavenging activity and hence the antioxidant activity of the leaves of *P. afra*.[17] Varying volumes (10–50 µL) of the crude plant extract were placed in capped test tubes and mixed with 700 µL of the DPPH work solution. In addition, 80% methanol was used to create a mixture of 1 mL. The samples were incubated in the dark, at ambient temperature for 45 min. Following the incubation period, the mixtures were placed into cuvettes that were placed within a spectrophotometer (Genesys 10S UV-VIS). A cuvette containing 1 mL of 80% methanol was used as the 'blank' to zero the spectrophotometer. The absorbance was measured at 517 nm against a blank. The test was done in triplicate and the percentage of inhibition was calculated using Equation 3:

$$\% \text{ Inhibition of DPPH radical} = \left([A_{br} - A_{ar}]/A_{br}\right) \times 100 \hspace{1cm} \text{Equation 3}$$

where A_{br} is the absorbance of the control and A_{ar} is the absorbance of the sample. The concentrations were then plotted against percentage inhibition values. Microsoft Excel was used to perform a linear regression to obtain the IC_{50}. R-studio was used to perform a repeated measures ANOVA. All statistical tests were conducted at a significance level of 0.05.

H_2O_2 scavenging activity

The free radical H_2O_2 was used to determine the scavenging activity and hence the antioxidant activity of leaves of *P. afra*.[17] A 40 mM solution of H_2O_2 (hydrogen peroxide) was prepared in a phosphate buffer (pH 7.4). The crude plant extracts (10–50 µL) were placed in vials and mixed with 600 µL of the hydrogen peroxide solution. The vials were then placed in the dark and incubated for 10 min. The absorbance of the mixtures was determined using a spectrophotometer (Genesys 10S UV-VIS) at 230 nm. A cuvette containing 1 mL of the phosphate buffer without hydrogen peroxide was used as the 'blank' to zero the spectrophotometer. This test was undertaken in triplicate. The different absorbances were recorded. The percentage of hydrogen peroxide scavenging activities was calculated using Equation 4:

$$\% \text{ Scavenged } [H_2O_2] = \left([A_{br} - A_{ar}]/A_{br}\right) \times 100 \hspace{1cm} \text{Equation 4}$$

where A_{br} is the absorbance of the control and A_{ar} is the absorbance of the sample. The concentrations were then plotted against percentage inhibition values. Microsoft Excel was used to perform a linear regression to obtain the IC_{50}. R-studio was used to perform a repeated measures ANOVA. All statistical tests were conducted at a significance level of 0.05.

Antimicrobial assay

Culture preparations

Bacterial microorganisms selected for the study were *Staphylococcus aureus* (ATCC 25923), *Staphylococcus epidermidis* (ATCC 1228), *Cutibacterium acnes* (ATCC 29212), *Pseudomonas aeruginosa* (ATCC 27853), and *Klebsiella aerogenes* (ATCC 13048). The yeast microorganism was *Candida albicans* (ATCC 10231). These microorganism strains are all pathogens pertaining to skin infections and were obtained from the American Type Culture Collection from the Department of Pharmacy and Pharmacology, University of the Witwatersrand, Johannesburg.

Minimum inhibitory concentration antimicrobial assay

The twofold serial dilution microdilution technique was used to determine the lowest concentration of the plant extracts that would inhibit the growth of the selected microorganisms, that is, the minimum inhibitory concentration (MIC). A 96-well microtitre plate was prepared by adding 100 μL of the respective broth into the wells. A volume of 100 μL of the plant extract was placed into the first well (A) using an aseptic technique. A positive control (ciprofloxacin for bacteria or nystatin for the yeast), negative control (dichloromethane: methanol) and culture control (TSB or TGB) were included. The positive control was used to detect the microbial susceptibility of the pathogens. The negative control was used to detect whether the pathogen was reacting to the solvent. The culture control was to ensure that microbial growth did occur.

A serial dilution was conducted at concentrations of 8, 4, 2, 0.5, 0.25, 0.125 and 0.0625 mg/mL. A 0.5 McFarland turbidity standard was created by mixing the bacterial culture with the broth at a ratio of 1:100. A volume of 100 μL of the culture was then placed in each well. The microtitre plates were sealed with a sterile adhesive film and incubated at the respective optimum conditions. The optimum conditions for bacterial growth were 37 °C for 24 h, with the exception of *C. acnes* which was incubated at 37 °C for 96 h. *C. albicans* was incubated at 37 °C for 48 h.

Following the incubation periods, 40 μL of *p*-iodonitrotetrazolium violet solution (INT) was added to all the wells in the microtitre plate. The wells that exhibited microbial growth (shown by a purple-pink colour) were recorded, and the MIC was taken as the lowest concentration inhibiting microbial growth.

Results and discussion

Weight and chlorophyll content

The weight of the leaves increased in the plants that were exposed to both 420 ppm and 600 ppm of CO_2 (Table 1). However, the leaves exposed to 420 ppm exhibited a weight increase of 23.6% ($p > 0.05$), whereas the weight of the leaves exposed to an elevated CO_2 exhibited a 65.7% increase in weight between harvest 1 and harvest 3 ($p < 0.05$). In a previous study, it was determined that Crassulacean acid metabolism (CAM) species (which include *P. afra*) exhibited an increase in plant biomass when exposed to CO_2 concentrations of 650–750 ppm for 3 months.[18] The results of this study are similar to those in the literature. This confirms the general trend that increased CO_2 may have on the physiological properties of *P. afra.*

Chlorophyll has a wide range of nutritional and health-promoting benefits, as a result of it being rich in vitamins and minerals, as well as exhibiting antioxidant and anti-inflammatory properties.[19] The chlorophyll content of the leaves of *P. afra* remained relatively constant in both treatments, across all three harvests ($p > 0.05$). This suggests that the plant may offer health-promoting benefits despite being exposed to elevated CO_2 concentrations.

Phytochemical profile

The control sample (greenhouse) of *P. afra* exhibited the presence of 8 out of the 10 phytochemicals tested (Table 2). A high level of coumarins was detected in control plant samples. In contrast, the test for flavonoids and phlobatannins yielded an 'absent' result, indicating no presence of these phytochemicals or a presence too low to be detected by these

phytochemical analyses. The plant samples that were exposed to a CO_2 concentration of 420 ppm exhibited a phytochemical profile similar to that of the control samples.

The samples exposed to 420 ppm experienced a general decrease in phytochemical presence. After harvest 1, the saponins, phenolics and coumarins exhibited an increase in presence. However, after harvests 2 and 3, the presence was similar to or lower than the greenhouse samples. The glycosides were the only group to have a greater presence in harvests 1, 2 and 3 than the control. Similar to the greenhouse extracts, flavonoids and phlobatannins were recorded as absent throughout all three harvests. The samples that were exposed to 420 ppm exhibited a lower phytochemical presence than the samples exposed to 600 ppm for 5 out of the 10 phytochemicals analysed.

The plants that were exposed to a CO_2 concentration of 600 ppm exhibited a general increase in phytochemical presence compared to the control and samples exposed to 420 ppm. Terpenoids were the only phytochemical group that decreased from present to absent in harvests 2 and 3. Other phytochemicals, such as saponins, flavonoids, glycosides, tannins, coumarins and volatile oils, exhibited a stronger phytochemical presence.

The most notable increase was in the flavonoid group, which was documented as absent in the greenhouse and 420-ppm samples, but present and moderately present in the methanolic and hot water samples that were exposed to 600 ppm. Upon further investigation, the quantification of flavonoids revealed that flavonoids were not absent in the greenhouse and 420-ppm extracts, just present in low quantities (Table 3).

The leaf samples that were exposed to elevated CO_2 concentrations exhibited a significant increase in flavonoid content in the methanolic, hexanic and hot water extracts ($p < 0.05$). The highest flavonoid content was recorded in the methanolic leaf extracts that had been exposed to 600 ppm (19.27 ± 0.10 mg QE/g). All four 420-ppm treatment extracts exhibited an initial decrease in flavonoid content in harvest 1. Similar to the phytochemical presence of saponins, phenolics and coumarins, harvests 2 and 3 produced flavonoid contents similar to that of the greenhouse samples. This change after harvest 1 may allude to an adaptation period for the plant. The flavonoid content in the leaves that were exposed to 600 ppm continued to increase significantly throughout harvests 2 and 3.

Flavonoids are chemically important molecules in plants, as they contribute to the colour and scent of plants, and they also protect plants by acting as a signal molecule in times of stress. The molecule also exhibits strong therapeutic potential[20,21]; thus, the increase in flavonoids in response to elevated CO_2 concentration may allude to the potential increase in therapeutic potential of the plant. This phytochemical group is also an important reactive oxygen species scavenger.[21]

Antioxidant activity

The leaves of *P. afra* exhibited strong scavenging activity against DPPH (Table 4) and H_2O_2 (Table 5). The efficacy of the extract was determined by the IC_{50} value. An extract was considered effective if it was below the upper limit of 10 μg/mL.[22]

The leaf extracts that were exposed to 420 ppm exhibited moderate scavenging activity against DPPH, with all harvests exhibiting an IC_{50} value below the upper limit, in contrast to the hexane and dichloromethane extracts that were consistently above the upper limit. No statistical

Table 1: The recorded weight and chlorophyll content of the leaves of *Portulacaria afra* after each harvest

	Harvest 1		Harvest 2		Harvest 3	
	420 ppm	600 ppm	420 ppm	600 ppm	420 ppm	600 ppm
Weight (g) of wet leaves (mean ± SD)	0.72 ± 0.43	1.08 ± 0.63	0.78 ± 0.03	1.65 ± 0.71	0.89 ± 0.07	1.79 ± 0.14
Chlorophyll content (mg/g FW)	5.95	5.99	5.90	6.83	6.44	6.68

Table 2: A heat map representing the presence of the different phytochemical groups in the leaves of *Portulacaria afar* in the greenhouse (control) and those under 420 ppm and 600 ppm treatment

Phytochemical		Leaves							
		420 ppm				600 ppm			
		Hex	Di	Meth	HW	Hex	Di	Meth	HW
Saponins	Greenhouse								
	Harvest 1								
	Harvest 2								
	Harvest 3								
Flavonoids	Greenhouse								
	Harvest 1								
	Harvest 2								
	Harvest 3								
Glycosides	Greenhouse								
	Harvest 1								
	Harvest 2								
	Harvest 3								
Phenolics	Greenhouse								
	Harvest 1								
	Harvest 2								
	Harvest 3								
Tannins	Greenhouse								
	Harvest 1								
	Harvest 2								
	Harvest 3								
Terpenoids	Greenhouse								
	Harvest 1								
	Harvest 2								
	Harvest 3								
Steroids	Greenhouse								
	Harvest 1								
	Harvest 2								
	Harvest 3								
Coumarins	Greenhouse								
	Harvest 1								
	Harvest 2								
	Harvest 3								
Phlobatannins	Greenhouse								
	Harvest 1								
	Harvest 2								
	Harvest 3								

Table 2 continued...

		Leaves							
Phytochemical		420 ppm				600 ppm			
Volatile oils	Greenhouse								
	Harvest 1								
	Harvest 2								
	Harvest 3								

Absent	Present	Moderate presence	High presence

Table 3: Total flavonoid content in the leaf extracts of *Portulacaria afra* which were exposed to 420 ppm and 600 ppm treatments in comparison to the control (greenhouse)

Flavonoid content (mg QE/g)								
Leaves	Methanol		Dichloromethane		Hexane		Hot water	
	420 ppm	600 ppm	420 ppm	600 ppm	420 ppm	600 ppm	420 ppm	600 ppm
Greenhouse	6.48 ± 0.32	6.48 ± 0.41	2.12 ± 0.87	2.12 ± 0.87	1.24 ± 0.54	1.24 ± 0.54	2.51 ± 0.43	2.51 ± 0.43
Harvest 1	3.79 ± 0.19	9.38 ± 0.45	1.62 ± 0.39	6.33 ± 0.77	1.09 ± 0.37	2.54 ± 0.72	2.20 ± 0.67	12.65 ± 0.19*
Harvest 2	4.88 ± 0.43	13.80 ± 0.29*	3.45 ± 0.67	6.87 ± 0.68	1.92 ± 0.34	4.68 ± 0.69	3.79 ± 0.12	14.28 ± 0.31*
Harvest 3	4.79 ± 0.11	19.27 ± 0.10*	3.69 ± 0.34	4.65 ± 0.61	3.51 ± 0.39	7.20 ± 0.83*	4.31 ± 0.19	1.81 ± 0.18*

statistically significantly different from the greenhouse extracts in a Tukey ad-hoc test

Table 4: The IC_{50} of the various leaf extracts exposed to 420 ppm and 600 ppm treatments in comparison to the control (greenhouse) against free radical DPPH

IC_{50} value (µg/mL) ± SD								
Leaves	Methanol		Dichloromethane		Hexane		Hot water	
	420 ppm	600 ppm	420 ppm	600 ppm	420 ppm	600 ppm	420 ppm	600 ppm
Greenhouse	4.7 ± 0.4	4.7 ± 0.4	15.7 ± 2.3	15.7 ± 2.3	14.8 ± 0.9	14.8 ± 0.9	3.7 ± 0.3	3.7 ± 0.3
Harvest 1	1.7 ± 0.6	2.8 ± 0.9	18.3 ± 1.8	15.0 ± 0.6	11.7 ± 1.6	2.7 ± 1.3*	6.0 ± 0.1	4.3 ± 0.8
Harvest 2	3.3 ± 0.7	1.4 ± 0.2	11.4 ± 0.6	13.6 ± 1.5	10.4 ± 1.1	9.3 ± 0.9*	9.4 ± 0.3	3.1 ± 0.2
Harvest 3	2.2 ± 0.6	2.1 ± 0.4	12.9 ± 1.1	8.8 ± 2.5*	12.3 ± 1.2	12.0 ± 1.5	4.6 ± 0.6	2.7 ± 0.5

statistically significantly different from the greenhouse extracts in a Tukey ad-hoc test ($p < 0.05$)

Table 5: The IC_{50} of the various leaf extracts exposed to 420 ppm and 600 ppm treatments in comparison to the control (greenhouse) against free radical H_2O_2

IC_{50} value (µg/mL) ± SD								
Leaves	Methanol		Dichloromethane		Hexane		Hot water	
	420 ppm	600 ppm	420 ppm	600 ppm	420 ppm	600 ppm	420 ppm	600 ppm
Greenhouse	0.3 ± 0.2	0.3 ± 0.2	18.7 ± 0.8	18.7 ± 0.8	3.9 ± 0.8	3.9 ± 0.8	0.2 ± 0.6	0.2 ± 0.6
Harvest 1	4.1 ± 0.4*	0.5 ± 0.6	24.6 ± 1.6	11.9 ± 0.5	5.0 ± 0.9	5.3 ± 0.6	4.2 ± 0.2	0.8 ± 0.2
Harvest 2	2.0 ± 0.3	0.9 ± 0.3	13.9 ± 1.3	9.8 ± 0.8	5.8 ± 1.3	5.8 ± 0.9	4.4 ± 0.3	0.9 ± 0.1
Harvest 3	0.4 ± 0.8	0.2 ± 0.7	10.0 ± 1.7	7.8 ± 1.3*	1.8 ± 0.9	7.2 ± 1.3*	0.7 ± 0.9	1.0 ± 0.5

statistically significantly different from the greenhouse extracts in a Tukey ad-hoc test

significance was determined for either extract exposed to 420 ppm against the greenhouse extracts. The strongest scavenging activity was exhibited by the methanolic sample collected in harvest 1 (1.7 ± 0.6 µg/mL). The samples that were exposed to 600 ppm exhibited an overall stronger scavenging activity against DPPH than the 420 ppm-treated counterparts.

The methanolic and hot water extracts exposed to 600 ppm exhibited a consistently strong scavenging activity with IC_{50} values below 5 µg/mL across all three harvests. The strongest scavenging activity was exhibited by the methanolic sample collected in harvest 2. There was a general increase in scavenging activity across the various harvests in comparison to the greenhouse extracts. The hexanic leaf extracts that were exposed to 600 ppm exhibited significantly stronger scavenging activity against the greenhouse extract ($F4,21:4.3$, $p<0.05$). The scavenging activity of the dichloromethane (harvest 3) and hexane (harvest 2) extracts was also significantly lower than that of the greenhouse leaf samples.

The leaves of *P. afra* exhibited a generally stronger scavenging activity against H_2O_2 than DPPH. The leaves exposed to 420 ppm exhibited a large variation in scavenging activity between the harvests. The dichloromethane extracts in harvest 1 had an IC_{50} value of 24.6 ± 1.6 µg/mL, whereas that for harvest 3 was 10.0 ± 1.7 µg/mL. These large fluctuations can also be seen in the methanolic and hot water 420 ppm extracts. There were also large differences between the 420 ppm extracts and their 600 ppm counterparts.

The leaves exposed to an elevated CO_2 concentration appear very effective as an antioxidant against H_2O_2, with an IC_{50} that reached a low of 0.2 ± 0.7 µg/mL in the methanolic extracts of harvest 3. There was a general increase in scavenging activity in leaves exposed to 600 ppm. The only exception to this finding was the hexane extracts that exhibited an IC_{50} value of 3.9 ± 0.8 µg/mL in the greenhouse extracts and reached 3.9 ± 0.8 µg/mL in leaves exposed to 600 ppm after harvest 3.

The skin is particularly susceptible to reactive oxygen species attacks. This is a result of the skin's direct exposure to UV radiation, environmental pollutants and high pressure of oxygen molecules.[23] Hence, the strong antioxidant activity exhibited by the plant may contribute to the relief of skin infections, which is what the plant is traditionally used for. The relatively steady antioxidant activity exhibited by the leaves indicates that the plant may retain its antioxidant activity in the future, despite the rising atmospheric CO_2 concentration. The phytochemicals that contribute to antioxidant activity in plants may also be effective at killing and hindering the growth of bacteria, viruses and fungi.[24]

Antimicrobial activity

The antimicrobial activity of the leaves of *P. afra* was tested against five microorganisms commonly associated with skin conditions. The antimicrobial activity of an extract was considered noteworthy if the MIC value was between 160 µg/mL and 1000 µg/mL.[25]

The leaf extracts of *P. afra* mostly exhibited MIC values above 1000 µg/mL (Table 6), and hence, the antimicrobial activity is considered relatively low. Although the antimicrobial activity was low, it was similar among the control and treated extracts. In samples exposed to an elevated CO_2 concentration, the extracts either showed a negligible increase in microbial activity (one serial dilution) or exhibited no change. A noteworthy result, however, was observed with *C. acnes*. The greenhouse plant extracts exhibited weak antimicrobial activity against *C. acnes*, but activity noticeably increased from an inhibitory concentration of 8000 µg/mL to 2000 µg/mL in plant extracts exposed to an elevated CO_2 concentration.

C. acnes are pleomorphic rod-shaped bacteria and form a vital component of the skin's microbiota. This opportunistic pathogen has been identified as a key component of acne vulgaris.[26] Thus, the leaves may continue to provide relief against this skin condition when the atmospheric CO_2 concentration reaches 600 ppm.

Conclusion

There is a concern amongst scientists that the rising atmospheric CO_2 concentration may affect the medicinal properties of plants. *P. afra* is a medicinally important plant used to treat skin infections. After exposure to an elevated CO_2 concentration, the plant exhibited a general increase in phytochemical presence. The antioxidant activity remained high in the methanolic and hot water extracts. The antimicrobial activity, while relatively weak, remained constant and the efficacy against *C. acnes* increased with increasing atmospheric CO_2 concentration.

The leaves of *P. afra* are resilient against exposure to an increased CO_2 concentration and may continue to provide relief against certain ailments in the future.

Table 6: Mean inhibitory concentrations (MIC) of the leaf extracts of the control (greenhouse) and those which were subjected to ambient and elevated CO_2 concentrations (420 ppm and 600 ppm, respectively)

Plant extract	Mean MIC values (µg/mL)					
	Gram-positive bacteria			Gram-negative bacteria		Yeast
Sample	*S. aureus*	*S. epidermidis*	*C. acnes*	*P. aeruginosa*	*K. aerogenes*	*C. albicans*
Greenhouse	4000	4000	8000	6000	2000	8000
420 (1)	4000	4000	8000	4000	2000	8000
420 (2)	4000	4000	4000	4000	2000	8000
420 (3)	4000	4000	4000	4000	2000	8000
600 (1)	4000	4000	4000	4000	2000	8000
600 (2)	3000	4000	2000	4000	2000	8000
600 (3)	2000	4000	2000	4000	2000	4000
Ciprofloxacin (negative control)	0.625	0.625	1.25	1.25	0.625	N/A
Nystatin (positive control)	N/A	N/A	N/A	N/A	N/A	2.50
Negative control	8000	>8000	8000	>8000	8000	>8000
Culture control	>8000	>8000	>8000	>8000	>8000	>8000

Acknowledgements

Competing interests

We have no competing interests to declare.

Authors' contributions

D.C.B.: Data collection; sample analysis; data curation; writing – the initial draft; writing – revisions. S.v.V.: Methodology; validation; writing – revisions; student supervision. I.M.R.: Conceptualisation; validation; writing – revisions; student supervision; funding acquisition.

References

1. Ganopolski A, Winkelmann R, Schellnhuber HJ. Critical insolation-CO$_2$ relation for diagnosing past and future glacial inception. Nature. 2016;529(7585):200–203. https://doi.org/10.1038/nature16494

2. Tan D, Lee W, Kim YE, Ko YN, Youn MH, Jeon YE, et al. In-Bi electrocatalyst for the reduction of CO$_2$ to formate in a wide potential window. ACS Appl Mater Interfaces. 2022;14(25):28890–28899. https://doi.org/10.1021/acsami.2c05596

3. Rajashekar CB. Elevated CO$_2$ levels affect the phytochemicals and nutritional quality of food crops. Am J Plant Sci. 2018;9(2):150–162. https://doi.org/10.4236/ajps.2018.92013

4. Ibrahim MH, Jaafar Hawa ZE. Increased carbon dioxide concentration improves the antioxidative properties of the Malaysian herb Kacip Fatimah (*Labisia pumila* Blume). Molecules. 2011;16(7):6068–6081. https://doi.org/10.3390/molecules16076068

5. Applequist WL, Wallace RS. Expanded circumscription of Didiereaceae and its division into three subfamilies. Adansonia. 2003;25(1):13–16.

6. Decker LE, de Roode JC, Hunter MD. Elevated atmospheric concentrations of carbon dioxide reduce monarch tolerance and increase parasite virulence by altering the medicinal properties of milkweeds. Ecol Lett. 2018;21(9):1353–1363. https://doi.org/10.1111/ele.13101

7. Weiss I, Mizrahi Y, Raveh E. Effect of elevated CO$_2$ on vegetative and reproductive growth characteristics of the CAM plants *Hylocereus undatus* and *Selenicereus megalanthus*. Sci Hortic. 2010;123(4):531–536. https://doi.org/10.1016/j.scienta.2009.11.002

8. Van Jaarsveld E, Le Roux A. *Portulacaria afra* (L.) Jacq.: Variability and distribution. Bradleya. 2021;2021(39):138–152. https://doi.org/10.25223/brad.n39.2021.a13

9. Van der Vyver ML, Cowling RM, Mills AJ, Difford M. Spontaneous return of biodiversity in restored subtropical thicket: *Portulacaria afra* as an ecosystem engineer. Restor Ecol. 2013;21(6):736–744. https://doi.org/10.1111/rec.12000

10. De Wet H, Nciki S, Van Vuuren SF. Medicinal plants used for the treatment of various skin disorders by a rural community in northern Maputaland, South Africa. J Ethnobiol Ethnomedicine. 2013;9(1):1–10. https://doi.org/10.1186/1746-4269-9-51

11. Tabassum S, Ahmad S, Rehman Khan KU, Tabassum F, Khursheed A, Zaman QU, et al. Phytochemical profiling, anti-oxidant, anti-inflammatory, thrombolytic, hemolytic activity in vitro and in silico potential of *Portulacaria afra*. Molecules. 2022;27(8):2377. https://doi.org/10.3390/molecules27082377

12. Liu X, Hong L, Li XY, Yao Y, Hu B, Li L. Improved drought and salt tolerance in transgenic *Arabidopsis* overexpressing a NAC transcriptional factor from *Arachis hypogaea*. Biosci Biotechnol Biochem. 2011;75(3):443–450. https://doi.org/10.1271/bbb.100614

13. Hassan MA, Omer AM, Abbas E, Baset W, Tamer TM. Preparation, physicochemical characterization and antimicrobial activities of novel two phenolic chitosan Schiff base derivatives. Sci Rep. 2018;8(1):1–4. https://doi.org/10.1038/s41598-018-29650-w

14. Pandey A, Tripathi S. Concept of standardization, extraction and pre phytochemical screening strategies for herbal drug. J Pharmacogn Phytochem. 2014;2(5):115–119.

15. Ahmed S, Moni BM, Ahmed S, Gomes DJ, Shohael AM. Comparative phytochemical, anti-oxidant, and antibacterial study of different parts of Doigota plants (*Bixa orellana* L.). Bull Natl Res Cent. 2020;44(1):1–10. https://doi.org/10.1186/s42269-020-00349-1

16. Pakade V, Cukrowska E, Chimuka L. Comparison of anti-oxidant activity of *Moringa oleifera* and selected vegetables. S Afr J Sci. 2013;109(3), Art. #1154. https://doi.org/10.1590/sajs.2013/1154

17. De Torre MP, Cavero RY, Calvo MI, Vizmanos JL. A simple and a reliable method to quantify anti-oxidant activity in vivo. Antioxidants. 2019;8(5), Art. #142. https://doi.org/10.3390/antiox8050142

18. Drennan PM, Nobel PS. Responses of CAM species to increasing atmospheric CO$_2$ concentrations. Plant Cell Environ. 2000;23(8):767–781. https://doi.org/10.1046/j.1365-3040.2000.00588.x

19. Ebrahimi P, Shokramraji Z, Tavakkoli S, Mihaylova D, Lante A. Chlorophylls as natural bioactive compounds existing in food by-products: A critical review. Plants. 2023;12(7), Art. #1533. https://doi.org/10.3390/plants12071533

20. Panche AN, Diwan AD, Chandra SR. Flavonoids: An overview. J Nutr Sci. 2016;5, e47. https://doi.org/10.1017/jns.2016.41

21. Samanta A, Das G, Das SK. Roles of flavonoids in plants. Carbon. 2011;100(6):12–35.

22. Brighente IM, Dias M, Verdi LG, Pizzolatti MG. Antioxidant activity and total phenolic content of some Brazilian species. Pharm Biol. 2007;45(2):156–161. https://doi.org/10.1080/13880200601113131

23. Kruk J, Duchnik E. Oxidative stress and skin diseases: Possible role of physical activity. Asian Pac J Cancer Prev. 2014;15(2):561–568. https://doi.org/10.7314/APJCP.2014.15.2.561

24. Palaniappan K, Holley RA. Use of natural antimicrobials to increase antibiotic susceptibility of drug resistant bacteria. Int J Food Microbiol. 2010;140(2–3):164–168. https://doi.org/10.1016/j.ijfoodmicro.2010.04.001

25. Van Vuuren S, Holl D. Antimicrobial natural product research: A review from a South African perspective for the years 2009-2016. J Ethnopharmacol. 2017;208:236–252. https://doi.org/10.1016/j.jep.2017.07.011

26. Dréno B, Pécastaings S, Corvec S, Veraldi S, Khammari A, Roques C. *Cutibacterium acnes (Propionibacterium acnes)* and acne vulgaris: A brief look at the latest updates. J Eur Acad Dermatol Venereol. 2018;32:5–14. https://doi.org/10.1111/jdv.15043

AUTHORS:
Chanel K. Daniel-Swartland[1,2] (ID)
Julia C. Meitz-Hopkins[1] (ID)
Filicity A. Vries[2]
Cheryl L. Lennox[1]

AFFILIATIONS:
[1]Fruit and Postharvest Pathology Research Programme, Department of Plant Pathology, Stellenbosch University, Stellenbosch, South Africa
[2]Tessara (Pty) Ltd, Cape Town, South Africa

CORRESPONDENCE TO:
Chanel Daniel-Swartland

EMAIL:
ChanelD@tessara.co.za

2 ZERO HUNGER

Use of sulfur dioxide to reduce postharvest decay and preserve the quality of fresh tomatoes

Tomatoes are an important and versatile crop with a short shelf life. Postharvest losses due to fruit decay and handling are reportedly as high as 30–50% globally; therefore, the agricultural sector would benefit from solutions that target the preservation of crops such as tomatoes. In this study, we investigated the potential use of sulfur dioxide (SO_2) to provide postharvest protection against fungal decay in tomatoes whilst maintaining the quality of tomato fruit. Three tomato varieties packed as bulk (3–5 kg) cartons were exposed to SO_2-generating sheets applied as either a top sheet over the fruit or a bottom sheet placed on the base of the carton before packing the fruit on the top. The results show that the application of SO_2-generating sheets reduced the natural progression of decay on 'Roma' tomatoes by up to 60% and up to 80% on 'Rosa' tomatoes. Only marginal decay control was observed on 'Round' tomatoes. The top sheet application rendered the best results, as the application of the bottom sheet resulted in phytotoxicity in the form of SO_2 damage. SO_2 application reduced fruit shrivelling, especially when tomatoes were stored at higher temperatures during their shelf life. The application of SO_2 on the tomatoes did not have any negative effects on fruit firmness. The results of this study provide a case to further explore the use of SO_2-generating sheets on tomatoes to prevent postharvest decay. Differences in varietal physiology may be key to the successful application of this technology.

Significance:

Tomatoes are a crop that is highly susceptible to postharvest decay. These effects lie mainly with the end consumer. Besides the monetary loss of buying a commodity that rots quickly, pathogens infecting tomatoes are known to produce mycotoxins that pose a risk to human health. With food availability and safety concerns, the application of a product that could reduce these concerns would be beneficial to the agricultural sector.

Introduction

Tomatoes are a lucrative crop globally, with production exceeding one billion tons in 2021.[1] Unfortunately, total production does not translate to total consumption due to losses that occur along the distribution chain.[2] In South Africa, tomatoes rank second in relation to potatoes as important commodities; they contribute approximately 24% of the country's annual vegetable production.[3] Commercial tomato-growing regions in South Africa include the provinces of Limpopo, Mpumalanga, KwaZulu-Natal, Eastern Cape and Western Cape, but the Limpopo region remains the largest grower of tomatoes. In 2021, the estimated production of tomatoes was around 530 834 tons[1], with an estimated gross production value of about ZAR3 031 644.[1] Tomatoes in South Africa are sold as fresh produce through direct sales and exported to neighbouring African countries like Mozambique, Angola and Zambia.[4]

Tomatoes ripen and deteriorate quickly. Ripe tomatoes could last for as long as 2 weeks, but the shelf life is further reduced by decay, injuries and general poor handling.[5] Tomatoes are prone to postharvest decay caused by a complex of fungi such as *Rhizopus stolonifer* (Ehrenb.) Vuill. 1902, *Botrytis cinerea* Pers. 1797, *Alternaria alternata* (Fr.) Keissl. 1912 and *Colletotrichum coccodes* (Wallr.) S. Hughes.[6] Bacterial and viral infections are also common. Statistics relating to postharvest decay of tomatoes are rarely reported as it mostly appears at consumer level. Studies that have investigated the progression of the disease over time report that 74% of ripe fruit will exhibit decay after 2 weeks of storage.[7] Due to quick deterioration, tomatoes cannot be stored for long periods, which puts pressure on producers and retailers to get the fruit to the market as soon as possible. This limits market access and the potential revenue that could be generated if the fruit was able to stay fresh for longer.

Current postharvest handling of tomatoes in South Africa involves handpicking in the early to mid-mornings and then transporting the tomatoes in open plastic crates to the packhouse. Where infrastructure allows, bins are tipped into a dump tank where the tomatoes are washed in chlorinated water at a dose of about 50–200 ppm. Tomatoes then pass over rollers with light brushes that dry the fruit before being tipped onto the packing tables to be packed either as loose or bagged product, as per retail requirements. Postharvest decay control with fungicides is rare, because application of any product would have to comply with maximum residue levels for safe human consumption. The short shelf life of tomatoes makes it difficult to identify such a postharvest chemical. The chlorine wash provides an effective sanitation method to prevent new infections; however, it does not prevent the proliferation of pathogens which may have already been established through the stem scar.[8]

Sulfur dioxide (SO_2) is extensively used in the food and beverage industry as an antioxidant and preservative[9] in dried fruits, soft drinks and alcoholic beverages. SO_2 applications to prevent enzymatic browning in bananas, lemons and apples have been reported.[10] SO_2 application is also used on fresh fruit such as table grapes to inhibit fungal growth caused by *B. cinerea*.[11,12] Cantin et al.[12] reported good inhibition of various postharvest pathogens on fresh figs when SO_2 was applied via fumigation or using SO_2 pads. Efficacy of SO_2 in conjunction with controlled atmosphere storage has also shown promising results in reducing decay, extending the shelf life and maintaining the nutritional value of fresh blueberries.[13]

The most common drawback of using SO_2 on soft fruit is the bleaching that may occur. Some studies report that SO_2-generating pads increased the occurrence of bleaching on the skins of figs[12] and grapes[10]. Usually this happens at wound sites or fresh abscission sites, but excessive exposure can also cause SO_2 damage in the form of sunken areas on fruit surfaces, as well as contribute to premature browning of the grape stems.[12]

SO_2 has long been considered an acceptable food additive; however, with the recent trends of moving away from chemical usage in foodstuffs, SO_2 usage is questioned. Although mostly harmless, exposure to SO_2 could be problematic for people who have sulfite allergies.[9,14] As requested by the European commission, the European Food Safety Authority published a review that re-evaluated the use of SO_2 and other sulfites as food additives.[15] The panel noted that the current acceptable daily intake of 0.7 mg/kg of body weight could be subjected to review as the actual intake was higher amongst population groups. The panel did not express any concerns of harmful genotoxic, chronic, carcinogenic or negative reproductive effects based on studies done regarding oral exposure; however, a later review in 2022 reported a lower acceptable daily intake of 0.38 mg/kg of body weight.[15]

Currently, there are no SO_2 products registered for postharvest use on tomatoes, globally. The preharvest application of sulfur as a nutrient for tomato plant growth is known. The exposure of tomato plants to SO_2 gas has been studied[16] with negative findings[17], but research on the postharvest application of SO_2 on fresh tomatoes is a novel concept. In this study, we investigated the potential use of SO_2-generating sheets to provide postharvest protection against fungal organisms that cause decay on tomatoes, whilst maintaining the quality of tomato fruit.

Materials and methods

Fruit

Two varieties of long-life tomatoes, 'Round' and 'Roma', and one variety of specialty tomato, 'Rosa', were procured from packhouses in the Western Cape, South Africa. The fruit was picked and packed into 5 kg cartons for the larger varieties and 3 kg cartons for the smaller 'Rosa' tomatoes. Normal packhouse procedures of fruit washing and grading per class and variety were followed. As per these normal procedures, the variety 'Rosa' was not subjected to postharvest fruit washing. Class 1 fruit was used for the study.

Sulfur dioxide and decay control

The ability of SO_2 to reduce the decay that occurs naturally in tomatoes was investigated. A SO_2-generating sheet was manufactured by Tessara (Pty) Ltd, Cape Town, South Africa, to be used in these experiments. The sheet provided for testing was formulated as a dual-release sheet, with a high first phase of SO_2 emissions (gas levels), followed by a second low-emission phase.

Treatments applied in this trial were as follows: (1) Top sheet – SO_2-generating sheet positioned over the tomatoes in the carton, (2) Bottom sheet – SO_2-generating sheet placed inside on the base of the carton with fruit packed on the top, and (3) Control - without any SO_2 sheets in the cartons. A total of six cartons were packed for each treatment group to satisfy statistical requirements. The cartons were stored at 10 °C for either 14 or 21 days, after which quality evaluations

were done. For both these time intervals, fruit boxes were evaluated, the decayed and poor-quality fruit was removed, and the remaining good fruit was moved to 18 °C for a further 7 days to mimic shelf-life storage. The trials were done in duplicate for the 'Round' and 'Roma' tomatoes and in triplicate for the 'Rosa' tomatoes over a space of 2 years.

At each evaluation, the amount of fungal decay, SO_2 damage and shrivel was recorded by fruit count for 'Round' and 'Roma' tomatoes and by fruit weight for the 'Rosa' tomatoes. Symptomatic fruits were subjected to isolation and molecular identification of a subset of isolates by sequencing the ITS gene region as described in the literature.[18] Percentage defects were calculated relative to the original numbers of fruit in the carton or the total box weight. In addition, a sample of 10 fruits was taken from each treatment group to evaluate firmness, which was done using a fruit texture analyser. The firmness test was conducted using an 11-mm probe for the large tomatoes and a 3-mm probe for the small tomatoes. Each fruit was pierced once at opposite ends to render a total of two readings per fruit. Average firmness is reported.

Statistical analysis

The analysis was conducted independently by the Agricultural Research Council (ARC)-Infruitec Nietvoorbij, Stellenbosch, South Africa. The experimental design was randomised with six replicates for each treatment combination. An analysis of variance (ANOVA) was applied on continuous variables and the output evaluated using SAS[19] statistical software. The four main parameters evaluated were fungal decay, shrivel, fruit firmness and SO_2 damage. A Shapiro–Wilk[20] test was used to determine deviation from normality. A Fisher's least significant difference was used to compare treatment means, where a probability level of 5% was deemed significant.[21]

Results

A combined average over two trials was calculated for fungal decay, shrivel, fruit firmness and SO_2 damage for the 'Round' and 'Roma' tomatoes. Three trials were conducted for the 'Rosa' tomatoes, with all being statistically different; therefore, the results are discussed separately. Varieties were not compared and will be discussed separately. All data reported for shelf life are cumulative.

Decay control

ANOVA interactions were significant for parameters of decay ($p \leq 0.0001$) and decay inhibition ($p = 0.04$) in 'Roma' tomatoes (Table 1). The control sets throughout the storage period had significantly higher decay incidence (21–67%) than tomatoes treated with the SO_2 sheets (8–55% decay incidence; Figure 1). The effect of the treatments continued for up to 7 days after cold storage, even after the sheets had been removed, indicating that the initial SO_2 application has an ongoing effect and that the protection against decay will continue in the absence of SO_2 exposure, thereby extending the shelf life of 'Roma' tomatoes. With regard to sheet placement, it was noted that both the top and bottom placed sheets effectively reduced the decay. The results infer that a bottom sheet will be more beneficial with longer storage time, as at 21 days of cold storage, tomatoes treated with a bottom placed sheet had significantly less decay than the tomatoes treated with a top sheet.

Table 1: *P*-values generated from analysis of variance (ANOVA) for each of the evaluation parameters

Treatment x Position x Storage x Shelf	'Roma' *	'Round' *	'Rosa' Trial 1	'Rosa' Trial 2	'Rosa' Trial 3
Decay	<0.0001	0.39	0.00	0.07	0.32
Decay inhibition	0.04	0.31	0.30	0.21	0.34
SO_2 damage	0.32	0.40	0.00	0.67	0.36
Shrivel	0.94	0.30	0.00	0.97	0.29

Average of two trials. $p \leq 0.05$ indicates significant interactions between evaluation parameters

An ANOVA for 'Round' tomatoes indicated no significant differences in decay (Table 1). The 'Round' tomatoes were presented with high decay levels throughout the trial. After 14 days at 10 °C, decay recorded on the 'Round' tomatoes was between 22% and 25% (Figure 1). At this point, there was no statistical difference between the treated and control fruit. When moved to shelf life, the tomatoes treated with a top placed sheet had significantly lower decay than the control fruit; however, the decay was high (62%). After extended storage (21 days) at 10 °C, the fruit treated with the bottom sheet showed the least amount of decay, but this was not statistically different to the other groups.

The ANOVA for 'Rosa' tomatoes showed that all three trials conducted were different; therefore, the results could not be combined, and the trials need to be discussed separately (Table 1). In the first trial, there

was no significant interaction between all parameters evaluated during initial storage; however, the interaction was significant when tomatoes were moved to shelf life. In the second and third trials, interactions were not significant for both initial storage and shelf life.

The trials for 'Rosa' tomatoes differed significantly with the amount of decay observed (Figure 2). The first trial showed a fair amount of decay with the controls reaching up to 30% decay during the initial storage, whilst Trials 2 and 3 recorded between 0.9% and 7% decay during the same time. In Trial 1, there was no significant difference in decay between SO_2 treatments and the control during the initial 7-day storage evaluation and subsequent shelf life, although tomatoes treated with SO_2 had about 10% less decay than the control during initial storage. Exposure to SO_2 sheets for up to 14 days showed that treated tomatoes had significantly less decay than the control. The bottom sheet proved
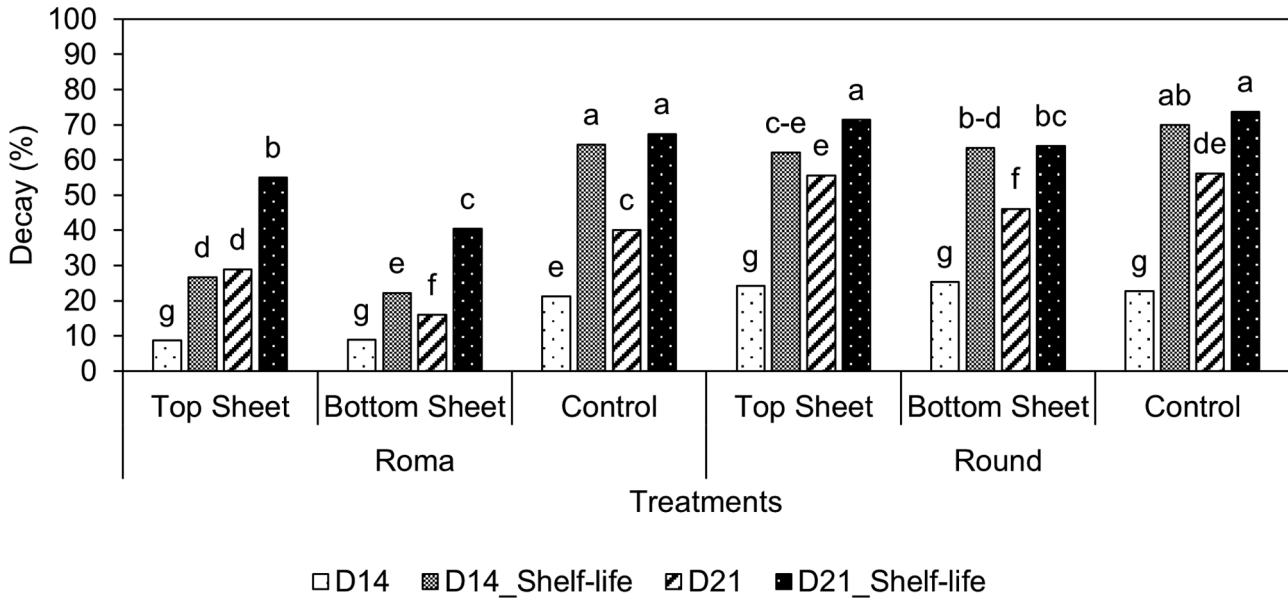
Figure 1: Percentage decay recorded on 'Roma' and 'Round' tomatoes after 14 (D14) and 21 (D21) days in cold storage at 10 °C and a further 7 days at shelf-life temperature of 18 °C. Bars with the same lettering do not differ significantly ($p > 0.05$) according to Fisher's least significant difference test. Data presented represent an average of two trials conducted.
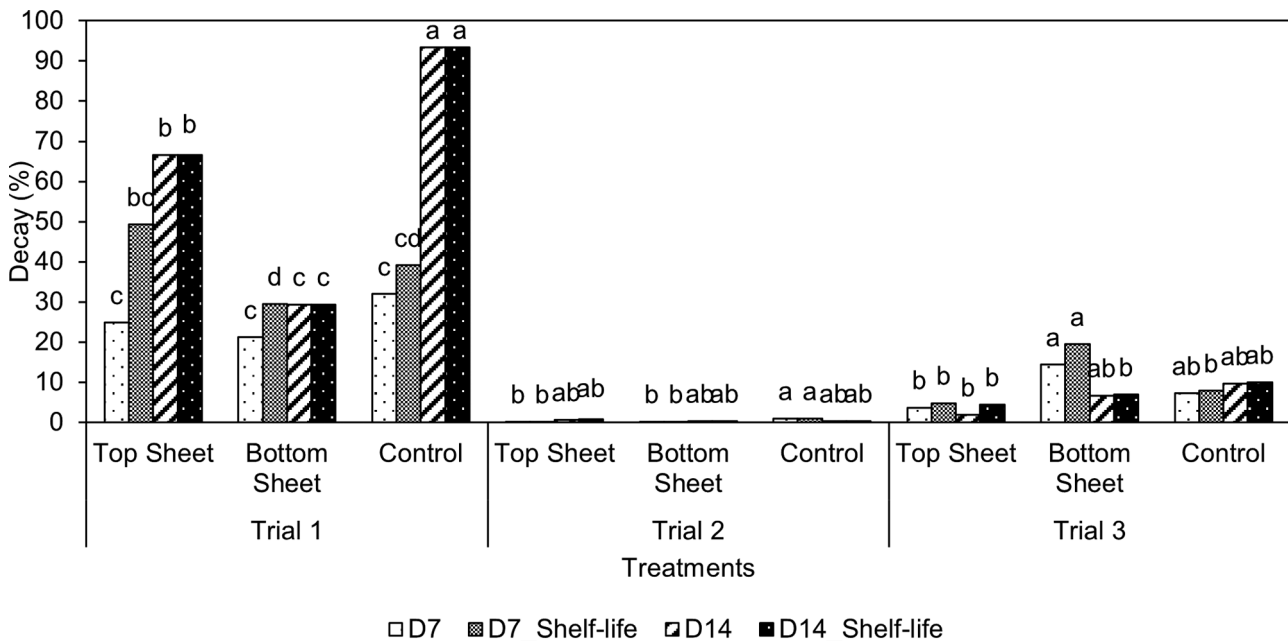

Figure 2: Percentage decay recorded on 'Rosa' tomatoes after 7 (D7) and 14 (D14) days of initial storage with SO_2 sheets and a further 7 days at shelf life. Temperature was kept constant at 18 °C for the duration of the trial. Bars with the same lettering do not differ significantly ($p > 0.05$) according to Fisher's least significant difference test.

to be the best treatment, differing significantly from the top sheet. The results for the 14 days plus shelf life evaluation were not different to the initial evaluation, as decay did not develop further.

In Trial 2, after 7 days in storage, the control had less than 1% decay. Still, the results indicate a significant reduction in decay when the tomatoes were treated with SO_2 sheets. The sheet placements did not differ from each other. Decay was so low that no further development occurred during shelf life. In this trial, there were no significant differences between SO_2 treatments and controls after 14 days in storage. Decay recorded in the third trial was less than 10% in the control, and for both storage periods and the respective shelf life, the treatments and controls did not differ significantly, although the top sheet treatment did have less actual decay than the control.

Percentage decay inhibition was calculated for all varieties. This determines which concentrations would be necessary to kill off at least 50% of the fungal growth on tomatoes. After 14 days at 10 °C and subsequent shelf life at 18 °C, decay on 'Roma' tomatoes was inhibited by approximately 63% and 58%, respectively, when a top sheet was applied (Table 2). A bottom sheet application resulted in an inhibition range of 54–65% over the same period. Long-term cold storage (21 days) with a top sheet shows that decay is inhibited by about 30% and 17% inhibition at shelf life. The bottom sheet resulted in 60% inhibition after 21 days in cold storage and 38% after shelf life.

For 'Round' tomatoes, there was a negative correlation when percentage decay inhibition was calculated (Table 2). Only marginal inhibition was noted with the application of the bottom sheet over the long-term storage period. This value is less than 50%; therefore, SO_2 treatments on 'Round' tomatoes cannot be recommended as a viable treatment option at present.

Decay inhibition differed between trials conducted on 'Rosa' tomatoes. In the first trial, application of a top sheet inhibited decay after 7 days by 22% and by 28% after 14 days. A bottom sheet inhibited decay after 7 days by 33% and by 68% after storage for 14 days. In the second trial, SO_2 application was most successful during the 7-day storage period. During this time, a top sheet reduced decay by 81% relative to the control and a bottom sheet reduced decay on 'Rosa' tomatoes by 78%. The sheets did not seem to be effective after 14 days. In the third trial, the top sheet SO_2 application was most successful, inhibiting decay during the first 7 days by 48% and then by 80% when used for a period of 14 days (Table 3). Isolates from all three varieties that were subjected to molecular identification were found to be identical to GenBank references of *Alternaria* spp. (KP125281).

SO_2 damage

When the concentration of SO_2 gas in the carton is too high, or the gas becomes localised to certain areas, it can cause SO_2 bleaching or damage to the fruit.[22] The percentage of SO_2 damage that was recorded on 'Roma' and 'Round' tomatoes after the 14- and 21-day cold storage periods and subsequent shelf life is presented in Figure 3. A significant amount of SO_2 damage was recorded in cartons with a bottom sheet when compared to other treatments. Although significantly higher than

Table 2: Percentage decay inhibition by SO_2 treatments on 'Roma' and 'Round' tomatoes after 14 (D14) and 21 (D21) days in cold storage at 10 °C and the respective 7-day shelf-life period at 18 °C

Variety	Treatment / storage	Percentage decay inhibition			
		D14 at 10 °C	D14 + D7 at 18 °C	D21 at 10 °C	D21 + D7 at 18 °C
'Roma'	Top sheet	63.96 a	58.94 a	30.32 bc	17.59 c
	Bottom sheet	54.22 a	65.59 a	60.67 a	38.70 b
	Control	0.00 d	0.00 d	0.00 d	0.00 d
'Round'	Top sheet	−4.90 bc	13.24 ab	−3.11 bc	2.21 bc
	Bottom sheet	−10.19 c	10.89 ab	22.57 a	13.70 ab
	Control	0.00 bc	0.00 bc	0.00 bc	0.00 bc

The same lettering indicates no significant difference (p > 0.05) according to Fisher's least significant difference test. Data presented represent the average of two trials.

Table 3: Percentage decay inhibition by SO_2 treatments on 'Rosa' tomatoes after 7 (D7) and 14 (D14) days in storage and the respective 7-day shelf-life period at 18 °C

Trial number	Treatment / storage	Percentage decay inhibition			
		D7	D7_Shelf life	D14	D14_Shelf life
Trial 1	Top sheet	22.26 b	−26.05 d	28.62 b	28.62 b
	Bottom sheet	33.72 b	24.59 bc	68.50 a	68.50 a
	Control	0.00 b	0.00 cd	0.00 b	0.00 cd
Trial 2	Top Sheet	81.07 a	81.07 a	−98.70 b	−125.7 b
	Bottom sheet	78.16 a	78.16 a	9.94 ab	9.94 ab
	Control	0.00 ab	0.00 ab	0.00 ab	0.00 ab
Trial 3	Top sheet	48.70 a	38.61 a	80.53 a	55.71 a
	Bottom sheet	−98.98 b	−147.10 b	30.70 ab	30.88 a
	Control	0.00 ab	0.00 a	0.00 ab	0.00 a

The same lettering indicates no significant difference (p > 0.05) according to Fisher's least significant difference test.

the top sheet treatments, it must be noted that the overall value of SO_2 damage at its highest was a mere 5.4% on 'Roma' tomatoes.

The SO_2 damage was more pronounced on the 'Round' tomatoes (Figure 3). After 14 days, the incidence of SO_2 damage ranged from 0.89% for the top sheet application to 16.21% for the bottom sheet application. A significant increase in SO_2 damage was noted on tomatoes that were treated with a bottom sheet. Inexplicably though, the continued exposure to the bottom pads after 21 days in cold storage yielded less SO_2 damage on fruit than at 14 days.

The results for SO_2 damage on 'Rosa' tomatoes echo that of the other two varieties. Minimal SO_2 damage was observed with application of a top sheet, and high SO_2 damage (up to 18%) was observed when a bottom sheet was used across the three trials (Figure 4). For all varieties,

the SO_2 damage that was observed during the initial storage did not develop further after shelf life. In cases where increased amounts during shelf life are noted, this could be due to inadequate sorting during the initial inspection.

Shrivel

Following the 2 weeks at a 10 °C storage regime, the amount of shrivel on 'Roma' tomatoes did not differ between treatments; however, when moved to shelf life, the SO_2-treated tomatoes had more shrivel than the control tomatoes (Figure 5). Over the 21-day storage period, however, SO_2-treated tomatoes had significantly less shrivel than the control fruit; and during the respective shelf life, tomatoes treated with a top sheet showed significantly less shrivel than the bottom sheet treated tomatoes and the control tomatoes.
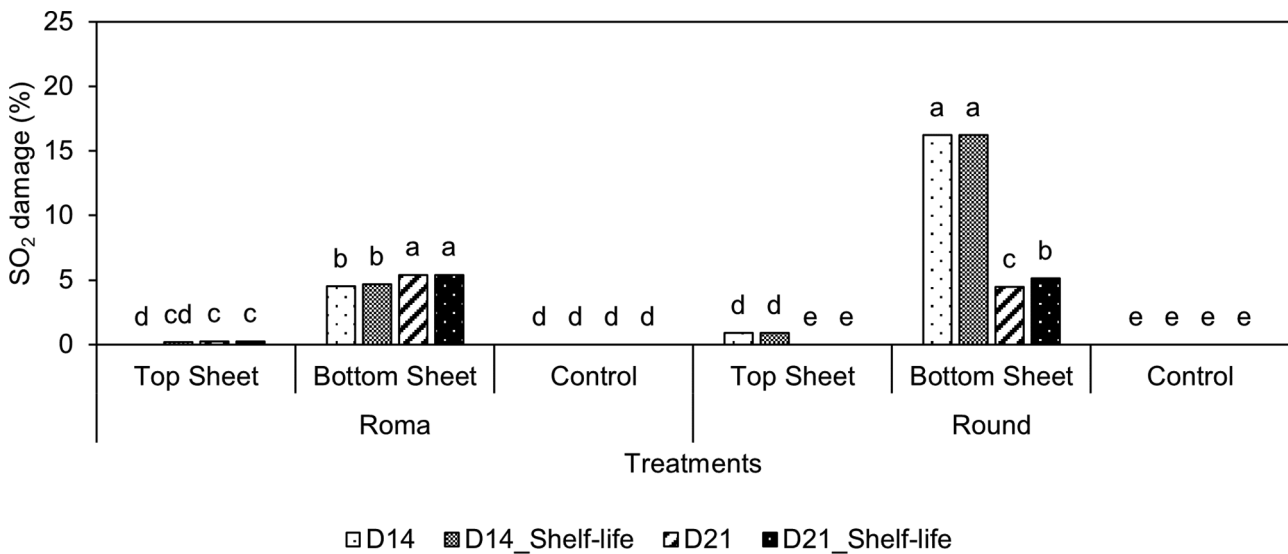

Figure 3: Percentage sulfur dioxide damage observed on 'Roma' and 'Round' tomatoes after 14 (D14) and 21 (D21) days in cold storage at 10 °C and the respective 7-day shelf-life period at 18 °C. Bars with the same lettering do not differ significantly ($p > 0.05$) according to Fisher's least significant difference test. Data presented are the average of two trials.

Figure 4: Percentage sulfur dioxide damage recorded on 'Rosa' tomatoes after 7 (D7) and 14 (D14) days of initial storage with SO_2 sheets and a further 7 days at shelf life. Temperature was kept constant at 18 °C for the duration of the trials. Bars with the same lettering do not differ significantly ($p > 0.05$) according to Fisher's least significant difference test.

For 'Round' tomatoes, the results were mostly insignificant except for the 21-day shelf-life period, where tomatoes treated with SO_2 applied as a top sheet were significantly less shrivelled than the bottom sheet treated and control tomatoes. Overall, the shrivel on 'Round' tomatoes was minimal throughout the trial (Figure 5).

For 'Rosa' tomatoes, the results varied between trials (Figure 6). In the first trial, no significant difference in shrivel was observed among treatments after the first 7 days; however, during the shelf life, tomatoes treated with a top sheet had significantly less shrivel than the bottom sheet treated and control tomatoes. After 14 days, the converse was shown, with the controls having the least amount of shrivel. In Trial 2, there was no difference in shrivel between SO_2 treatments and the respective controls throughout the trial period. In the third trial, tomatoes treated with SO_2 consistently had less shrivel than the controls throughout the trial.

Firmness

Prior to the initial cold storage, samples were taken for firmness testing to establish a baseline representing the firmness of the fruit on arrival. Measurements at each initial evaluation were then compared to this baseline. For all varieties, a natural decline in firmness was observed when comparing the readings for the control to the arrival reading (Table 4).

For 'Roma' tomatoes, firmness values for all treatments were similar throughout the trial; however, in the 21-day cold storage, tomatoes treated with a top sheet were significantly less firm than the controls. For the 'Round' tomatoes, there was also a natural decline in firmness relative to the arrival readings, as noted in the controls. However, it was observed that tomatoes treated with a bottom sheet maintained readings relative to that of the arrival quality for up to 21 days in cold storage. 'Rosa' tomatoes held their firmness throughout the trial period of
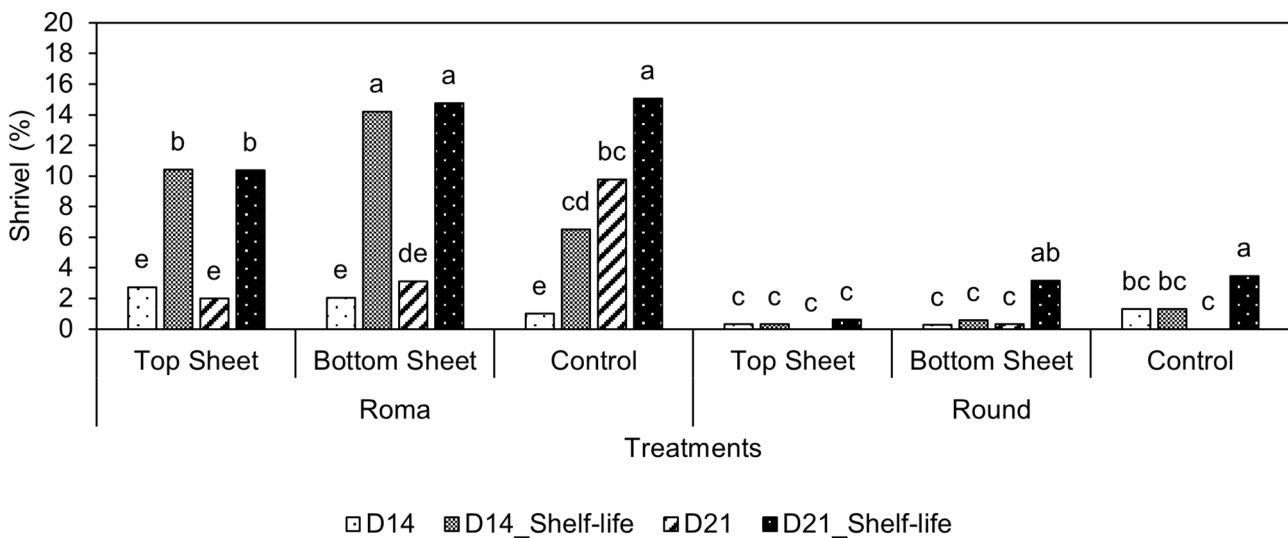
Figure 5: Percentage shrivel recorded on 'Roma' and 'Round' tomatoes after 14 (D14) and 21 (D21) days in cold storage at 10 °C and the respective 7-day shelf-life period at 18 °C. Bars with the same lettering do not differ significantly ($p > 0.05$) according to Fisher's least significant difference test. Data presented are the average of two trials.
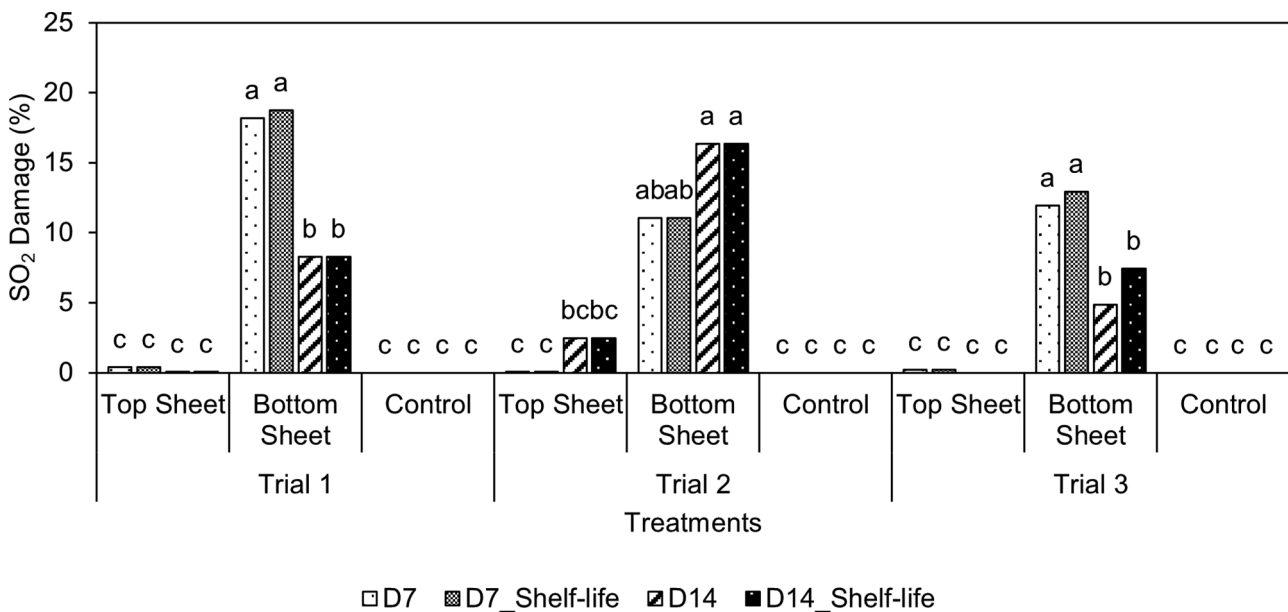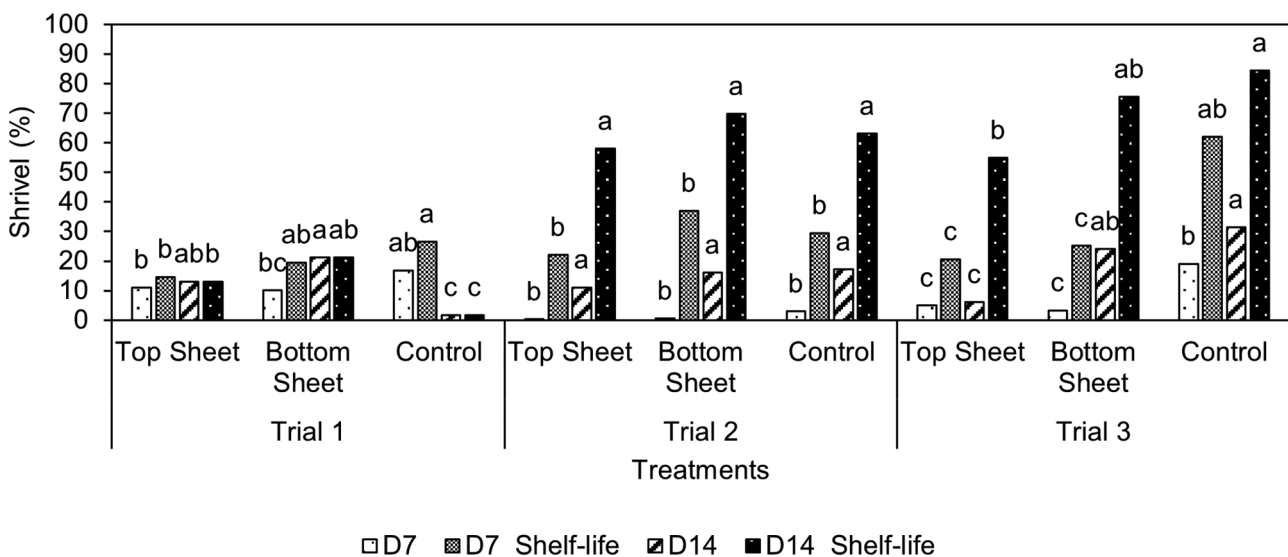

Figure 6: Percentage shrivel recorded on 'Rosa' tomatoes after 7 (D7) and 14 (D14) days of initial storage with SO_2 sheets and a further 7 days at shelf life. Temperature was kept constant at 18 °C for the duration of the trials. Bars with the same lettering do not differ significantly ($p > 0.05$) according to Fisher's least significant difference test.

Table 4: Average firmness (N) recorded on 'Roma', 'Round' and 'Rosa' tomatoes after initial storage with the SO$_2$ sheets

Variety	Treatment / storage	Average firmness (N)		
		D0 (Arrival)	D14 at 10 °C	D21 at 10 °C
'Roma'	Top sheet	22.45 a	17.35 bc	16.37 c
	Bottom sheet		15.98 c	17.06 bc
	Control		16.76 bc	18.82 b
'Round'	Top sheet	16.27 a	11.86 c	10.68 c
	Bottom sheet		15.59 ab	13.72 a–c
	Control		11.76 c	11.66 c
		D0 (Arrival)	D7 at 18 °C	D14 at 18 °C
'Rosa'	Top sheet	9.80 bc	10.10 a–c	10.68 a–c
	Bottom sheet		9.70 c	11.57 a–c
	Control		10.00 a–c	11.27 a

The same lettering indicates no significant difference (p > 0.05) according to Fisher's least significant difference test. Data presented are the average of two trials for 'Roma' and 'Round' tomatoes and three trials for 'Rosa' tomatoes.

14 days. There was no significant drop in firmness relative to the baseline reading, and there was no difference between the control and SO$_2$-treated tomatoes.

Discussion

Tomatoes are picked, packed, and distributed to retailers within a period of 2 to 3 days. This is primarily due to tomatoes ripening and deteriorating quickly. This quick turnaround time puts pressure on the packhouse and retailer to get the fruit out. Based on information gathered from industry stakeholders on preferred storage times, the 'Roma' and 'Round' tomatoes were stored for up to 14 and 21 days in cold storage. The 'Rosa' tomatoes were stored for up to 7 and 14 days. These time frames could open opportunities for transporting tomatoes to new markets located further away than those currently supplied, or even for storing tomatoes for longer when production volumes are high.

Spoilage of tomatoes in storage due to pathogens is an important aspect to consider for the postharvest preservation of the crop. The resulting postharvest diseases contribute to losses in both quantity and quality of the crop.[23] The application of SO$_2$ as a postharvest decay control strategy was investigated in this study to prevent pathogen proliferation and preserve the quality of tomatoes. For years, SO$_2$ has been successfully applied to grapes. In the USA, it is a widespread practice to first fumigate harvested grapes with a high dosage (1000–5000 ppm) of SO$_2$, followed by lower dosages over time, gradually decreasing the dose with each application.[24] Grape farmers in the USA still use SO$_2$ fumigation; however, fumigation during storage is now being done by treatment with a SO$_2$ pad within the grape carton. The idea around this methodology is that the initial high SO$_2$ dosage will sterilise the surface of the grapes, and the in-package treatment will allow for a lower dose of SO$_2$ to be constantly released to offer added protection against decay development over the storage period. In other areas of the world, like South Africa, table grape farmers rely solely on the SO$_2$ packaging insert to control decay development in their grapes.

In this study, we used a dual-release SO$_2$ sheet. The active ingredient, sodium metabisulfite (Na$_2$S$_2$O$_5$), is encapsulated between layers of non-woven material and plastic. The layers that make up the dual-release system are designed such that when the Na$_2$S$_2$O$_5$ encounters moisture from the respiration of the fruit and humidity of the environment, SO$_2$ gas is released. The first stage releases a high dose (mimicking fumigation), and the second stage releases a lower dose of SO$_2$ to keep the fruit surface sterile and prevent decay.

Decay symptoms observed in these trials were presented primarily as water-soaked and black sunken lesions that are consistent with black mould infection caused by *Alternaria* spp. Decay on 'Roma' tomatoes was reduced by both the top and bottom sheet applications. The results suggest that short-term storage (14 days) of 'Roma' tomatoes benefits from using a top SO$_2$ sheet, whilst for long storage periods (21 days), a bottom sheet would be better. However, use of a bottom sheet does present complications because of SO$_2$ damage that can occur. The 'Round' tomatoes used in this study seemed to be sensitive to pathogen attack. Considerable amounts of decay were observed throughout storage and shelf life. Decay on 'Round' tomatoes was not shown to be statistically reduced with the application of a SO$_2$ sheet, even though numerical values of the treatments were lower than the control. Decay on 'Rosa' tomatoes differed between trials conducted. This shows the natural variation that occurs throughout the year and between seasons. Due to the availability of tomatoes, the trials were conducted over a period of 1 year (2020–2021). Tomatoes for the first trial were sourced in November, which is the start of rising summer temperatures. The conditions in the Western Cape over this time are often conducive to pathogen development, which could be what was observed. The second trial was conducted in September (early spring). These were early season fruit and may have been more steady and less prone to pathogen attack. The third trial was conducted in November 2021, and whilst decay was present, the overall inoculum level was low. Literature suggests that the time of harvest has a significant impact on fruit quality and that there are definitive differences between tomatoes harvested in autumn and spring.[25] In addition to the timing of harvest, the reduced inoculum pressure over the course of the year could be attributed to preharvest practices on the farm. Natural variation in decay over time is a possibility, and played a role in these trials. Overall, though, when looking at the effect of an SO$_2$-generating sheet on the reduction of natural decay on 'Rosa' tomatoes, both the top and bottom SO$_2$ sheet applications were useful in reducing decay, with results favouring the use of a top sheet application for this variety.

We have shown that the application of a SO$_2$ sheet during storage also had an impact on the decay that developed after shelf life when the SO$_2$ sheet was removed. This was consistent for all varieties. Whilst decay did develop, it was much less than what was recorded on the controls. This is explained in that SO$_2$ sterilises the fruit surface and can kill off spores and mycelia on the fruit surface; however, it has less impact on latent infections[26] which come into play in the absence of SO$_2$ (removal of sheets). The further development of decay during the shelf life indicates that residue left on the fruit surface after sheets are removed is not adequate to prevent the proliferation of fungi.

We did not directly compare varieties; however, the variations were apparent. Variations in decay amongst tomato varieties have been previously reported. Sinha et al.[23] reported that, under ambient conditions after 16 days, the variety 'Sofol' showed 21% decay whilst the variety 'Roma' had 43% decay. 'Sofol' is a round-type tomato, and with that in mind, we see the opposite occurrence with regard to decay patterns in this study, perhaps due to differences in cultivation practices and climatic regions.

An accumulation or localisation of SO$_2$ gas can cause SO$_2$ bleaching or damage to the fruit.[22] Through the sorting process, all defects observed at the cold storage evaluation were removed before placing the fruit at shelf life. When the fruit is still cold, the condensation that forms on the fruit surface makes it hard to see slight SO$_2$ damage. Therefore, any SO$_2$-damaged tomatoes observed during the shelf life can be attributed to improper sorting at the initial cold storage evaluation.

A key finding from these trials is that the bottom sheet yielded more SO$_2$ damage than the top sheet. The damage observed from the bottom sheet could be attributed to the increased pressure of fruits sitting on top of each other and pressing onto the bottom sheet. This is the most obvious explanation, as damage seen with the top sheet application was not as pronounced. With the application of a top sheet, the fruits are not pressing onto the sheet, even if

contact is made. For all varieties, the amount of SO_2 damage observed with the application of a top sheet was negligible (less than 1%); in comparison, SO_2 damage from a bottom sheet was up to 5% on 'Roma' tomatoes, up to 16% on 'Round' tomatoes and up to 18% on 'Rosa' tomatoes. SO_2 damage can have various manifestations on fruit, the most common being bleaching (discolouration) and fruit pitting.[27] Both these symptoms were observed on tomatoes to varying degrees. In some instances, only a slight discolouration around the stem scar was visible; however, where pressure was applied, as with the application of a bottom sheet, the fruit showed signs of pitting. Unfortunately, whilst a slight discolouration can be overlooked, pitting would render the tomatoes unmarketable due to appearance.

Tomatoes have a high water content and are prone to shrinkage due to moisture loss after harvest.[28] Of the three varieties used in this trial, 'Roma' and 'Rosa' tomatoes were most affected by shrivelling, during the shelf-life storage phase. In contrast, shrivel on 'Round' tomatoes was minimal. This could be due to the initial water content of the varieties and the respective respiration rates. With 'Roma' tomatoes, short-term exposure to SO_2 had no effect on shrivelling; however, long-term exposure to SO_2 did reduce the incidence of shrivel. A similar pattern was observed with the 'Round' tomatoes. 'Rosa' tomatoes benefitted from SO_2 exposure throughout the trial period.

In conclusion to a study on the firmness of tomatoes, Batu[29] proposed two firmness readings for large tomatoes. The first is for fruit marketed at retail level, which should render a firmness reading greater than 1.45 N/mm (0.15 kg), and the second value is for tomatoes for home consumption, which should be greater than 1.28 N/mm (0.13 kg). The firmness of both large tomato varieties recorded in this study was well above these readings, even for the control fruit; the 'Roma' tomatoes were firmer than the 'Round' tomatoes. This finding corresponds with the literature that states that processing tomatoes like 'Roma' tend to be firmer and last longer than fresh market tomatoes. Varietal differences in firmness are accounted for by properties such as cell wall composition.[30] A gradual decline in firmness was observed across all three varieties used in this study, and the results obtained show that SO_2 exposure did not negatively affect this quality parameter.

Conclusion

The results confirm that a SO_2 sheet applied as a top sheet can be used to effectively reduce decay by a minimum of 58% on 'Roma' tomatoes when refrigerated at 10 °C for up to 14 days and then exposed to ambient temperatures for another 7 days. Alternatively, a bottom sheet application can reduce decay by 60% if tomatoes are kept refrigerated for 21 days, but the potential SO_2 damage must be considered. 'Round' tomatoes in this study only benefitted from a marginal reduction in decay by the application of a bottom sheet. Continuous refinement of this product for use on 'Round' tomatoes will need to focus on reducing the SO_2 damage. The potential for decay control in 'Rosa' tomatoes is high because, even though trials differed, a decay inhibition of up to 80% was obtained at different intervals.

The outcome of this study builds a strong case for the use of SO_2-generating sheets to be used on certain tomato varieties under commercial conditions for the purpose of decay reduction during storage. As an added benefit, the use of SO_2-generating sheets may assist with reduced shrivelling of tomatoes during storage. Furthermore, SO_2 does not negatively affect the firmness of tomatoes during storage.

Acknowledgements

Competing interests

We have no competing interests to declare.

Authors' contributions

C.K.D.-S.: Conceptualisation; methodology; data collection; sample analysis; validation; writing; project management. J.C.M.-H.: Student supervision; editing of manuscript. F.A.V.: Student supervision; project leadership; funding acquisition. C.L.L.: Student supervision.

References

1. Food and Agriculture Organization of the United Nations (FAO). FAOSTAT – The Food and Agriculture Organization Corporate Statistical Database [webpage on the Internet]. c2021 [cited 2023 Aug 05]. Available from: https://www.fao.org/faostat/en/#data/Q

2. Bahammou N, Cherifi O, Bouamama H, Cherifi K, Moubchir T, Bertrand M. Postharvest control of gray mould of tomato using seaweed extracts. J Mater Environ Sci. 2017;8(3):831–836. http://www.jmaterenvironsci.com/

3. South African Department of Agriculture, Forestry and Fisheries (DAFF). Production guidelines for tomato [document on the Internet]. c2014 [cited 2023 Aug 05]. Available from: https://www.dalrrd.gov.za/Portals/0/Brochures%20and%20Production%20guidelines/Production%20Guidelines%20Tomato.pdf

4. South African Department of Agriculture, Forestry and Fisheries (DAFF). Trends in the agricultural sector 2017: Tomatoes. Pretoria: DAFF; 2017. p. 59–60.

5. Pinheiro J, Goncalves EM, Silva CLM. Alternative technologies for tomato postharvest quality preservation. CAB Rev. 2013;8:61. https://doi.org/10.1079/PAVSNNR20138061

6. Bartz JA, Sargent SA, Scott JW. Postharvest quality and decay incidence among tomato fruit as affected by weather and cultural practices. UF/IFAS Extension; 2015. Available from: http://edis.ifas.ufl.edu/

7. Moneruzzaman KK, Hossain ABMS, Sani W, Saifuddin M. Effect of stages of maturity and ripening conditions on the physical characteristics of tomato. Am J Biochem Biotechnol. 2008;4(4):329–335. https://doi.org/10.3844/ajbbsp.2008.329.335

8. Hurst W. Harvest, handling and sanitation. In: Commercial tomato production handbook. University of Georgia Extension Bulletin 1312. 2017. p 38–42. https://extension.uga.edu

9. Freedman BJ. Sulphur dioxide in foods and beverages: Its use as a preservative and its effects on asthma. Br J Dis Chest. 1980;74(2):128–134. https://doi.org/10.1016/0007-0971(80)90023-6

10. Sandarani MDJC, Dasanayaka DCMCK, Jayasinghe CVL. Strategies used to prolong the shelf life of fresh commodities. J Agri Sci Food Res. 2018;9(1):206.

11. Gabler FM, Mercier J, Jimenez JI, Smilanick JL. Integration of continuous biofumigation with *Muscodor albus* with pre-cooling fumigation with ozone or sulfur dioxide to control postharvest gray mold of table grapes. Postharvest Biol Technol. 2010;55:78–84. https://doi.org/10.1016/j.postharvbio.2009.07.012

12. Cantin CM, Palou L, Bremer V, Michailides TJ, Crisosto CH. Evaluation of the use of sulfur dioxide to reduce postharvest losses on dark and green figs. Postharvest Biol Technol. 2011;59:150–158. https://doi.org/10.1016/j.postharvbio.2010.09.016

13. Cantin CM, Minas IS, Jimenez M, Manganaris GA, Michailides TJ, Crisosto CH. Sulfur dioxide fumigation alone or in combination with CO_2 – enriched atmosphere extends the market life of highbush blueberry fruit. Postharvest Biol Technol. 2012;67:84–91. https://doi.org/10.1016/j.postharvbio.2011.12.006

14. Garcia-Fuentes AR, Wirtz S, Vos E, Verhagen H. Short review of sulphites as food additives. Eur J Nutr Food Saf. 2015;5(2):113–120. https://doi.org/10.9734/EJNFS/2015/11557

15. EFSA Panel on Food additives and Nutrient Sources added to Food (ANS). Scientific Opinion Follow up on the re-evaluation of sulfur dioxide (E 220), sodium sulfite (E 221), sodium bisulfite (E 222), sodium metabisulfite (E 223), potassium metabisulfite (E 224), calcium sulfite (E 226), calcium bisulfite (E 227) and potassium bisulfite (E 228) as food additive. EFSA J. 2022;14(4):4438. https://doi.org/10.2903/j.efsa.2022.7594

16. Leone IA, Brennan E. Modification of sulfur dioxide injury to tobacco and tomato by varying nitrogen and sulfur nutrition. J Air Pollut Control Assoc. 1972;22(7):544–547. https://doi.org/10.1080/00022470.1972.10469677

17. Padhi SK, Dash M, Swain SC. Effect of sulphur dioxide on growth, chlorophyll, and sulphur content of tomato (*Solanum lycopersicum* L.). Eur Sci J. 2013;9:36.

18. Basson E, Meitz-Hipkins JC, Lennox CL. Morphological and molecular identification of fungi associated with South African apple core rot. Eur J Plant Pathol. 2019;153:849–868. https://doi.org/10.1007/s10658-018-1601-x

19. SAS Institute, Inc. SAS version 9.4. Cary, NC: SAS Institute Inc.; 2023.

20. Shapiro SS, Wilk MB. An analysis of variance test for normality (complete samples). Biometrika. 1965;52:591–611. https://doi.org/10.1093/biomet/52.3-4.591

21. Ott R. An introduction to statistical methods and data analysis. Belmont, CA: Duxbury Press; 1998. p. 807–837.

22. Ahmed S, Roberto SR, Domingues AR, Shahab M, Chaves Junior OJ, Sumida CH, et al. Effects of different sulfur dioxide pads on *Botrytis* mold in 'Italia' table grapes under cold storage. Horticulturae. 2018;4:29. https://doi.org/10.3390/horticulturae4040029

23. Sinha SR, Singha A, Faruquee M, Jiku MAS, Rahaman MA, Alam MA, et al. Post-harvest assessment of fruit quality and shelf life of two elite tomato varieties cultivated in Bangladesh. Bull Natl Res Cent. 2019;43:185. https://doi.org/10.1186/s42269-019-0232-5

24. Smilanick JL, Henson DJ. Minimum gaseous sulfur dioxide concentrations and exposure periods to control *Botrytis cinerea*. Crop Prot. 1992;11:535–540. https://doi.org/10.1016/0261-2194(92)90171-Z

25. Kasampalis D, Tsouvaltzis, P, Siomos A. Tomato fruit quality in relation to growing season, harvest period, ripening stage and postharvest storage. Emir J Food Agric. 2021;33(2):130–138. https://doi.org/10.9755/ejfa.2021.v33.i2.2176

26. Smilanick JL, Hartsell PI, Henson D, Fouse DC, Assemi M, Harris CM. Inhibitory activity of sulfur dioxide on the germination of spores of *Botrytis cinerea*. Phytopathology. 1990;80(2):217–220. https://doi.org/10.1094/Phyto-80-217

27. Zaffoli JP, Latorre BA, Naranjo P. Hairline, a postharvest cracking disorder in table grapes induced by sulfur dioxide. Postharvest Biol Technol. 2008;47:90–97. https://doi.org/10.1016/j.postharvbio.2008.06.013

28. Arah IK, Amaglo H, Kumah EK, Ofori H. Preharvest and postharvest factors affecting the quality and shelf life of harvested tomatoes: A mini review. Int J Agron. 2015;2015, Art. #478041. http://dx.doi.org/10.1155/2015/478041

29. Batu A. Determination of acceptable firmness and colour values of tomatoes. J Food Eng. 2004;61:471–475. https://doi.org/10.1016/S0260-8774(03)00141-9

30. Tigist M, Workneh TS, Woldetsadik K. Effects of variety on the quality of tomato stored under ambient conditions. J Food Sci Technol. 2013;50(3):477–486. https://doi.org/10.1007/s13197-011-0378-0

AUTHORS:
John P. Rheeder[1] (iD)
Liana van der Westhuizen[2] (iD)

AFFILIATIONS:
[1]Biotechnology and Consumer Sciences, Faculty of Applied Sciences, Cape Peninsula University of Technology, Cape Town, South Africa

[2]Applied Microbial and Health Biotechnology Institute, Cape Peninsula University of Technology, Cape Town, South Africa

CORRESPONDENCE TO:
Liana van der Westhuizen

EMAIL:
vanderwesthuizenL@cput.ac.za

Fusarium and fumonisin in GM maize grown by small-scale farmers in KwaZulu-Natal, South Africa

The genetic modification (GM) of maize to contain proteins that act to control insects has become a widespread agricultural practice. Although the reduction of insect damage to maize ears could potentially increase crop yield, rural small-scale farmers might be reluctant to buy expensive GM seed every season even when the lower fungal infection of the GM maize might also result in health benefits. This study was conducted over 5 years in three districts of northern KwaZulu-Natal, South Africa, to study the *Fusarium verticillioides* and *F. proliferatum* infection and fumonisin contamination levels in *Bt* maize, Roundup-Ready® maize, conventional commercial maize and traditional landrace maize planted by rural farmers following their traditional agricultural practices. Mean *Fusarium* infection rates varied between 3.0% and 38.3% with large standard deviations. *Fusarium* infection was not significantly different ($p>0.05$) between the various genotypes, possibly due to the wide variation in results and low sample numbers. Although the fumonisin results also showed wide variation, the trend of contamination was lower in *Bt* maize compared to conventional commercial genotypes. The mean fumonisin levels in *Bt* hybrids were mostly <300 µg/kg, ensuring a safe maize supply in populations consuming maize as a dietary staple. The wide variations in *Fusarium* and fumonisin levels within each district point to the influence of local agricultural practices, local environmental conditions, and seasonal variations. Reducing exposure to fumonisins in these communities requires both further attention to the possible influence of these factors, as well as the use of appropriate post-harvest strategies.

Significance:

This study was the first in South Africa to follow rural farmers planting *Bt* and commercial maize hybrids as well as open-pollinated landrace maize, according to their traditional agricultural practices. The results show that in some instances the *Bt* maize had the lowest fumonisin levels.

Introduction

Bacillus thuringiensis (*Bt*) genes express an insecticidal crystal protein that has been shown to be a powerful deterrent of lepidopteran and coleopteran pests in maize, thereby reducing damage caused by these insect pests.[1-3] Insect damage in maize, particularly by stalk borers, is the main mechanism of fungal infection which can cause reduced crop yields. Hence, efficient insect control can have the additional benefit of a concomitant reduction in *Fusarium* infection rates, particularly that of the fumonisin mycotoxin producer, *Fusarium verticillioides*.[4,5]

South Africa has an estimated 50 000 large-scale conventional commercial farmers, 240 000 small-scale farmers and more than 1 million farmers who produce food on a subsistence level.[6,7] Insect-resistant, transgenic (*Bt*) maize has seen impressive adoption rates by conventional maize farmers since its introduction in the 1998/1999 crop season. A survey among maize farmers planting *Bt* maize in South Africa showed average maize crop yield improvements of 11% and reductions in pesticide use of between 50% and 80%.[8] At a subsistence level, the predominant practice remains the planting of open-pollinated landrace varieties using seed held over from the previous crop. Between these two groups, small-scale farmers plant either open-pollinated or conventional commercial varieties, but yields can be low.[6] However, the use of the improved *Bt* hybrids by small-scale farmers can be problematic. Both subsistence and small-scale farmers in South Africa may not be able to make use of the technology if it is not appropriate to their needs, if it is too expensive, or if it is not available to them because there is no government infrastructure or incentives for companies to sell genetically modified (GM) seed to small-scale farmers. Nonetheless, more efficient maize production methods suitable for small-scale farmers would bring economic benefits and add to the level of food security of the whole southern African region, where maize is the staple food, by increasing the availability of maize in the countryside.

In this study, we investigated the possible economic benefits that would accrue to smallholder farmers in South Africa through the adoption of *Bt* maize. The University of Pretoria, with funding from the Rockefeller Foundation, initiated this multi-year study of crop yields among farmers in three districts of the KwaZulu-Natal Province, South Africa.[6] For farmers adopting *Bt* technology, the data collected during the study years show economic benefits from reduced insecticide application and reduced harvest loss due to stalk borer in districts and/or years in which there was significant insect pressure.

Given the perceived benefits of GM maize, the aim of the multi-year study reported here, as an adjunct to the above economics study, was to measure the effect on *Fusarium* infection and fumonisin contamination of introducing new GM technologies in three rural maize growing districts in northern KwaZulu-Natal, South Africa. Given the improvement in economic parameters shown by the study, it was hoped that there would also be reduced fumonisin exposure and thus a health advantage to the adoption of these GM varieties by smallholder farmers, who also consume a portion of their crop in a subsistence scenario.

Materials and methods

Study areas and sampling

This study was conducted over five successive crop seasons among dry land small-scale farmers in the districts of Simdlangentsha and Hlabisa, northern KwaZulu-Natal Province, South Africa. The study was extended to the Dumbe district for the last three seasons. Workshops were held, 2 years prior to the start of the trial, to introduce *Bt* maize to smallholder farmers during which they were provided with small bags of *Bt* maize seed for free. Thereafter, farmers who wanted to plant the new seed were required to purchase it. Further, for season 2, the Roundup-Ready® (RR) maize seed, genetically modified to enable the use of Roundup® herbicide, was also made available. In addition, farmers in the study areas planted conventional commercial varieties and some planted open-pollinated traditional maize.

Problems associated with maize sampling in rural subsistence farming areas have been previously noted.[9] Samples of all the above maize types were collected from farmers participating in the agricultural economics study of Gouse et al.[6] A minimum of 2 kg of shelled maize held by the farmer was obtained to be as representative of the harvest as possible. As the study did not involve organised field trials but rather was aimed at monitoring the maize harvested by farmers using their traditional practices and irrespective of their seed choice, sample numbers for traditional maize, conventional commercial maize, *Bt* maize and RR maize varied widely between districts and between years, depending on the maize type favoured in any one crop year, crop yield, farmer availability to field workers and farmer acceptance of participation in the economic study of the University of Pretoria.[6] Collected maize samples were stored in linen bags at 4 °C until analysed for fumonisins and *Fusarium* infection.

Fumonisin determination

Fumonisins B$_1$ (FB$_1$), B$_2$ (FB$_2$) and B$_3$ (FB$_3$) were determined by high-performance liquid chromatography as previously described.[10,11] Briefly, maize was ground in a laboratory mill to a fine meal and extracted with 100 mL methanol/water by homogenisation. A 10 mL aliquot was applied to a strong anion exchange solid phase extraction cartridge and the fumonisins were eluted with acetic acid in methanol. The purified extract was evaporated to dryness with nitrogen gas at 60 °C and the dried residue was stored at 4 °C whilst awaiting chromatographic analysis. The derivatised extracts were separated on a reversed-phase Luna 4 μm C18 (2) (150 x 4.6 mm I.D.) column (Phenomenex, Torrance, CA, USA) and the isocratic mobile phase of methanol/0.1 M sodium dihydrogen phosphate (pH 3.35) (77:23) was pumped at a flow rate of 1 mL/min. The chromatographic system consisted of a Rheodyne 7725i injector (Cotati, CA, USA), Waters Model 510 solvent delivery system (Milford, MA, USA), Borwin Chromatography Integration Software (Varian JBMS Developements, Le Fontanil, France) and Waters Fluorescence 474 detector (excitation – 335 nm and emission – 440 nm). Fumonisin data are indicated throughout the current study as total fumonisin levels (FB$_1$+FB$_2$+FB$_3$) and were not corrected for recovery.[10]

Fusarium determination

As *Fusarium verticillioides* and *F. proliferatum* are the main species contributing to fumonisin contamination in South African studies, they were identified by morphology and quantified as percent infected kernels as previously described.[7,12-15] Briefly, a subsample (approximately 150 g) from each well mixed sample was surface-disinfected for 1 min in a 3.5% sodium hypochlorite solution and rinsed twice in sterile water. One hundred kernels per subsample were plated (five kernels per petri dish) onto 1.5% malt extract agar (MEA) containing 150 mg/L novobiocin to minimise bacterial growth. The MEA plates were incubated in the dark at 25 °C for 5–7 days. *Fusarium* species that developed from the kernels were then identified according to their morphological characteristics.[14]

Statistical analyses

Statistical analysis was done on the mycology and fumonisin data using SPSS software (Chicago, IL, USA) to determine any significant differences between the maize categories/groups (i.e. traditional, conventional, *Bt*, and Roundup-Ready). Locations were not compared with each other, neither were seasons. Analyses were conducted on the '*F. verticillioides* + *F. proliferatum*' FvFp and total fumonisin variables only, with natural log transformation of the latter variable. Pearson's correlation coefficients, between the FvFp fungal and ln-transformed total fumonisin levels, were calculated for each location.

Results and discussion

This study was intended to monitor the *Fusarium* and fumonisin contamination of the maize harvest of small-scale farmers planting a variety of maize genotypes, where the farmers themselves selected the varieties to plant and conducted their own normal agricultural practices.[16] A summary of total fumonisins is given in Table 1. Detailed data (number of samples, number of positive samples, mean, standard deviation and range) of *F. verticillioides, F. proliferatum*, their sum, FB$_1$, FB$_2$, FB$_3$ and their sum as total fumonisins are shown in Supplementary table 1. Consequently, sample numbers varied greatly depending on cooperation of the farmers and the actual uptake of the new GM maize varieties. As the results show, standard deviations of the mean data for *Fusarium* infection and total fumonisins were large, and consequently, only a limited number of results were statistically significantly different at the 5% level.

Of the two fumonisin-producing *Fusarium* species analysed, *F. verticillioides* predominated by far and in many cases *F. proliferatum* was not found in the samples. Mean infection rates for *F. verticillioides* ranged from 3.0% (Simdlangentsha *Bt* maize season 3) to 38.3% (Dumbe Roundup-Ready® season 3). The low sample numbers obtained in these results precluded statistical analysis. The infection rates varied widely within years, and, in general, no consistent trend was observed between the different maize genotypes. In only one instance was there a significant difference in mean *F. verticillioides* and *F. proliferatum* infection rates, namely in Simdlangentsha in season 5. In this instance, *Bt* maize (mean 11.4 ± 12.3%) and RR maize (mean 10.7 ± 9.9%) differed significantly ($p<0.05$) from conventional varieties (mean 30.4 ± 24.5%). The low incidence of *F. proliferatum* is similar to mycological results from the Centane and Mbizana areas of the former Transkei in a study conducted over 3 years.[14] In that study of good and mouldy maize harvested by rural farmers, no *F. proliferatum* was detected using the same analytical method. The traditional practice by the subsistence farmers of keeping the best cobs of the harvest for use as seed the following season might have acted as a natural breeding programme, selecting out the hardiest plant material.

Total fumonisin levels varied widely with high standard deviations (Table 1) and no correlation was found with *Fusarium* infection rates. Significant statistical differences between hybrids within individual areas were limited due to the high standard deviations. Nevertheless, Table 1 shows the five cases in which statistical differences were obtained where the *Bt* hybrid was statistically ($p<0.05$) less contaminated than conventional varieties. Overall, *Bt* maize performed well with low fumonisin levels, apart from a few notable exceptions, such as Hlabisa in season 1. In this case, it was reported that the *Bt* hybrid planted produced a poor ear morphology, with sheath leaves that did not cover the tip of the ear. In addition, these ears did not droop at maturity, so were exposed to the unseasonal rainfall at Hlabisa in 2004. This explains the high fumonisin levels in the Bt hybrid in that particular year.[6] This poor ear morphology was corrected by the relevant seed company in the following years. The maximum level set for fumonisin mycotoxins for raw maize (4000 μg/kg) by the Codex Alimentarius Commission have been incorporated into the South African health regulations.[17] Although none of the mean total fumonisin levels in the Supplementary table 1 exceeded 4000 μg/kg, there were several of the highest levels of the ranges of the total fumonisin levels exceeding this set maximum level for fumonisin mycotoxins for raw maize. In general, the *Bt* hybrid achieved mean contamination levels of the order of or less than 300 μg/kg – a level that has been suggested to achieve a safe maize supply in rural areas, where a large daily consumption of maize is part of the typical diet.[18] Consequently, these lower fumonisin levels carry health benefits in that they reduce exposure to these carcinogenic mycotoxins in rural areas.[9,19]

Table 1: Mean values of all samples ($\pm$ standard deviation) of total fumonisin (sum of FB$_1$, FB$_2$ and FB$_3$) contamination (µg/kg) of maize varieties collected in three districts (Simdlangentsha, Hlabisa and Dumbe) of KwaZulu-Natal, South Africa over 5 years

Maize type	Total fumonisins $\pm$ standard deviation (µg/kg)		
	Simdlangentsha	**Hlabisa**	**Dumbe**
Season 1			
Traditional	**753 $\pm$ 814 (5)a***	159 $\pm$ 91 **(4)**	No sample
Conventional	**623 $\pm$ 917 (8)a***	450 $\pm$ 627 **(11)**	No sample
Bt	**239 $\pm$ 411 (7)b***	1150 $\pm$ 1430 **(8)**	No sample
Season 2			
Traditional	271 $\pm$ 352 **(6)**	250 $\pm$ 353 **(2)**	No sample
Conventional	815 $\pm$ 1150 **(12)**	**472 $\pm$ 506 (15)a****	No sample
Bt	396 $\pm$ 395 **(7)**	**22 $\pm$ 25 (11)b****	No sample
Season 3			
Traditional	996 $\pm$ 1290 **(5)**	2070 $\pm$ 3230 **(9)**	**426 $\pm$ 606 (6)a***
Conventional	No sample	1380 $\pm$ 2820 **(8)**	No sample
Bt	152 $\pm$ 64 **(2)**	804 $\pm$ 989 **(13)**	**110 $\pm$ 139 (3)a***
Roundup-Ready®	1200 $\pm$ 949 **(6)**	870 $\pm$ 1600 **(12)**	**2600 $\pm$ 4140 (3)b***
Season 4			
Traditional	No sample	No sample	3390 $\pm$ 5240 **(6)**
Conventional	**1580 $\pm$ 2590 (9)a***	66 $\pm$ 103 **(5)**	932 $\pm$ 1510 **(16)**
Bt	**51 $\pm$ 70 (6)b***	129 $\pm$ 138 **(10)**	501 $\pm$ 927 **(4)**
Roundup-Ready®	**2230 $\pm$ 4490 (5)a***	466 $\pm$ 1370 **(12)**	179 $\pm$ 223 **(2)**
Season 5			
Traditional	No sample	No sample	No sample
Conventional	**500 $\pm$ 962 (10)a***	16 $\pm$ 26 **(10)**	1690 $\pm$ 4170 **(8)**
Bt	**55 $\pm$ 169 (11)b***	45 $\pm$ 117 **(13)**	1140 $\pm$ 2130 **(15)**
Roundup-Ready®	**2 $\pm$ 3 (11)b***	101 $\pm$ 245 **(10)**	674 $\pm$ 1440 **(12)**

Numbers in brackets indicate the number of samples collected; significant differences between maize types within a year within a location are indicated by bold type; values with different lowercase letters are significantly different, while those with the same lowercase letters are not significantly different:

*p < 0.05

**p < 0.001

Studies conducted in the former Transkei over a number of years have also shown that fumonisin contamination levels can vary widely within a district.[14,15] Thus, improvements in food safety by reducing fumonisin levels may lie in targeting individual agronomic practices and understanding seasonal environmental factors.[9,19] Studies conducted in Argentina and the Philippines to model the effects of environment, insect damage and maize genotype on fumonisin contamination concluded that most of the variability in fumonisin levels could be ascribed to location and weather, rather than to the maize genotype.[1]

The results achieved with RR maize suggest that this hybrid type carries no advantage over *Bt* maize planted in these districts. Indeed, in seasons 3 and 4, RR maize was significantly ($p<0.05$) more contaminated than the *Bt* hybrids planted, resulting in some of the highest mean contamination levels over all 5 years of the study period. Only a limited number of samples of traditional maize were obtained. The results in Table 1 show that the mean fumonisin contamination levels of traditional maize compare mostly favourably with those of conventional commercial hybrids grown in the same district in the same year. The traditional practice of keeping the best cobs of the harvest for use as seed the following season might have acted as an adaptation mechanism in the form of a natural breeding programme. The variation in fumonisin levels and their reduction may best be approached through studies of individual agronomic practices.

Over the past 20 years several field trials have indicated that, together with reduced *Fusarium* infection rates, a reduction in the levels of the fumonisin mycotoxins was observed.[20-26] Meta-analyses of data contained in 21 publications have confirmed the positive effect of *Bt* genetic modification on fumonisin levels, with decreases of 14–67%, although the degree of reduction reported can depend on the type of statistical approach followed.[27] Thus, the use of GM maize seed that expresses insecticidal proteins has become a common maize production practice.[28]

Conclusions

This study was the first in South Africa to follow rural farmers who planted various maize hybrids as well as open-pollinated landrace maize according to their traditional agricultural practices. The maize harvested by these smallholder/subsistence farmers, irrespective of their seed choice of traditional, commercial, *Bt* or Roundup-Ready® maize, showed lower fumonisin contamination in general and the *Bt* maize in some instances showed the lowest fumonisin levels. However, the sustainability of planting *Bt* maize, which is a technology aimed at reducing insect damage and not directly at harvest yield, has been questioned[29], as it depends on the perceptions of the rural farmers as to whether they are prepared to pay the additional cost of this technology.

The large variations in fumonisin levels observed within each area for each maize variety planted indicate the impact of local farming practices, plus the immense role that environmental and climatic conditions played in the outcome of this study. Thus, more controlled studies by researchers (rather than smallholder farmers) with larger sample sizes would be required to determine if GM maize would be less susceptible to *Fusarium* infection and fumonisin contamination under South African conditions. It should also be kept in mind that traditionally the smallholder/subsistence farmers select their best seed visually when they harvest their crops for planting in the following season. Thus, to ensure that fumonisin levels are kept as low as possible, both improved agricultural practices as well as appropriate post-harvest interventions are required to achieve a safe rural maize supply.[30]

Acknowledgements

Competing interests

We have no competing interests to declare.

Authors' contributions

J.P.R.: Conceptualisation; methodology; data collection; data analysis; writing – the initial draft. L.v.d.W.: Methodology; data analysis; writing – revisions.

References

1. De la Campa RD, Hooker DC, Miller JD, Schaafsma AW, Hammond BG. Modeling effects of environment, insect damage, and *Bt* genotypes on fumonisin accumulation in maize in Argentina and the Philippines. Mycopathologia. 2005;159(4):539–552. https://doi.org/10.1007/s11046-005-2150-3

2. Dowd PF. Biotic and abiotic factors limiting efficiency of *Bt* corn in indirectly reducing mycotoxin levels in conventional fields. J Econ Entomol. 2001;94(5):1067–1074. https://doi.org/10.1603/0022-0493-94.5.1067

3. Munkvold GP, Hellmich RL, Rice LG. Comparison of fumonisin concentrations in kernels of transgenic *Bt* maize hybrids and nontransgenic hybrids. Plant Dis. 1999;83(2):130–138. https://doi.org/10.1094/PDIS.1999.83.2.130

4. Díaz-Gómez J, Marin S, Capell T, Sanchis V, Ramos AJ. The impact of *Bacillus thuringiensis* technology on the occurrence of fumonisins and other mycotoxins in maize. World Mycotoxin J. 2016;9:475–486. https://doi.org/10.3920/WMJ2015.1960

5. Hammond BG, Campbell KW, Pilcher CD, Degooyer TA, Robinson AE, McMillen BL, et al. Lower fumonisin mycotoxin levels in the grain of *Bt* corn grown in the United States in 2000-2002. J Agric Food Chem. 2004;52(5):1390–1397. https://doi.org/10.1021/jf030441c

6. Gouse M, Kirsten JF, Van Der Walt WJ. *Bt* Cotton and *Bt* maize: An evaluation of direct and indirect impact on the cotton and maize farming sectors in South Africa [document on the Internet]. c2008 [cited 2023 Feb 20]. Available from: http://biotechbenefits.croplife.org/paper/bt-cotton-and-bt-maize-an-evaluation-of-direct-and-indirect-impact-on-the-cotton-and-maize-farming-sectors-in-south-africa/

7. Ncube E, Flett BC, Van den Berg J, Erasmus A, Viljoen A. *Fusarium* ear rot and fumonisins in maize kernels when comparing a *Bt* hybrid with its non-*Bt* isohybrid and under conventional insecticide control of *Busseola fusca* infestations. Crop Prot. 2018;110:183–190. https://doi.org/10.1016/10.1016/j.cropro.2017.09.015

8. Gouse M, Pray CE, Kirsten J, Schimmelpfennig D. A GM subsistence crop in Africa: The case of *Bt* white maize in South Africa. Int J Biotechnol. 2005;7(1–3):84–94. https://doi.org/10.1504/IJBT.2005.006447

9. Wild CP, Baan RA, Gelderblom WC, Miller JD, Riley RT, Wu F. Sampling and sample preparation methods for determining concentrations of mycotoxins in foods and feeds. In: Pitt JI, editor. Improving public health through mycotoxin control. Scientific Publications Series, 158. Lyon: International Agency for Research on Cancer; 2012. p. 39–51. https://publications.iarc.fr/_publications/media/download/1372/292f829f75e2d9bd641611e2768b3db005839ad2.pdf

10. Sydenham EW, Shephard GS, Thiel PG, Stockenström S, Snijman PW, Van Schalkwyk DJ. Liquid chromatographic determination of fumonisins B$_1$, B$_2$, and B$_3$ in corn: AOAC/IUPAC collaborative study. J AOAC Int. 1996;79(3):688–696. https://doi.org/10.1093/jaoac/79.3.688

11. Vismer HF, Shephard GS, Rheeder JP, Van der Westhuizen L, Bandyopadhyay R. Relative severity of fumonisin contamination of cereal crops in West Africa. Food Addit Contam Part A Chem Anal Control Expo Risk Assess. 2015;32(11):1952–1958. https://doi.org/10.1080/19440049.2015.1084654

12. Shephard GS, Thiel PG, Stockenström S, Sydenham EW. Worldwide survey of fumonisin contamination of corn and corn-based products. J AOAC Int. 1996;79(3):671–687. https://doi.org/10.1093/jaoac/79.3.671

13. Rheeder JP, Marasas WF, Vismer HF. Production of fumonisin analogs by *Fusarium* species. Appl Environ Microbiol. 2002;68(5):2101–2105. https://doi.org/10.1128/AEM.68.5.2101–2105.2002

14. Rheeder JP, Van der Westhuizen L, Imrie G, Shephard GS. *Fusarium* species and fumonisins in subsistence maize in the former Transkei region, South Africa: A multi-year study in rural villages. Food Addit Contam Part B Surveill. 2016;9(3):176–184. https://doi.org/10.1080/19393210.2016.1154612

15. Marasas WF. Discovery and occurrence of the fumonisins: A historical perspective. Environ Health Perspect. 2001;109(suppl 2):239–243. https://doi.org/10.1289/ehp.01109s2239

16. Phokane S, Flett BC, Ncube E, Rheeder JP, Rose LJ. Agricultural practices and their potential role in mycotoxin contamination of maize and groundnut subsistence farming. S Afr J Sci. 2019;115(9/10), Art. #6221. http://dx.doi.org/10.17159/sajs.2019/6221

17. South African National Department of Health (DoH). Regulations governing tolerances for fungus-produced toxins in foodstuffs: Amendment. Government Notices, Government Gazette, Vol. 40250, 4–5 [document on the Internet]. c2016 [cited 2023 Sep 17]. Available from: https://www.gov.za/sites/default/files/gcis_document/201609/40250gon987.pdf

18. Shephard GS, Burger HM, Rheeder JP, Alberts JF, Gelderblom WC. The effectiveness of regulatory maximum levels for fumonisin mycotoxins in commercial and subsistence maize crops in South Africa. Food Control. 2019;97:77–80. https://doi.org/10.1016/j.foodcont.2018.10.004

19. Pray CE, Rheeder JP, Gouse M, Volkwyn Y, Van der Westhuizen L, Shephard GS. *Bt* maize and fumonisin reduction in South Africa: Potential health impacts. In: Falck-Zepeda J, Gruere G, Sithole-Niang I, editors. Genetically modified crops in Africa: Economic and policy lessons from countries south of the Sahara. Washington DC: International Food Policy Research Institute (IFPRI); 2013. p. 43–57. https://doi.org/10.2499/9780896297951

20. Barroso VM, Rocha LO, Reis TA, Reis GM, Duarte AP, Michelotto MD, et al. *Fusarium verticillioides* and fumonisin contamination in *Bt* and non-*Bt* maize cultivated in Brazil. Mycotoxin Res. 2017;33(2):121–127. https://doi.org/10.1007/s12550-017-0271-4

21. Bordini JG, Ono MA, Garcia GT, Vizoni É, Amador IR, Hirozawa MT, et al. Transgenic versus conventional corn: Fate of fumonisins during industrial dry milling. Mycotoxin Res. 2019;35(2):169–176. https://doi.org/10.1007/s12550-019-00343-1

22. Bowers E, Hellmich R, Munkvold G. Comparison of fumonisin contamination using HPLC and ELISA methods in Bt and near-isogenic maize hybrids infested with European corn borer or Western bean cutworm. J Agric Food Chem. 2014;62(27):6463–6472.

23. Gasperini AM, Garcia-Cela E, Sulyok M, Medina A, Magan N. Fungal diversity and metabolomic profiles in GM and isogenic non-GM maize cultivars from Brazil. Mycotoxin Res. 2021;37(1):39–48. https://doi.org/10.1007/s12550-020-00414-8

24. Jacobson K, Myhr AI. GM crops and smallholders: Biosafety and local practice. J Environ Dev. 2013;22(1):104–124. https://doi.org/10.1177/1070496512466856

25. Ostry V, Ovesna J, Skarkova J, Pouchova V, Ruprich J. A review on comparative data concerning *Fusarium* mycotoxins in *Bt* maize and non-*Bt* isogenic maize. Mycotoxin Res. 2010;26(3):141–145. https://doi.org/10.1007/s12550-010-0056-5

26. Weaver MA, Abbas HK, Brewer MJ, Pruter LS, Little NS. Integration of biological control and transgenic insect protection for mitigation of mycotoxins in corn. Crop Prot. 2017;98:108–115. https://doi.org/10.1016/j.cropro.2017.03.020

27. Cappelle KM, Munkvold GP, Wolt JD. Meta-effect of insect resistant maize on fumonisins B_1 in grain estimated by variance-weighted and replication-weighted analyses. World Mycotoxin J. 2019;12:141–151. https://doi.org/10.3920/WMJ2018.2387

28. Pellegrino E, Bedini S, Nuti M, Ercoli L. Impact of genetically engineered maize on agronomic, environmental and toxicological traits: A meta-analysis of 21 years of field data. Sci Rep. 2018;8(1):3113. https://doi.org/10.1038/s41598-018-21284-2

29. Fischer K. Why Africa's new green revolution is failing – Maize as a commodity and anti-commodity in South Africa. Geoforum. 2022;130:96–104. https://doi.org/10.1016/j.geoforum.2021.08.001

30. Van der Westhuizen L, Shephard GS, Gelderblom WC, Torres O, Riley RT. Fumonisin biomarkers in maize eaters and implications for human disease. World Mycotoxin J. 2013;6(3):223–232. https://doi.org/10.3920/WMJ2013.1589

AUTHORS:
Rembuluwani P. Tshubwana[1] (iD)
Chandima Gomes[2] (iD)
John van Coller[2] (iD)
Thierry J.C. Regnier[3] (iD)

AFFILIATIONS:
[1]Department of Electrical Engineering, Tshwane University of Technology, Pretoria, South Africa

[2]School of Electrical and Information Engineering, University of the Witwatersrand, Johannesburg, South Africa

[3]Department of Biotechnology and Food Technology, Tshwane University of Technology, Pretoria, South Africa

CORRESPONDENCE TO:
Rembuluwani Tshubwana

EMAIL:
tshubwanarp@tut.ac.za

Microbiological growth analysis on a 275 kV transmission line composite insulator in South Africa

Transmission line silicon rubber insulators are frequently subjected to harsh environmental conditions that can change their surface characteristics and result in a transient or permanent loss of hydrophobicity. Algae, fungi, mould, and lichen are examples of biological growth that can compromise the power system's ability to operate safely by lowering the insulator's flashover voltage. We evaluated and identified the major microorganisms that could be responsible for the flashovers of insulators in South Africa. Due to the difficulty of accessing insulators currently in use within the electricity network, only two insulators, from two provinces in South Africa, were used in the study. Although algae were not found on the insulators, two major filamentous fungi identified as *Curvularia* sp. and *Aspergillus* sp. were isolated. The absence of algae could be attributed to the weather pattern of the two locations where the insulators were placed during their use as part of the network. According to previous studies worldwide, the high occurrence of these fungi could be linked to the reduction of flashover voltage of the composite insulators. Although a larger survey including more insulators from all provinces in South Africa should be conducted, the current study demonstrates the need for a cleaning programme for insulators using cost-effective eco-friendly antimicrobial agents.

Significance:

- Degradation and loss of hydrophobicity on the overhead line composite due to biological growth were found to be an impediment to the safe operation of high-voltage insulators. Fungi were found to be the main factor causing the degradation of overhead transmission lines.

- With these data, a model can be developed to predict the rate of biological growth on a given insulator, to estimate the degradation of the surface conditions with time for the utility to know when to remove them from the line. Such a model could lead to a more reliable and efficient power system and reduce the engineer's workload.

Introduction

Transmission line silicon rubber insulators can be degraded when exposed to extreme environmental conditions. This degradation could be due to biological growth. Several studies on the biological growth of microorganisms on insulators, especially on porcelain and silicon rubber insulators, have been conducted in various countries.[1] In Sri Lanka and Tanzania, insulators exposed to tropical climates (with medium to high humidity) were colonised by algae and mould, under both polluted and clean conditions. In general, microorganisms were found on 33 kV porcelain and glass insulators.[2] Algae and lichen were retrieved from porcelain insulators in areas characterised by low pollution, high rainfall, and high humidity.[3] Colonisation of insulators by green algae has been reported from Asia, Australia, and even Europe, making it the most undesirable microorganism to be targeted when a cleaning programme is implemented.[4] Under wet conditions and biological growth, the flashover voltage decreases by 25–75%. Due to the inorganic components that comprise silicone rubber insulators, which fungi and other microbes cannot digest, biological growth is significantly lower than on porcelain insulators.[4] Biological growth such as moss and algae was discovered on insulators placed in muddy wetlands in New Orleans[5], and even years after silicone rubber insulators were installed, biological growth was found in nations like Germany and China[6]. According to the reports, a wet, contaminated surface had a three times higher leakage current than a dry or clean surface. Also highlighted was the connection between biological growth, which covered one third of the insulator, and the leakage current level of 50 μA at the specified 127 kV phase-to-ground voltage. It has been reported that countries like Sweden and India, with temperate and tropical climates respectively, regularly find biological growth, including algae, fungi, and bacteria, on the surfaces of insulators.[7] Limited reports are available regarding the occurrence of microorganisms on composite insulators in South Africa. Therefore, this preliminary study focused on determining and identifying the major microorganisms present on different parts of the insulator.

Properties of composite insulators

Composite insulators are made from cross-linked polydimethylsiloxane (PDMS). These materials have several properties, such as:

- Hydrophobic surfaces

- Low surface free energy

- High surface mobility

- Good dielectric properties

- Non-adhering surfaces

- Excellent resistance to weathering

- Insolubility in water

- Low reactivity and toxicity

- Low glass transition temperature

They also have the advantage of being lightweight, which makes them uncomplicated to install. Composite insulators have excellent surface hydrophobicity. Polymeric insulators have replaced ceramic units for a wide range of reasons, including lower weight, less cost, high mechanical strength, and resistance to contamination.

Hydrophobicity of silicone rubber insulators

The hydrophobic property plays a crucial role in ensuring the safe and stable operation of silicone rubber insulators. However, the hydrophobic properties of high-voltage insulators are diminished by the presence of biological organisms such as fungi, mould, and lichen. The hydrophobicity of the silicone rubber insulator was found to be significantly compromised due to the presence of algae contamination, particularly in a humid environment. Silicone rubber materials have been seen to facilitate the provision of nutrients necessary for the growth of fungi, potentially leading to modifications in the electrical characteristics of the insulator. These phenomena typically occur in regions characterised by tropical and subtropical climates, where the temperature spans from 0 °C to 40 °C.[8]

Composition of biofilms

Microbial communities that inhabit a substrate that is resistant to biological degradation have a propensity to develop a cohesive layer on the material's surface. A biofilm consists of bacteria that are enclosed inside a matrix of extracellular polymeric substances, mostly polysaccharides and proteins, and are highly hydrated. The film frequently contains a combination of bacterial, fungal, protozoan, and algal communities.[9] Nevertheless, the proportion of viable cells within a fully developed biofilm is frequently minimal. In most biofilms, water is the prevailing ingredient in terms of abundance, while the organic component of the film is frequently characterised by a dominance of extracellular polymeric substances. Furthermore, the mature biofilm may contain particle debris, such as clay, humic compounds, and corrosion products. When attempting to identify the root cause of a biofouling issue, there is a tendency to underestimate the significance of the microbial population within the biofilm, inadvertently disregarding its role in the fouling process.[10]

Adherence and growth of biofilms

Biofilms are produced by microorganisms that exhibit the expression of genes responsible for the synthesis of surface features, such as the production of extracellular polymeric substances, as well as the regulation of growth rate.

Physicochemical considerations encompass various parameters that influence biological processes. These elements include substrate composition and roughness, temperature, pH, water potential, and oxygen supply.

Stochastic processes encompass the initial colonisation and subsequent random fluctuations in both biotic and abiotic variables. Deterministic phenomena refer to the occurrence of specific interactions between species, such as competition, cooperation, and predation. The mechanical mechanisms involved in this study include the application of shear forces resulting from the interaction of water or airflow with the substrate. The configuration of the biofilm is contingent upon the interactions between the bacteria and their immediate surroundings. Nevertheless, most biofilms have a structure like that of a sponge, characterised by the presence of interconnected pores and channels. This intricate network allows for the unrestricted movement of a liquid phase inside the biofilm, facilitating the efficient transfer of oxygen and nutrients throughout various regions of the biofilm.[11]

Materials and methods

Composite insulators

Figure 1a–d illustrates the two composite insulators that were tested and their locations. The two 275 kV composite insulators were donated by Eskom and were removed from the line because they flashed over during normal operation. As both insulators had been in operation for more than 10 years, they had been exposed to elements that would have decreased the hydrophobicity of the shed substance.

Power frequency current leakage tests

The insulators were hung on the test apparatus in the same way in which they were attached to the line. Each insulator's two electrodes were each linked to a voltage source terminal with variable power and frequency. Insulators were sprayed with water at a 45° angle so that the maximum precipitation rate under wet conditions would be 5.08 mm per minute. According to IEC 600112 specifications, the resistance of the water used for spraying was determined to be 9.45 k per cm^3 at standard ambient pressure and temperature. Rain was simulated on the insulator for 15 min. The power frequency voltage was applied and raised gradually until it reached the desired level. The humidity was 40%, the pressure was 849 kpa, and the temperature was 13 °C. Figures 2 and 3 show the outcomes for both dry and rainy conditions.

### Environmental conditions of the locations

The Richard's Bay climate is tropical with summer rainfall and cold weather in winter. The annual rainfall is 1228 mm. The annual average temperature is 21.5 °C.

Lichtenberg has a semi-arid climate; the annual average temperature is 17.3 °C and approximately 609 mm of precipitation occurs every year.

The environmental conditions for insulator A and B are reported in Supplementary figures 1–8. All weather data sets were provided by the South African Weather Service.

Microbiological analysis

Swabs were used to detect the presence of microorganisms on three randomly selected discs ($n = 16$ per insulator). Both surfaces ($n = 2$ per top and $n = 2$ per bottom) of each chosen disc were sampled because exposure to rain and sunlight could affect the biodiversity of the microorganisms present. The samples were packed in a cooler box, and transported to Tshwane University of Technology (Pretoria, South Africa) where they were stored at 4 °C until further use. Each swap was aseptically placed into a test tube containing 10 mL of sterile Ringer's solution (Merck, Johannesburg, South Africa). Up to 10^6 sequential decimal dilutions were generated. With 100 μL of each dilution, the spread plate technique was used to conduct microbiological analysis. Following the incubation of inoculated Petri plates with Plates Count Agar (PCA, Biolab, Johannesburg, South Africa) for 48 h at 30 °C, the presence of total mesophilic aerobic bacterial colonies was noted. After five days of incubation at 25 °C, mould and yeast counts were performed using Dichloran Rose Bengal Chloramphenicol Agar (DRBC, Biolab, Johannesburg, South Africa). Based on the presence of the colonies, pure isolates were obtained from a single spore as described by the authors.[10] These pure isolates were placed in the culture collection and identified using the molecular technique. The partial gene sequences of translation elongation factor 1α (EF1α) were examined to genetically identify five major isolates from insulator A and eight isolates from insulator B. A ZR Fungal/Bacterial DNA MiniPrep™ extraction kit (Zymo Research) was used to extract genomic DNA from pure cultures of each fungal isolate using about 100 mg of mycelium. Gel electrophoresis was used to verify the isolation of the DNA. The primers EF1 + EF2 [5'-ATGGGTAAGGAGGACAAGAC-3' and 5'-GGAAGTACCAGTGATCATGTT-3', respectively] were purchased from Inqaba Biotec (Pretoria, South Africa) for use in polymerase chain reaction (PCR) amplification. The identity of the fungal isolates was confirmed by comparing the EF1α nucleotide sequences to fungal sequences in the genetic database.
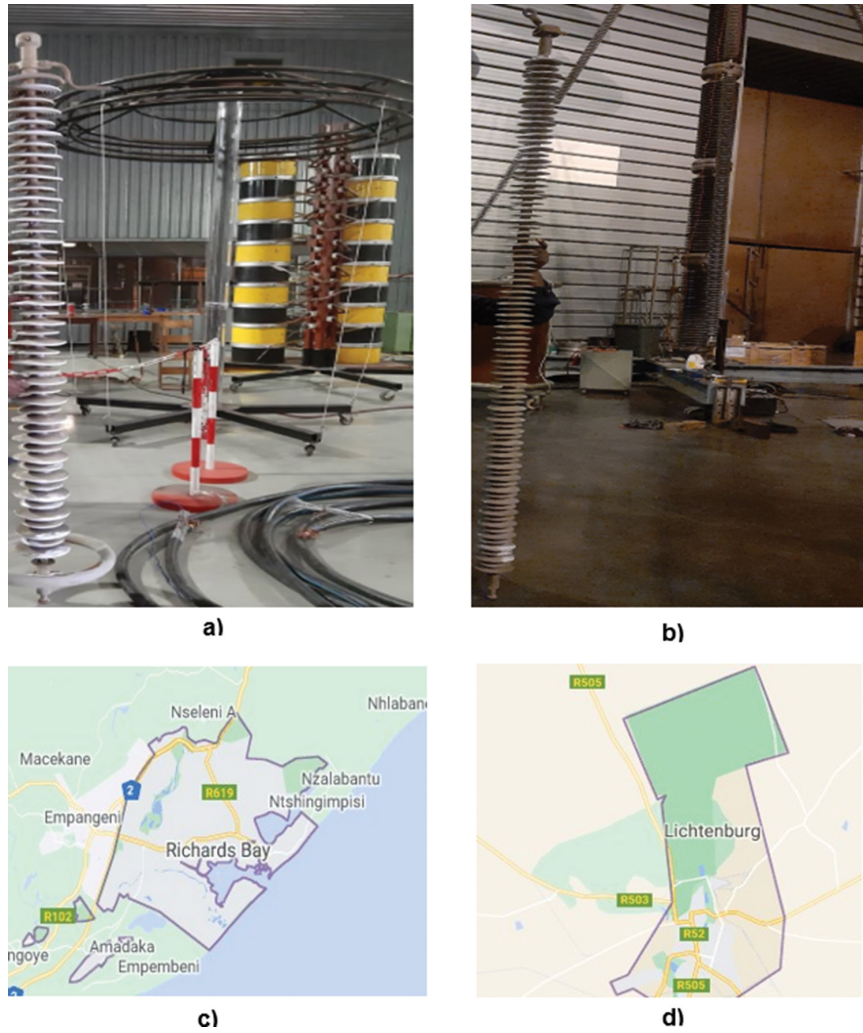
Source (c) and (d): Google Maps

Figure 1: (a) 275 kV insulator A from KwaZulu-Natal. (b) 275 kV insulator B from Lichtenburg. (c) The geographical location of insulator A. (d) The geographical location of insulator B.
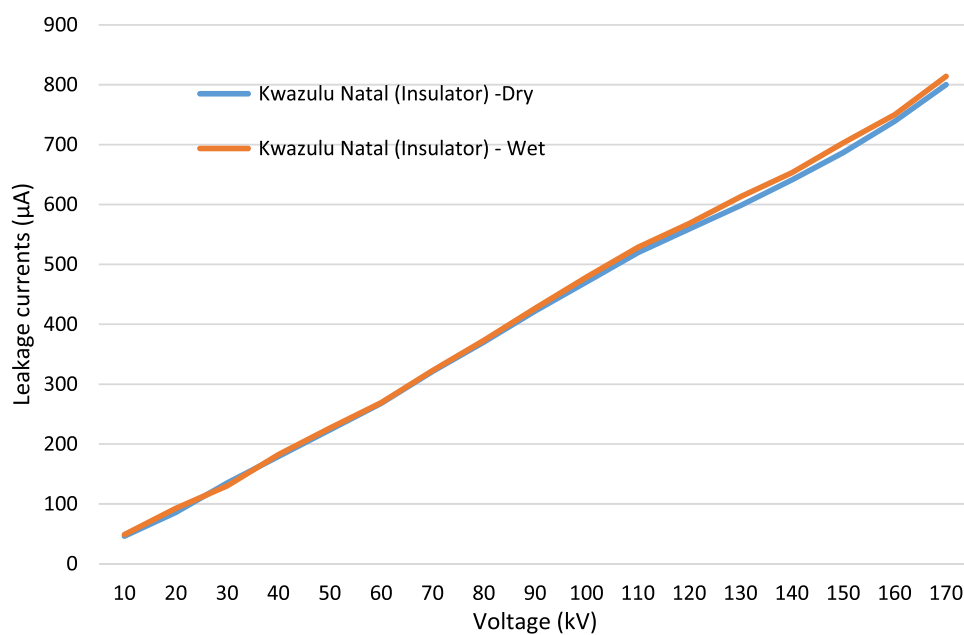

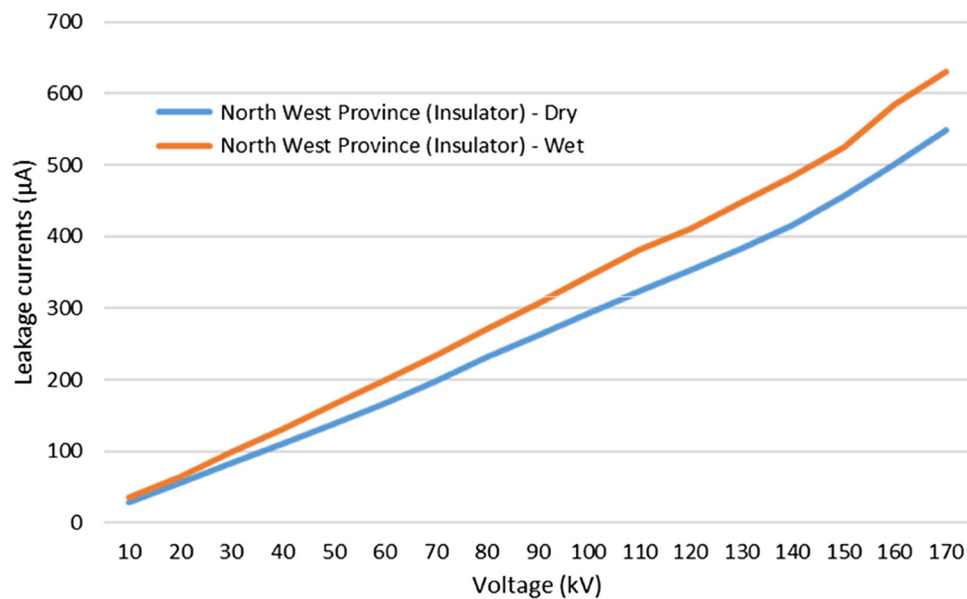**Figure 2:** Insulator A's leakage current in dry and wet environments.

Figure 3: Insulator B's leakage current in dry and wet environments.

Results and discussion

Microbial population

As anticipated, there was no pattern in the distribution of microorganisms on the insulators within the discs, regardless of where the swabs were taken. Disc 1 did not support the development of any microorganisms (Table 1), despite its location on the top portion of the insulator and exposure to sunlight. On the other hand, more microorganisms, primarily filamentous fungi, were present on the discs in between. Surprisingly, none of the discs examined could be used to isolate yeast or algae, not even on the insulator from Richard's Bay, a coastal region on the Indian Ocean, only filamentous fungi.

The identification of the filamentous fungi was emphasised considering the discs' mould contamination. Numerous species have been documented to grow on composite insulators and the variety and quantity of the microorganisms, primarily filamentous fungi, differed amongst insulators depending on ambient factors, materials used, and shed design.[11] Although there have been earlier reports of filamentous fungi on insulators in Florida and other parts of the world, to our knowledge, no studies have been conducted in South Africa to identify the fungal populations on the insulators currently used in the country. *Curvularia* sp. and *Aspergillus* sp. were prevalent filamentous fungi that developed on the insulators (Table 2).

In a study on the potential use of laser-induced fluorescence for the detection of fungus covering insulator materials, *Curvularia* sp. and *Aspergillus* sp. were previously isolated from insulators in different countries.[11] Despite reports of green algae microorganisms on 500 kV transmission line insulators, none of the two insulators – particularly insulator A, which was based near the coast and subject to subtropical weather and high humidity – had any isolated algae. However, a larger number of insulators must be selected to verify the occurrence of filamentous fungi and the absence of algae and yeast contamination.[12] For this reason, further studies should be done using a larger number of insulators, including those from the coastal areas of the Eastern and Western Cape Provinces of South Africa. The presence of filamentous fungi, which might cause the sheds' hydrophobicity to decrease, may be the cause of the leaking currents seen on insulators A and B. This could explain why these insulators' flashing is over-recorded. Based on previous studies, natural essential oils were tested on ceramic tiles, for their ability to suppress several filamentous fungi such as *Curvularia* sp. and *Aspergillus* sp.[13] Several oils, including those from

Mentha piperita L. and thyme oil, were recovered using exchanges.[14] Strong antibacterial action was shown by the essential oil thymol and by R- (-)-carvone to suppress this fungus. However, biocidal antimicrobial molecular barrier treatments are also recognised for their ability to reduce the growth of bacteria on surfaces in industrial settings, necessitating more research and testing of such solutions' potential for use on insulators.[15] Additionally, wiping after washing the insulator in water, wiping the dry insulator with cotton swabs dipped in salt water or tap water, or cleaning de-energised insulators can be done in a variety of ways, including through vigorous hand wiping of the dry insulators. However, further research is required to determine how to remove mould in electrified environments.

Conclusions

This article underlines the importance of investigating the presence of microorganisms on insulators. The occurrence of these microorganisms is scientifically poorly known in South Africa. The findings of this paper show that filamentous fungus colonised insulators used in South Africa over time, which led to flashover. In that sense, this preliminary research demonstrates the need for further investigation in South Africa. Although perhaps optimistic, investigation of the potential prophylactic use of essential oils and therapeutic solutions is encouraging as a cost-effective alternative for removing contaminants on the insulators and must be investigated further using more microorganisms (including fungi) isolated from a larger number of insulators.

Competing interests

We have no competing interests to declare.

Authors' contributions

R.P.T.: Conceptualisation; methodology; investigation and formal analysis; writing – original draft preparation; writing – review and editing. C.G.: Conceptualisation; investigation and formal analysis; writing – review and editing. J.v.C.: Writing – review and editing. T.J.C.R.: Methodology; writing – review and editing.

Table 1: Presence of microorganisms on two different insulators

Sample description	Total count (CFU/mL)	Yeast and mould (CFU/mL)
Disc 1 Top - Insulator A	TFTC	TFTC
Disc 1 Top - Insulator B	Nothing	Nothing
Disc 1 Bottom - Insulator A	1190.00 ± 91.33	1113.33 ± 85.05
Disc 1 Bottom - Insulator B	Nothing	Nothing
Disc 2 Top - Insulator A	Nothing	Nothing
Disc 2 Top - Insulator B	Nothing	Nothing
Disc 2 Bottom - Insulator A	344.33 ± 35.12	138.00 ± 19.47
Disc 2 Bottom - Insulator B	Nothing	Nothing
Disc 3 Top - Insulator A	18 500.00 ± 458.26	16 266.67 ± 1305.12
Disc 3 Top - Insulator B	171 333.30 ± 19 731.53	Nothing
Disc 5 Bottom - Insulator A	TFTC	Nothing
Disc 5 Bottom - Insulator B	22 800.00 ± 793.73	22 366.67 ± 929.16
Disc 23 Top - Insulator A	Nothing	Nothing
Disc 23 Top - Insulator B	Nothing	Nothing
Disc 24 Top - Insulator A	Nothing	Nothing
Disc 24 Top - Insulator B	TFTC	Nothing
Disc 25 Top - Insulator A	Nothing	Nothing
Disc 25 Top - Insulator B	Nothing	Nothing
Disc 45 Top - Insulator A	Nothing	Nothing
Disc 45 Top - Insulator B	Nothing	Nothing
Disc 45 Bottom - Insulator A	2363.33 ± 75.72	1846.67 ± 335.01
Disc 45 Bottom - Insulator B	46.33 ± 4.16	TFTC
Disc 46 Top - Insulator A	Nothing	Nothing
Disc 46 Top - Insulator B	Nothing	Nothing
Disc 46 Bottom - Insulator A	136 333.30 ± 1154.70	22 300 ± 700
Disc 46 Bottom - Insulator B	15 066.67 ± 776.75	2446.67 ± 15.28
Disc 47 Top - Insulator A	Nothing	TFTC
Disc 47 Top - Insulator B	Nothing	TFTC
Disc 47 Bottom - Insulator A	Nothing	Nothing
Disc 47 Bottom - Insulator B	546.67 ± 80.83	540.00 ± 52.92

Values are the mean±SD of 30 plates

TFTC, too few to count; CFU, colony forming units

Table 2: Significant filamentous fungus identified from the two insulators

Insulator A	Insulator B
Curvularia trifolii	*Thielavia terricola*
Curvularia protuberata	*Sordaria tameness*
Lecythophora sp.	*Curvularia lunata*
Aspergillus fumigatus	*Trichoderma citrinoviride*
	Lecythophora sp.
	Aspergillus fumigatus
	Aspergillus niger

References

1. McAfee RD, Heaton RD, King JM, Falster AU. A study of biological contaminants on HV porcelain insulators. Electr Power Syst Res. 1997;42:35–39. https://doi.org/10.1016/S0378-7796(96)01174-1

2. Gubanski SM, Fernando MARM, Pietr SJ, Matula J, Kyaruzi A. Effects of biological contamination on insulator performance. In: 6th International Conference on Properties and Applications of Dielectric Materials; 2000 June 21–26; Xi'an, China. IEEE; 2000. p. 797–801. https://api.semanticscholar.org/CorpusID:129600229

3. Flemming HC. Relevance of biofilms for the biodeterioration of surfaces of polymeric materials. Polym Degrad Stab. 1998;59:309–315. https://doi.org/10.1016/S0141-3910(97)00189-4

4. Wallstrom S. Biofilms on silicone rubber materials for outdoor high voltage insulation [PhD thesis]. Stockholm: Royal Institute of Technology; 2005. https://www.diva-portal.org/smash/get/diva2:7667/FULLTEXT01.pdf

5. Wallstrom S, Karlsson S. Biofilms on silicone rubber insulators; microbial composition and diagnostics of removal by use of ESEM/EDS: Composition of biofilms infecting silicone rubber insulators. Polym Stab. 2004;85(2):841–846. https://doi.org/10.1016/j.polymdegradstab.2004.02.014

6. Kumagai S. Influence of algal fouling on hydrophobicity and leakage current on silicone rubber. IEEE Trans Dielectr Electr Insul. 2007;14(5):1201–1206. https://doi.org/10.1109/TDEI.2007.4339480

7. Gorur RS, Montesinos J, Roberson R, Burnham J, Hill R. Mold growth on nonceramic insulators and its impact on electrical performance. IEEE Trans Power Deliv. 2003;18:559–563. https://doi.org/10.1109/TPWRD.2003.809681

8. Gorur RS, Bernstein BS. Field, and laboratory aging of polymeric distribution cable terminations. I. Field aging. IEEE Trans Power Deliv. 1998;13:316–322. https://doi.org/10.1109/61.660896

9. Dinesh MN, Vasudev N, Nambudri PVV, Suryanarayana K. The performance of composite insulators with and without biocontamination. In: International Conference on High Voltage Engineering and Application; 2008 November 09–12; Chongqing, China. IEEE; 2008. p. 124–127. https://doi.org/10.1109/ICHVE.2008.4773889

10. Noman E, Al-Gheethi AA, Rahman NK, Talip B, Mohamed R, Kadir OA. Single spore isolation as a simple and efficient technique to obtain a fungal pure culture. IOP Conf Ser: Earth Environ Sci. 2018;140, 012055. https://doi.org/10.1088/1755-1315/140/1/012055

11. Ouyang X, Yin F, Jia Z, Yang S, Wang Y, Bai H, et al. Research of biological contamination and its effect on the properties of RTV-coated insulators. Electr Power Syst Res. 2019;167:138–149. https://doi.org/10.1016/j.epsr.2018.10.025

12. Xie C, Li C, Zhu W, Gan Y. Analysis of green algae contamination on 500 kV silicon rubber composite insulators. In: First International Conference on Electrical Material and Power Equipment; 2017 May 14–17; Xi'an, China. IEEE; 2017. p. 664–666. https://doi.org/10.1109/ICEMPE.2017.7982184

13. Inouye S, Tsuruoka T, Watanabe M, Takeo K, Akao M, Nishiyama Y, et al. Inhibitory effect of essential oils on apical growth of *Aspergillus fumigatus* by vapour contact. Mycoses. 2000;43:17–23. https://doi.org/10.1046/j.1439-0507.2000.00538.x

14. Combrinck S, Regnier T, Kamatou GPP. In vitro activity of eighteen essential oils and some major components against common postharvest fungal pathogens of fruit. Ind Crop Prod. 2011;33:344–349. https://doi.org/10.1016/j.indcrop.2010.11.011

15. Özcan S, ÇalışAçıkbaş N, Açıkbaş G. Formation of antibacterial effect on ceramic tile surfaces. Anadolu Univ J Sci Technol A: Appl Sci Eng. 2017;18(1):122–130. https://doi.org/10.18038/aubtda.300424

AUTHORS:
Amore E. Nel[1]
Achim Morschhauser[2]
Foteini Vervelidou[3]
Jürgen Matzka[2]

AFFILIATIONS:
[1]South African National Space Agency, Hermanus, South Africa
[2]German Research Centre for Geosciences, Potsdam, Germany
[3]Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

CORRESPONDENCE TO:
Amoré Nel

EMAIL:
anel@sansa.org.za

A new high-resolution geomagnetic field model for southern Africa

Earth's magnetic field is a dynamic, changing phenomenon. The geomagnetic field consists of contributions from several sources, of which the main field originating in Earth's core makes up the bulk. On regional and local scales at Earth's surface, the lithospheric field can make a substantial contribution to the overall field and therefore needs to be considered in field models. A locally derived regional core field model, named HMOREG, has been shown to give accurate predictions of the southern African region. In this study, a new regional field model called the South African Regional Core and Crust model (SARCC) is introduced. This is the first time that a local lithospheric model, estimated by employing the revised spherical cap harmonic analysis modelling method, has been combined with the core component of CHAOS-6, a global field model. It is compared here with the existing regional field model as well as with global core field models. The SARCC model shows small-scale variations that are not present in the other three models. Including a lithospheric magnetic field component likely contributed to the better performance of the SARCC model when compared to other global and local field models. The SARCC model showed a 33% reduction in error compared to surface observations obtained from field surveys and INTERMAGNET stations in the Y component, and HMOREG showed a 7% reduction in error compared to the global field models. The new model can easily be updated with global geomagnetic models that incorporate the most recent, state-of-the-art core and magnetospheric field models.

Significance:

Earth's magnetic field is an integral part of many current navigational methods in use. Updates of geomagnetic field models are required to ensure the accuracy of maps, navigation, and positioning information. The SARCC regional geomagnetic field model introduced here was compared with global geomagnetic field models, and the inclusion of a lithospheric magnetic field component likely contributed to the better performance of the SARCC model. This regional model of southern Africa could easily be updated on a regular basis, and used for high-resolution information on the Earth's magnetic field for the wider scientific community.

Introduction

Earth is surrounded by its geomagnetic field, which protects us from the harmful effects of space weather.[1] The geomagnetic field also plays a role in navigation and mapping applications. The geomagnetic field originates primarily from the core due to the dynamo process occurring therein. Another geomagnetic field source is the magnetised part of Earth's crust.[2,3] This magnetic field component is known as the lithospheric magnetic field. The core field shows temporal variability on timescales of 1 year and longer[4], while the lithospheric field is, to a good approximation, time-invariant. Another difference between these two components of the geomagnetic field is spatial variation.[2] The core field dominates over large wavelengths (several thousands of kilometres) while the lithospheric field dominates over short (less than 500 km) wavelengths of the observed geomagnetic field. Geomagnetic field models can be expanded in spherical harmonic functions, solving the Laplace equation in spherical coordinates.[4] There is also an external field, formed by the interaction of the solar wind with the Earth's magnetosphere, that makes up the external contribution to the geomagnetic field from near-Earth current systems. This field has time scales in the order of seconds to days, much less than secular variation.[3]

The southern African region is known for its temporally and spatially highly variable geomagnetic field.[4,5] Due to the small-scale variations, the current available global field models are unable to describe this region accurately. A model called Southern African Model made of Splines (SAMS), using harmonic splines, was derived for this region, which consisted only of a core component.[3] In this study, core and lithospheric components were combined to create a new regional model. The lithospheric field component was obtained by employing a spherical harmonic degree of 16 and larger in the model. This component was derived by using ground measurements from repeat stations, which enabled the addition of small-scale spatial variations in the model predictions. For the core component, we used spherical harmonic degrees from 1 to 15. Combining both contributions into one model can improve the accuracy of the magnetic field model predictions. Two global field models, the CHAOS-7 and the IGRF-13 model, which will be discussed in more detail in the sections to follow, were compared with the two southern African regional field models, HMOREG and SARCC.

Global field models

IGRF-13

The International Geomagnetic Reference Field (IGRF) is a spherical harmonic model used to describe the large-scale main (core) field globally and is updated every 5 years using observatory and satellite mission data sets.[6] The 13th generation used in this study covers the period from 1900 to 2025.[7] The current version was expanded to a maximum spherical harmonic degree of 13. Neither this model nor the CHAOS model incorporates field survey data from southern Africa.

CHAOS-6

CHAOS is a series of core field models derived by the Technical University of Denmark.[8] Along with the basic Ørsted, CHAMP and SAC-C satellite mission data sets used in the earlier CHAOS models, CHAOS-6 also uses monthly means derived from the hourly mean values of 160 ground observatory magnetic measurements and over 2 years of Swarm data available up until March 2016.[9] CHAOS-6 includes along-track differences between two Swarm satellites, Alpha and Charlie. It covers the epoch from 1997 to 2018.[8] Only the internal magnetic field contributions, the core and crust components, were used in this study.

CHAOS-7

CHAOS-7 is the latest of the CHAOS series and was released at the end of 2019; it includes both core and crustal components. The model is based on Ørsted, CHAMP, SACC, Cryosat2, and Swarm satellite data, as well as ground geomagnetic observatory data.

Regional field models

HMOREG

The Hermanus Magnetic Observatory (34°25.5' S, 19°13.5' E) was established in 1941 and today falls under the South African National Space Agency (SANSA).[10] Between 1960 and 2005 there was an undertaking to establish other magnetic observatories in southern Africa. A site was identified outside Tsumeb (19°°12' S, 1735' E) in Namibia in 1964 on the premises of a permanent ionospheric observation station of the *Max Planck Institut für Aeronomie*. Another was established at Hartebeesthoek (25°°52.9' S, 27°42.4' E) in 1972. These, along with the observatory at Hermanus, are members of the INTERMAGNET network.[11]

Since 2005, two sets of 20 repeat stations have been measured bi-annually in survey campaigns across southern Africa, which include South Africa, Namibia, Zimbabwe, and Botswana.[10] Along with the continuous recording output of the three aforementioned INTERMAGNET stations, these data were used to derive a polynomial-based core field regional model. This mathematical model was derived using third-order polynomials with 10 statistically significant coefficients:

$$p(X, Y) = A + BX^3 + CYX^2 + DXY^2 + EY^3$$
$$+ FX^2 + GXY + HY^2 + IX + JY,$$

Equation 1

where p refers to the main field components (F, H, D, Z), X = 26° latitude, and Y = 24° longitude. The coefficients A–J were estimated from the field survey data using a least-squares fit for each separate component. This model is referred to as HMOREG. In this study, only input grid points that fell within the boundaries defined by the location of the outermost repeat stations were considered. This model is updated annually using the latest field surveys. Field survey locations for 2016 are shown in Figure 1. Since 2016, the number of repeat stations has been reduced to 19.

SARCC

The latest model for southern Africa that we describe in this paper is a novel coupling of the CHAOS-6 core model (for spherical harmonic degrees 1 to 15) and the lithospheric field model of Vervelidou et al.[12] The combined model is referred to as the South African Regional Core and Crust (SARCC) model henceforth. It accepts as input an epoch, and the altitude and geodetic coordinates of whichever grid points are desired. As output, it calculates the north, east and downward components of the magnetic field as well as the magnetic field declination. The model only accepts grid points within the lithospheric model of Vervelidou et al.[12] The SARCC model can be updated by adding future global field model versions instead of the CHAOS-6 model.

The SARCC is a high-resolution model based on satellite, near-surface, and ground magnetic field measurements. It is parameterised by the revised spherical cap harmonic analysis (R-SCHA) modelling method. The latter is a regional analysis technique used to obtain magnetic field measurements at different altitudes.[13] The lithospheric magnetic field is modelled inside the
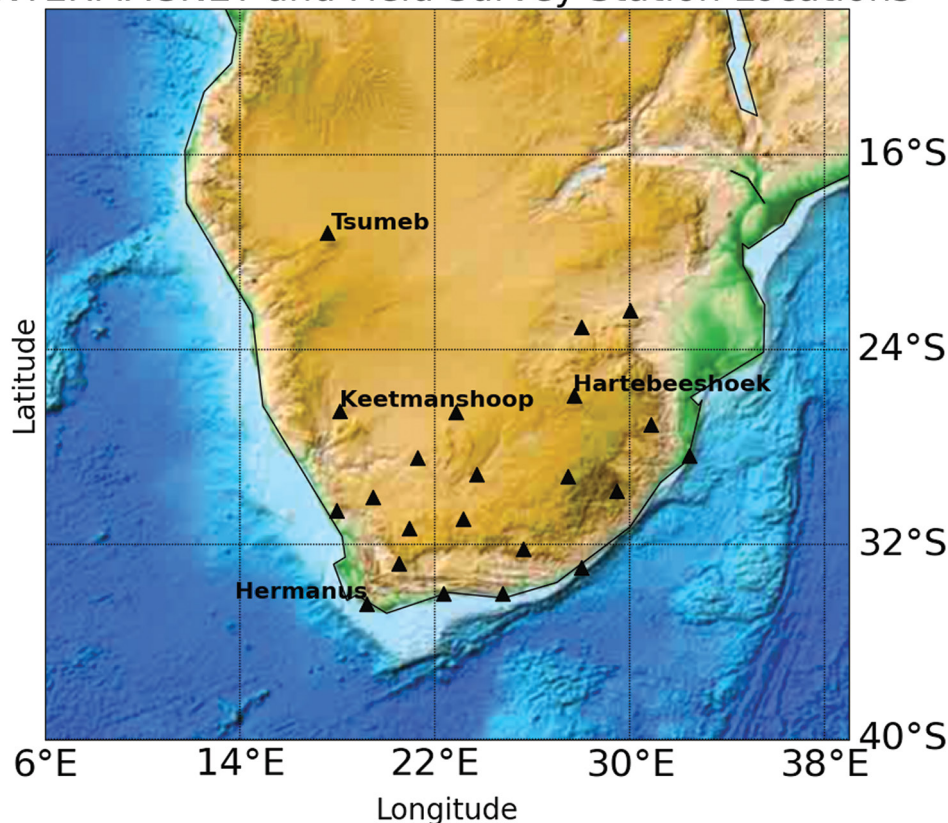

Figure 1: Location of the 19 repeat stations during 2016 field surveys and the four INTERMAGNET stations Hermanus, Tsumeb, Keetmanshoop, and Hartebeesthoek.

volume of a spherical cone over southern Africa with a half-angle aperture of $\theta_0 = 15$ centred at a geocentric longitude of 22.5° E and latitude of 25° S, excluding a thin ring of 0.1° from the cap's volume using the R-SCHA modelling method. The spherical harmonic degree starts at 16 (which corresponds to wavelengths of about 2500 km) and goes up to degree and order 1400. It can predict the geomagnetic field at altitudes ranging from 0 to 500 km. The R-SCHA result relies on CHAMP and Swarm satellite magnetic field measurements and on a version of the World Digital Magnetic Anomaly Map (WDMAM) that relies on aeromagnetic and marine measurements taken over southern Africa since 1970, and it has a spatial resolution of 0.1°. The model also includes data from repeat station measurements conducted annually throughout 2005–2009. Finally, the annual means of three geomagnetic observatories – Hermanus and Hartebeesthoek in South Africa and Tsumeb in Namibia (Figure 1) – were also part of the input data set.

Methods

Comparison methodology

The global IGRF-13 and CHAOS-7 models, and the two southern African regional field models SARCC and HMOREG, were compared with each other. It is important to note that IGRF-13 and HMOREG do not take into account the lithospheric field, whereas SARCC and CHAOS-7 do.

The model predictions were also compared to the results of the repeat station surveys. As mentioned in the section describing the HMOREG regional field model, surveys have been conducted at several repeat stations across southern Africa since 2000.

These surveys take place between middle September to middle December and are split between two fieldwork legs: one leg covers the stations on the eastern side of South Africa and the other covers the western side. Each repeat station is marked by concrete beacons, which ensures that the location of observation points stays constant for successive surveys. Geomagnetic field observations are done in the early evening and early morning, with the variometer operating continuously during the night.[14] The data are augmented by standard observatory annual means and centred in the middle of each year. The reduction methodology is the same for both the eastern and western leg data sets. The data are not corrected for external or lithospheric signals.[15]

Statistical method

In this study, the field surveys from 2015.5, 2016.5, 2017.5, and 2018.5 epochs were combined with recordings from four INTERMAGNET observatories: Hermanus (HER), Hartebeesthoek (HBK), Keetmanshoop (KMH), and Tsumeb (TSU). These values were then compared with the values predicted by the four models introduced earlier.

For each of the four epochs considered in this study, the magnetic field components X, Y, Z (which represent the three orthogonal component field directions for local geodetic northward, eastward, and vertically down, respectively) and the declination D were determined at the survey location points using the four models – SARCC, HMOREG, CHAOS-7, and IGRF-13. The total number of locations was 40 for 2015.5, 22 for 2016.5, 24 for 2017.5, and 22 for 2018.5. Only the 19 points common to all field surveys were used for the comparison. From this, the mean difference was determined for each component of each model by comparison with the field survey data over all the years. To get an idea of each model's error distribution characteristics, the standard deviation and skewness were estimated from the average value over all available years. Lastly, the secular variation over all available data sets
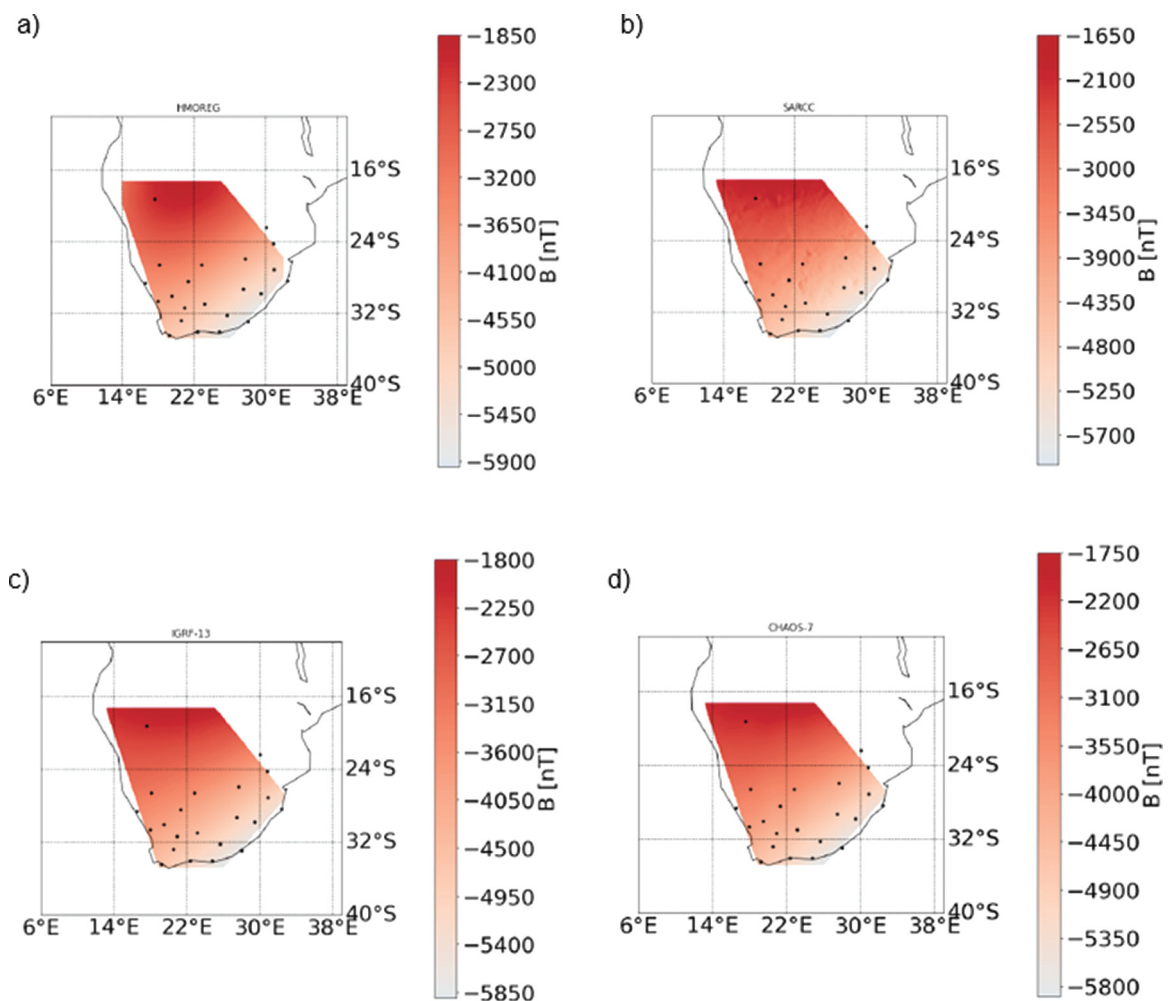

Figure 2: The predicted Y component of the respective models, epoch = 2018.5. The top panel shows the regional field models: (a) HMOREG and (b) SARCC. The bottom panels show the global field models: (c) IGRF-13 and (d) CHAOS-7.

was determined using each model and compared with the field survey results. Secular variation at time t was taken as the differences of the X, Y, and Z values at time $t + 6$ months and $t - 6$ months.

Results and discussion

For the sake of comparison, a 0.2° resolution grid for the predicted Y component over southern Africa is shown for each model in Figure 2 within the borders of each respective model. The SARCC predicted grid in the top right figure shows small-scale variations and gives more information about local anomalies on the surface. The lithospheric model predictions are also shown separately for the X, Y, and Z components at Earth's mean reference radius using the SARCC model for 2018.5 in Figure 3. The SARCC model lithospheric field signal includes many fine-scaled magnetic field features at Earth's mean reference radius that are not present in the other models.

Comparisons were made between differences at every field survey location and INTERMAGNET observatories. A visual representation of these differences for the Y component (in nT) is shown in Figure 4 for all available years. The standard error of the mean (SEM) is calculated by Equation 2:

$$(SEM = \frac{\sigma}{\sqrt{n}}) \hspace{3cm} \text{Equation 2}$$

where σ is the standard deviation and n is the sample size. SEM was calculated as 6.8 nT, 9.4 nT, 10.1 nT, and 10.1 nT for SARCC, HMOREG, IGRF-13, and CHAOS-7, respectively. The SEM gives an indication of how much discrepancy is likely in the sample mean compared with the population mean. In this case, the population mean was the mean derived from the respective model outputs, and the sample mean was calculated from the field survey data. SARCC shows the least amount of difference between field survey measurements and prediction, with a ~32.7% reduction in error compared to the global field models. HMOREG also shows smaller differences compared to the global field models IGRF-13 and CHAOS-7, with a ~6.9% reduction in error. Similar prediction performance is seen for the X component. For the X component, the SEM was calculated as 7.0 nT, 8.8 nT, 10.8 nT, and 12.2 nT for SARCC, HMOREG, IGRF-13, and CHAOS-7, respectively, which is a 39.1% and 23.5% reduction in error for SARCC and HMOREG, respectively, when compared to global field models. For the Z component, the SEM was calculated as 12.4 nT, 21.4 nT, 22.8 nT, and 18.8 nT for SARCC, HMOREG, IGRF-13, and CHAOS-7, which gave a 40.4% reduction in error and a 2.9% increase in error for SARCC and HMOREG, respectively, when compared to global field models. Of all the models, the SARCC shows the least amount of difference between prediction and field survey measurements. The histograms of the differences between the 19 repeat station measurements and the model predictions for all available years are shown in Figure 5 for the Y component, with the mean and variance indicated in the legends of each plot. The standard deviation (σ) of the SARCC is lower than those of the other three models for all components (Table 1).

Results of the statistical comparisons of the differences between the measured and modelled components for 2015.5–2018.5 are also shown in Table 1. The standard deviation measures the variability of the data
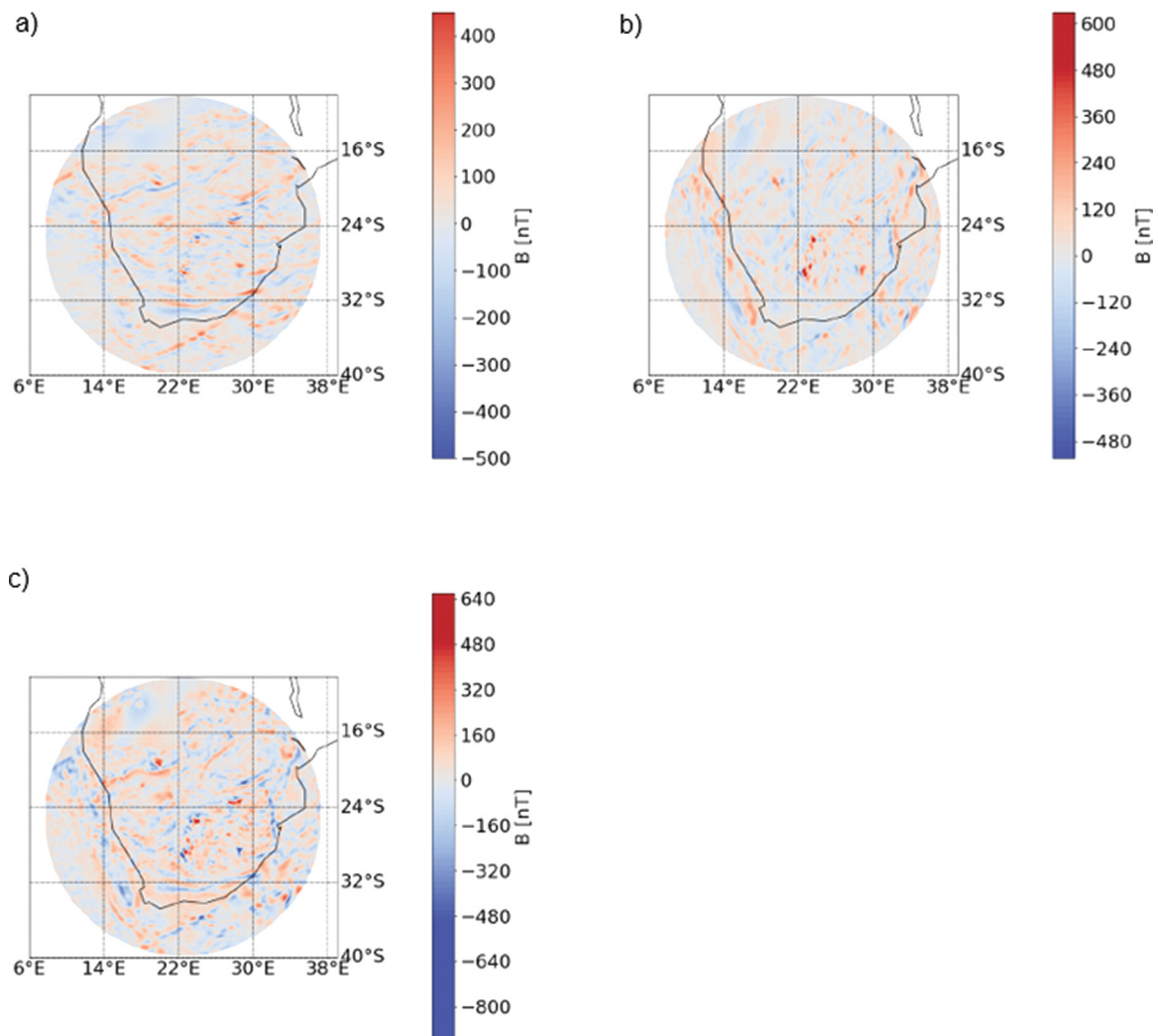

Figure 3: Lithospheric field model components at Earth's mean reference radius as predicted by the SARCC model adapted from the R-SCHA. The X and Y components are shown in (a) and (b), respectively, and (c) depicts the Z component.
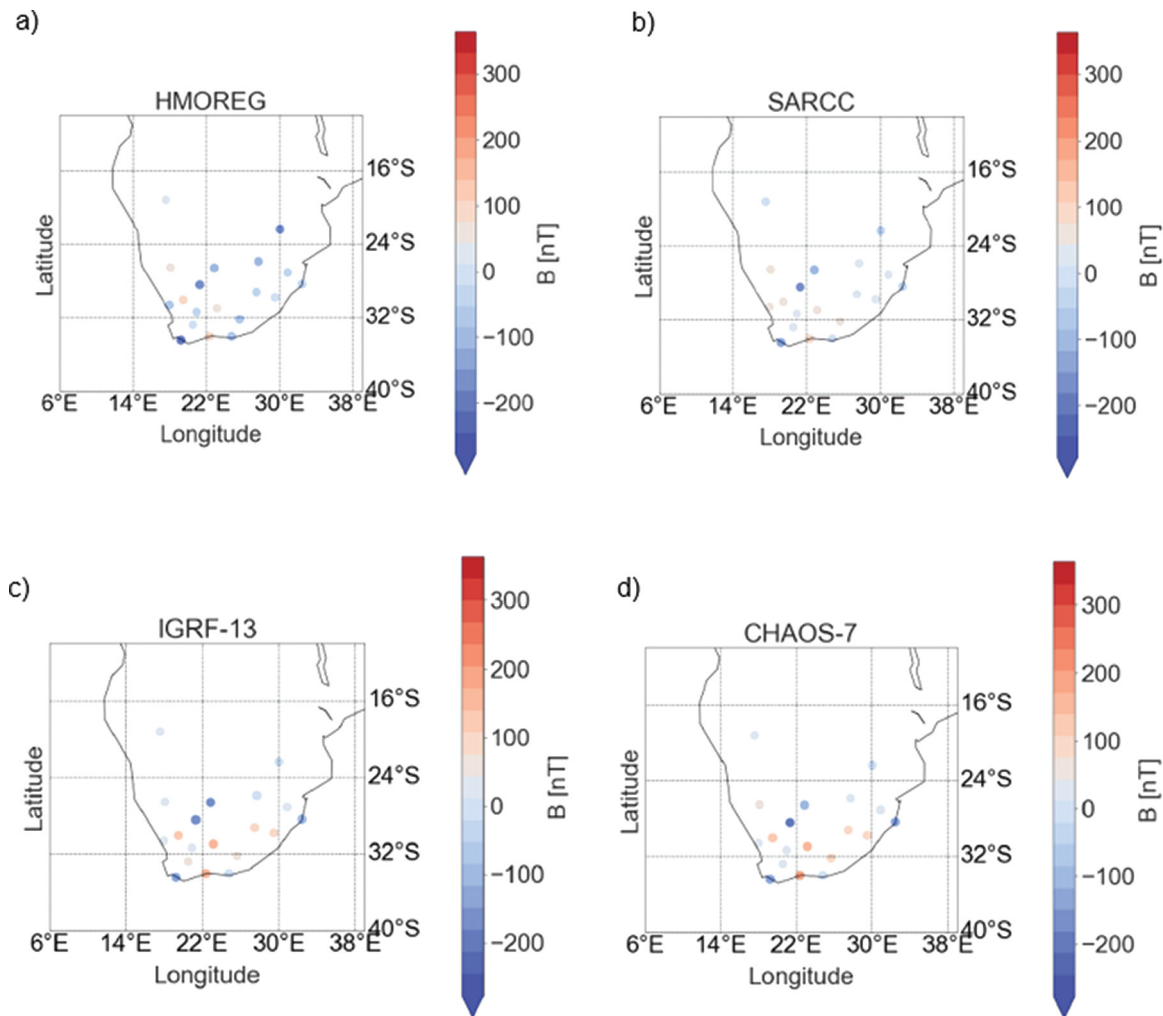
Figure 4: The differences between the Y component of the geomagnetic field predicted by the models considered in this study and the respective field survey values, shown at each of the 19 survey locations. The top panel shows the regional field models: (a) HMOREG and (b) SARCC. The bottom panel shows the global field models: (c) IGRF-13 and (d) CHAOS-7.

set to the mean. The standard deviation does give an indication of the predictive abilities of each model, but given the size of the area from which these statistical measurements are derived, and considering local, temporal, magnetic anomalies, using the standard deviation for a baseline correction could cause an erroneous result. The SARCC model standard deviations are, on average, 39%, 33%, 43%, and 32% lower than the global field models for X, Y, Z, and declination, respectively. Thus, SARCC shows the least deviation for all examples in all components. HMOREG's performance varies, as can be seen in Figure 5 and Table 1. HMOREG's standard deviation is on average 24%, 7%, and 9% lower than the global field models for X, Y, and declination. HMOREG's standard deviation for the Z component is 4% more than that for the global field models. The standard deviation for HMOREG is lower for the X and Y components, as well as declination, compared to those of the global field models.

The skewness of a data set is defined as the distortion or deviation from a normal distribution. The distribution can be negatively or positively skewed, but here we refer only to the absolute skewness. If skewness is less than 0.5, the distribution is approximately symmetrical; if the skewness is between 0.5 and 1, it is moderately skewed; and if the skewness is greater than 1, the distribution is described as highly skewed. The SARCC's X component is moderately skewed, the Y component highly skewed, and the Z component approximately symmetrical. HMOREG's X and Y components are approximately symmetrical, and the Z component is moderately skewed.

The calculated secular variation for the total field for 2017/2018 is shown in Figure 6. This plot shows the difference between the secular variation measured during field surveys (referred to the middle of the respective year as described in the 'Statistical method' section), and the values determined by the respective models. All models show similar outputs, with the double peak visible in the total field histogram showing a deviation between the model predictions and field survey measurements. The field survey data were measured and collected in two phases: an east leg and a west leg. The total field values of the repeat stations for all available years are shown in Supplementary table 1. From these values, the secular variation is derived and shown in columns SV1516, SV1617, and SV1718. The locations that fall under the west leg are highlighted in yellow, and the locations for the east leg are in blue. From the average secular variation values (in rows labelled 'AVG W' and 'AVG E') in Supplementary table 2, it is clear that there is a significant increase in secular variation strength moving from east to west. Models in this study have been shown to overestimate the west leg data set and underestimate the east leg data set, albeit to a lesser extent than the ground observatory data. The data reduction methodologies used for the repeat stations are the same for both data sets, which leads to the conclusion that this double peak is likely because of the presence of the South Atlantic Anomaly and the most recent geomagnetic jerk, and the cause of this spatial gradient observed over southern Africa. The South Atlantic Anomaly is an area over the Atlantic Ocean where there is a significant depletion of magnetic field strength, and partially overlaps southern Africa. Supplementary table 2 shows the calculated CHAOS values for secular variation for the same times and locations as the repeat stations; these values were used to cross-correlate the results shown in
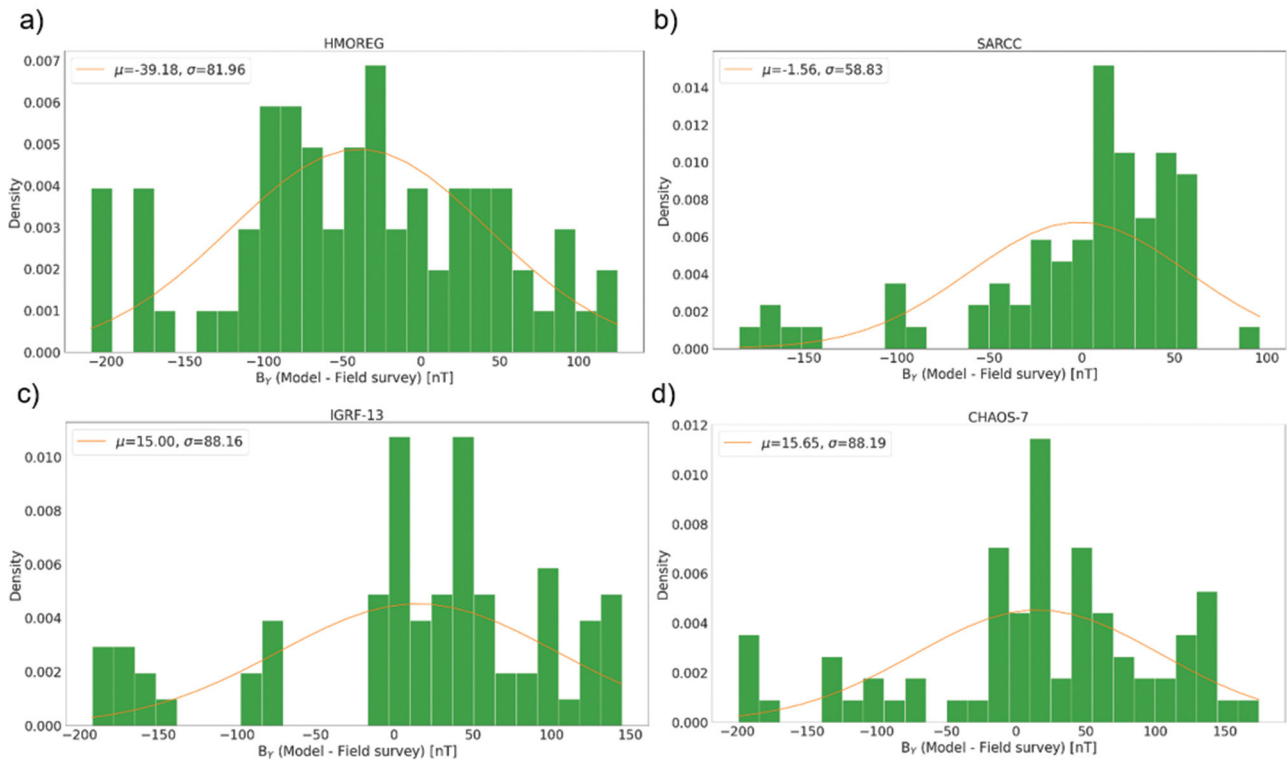
Figure 5: Histogram comparisons of the differences between field survey recordings and model predictions over all available years (2015.5–2018.5) for the Y component. The density curves show the normalised probability for each bin. The top panel shows the regional field models: (a) HMOREG and (b) SARCC. The bottom panel shows the global field models: (c) IGRF-13 and (d) CHAOS-7.

Table 1: Statistics for the average differences between the X, Y and Z components and declination (D) of the field survey measurements and the SARCC, HMOREG and IGRF-13 models over southern Africa, for all years. The X, Y and Z components are given in nanotesla, and D in arcminutes.

	SARCC	HMOREG	IGRF-12	CHAOS
X:				
Mean	-5.84	15.25	-2.66	-17.36
Standard deviation	61.28	76.42	93.87	106.01
Skewness	-0.83	0.08	-0.20	-0.12
Y:				
Mean	-1.56	-39.18	15.00	15.65
Standard deviation	58.83	81.96	88.16	88.19
Skewness	-1.50	-0.12	-0.84	-0.71
Z:				
Mean	-21.63	11.43	-18.44	-20.76
Standard deviation	54.68	100.44	102.86	91.31
Skewness	-0.38	0.67	1.11	1.02
D:				
Mean	-0.46	-8.06	4.41	-4.38
Standard deviation	20.52	27.67	30.57	29.94
Skewness	-1.64	-0.26	-0.95	-0.8

Supplementary table 1. Although there was also a predicted increase in secular variation, thus picking up the magnetic field gradient due to the South Atlantic Anomaly, the CHAOS model does not pick it up as intensely as the ground observatory measurements.

Rapid secular variation pulses, or jerks, have been identified across the globe. Previous studies pertaining to the southern African region have noted its rapid changes in secular variation pattern.[16-18] The 2014 jerk was observed in southern Africa and was analysed by
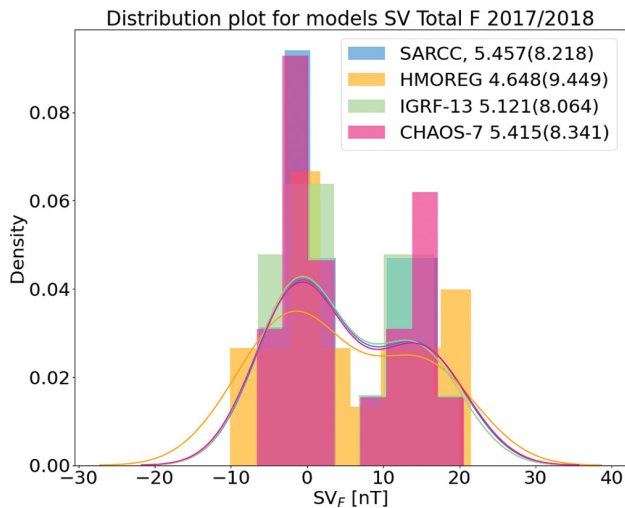
Figure 6: Histograms showing the difference between the calculated total field (in nT) yearly secular variation (SV) between the model data and the field survey data for 2017/2018. The values in the legend are the respective mean and standard deviation (in brackets) for each model.

Kotzé[19] using data from four observatories located at Hermanus, Hartebeesthoek, Keetmanshoop and Tsumeb. It was found that the data from all four of these observatories showed strong individual characteristics. These rapid changes in magnetic field cannot always be predicted by global geomagnetic field models[16] – an observation also seen in this study.

Using the CHAOS secular variation model, Mandea et al.[17] also showed that southern Africa is prone to rapidly changing secular variation. It was also clear from their study that the secular variation was more intense on the western side of South Africa than on the astern side. This finding can also be seen from the latest IGRF model output[20] (see Supplementary figure 1).

It should also be noted that jerks are not always observed simultaneously in all geomagnetic field components at a particular observatory[17], which should be taken into consideration when interpreting the results of this study where the secular variation of the total field is shown.

Conclusions

When looking at the total data set spanning over 4 years, both local field models (SARCC and HMOREG) deviated less from the mean than the global field models (CHAOS-7 and IGRF-13). Overall, SARCC showed a higher reduction in error than HMOREG and HMOREG showed a considerably lower skewness than SARCC.

When interpreting the secular variation results from this study, both the gradient of the secular variation strength observed over southern Africa (increasing in intensity from east to west) and the diverseness of the secular variation strength should be considered. The former is due to the presence of the South Atlantic Anomaly and is the cause of the spatial variability observed in this region. The latter is possibly due to the inhomogeneous occurrence of jerks at different locations, as well as the current limited prediction capabilities of global field models for rapid secular variation events.[5,10,16] A complete analysis and interpretation of this topic is beyond the scope of this paper and will be part of future work.

The new regional SARCC model consists of a core field from the CHAOS-6 global field model and a high-resolution lithospheric field component derived from several input data sets including regional field surveys. The novelty of this model is the inclusion of ground measurements from repeat station data for the lithospheric component, which enables the inclusion of small-scale spatial variations in the region usually not possible by global field models.

This model can be updated to include the latest version, CHAOS-7. It compares well with field survey data taken during four campaigns from 2015 to 2018, specifically when looking at the standard deviation for each respective component, as well as the distribution of each component. The SARCC showed on average a 33% reduction in error compared to the global field models. This is likely due to the inclusion of a lithospheric magnetic field component in the model, which includes regionally dependent, small-scale variations with a much weaker signal than the core component.[16] It is easy to produce updates for this model, and because of the availability of regional field surveys in southern Africa, the model can regularly be evaluated for accuracy.

Competing interests

We have no competing interests to declare.

Authors' contributions

A.E.N.: Methodology; data analysis; validation; writing – the initial draft; writing – revisions. A.M.: Conceptualisation. F.V.: Conceptualisation; validation. J.M.: Conceptualisation.

References

1. Mandea M, Chambodut A. Geomagnetic field processes and their implications for space weather. Surv Geophys. 2020;41:1611–1627. https://doi.org/10.1007/s10712-020-09598-1

2. Backus G, Parker R, Constable C. Foundations of geomagnetism. Cambridge: Cambridge University Press; 1996.

3. Geese A, Mandea M, Lesur V, Hayn M. Regional modelling of the southern African geomagnetic field using harmonic splines. Geophys J Int. 2010;181:1329–1342. https://doi.org/10.1111/j.1365-246X.2010.04575.x

4. Langel R, Hinze W. The magnetic field of the earth's lithosphere: The satellite perspective. Cambridge: Cambridge University Press; 2020. https://doi.org/10.1017/CBO9780511629549

5. Mandea M, Holme R, Pais A, Pinheiro K, Jackson A, Verbanac G. Geomagnetic jerks: Rapid core field variations and core dynamics. Space Sci Rev. 2010;155:147–175. https://doi.org/10.1007/s11214-010-9663-x

6. Thébault E, Finlay CC, Beggan CD, Alken P, Aubert J. Barrois O, et al. International geomagnetic reference field: The 12th generation. Earth Planet Sp. 2015;67, Art. #79. https://doi.org/10.1186/s40623-015-0228-9

7. Alken P. International Geomagnetic Reference Field [webpage on the Internet]. No date [updated 2019 Dec 19; cited 2020 Feb 24]. Available from: www.ngdc.noaa.gov/IAGA/vmod/igrf.html

8. Olsen N, Lühr H, Finlay CC, Sabaka TJ, Michaelis I, Rauberg J, et al. The CHAOS-4 geomagnetic field model. Geophys J Int. 2014;197(2):815–827. https://doi.org/10.1093/gji/ggu033

9. Finlay CC, Olsen N, Kotsiaros S, Gillet N, Clausen L. Recent geomagnetic secular variation from Swarm and ground observatories as estimated in the CHAOS-6 geomagnetic field model. Earth Planet Sp. 2016;68, Art. #112. https://doi.org/10.1186/s40623-016-0486-1

10. Kotzé PB. Hermanus magnetic observatory: A historical perspective of geomagnetism in southern Africa. Hist Geo Space Sci. 2018;9:125–131. https://doi.org/10.5194/hgss-9-125-2018

11. St-Louis B. INTERMAGNET Technical reference manual [webpage on the Internet]. c2012 [cited 2020 Jul]. Available from: https://intermagnet.org/docs/Technical-Manual/technical_manual.pdf

12. Vervelidou F, Thébault E, Korte M. A high-resolution lithospheric magnetic field model over southern Africa based on a joint inversion of CHAMP, Swarm, WDMAM, and ground magnetic field data. J Geophys Res Solid Earth. 2018;9:897–910. https://doi.org/10.5194/se-9-897-2018

13. Thébault E. Global lithospheric magnetic field modelling by successive regional analysis. Earth Planet Sp. 2006;58:485–495. https://doi.org/10.1186/BF03351944

14. Nahayo E, Kotzé PB, Korte M, Webb SJ. A harmonic spline magnetic main field model for southern Africa combining ground and satellite data to describe the evolution of the south Atlantic anomaly in this region between 2005 and 2010. Earth Planet Sp. 2018;70, Art. #30. https://doi.org/10.1186/s40623-018-0796-6

15. Kotzé PB, Korte M, Mandea M. Polynomial modelling of southern African secular variation observations since 2005. Data Sci J. 2011;10. https://doi.org/10.2481/dsj.IAGA-16

16. Kotzé PB, Korte M. Morphology of the southern African geomagetic field derived from observatory and repeat station survey observations: 2005-2014. Earth Planet Sp. 2016;68, Art. #23. https://doi.org/10.1186/s40623-016-0403-7

17. Kotzé PB. The time-varying geomagnetic field of southern Africa. Earth Planet Sp. 2003;55:111–116. https://doi.org/10.1186/BF03351738

18. Kotzé PB. The 2014 geomagnetic jerk as observed by southern African magnetic observatories. Earth Planet Sp. 2017;69, Art. #17. https://doi.org/10.1186/s40623-017-0605-7

19. Kotzé PB. Statistical evaluation of global geomagnetic field models over southern Africa during 2015. Earth Planet Sp. 2017;69, Art. #84. https://doi.org/10.1186/s40623-017-0670-y

20. Alken P, Thébault E, Beggan CD, Amit H, Aubert J, Baerenzung J, et al. International Geomagnetic Reference Field: The thirteenth generation. Earth Planet Sp. 2021;73, Art. #49. https://doi.org/10.1186/s40623-020-01288-x

AUTHORS:
Heidi Prozesky[1,2] (iD)
François van Schalkwyk[1,2] (iD)

AFFILIATIONS:
[1]DSI/NRF Centre of Excellence in Scientometrics and Science, Technology and Innovation Policy (SciSTIP), Stellenbosch University, Stellenbosch, South Africa
[2]Centre for Research on Evaluation, Science and Technology (CREST), Stellenbosch University, Stellenbosch, South Africa

CORRESPONDENCE TO:
Heidi Prozesky

EMAIL:
hep@sun.ac.za

The profile of postdoctoral research fellows in South Africa: Trends over the past two decades

The need for evidence on postdoctoral research fellows, or 'postdocs', has become increasingly evident globally, as many countries, including South Africa, lack information on these researchers. We used existing sources of quantitative data to provide a longitudinal profile of postdocs at South African public universities. These sources include national research-and-development surveys and data from the national Higher Education Management Information System. Our focus was on postdocs' representation and research intensity (i.e. time spent on research) within the national research system, as well as on two key demographic features, namely nationality and gender. We draw comparisons with findings reported for other countries to situate our results within the global science system. Interpretation of our results leads us to comment on global and national developments and policies relevant to postdocs.

Significance:
- This article provides, for the first time, longitudinal, quantitative evidence at the national level about postdocs in South Africa, a category of early-career researcher which has been growing in size and relevance but has remained largely invisible.
- The findings provide the basis for a critical reflection on current discourses and policies related to postdocs in South Africa.
- Interpretation of the findings allows us to identify priorities for future research on postdocs in South Africa that would be most useful to inform both national and institutional policy.

Introduction

In the winter of 2022, one South African university placed an advertisement for 200 new postdoctoral research fellows, or 'postdocs'. According to the same university's vice-chancellor, this would bring the number of postdocs at the institution to 650. To put this into perspective, in 2020, the same university employed 707 permanent academic staff with a doctoral degree, the equivalent qualification held by *all* postdocs. The advertisement refers to the university's rise in the global rankings and emphasises the research productivity of postdocs. The productivity requirement for postdocs at the university is contractual – postdocs are expected to "do research and to publish such research in accredited journals or conference proceedings. The norm is the publication of two accredited research units."

This example illustrates two global discourses concerning postdocs. According to the pipeline discourse, postdocs are important for sustaining the academic profession[1,2], that is, in reproducing academic labour[3]. In the human-capital discourse, postdocs are valued as highly skilled labour[1,4-6] for bolstering research and innovation. However, while the pipeline and capital discourses augur a productive future for the postdoc, many postdocs seem to face a professional cul-de-sac and the harsh realisation that a career in academia is unlikely.[1,7]

Another discourse relevant to postdocs is that of "science as a global market of competing world-class universities"[8]. The introduction of postdoctoral programmes by South African universities was, in part, underpinned by national priorities geared toward developing the nation's knowledge capital as globally competitive.[6] Therefore, the emergence of postdocs is, to some extent at least, indicative of the competitive behaviour of South African universities. Institutional policies emphasise the number of research publications produced by postdocs.[9] This emphasis is unsurprising in the South African context given that universities compete with each other for significant financial subsidies from the national government based on the number of publications produced in a given year. Combined with the fact that many postdocs are externally funded, postdocs are efficient income generators for South African universities.

At a more systemic level, national policy documents refer to the postdoc role as providing doctoral graduates with further experience in research and innovation[10], especially as academics in the university sector[11]. Postdocs are also seen as playing a crucial part in 'augmenting' doctoral supervisory capacity at universities.[12-14] This capacity is required to alleviate a supervisory 'bottleneck' created by the greater increase, from 1996 to 2014, in the number of doctoral enrolments (350%) compared with staff with a doctoral qualification (65%). South African universities face a conundrum: they need more academics with doctoral qualifications, and to do so, they need to produce more doctoral graduates. But for that, they need more staff with doctoral qualifications to supervise doctoral students.[15]

Postdocs partly solve this conundrum and are therefore considered functional for the South African university sector[16], but the postdoctoral fellowship as a social system may also be dysfunctional for postdocs. Concerns have been raised by the Organisation for Economic Co-operation and Development (OECD) about postdocs as part of the "research precariat […] who work in positions with little job security, poor compensation and an unclear path to a permanent post"[17]. In a recent national tracer survey of more than 6000 respondents who obtained a doctoral degree from a South African university during 2000–2018, nearly 30% had accepted a postdoctoral fellowship on completion of their studies because "other employment was not available"[18]. Recently, this percentage has increased, which suggests that the absorptive capacity of the South African science system to employ doctoral graduates and postdoctoral fellows has already reached a point of saturation[18] (also see Simmonds and Bitzer[11]).

Universities South Africa, in its 2013 national postdoc policy, was attuned to precarity and the risk of postdoc 'serialisation'. The organisation included a statement limiting the period of a postdoc fellowship to a maximum period of 5 years.[19] Whether universities adhere to this policy prescription is unknown; nor does it resolve the matter of postdocs facing an academic dead-end at the end of the 5-year period. An oversupply of postdocs relative to available faculty positions has led to calls for improved assessment of career outcomes.[20-22]

Considering the background provided above, it seems surprising that there has not yet been any large-scale, quantitative investigation of postdocs in South Africa. In its policy paper on reducing the precarity of academic research careers, the OECD identifies the need for more evidence on postdocs, noting that countries often do not have a good understanding of even the number of such researchers.[23] This lack of data about postdocs "has made it difficult to develop policies to address concerns about salaries, working conditions, diversity and career development, and to evaluate the impact of existing policies"[24] (also see Bankston and McDowell[25]).

Even though we lack generalisable data on South African postdocs that would allow us to answer questions on the effects of employment precarity on postdocs, some existing data are available and – when analysed in combination – provide an insightful profile of postdoctoral research fellows in South Africa over the past two decades. This is the focus of this article.

In South Africa, annual, national surveys of research and development (from here on referred to as the R&D surveys) are used to collect data on postdocs in the country's science system (including universities). However, until now, these figures have been reported without any further analysis and/or interpretation. We analysed this data source, in combination with data from the national Higher Education Management Information System (HEMIS), to sketch a national profile of postdocs, and how it has changed over the past two decades.

We calculated and examined postdocs' representation and research intensity (i.e. time spent on research) within the higher education sector, with a further focus on two demographic features: nationality and gender. Nationality is a key variable, given that postdoc cohorts typically comprise large numbers of international fellows, while in South Africa national policy has recently placed restrictions on the number of non-South-African postdocs that can be funded by the public purse[26], and there is a growing concern about xenophobia in South African universities[27]. Gender is included in our analysis because the postdoc has been described as "a critical time for women, when personal life transitions and professional ambitions collide"[28]. Thus, it may be a stage during which women begin 'leaking' from the academic 'pipeline' at an increasing rate, as research elsewhere[24,29,30] has shown. Finally, we compared, where such data exist, South African postdocs with those

elsewhere, as the first step towards addressing the lack of evidence required as a basis for designing policy geared towards the management and professionalisation of postdocs in South Africa.

Methodology

The main sources of data for our secondary analysis were the statistical reports covering the fiscal years 2007/2008 to 2019/2020 produced by the Human Sciences Research Council's Centre for Science, Technology and Innovation Indicators, on the basis of the annual R&D surveys. In the reports, figures are also available for the preceding 4 years, which required consulting only the 2011/2012, 2016/2017 and 2019/2020 reports for figures pertaining to the period 2007–2019.[31-33] We also included data on postdocs found in one of the earliest R&D survey reports (for 2003).[34] All these reports are freely available online.

The R&D surveys collect data on all public universities, using a census approach, although in the past few years, between five and seven of the universities did not respond. The reference period is the year prior to the survey (the first year of the two years in a report's title). The reports provide the following data for postdocs: headcounts, full-time equivalents as a percentage of headcounts (an indicator of time spent on research, or research intensity), nationality (South African versus non-South African), and gender.[31-33] Where relevant, data from the Department of Higher Education and Training (DHET)'s Higher Education Management Information System (HEMIS)[35] were also drawn upon.

Regarding comparative global figures for the past two decades, the largest recent survey of postdocs is the 2020 *Nature* survey, which collected data for a total of 7670 respondents, representing 93 countries. The anonymised data[36] have been deposited in an open access repository (see go.nature.com/3tmckuq). Only 54 of the respondents lived in South Africa at the time of the survey; we deleted these data from the data set, together with those of the 381 respondents who were not employed in the university sector, to increase comparability with the R&D survey data. We analysed the remaining data ($n = 7235$) to provide a benchmark with which to compare the South African figures. Other benchmarks took the form of results from national surveys of postdocs conducted in other countries, which are referenced in the 'Results' section. Given that we analysed existing data, ethical approval does not apply.

Results

The first record of the number of postdocs in South Africa suggests that in 1999 there were approximately 300 postdocs.[37] One of the earliest R&D surveys found that number to be 357 in 2003.[34] Since then, the number of postdocs has increased to 2867 in 2019. The steepest growth rate (close to 200%) is observed for the first half of the period (2007–2013), and as Figure 1 shows, the growth has decelerated since 2017.
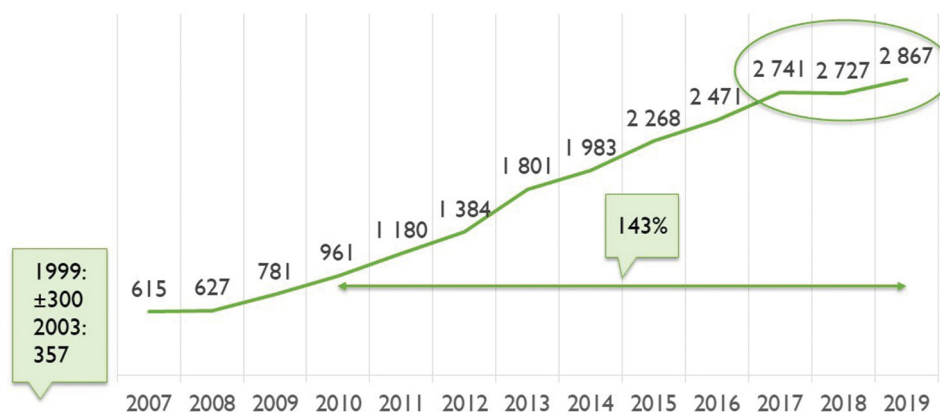

Figure 1: Number of postdocs at South African higher education institutions in 1999, 2003 and from 2007 to 2019, indicating decelerated growth since 2017, as well as the rate of increase from 2010 to 2019.

Comparative figures are rare, but those available suggest that South Africa's trend is aligned with those reported for at least two other countries. For example, from 2010 to 2019, the rate of increase of 143% (or 1180 to 2867) measured for South Africa is similar to the 144% increase recorded for Finland[23], and the 142% (from 10 559 to 25 514) for China over the same period[38].

Nationally, relational measures are useful to contextualise the increase in postdocs in South Africa. The first of such measures is provided by the R&D surveys in the form of researchers in the university sector with a doctoral degree or equivalent. Table 1 shows that, in 2007, postdocs constituted only 11% of such researchers. Six years later, this proportion almost doubled (to 21%), but has decreased slightly since then, to 18% in 2019.

From 2011 to 2019, the percentage of academic staff who hold a doctoral degree also increased quite dramatically, from 36% to 48%.[39] If this had not been the case, it is likely that the percentage of postdocs relative to doctoral-qualified academic staff may have at least remained stable or increased further.

A second national relational measure is the number of lecturers in South African public universities. We combined the R&D survey data with data from HEMIS on the number of permanent academic staff who hold the position lecturer (at all levels, i.e. junior lecturer, lecturer and senior lecturer), as this is the level at which one may assume postdocs, depending on their level of experience, would enter academe. We found that the number of those staff grew at only 37% from 2007 to 2019, compared with the almost ten times greater rate of growth (366%) in postdocs over the same period (Figure 2).

Postdocs do not only contribute to the human-resource base in absolute or relative numbers, but also in terms of their research intensity, i.e. the time they spend on research. The R&D surveys measure time spent on research in terms of research full-time equivalents or 'person years of effort' allocated to research, which are expressed as a percentage of headcounts. According to our analysis of these data, postdocs dedicate on average 92% of their time performing research (for the period 2007–2019). This is more than the 55% for doctoral students and almost four times the 24% reported for researchers (irrespective of highest qualification, as full-time equivalents are not disaggregated by that variable). As Figure 3 shows, these percentages fluctuate

noticeably for postdocs prior to 2012, but they have remained relatively stable since then and have not deviated by more than 2–4% from the averages for all three subgroups.

Next, we analysed South African postdocs according to two key demographic features. The R&D surveys disaggregate the headcounts of postdocs by nationality, but only from 2011 to 2016. On average over that period, 62% of the postdocs in South Africa's universities were not South African citizens. The percentage increased (by 13 points) from 54% in 2011 to 67% in 2014 but decreased again thereafter (Figure 4). The most recent available figure from the R&D surveys (2016) matches the 61% of respondents in the *Nature* survey who reported that they were not (in 2020) undertaking a postdoctoral fellowship in their native country.

Earlier research that involved, inter alia, the collection of data from the South African National Research Foundation (NRF) and from 13 universities, found that, from 2005 to 2010, women constituted 40% of postdocs in South Africa[40], while our analysis of the R&D surveys produced a slightly higher average for 2007 to 2019 of 42% of postdocs. These percentages are lower than those reported by cross-national surveys. In 2019, women's representation among postdocs in South Africa was measured at 41%, which is a notable 12 points lower than the percentage of female respondents (53% of the 7151 who self-identified as either male or female) in the global *Nature* survey. A 2017 survey of 898 postdocs at European universities found a similar, "higher number of responses from women (61%) than men"[41].

Figures reported by smaller, national surveys are more aligned with the South African ones. The 2013, 2016 and 2019 National Postdoctoral Association's surveys in the USA found women's representation to range between 43% and 44%.[42,43] The Canadian Association of Postdoctoral Scholars surveys of postdocs working in Canada and Canadians working internationally found female respondents in the 2009, 2013 and 2016 waves to be 44%, 46% and 48%, respectively.[3] Still, the percentage of South African women postdocs is consistently lower than those reported elsewhere.

To understand the extent of women's minority among postdocs in South Africa, a relational measure is again useful, this time taking into account career stage. Using HEMIS data[35], we determined that, from 2007 to 2019, women represented on average 43% of doctoral graduates

Table 1: Number of PhD-qualified researchers and postdocs in the higher education sector, 2007–2019. PhD-qualified researchers include permanent and contract (6 months or longer) positions, as well as emeritus professors and honorary fellows.

Year	PhD-qualified researchers	Postdocs	
	N	*N*	%
2007	5784	615	11%
2008	6392	627	10%
2009	7141	781	11%
2010	6578	961	15%
2011	7188	1180	16%
2012	7974	1384	17%
2013	9257	1801	19%
2014	9739	1983	20%
2015	10 638	2268	21%
2016	12099	2471	20%
2017	14 467	2741	19%
2018	14 735	2727	19%
2019	15 705	2867	18%
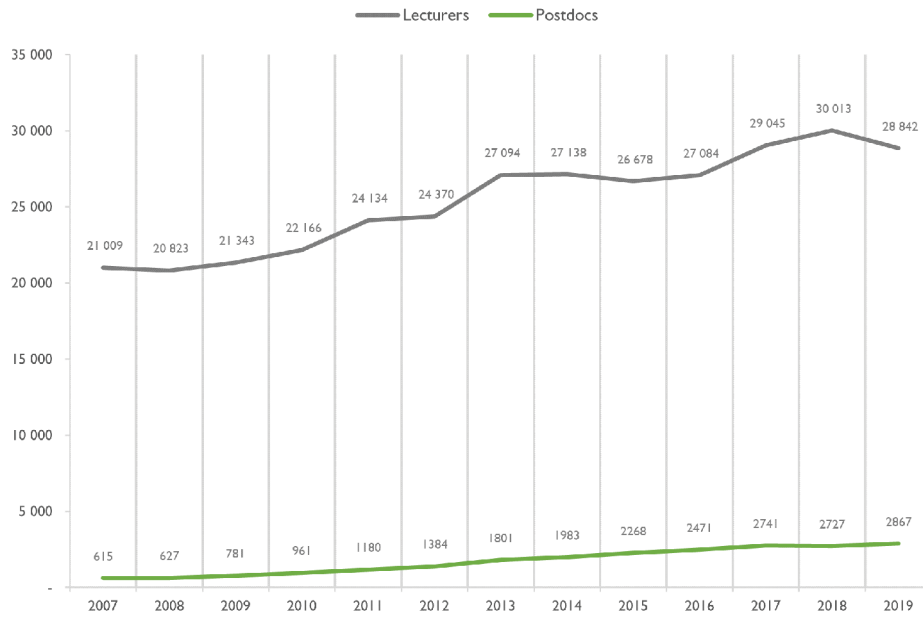
Figure 2: Number of postdocs compared to that of lecturers (junior lecturers, lecturers, and senior lecturers) at South African public universities, 2007–2019.
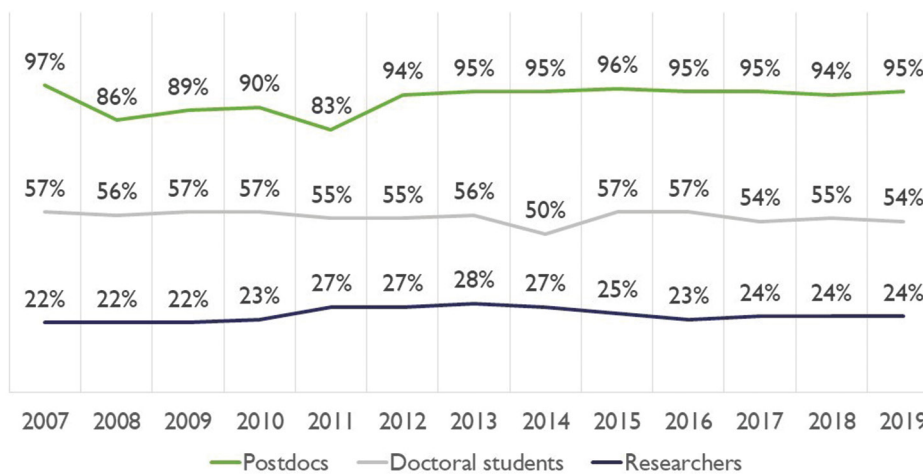
Figure 3: Research full-time equivalents as percentages of headcounts (a measure of time spent on research) for doctoral students, postdocs, and researchers in the university sector, 2007–2019.
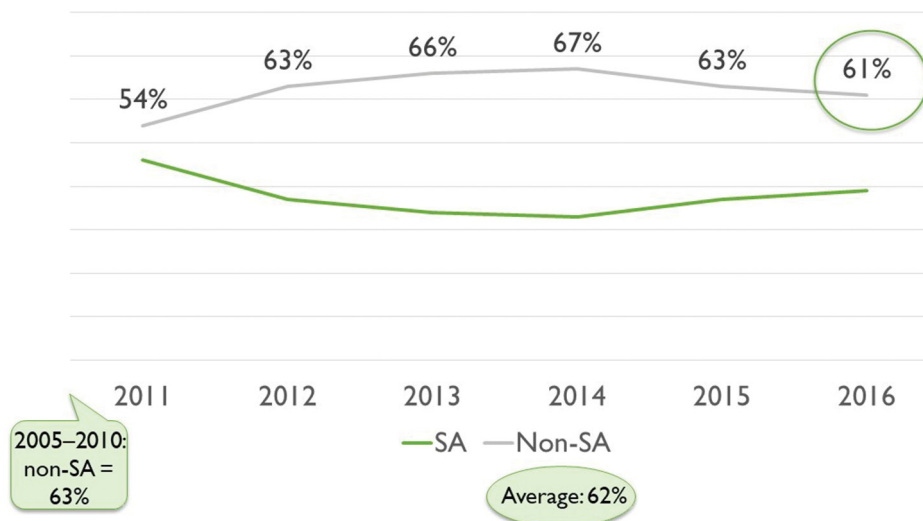
Figure 4: Percentage of postdoctoral fellows who are South African (SA) and non-SA nationals, 2011–2016.

from South African universities. The R&D survey data indicate that, among researchers at universities with a doctoral qualification, women constituted only 39%, on average, over the same period. Figure 5 shows that the percentages of women in the three career stages have begun to converge, especially since 2013.

Figure 5 further shows that in South Africa there has been no clear trend towards either an increase or decrease in the percentage of female postdocs from 2007 to 2019. Our results do not align with a steady growth in female representation among postdocs reported in Canada[3] and the USA[44] over similar periods.

Discussion and conclusion

Our longitudinal analysis shows a significant growth in the number of postdocs at universities in South Africa, both in absolute numbers and relative to the growth of doctoral-qualified researchers and permanently employed lecturers in the university sector. Since 2010, the growth has been comparable with that observed for at least two other countries for which data are available. We also show that postdocs are highly research intensive, as they dedicate much more of their time to research than do permanent academic staff and even doctoral students. Among the latter, 60% study while they are employed[18], while enrolment in degree programmes with a tuition component also reduces time spent on research. South African policy statements[12-14] and recent research[1,11] indicate that postdocs in the country are increasingly required to assist academics with other academic duties, in particular, doctoral supervision, but our results show that this has not (yet) impacted negatively on the high percentage of postdocs' time spent on research. Postdocs, therefore, can be expected to be a major contributor to the research output of their host universities, although the extent of their contribution is yet to be determined and should be a focus of future research.

In relation to the dominant discourses in higher education referred to in the introduction to this article, the findings support the notion of ascendancy of one particular global discourse (of four) proposed by Marginson[8], that is, of science as a global market of competing universities. Research, particularly the number of peer-reviewed publications produced, is a major indicator in the calculation of universities' scores in global rankings, and one of the primary drivers of the global reputations of both researchers and their host universities.[45] From a local perspective, South Africa's system of rewarding universities financially for publications means that the country's 26 public universities also compete nationally for finite government funding ringfenced for research-output-based subsidy allocations.[46] It is therefore likely that universities, in response, are appointing increasing numbers of postdocs at relatively low labour cost to sustain or increase a competitive advantage.

From a broader and theoretical perspective, this institutional response can be understood with reference to academic capitalism[47] and resource dependence theory[48]. The former describes a shift to enterprise modes of academic production (attributable to the spread of global capitalism), including the rise of competition regimes in the distribution of public funds for research. At the same time, because of changing priorities in government funding, universities in South Africa are becoming increasingly dependent on external sources of funding (i.e. so-called 'third-stream' funding)[49] to support their research activities. According to resource-dependence theory, this situation fosters organisational strategies that, for example, stimulate demand for postdocs as highly qualified labour over which universities have leverage (instead of creating positions for permanent academic staff protected by labour-friendly legislation), and which do not represent long-term commitments by universities. These developments have coincided with global and national increases in doctoral graduates, and an inadequate labour market demand for their skills, thereby providing a surplus supply of labour, in the form of postdocs.[50]

If postdocs provide the economically optimal research workers for universities in this context, they will be recruited, without regard for the number of career positions available.[50] In this way, the human-capital discourse is overridden by the imperatives of academic capitalism, as reflected by our finding that the growth in postdocs is almost ten times higher than the growth in staff in permanent positions available to postdocs. In such a context, one also observes a form of credential inflation[1]: the greater the mismatch between the supply of doctoral graduates and the availability of permanent academic positions, the more doctoral graduates are encouraged to pursue a postdoc in order to become more competitive and improve their chances in the job market, if they are strongly committed to an academic career.[21] Although postdocs may find career opportunities outside of universities, these are limited in South Africa for doctoral graduates[18], and, for most postdocs, the goal remains to secure a permanent position in academia[11,18].

Our findings therefore support concerns raised[1,9,18] about the casualisation of academic labour in the country, and further indicate that it may be gendered. Although there is variation over time in the representation of women among South African postdocs, it is within a relatively narrow range (compared to some other countries) and the variation seems to be decreasing. At the same time, the percentage of women among researchers with a doctoral qualification has increased, and there does not seem to be a 'leak' from the postdoctoral pipeline into permanent positions. Male and female doctoral graduates from South African universities have also been found to be equally likely to accept a postdoc fellowship, suggesting that female representation is not a major issue of concern in the case of postdocs.[18]
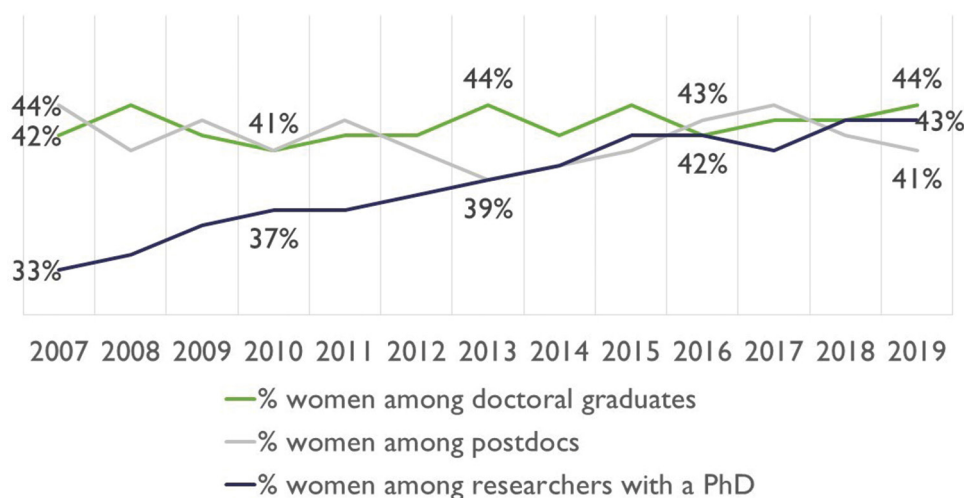

Figure 5: Percentage of women among postdocs compared to percentage of women among doctoral graduates and among researchers with a doctoral qualification (university sector only), 2007–2019.

It may be that employment policies favouring women doctoral graduates for permanent positions are reducing the percentage of women in postdoctoral fellowships. In other words, if women have a greater chance of securing scarce full-time positions than men, they may not need to increase their market value through postdoctoral training. However, further research on this issue is required, and as postdoctoral fellowships are not only precarious, but also offer various benefits to those who hold them, the percentage of female postdocs in South Africa needs to be carefully monitored.

It is important to note that an increase in postdocs is not necessarily indicative of a healthy national system of innovation. Growth in postdocs tends to be associated with economic recessions and limited growth in R&D funds[51], which, in turn, make it increasingly hard for postdocs to find permanent employment[52]. Postdocs' precarious contracts could lead to a potential loss of knowledge for a national system of innovation, as they switch from one contract to another, or exit research career paths altogether.[53] The latter outcome amounts to an inefficient use of resources and inefficiency in the production of knowledge associated with an increase in postdocs because

> as a low-cost complement to faculty and a substitute for other labour inputs there is little incentive to make full use of postdocs' capabilities or to make difficult decisions about the allocation of scarce resources.[50]

There would also be fewer incentives to increase the number of permanent faculty members[54], as our findings suggest is the case in South Africa.

While the number of postdocs in the South African university system has grown over the past two decades, the findings also show a recent deceleration in the growth of postdocs. Without more data, also on choices made by doctoral graduates and the experiences of postdocs, it is difficult to pinpoint reasons for the deceleration in the growth of postdocs in South Africa. The notion of credential inflation referred to above seems to discount, as a possible reason, an increasing realisation among doctoral graduates and first-time postdocs that postdoc positions are increasingly unlikely to lead to permanent academic or other work opportunities in South Africa. We would suggest, for future research, that the focus should be on a decline in the growth in international postdocs in South Africa, as our limited data seem to suggest is the case.

In this regard, 'demand' in host countries for international postdocs[55], as reflected in various aspects of public policy that attract the foreign born to study in a country[56], has been shown to play an important role. A notable signal of a reduced demand for international postdocs is that NRF funding for such postdocs has recently been reduced. The NRF's new funding framework prescribes that universities' applications for its postdoctoral funding must be aligned with the foundation's "equity target" of "80% of South African citizens and permanent residents"[57]. In 2018, before the implementation of the framework, 41% of the 799 NRF-funded postdocs were not South African citizens or permanent residents[58], and that percentage has most likely decreased since then.

The new framework has been described as a short-sighted development that does not appreciate the importance of attracting foreign talent to the country[18] and is likely to stifle the contribution of international postdocs to South Africa's science system and the country's development[26]. The contributions of foreign-born postdocs to their host countries have been detailed elsewhere[56], but Gaughan and Bozeman[59] caution that "dependence on the foreign-born" exposes a country to "the vicissitudes of highly political immigration policy debates". This observation is highly relevant to South Africa, where evidence[27] indicates that foreign-born researchers, especially from the rest of Africa, are facing various forms of xenophobia.

The postdoc therefore becomes a crucial site for the investigation of divergence or convergence of nationalist political priorities (including policy imperatives of local equity and inclusion), the national discourse of development (specifically, universities' role in development[52]),

and universities' institutional policies that promote the recruitment of postdocs within the discourse of science as a global market of competing universities. To do so constructively, the notion and expectation that postdocs contribute, via their research, to economic development needs to be interrogated critically.

In this article we have illustrated how relatively standard R&D survey data may be put to use to provide an evidence-based overview of postdocs in a country, especially if it is interpreted within the local policy context and compared with global trends. However, as is usually the case with the secondary analysis of existing data, our analyses remain limited in many respects. Further research is required to better understand postdocs' contribution to the national science system and to their host universities' research output, as well as their experiences as a relatively new type of knowledge worker that has become institutionalised in the South African university workforce.

Acknowledgements

Competing interests

We have no competing interests to declare.

Authors' contributions

H.P.: Conceptualisation; methodology; data collection; analysis; writing – the initial draft; writing – revisions. F.v.S.: Data collection; analysis; writing – revisions.

References

1. Kerr P. Academic pipeline or academic treadmill? Postdoctoral fellowships and the circular logic of 'development'. S Afr J High Educ. 2022;36(3):72–90. https://doi.org/10.20853/36-3-5080

2. Wei TE, Levin V, Sabik LM. A referral is worth a thousand ads: Job search methods and scientist outcomes in the market for postdoctoral scholars. Sci Public Policy. 2012;39:60–73. https://doi.org/10.1093/scipol/scr006

3. Van Benthem K, Adi MN, Corkery CT, Inoue J, Jadavji NM. The changing postdoc and key predictors of satisfaction with professional training. Stud Grad Postdr Educ. 2020;11:123–142. https://doi.org/10.1108/sgpe-06-2019-0055

4. Acs ZJ, Braunerhjelm P, Audretsch DB, Carlsson B. The knowledge spillover theory of entrepreneurship. Small Bus Econ. 2009;32:15–30. https://doi.org/10.1007/s11187-008-9157-3

5. Bozeman B, Dietz JS, Gaughan M. Scientific and technical human capital: An alternative model for research evaluation. Int J Technol Manag. 2001;22(7/8):716–740. https://doi.org/10.1504/IJTM.2001.002988

6. Holley K, Kuzhabekova A, Osbaldiston N, Cannizzo F, Mauri C, Simmonds S, et al. Global perspectives on the postdoctoral scholar experience. In: Jaeger AJ, Dinin AJ, editors. The postdoc landscape: The invisible scholars. London: Academic Press; 2018. p. 203–226. https://doi.org/10.1016/b978-0-12-813169-5.00009-4

7. Hayter CS, Parker MA. Factors that influence the transition of university postdocs to non-academic scientific careers: An exploratory study. Res Policy. 2019;48(3):556–570. https://doi.org/10.1016/j.respol.2018.09.009

8. Marginson S. What drives global science? The four competing narratives. Stud High Educ. 2022;47:1566–1584. https://doi.org/10.1080/03075079.2021.1942822

9. Van Schalkwyk FB. The 'academic precariat': Postdoctoral fellows in South African higher education. Brief Speak. 2022;20:1–14. Available from: https://www.che.ac.za/file/6493/download?token=sBaJ3iMj

10. South African National Research Foundation (NRF). Annual report 2017–2018. Pretoria: NRF; 2018. Available from: https://www.nrf.ac.za/wp-content/uploads/2021/05/NRF-Annual-Report-2017-18.pdf

11. Simmonds S, Bitzer E. The career trajectories of postdocs in their journeys of becoming researchers. In: Du Preez P, Simmonds S, editors. A scholarship of doctoral education: On becoming a researcher. Stellenbosch: African Sun Media; 2018. p. 273–293. https://doi.org/10.18820/9781928357933

12. South African Department of Science and Technology (DST). Human capital development strategy for research, innovation and scholarship. Pretoria: DST; 2016. Available from: https://www.dst.gov.za/images/Human-Capital-Development-Strategy-for-Research-Innovation-and-Scholarship.pdf

13. South African Department of Science and Technology (DST). White paper on science, technology and innovation. Pretoria: DST; 2019. Available from: https://www.dst.gov.za/images/2019/White_paper_web_copyv1.pdf

14. Academy of Science of South Africa (ASSAf). Status of postgraduate research training in engineering in South Africa. Pretoria: ASSAf; 2018. https://doi.org/10.17159/assaf.2018/0032

15. Breier M, Herman C. The PhD conundrum in South African academia. High Educ Q. 2017;71:352–368. https://doi.org/10.1111/hequ.12134

16. Merton RK. The Matthew effect in science. Science. 1968;159(3810):56–63. https://doi.org/10.1126/science.159.3810.56

17. Woolston C. Postdoc survey reveals disenchantment with working life. Nature. 2020;587:505–508. https://doi.org/10.1038/d41586-020-03191-7

18. Mouton J, Van Lill M, Prozesky H, Bailey T, Duncan M, Boshoff N, et al. A national tracer study of doctoral graduates in South Africa. Report of the DSI-NRF Centre of Excellence in Scientometrics and Science, Technology and Innovation Policy to the Department of Science and Innovation; 2022 [cited 2023 Mar 12]. Available from: https://www0.sun.ac.za/scistip/wp-content/uploads/2022/07/National-Tracer-Study-of-Doctoral-Graduates-in-SA_DSI-approved.pdf

19. Universities South Africa. Policy for postdoctoral research fellowships / postdoctoral research fellows. Available on request from corresponding author.

20. Gibbs KD, McGready J, Griffin K. Career development among American biomedical postdocs. CBE Life Sci Educ. 2015;14(4), Art. #44. https://doi.org/10.1187/cbe.15-03-0075

21. Sauermann H, Roach M. Why pursue the postdoc path? Science. 2016;352(6286):663–664. https://doi.org/10.1126/science.aaf2061

22. Silva EA, Des Jarlais C, Lindstaedt B, Rotman E, Watkins ES. Tracking career outcomes for postdoctoral scholars: A call to action. PLoS Biol. 2016;14(5), e1002458. https://doi.org/10.1371/journal.pbio.1002458

23. Organisation for Economic Co-operation and Development (OECD). Reducing the precarity of academic research careers. OECD science, technology and industry policy papers no. 113. Paris: OECD Publishing; 2021. https://doi.org/10.1787/0f8bd468-en

24. McConnell SC, Westerman EL, Pierre JF, Heckler EJ, Schwartz NB. United States National Postdoc Survey results and the interaction of gender, career choice and mentor impact. eLife. 2018;7, e40189. https://doi.org/10.7554/eLife.40189

25. Bankston A, McDowell GS. A review of postdoc reforms in the United States and the case of the Fair Labor Standards Act Updates of 2016. In: Jaeger AJ, Dinin AJ, editors. The postdoc landscape: The invisible scholars. London: Academic Press; 2018. p. 15–48. https://doi.org/10.1016/B978-0-12-813169-5.00002-1

26. Van Schalkwyk FB, Van Lill MH, Cloete N. Brain circuity: The case of South Africa as a hub for doctoral education. S Afr J Sci. 2021;117(9/10), Art. #10674. https://doi.org/10.17159/sajs.2021/10674

27. Jansen J. Corrupted: A study of chronically dysfunctional universities. Johannesburg: Wits University Press; 2023. https://doi.org/10.18772/12023027946

28. Ysseldyk R, Greenaway KH, Hassinger E, Zutrauen S, Lintz J, Bhatia MP, et al. A leak in the academic pipeline: Identity and health among postdoctoral women. Front Psychol. 2019;10:1297. https://doi.org/10.3389/fpsyg.2019.01297

29. Lambert WM, Wells MT, Cipriano MF, Sneva JN, Morris JA, Golightly LM. Career choices of underrepresented and female postdocs in the biomedical sciences. eLife. 2020;9, e48774. https://doi.org/10.7554/eLife.48774

30. Martinez ED, Botos J, Dohoney KM, Geiman TM, Kolla SS, Olivera A, et al. Falling off the academic bandwagon: Women are more likely to quit at the postdoc to principal investigator transition. EMBO Rep. 2007;8(11):977–981. https://doi.org/10.1038/sj.embor.7401110

31. Centre for Science, Technology and Innovation Indicators. South African national survey of research and experimental development: Statistical report 2011/12. Pretoria: Human Sciences Research Council; 2014. Available from: https://repository.hsrc.ac.za/bitstream/handle/20.500.11910/9395/9074.pdf?sequence=1&isAllowed=y

32. Centre for Science, Technology and Innovation Indicators. South African national survey of research and experimental development: Statistical report 2016/17. Pretoria: Human Sciences Research Council; 2018. Available from: https://www.dst.gov.za/images/2018/SA-RD-SURVEY-STATISTICAL-REPORT-201617_WEB.pdf

33. Centre for Science, Technology and Innovation Indicators. South African national survey of research and experimental development: Statistical report 2019/20. Pretoria: Human Sciences Research Council; 2021. Available from: https://www.dst.gov.za/images/RD_StatisticalReport2019-20__WEB.pdf

34. Centre for Science, Technology and Innovation Indicators. National survey of research & development 2003/2004: Provisional report on the sectors. Pretoria: Human Sciences Research Council; 2005. https://repository.hsrc.ac.za/handle/20.500.11910/9184

35. South African Department of Higher Education and Training (DHET). Higher Education Management Information System. April 2022 version. Available from: https://www.dhet.gov.za/SitePages/Higher-Education-Management-Information-System.aspx

36. Nature, Shift Learning. Nature_Postdoc survey 2020_anonymised data_v1 [data file]. c2020 [cited 2022 Apr 26]. Available from: https://go.nature.com/3tmckuq

37. Vranas L, Hendry J. Discovering the postdoctoral sector at UCT with a view to improving the PDRF experience. S Afr J High Educ. 2013;27(2):454–473.

38. Wang C, Yang L. Providing the research system with sufficient competent early-career researchers: The system of postdoctoral programs in China. HERB. 2020;2(23):15–17.

39. South African Department of Higher Education and Training (DHET). Post-school education and training monitor report: Macro-indicator trends March 2021. Pretoria: DHET; 2021. Available from: https://www.dhet.gov.za/Planning%20Monitoring%20and%20Evaluation%20Coordination/Post-School%20Education%20and%20Training%20Monitor%20-%20Macro-Indicator%20Trends%20-%20March%202021.pdf

40. Centre for Research on Evaluation, Science and Technology. Impact assessment of the Innovation Postdoctoral Fellowship Programme, 2005–2010. Report commissioned by the NRF. Available on request from corresponding author.

41. Ribeiro MJ, Fonseca A, Ramos MM, Costa M, Kilteni K, Andersen LM, et al. Postdoc X-ray in Europe 2017: Work conditions, productivity, institutional support and career outlooks [preprint]. bioRxiv; Version 2. https://doi.org/10.1101/523621

42. Ferguson K, McTighe M, Amlani B, Costello T. National Postdoctoral Association institutional policy report 2017: Supporting the needs of postdocs. Rockville, MD: National Postdoctoral Association; 2017. Available from: https://www.sigmaxi.org/docs/default-source/Publications-Documents/2017_supporting_the_needs_of_postdocs.pdf?sfvrsn=81c9b858_12

43. Ferguson K, Chen L, Costello T. Growing progress in supporting postdocs 2021: National Postdoctoral Association institutional policy report. Am Sci. 2021;109(3). Available from: https://www.sigmaxi.org/docs/npa-report-2021

44. Garrison HH, Justement LB, Gerbi SA. Biomedical science postdocs: An end to the era of expansion. FASEB J. 2016;30:41–44. https://doi.org/10.1096/fj.15-280552

45. Hazelkorn E. Impact of global rankings on higher education research and the production of knowledge. UNESCO Forum on Higher Education, Research and Knowledge. Occasional Paper No. 18; 2009. https://doi.org/10.21427/D7BG9J

46. Academy of Science of South Africa (ASSAf). Twelve years later: Second ASSAf report on research publishing in and from South Africa. Pretoria: ASSAf; 2019. https://doi.org/10.17159/assaf.2018/0030

47. Rhoades G. Postdoc identity, jurisdictional issues, ideologies, and unions: Considerations in organizing professionals. Labor Stud J. 2023;48:101–120. https://doi.org/10.1177/0160449X231155636

48. Cantwell B, Taylor BJ. Rise of the science and engineering postdoctorate and the restructuring of academic research. J High Educ. 2015;86:667–696. https://doi.org/10.1353/jhe.2015.0028

49. Swartz R, Ivancheva M, Czerniewicz L, Morris NP. Between a rock and a hard place: Dilemmas regarding the purpose of public universities in South Africa. High Educ. 2019;77:567–583. https://doi.org/10.1007/s10734-018-0291-9

50. Miller JM, Feldman MP. The sorcerer's postdoc apprentice: Uncertain funding and contingent highly skilled labour. Camb J Reg Econ Soc. 2014;7:289–305. https://doi.org/10.1093/cjres/rsu003

51. Powell K. The postdoc experience: High expectations grounded in reality. Science. 2012;337:992–996. Available from: https://www.science.org/do/10.5555/article.2380983/full/xscience1200121.pdf

52. Ahmed MZ, Plotkin D, Qiu B-L, Kawahara AY. Postdocs in science: A comparison between China and the United States. BioScience. 2015;65:1088–1095. https://doi.org/10.1093/biosci/biv125

53. Ma L. No metrics for postdocs: Precarious labour in science policy. In: Glänzel W, Heeffer S, Chi PS, Rousseau R, editors. 18th International Conference on Scientometrics & Informetrics; 2021 July 12–15; Leuven, Belgium. Leuven: International Society for Scientometrics and Informetrics; 2021. p. 711–715.

54. Andalib MA, Ghaffarzadegan N, Larson RC. The postdoc queue: A labour force in waiting. Syst Res Behav Sci. 2018;35:675–686. https://doi.org/10.1002/sres.2510

55. Cantwell B, Taylor BJ. Internationalization of the postdoctorate in the United States: Analyzing the demand for international postdoc labor. High Educ. 2013;66:551–567. https://doi.org/10.1007/s10734-013-9621-0

56. Stephan P, Scellato G, Franzoni C. International competition for PhDs and postdoctoral scholars: What does (and does not) matter. In: Kerr WR, Lerner J, Stern S, editors. Innovation policy and the economy. Vol. 15. Chicago, IL: University of Chicago Press; 2015. p. 73–113.

57. South African National Research Foundation (NRF). Freestanding, Innovation and Scarce Skills postdoctoral fellowships: Framework. Pretoria: NRF; 2021. Available from: https://www.up.ac.za/media/shared/748/fiss-postdoc-framework-2022-final.zp199160.pdf

58. South African Department of Science and Technology (DST). Annual report 2018–2019. Pretoria: DST; 2019. Available from: https://www.gov.za/sites/default/files/gcis_document/202002/dst-annual-report1.pdf

59. Gaughan M, Bozeman B. Institutionalized inequity in the USA: The case of postdoctoral researchers. Sci Public Policy. 2019;46:358–368. https://doi.org/10.1093/scipol/scy063

AUTHORS:
Jason Bantjes[1,2]
Sophia du Plessis[3]
Ada Jansen[3]
Philip Slabbert[4]

AFFILIATIONS:
[1]Mental Health, Alcohol, Substance Use and Tobacco (MAST) Research Unit, South African Medical Research Council, Cape Town, South Africa
[2]Department of Psychiatry and Mental Health, University of Cape Town, Cape Town, South Africa
[3]Department of Economics, Stellenbosch University, Stellenbosch, South Africa
[4]Clinical Psychologist, Tara Hospital, Johannesburg, South Africa

CORRESPONDENCE TO:
Sophia du Plessis

EMAIL:
sophia@sun.ac.za

Pedestrian safety: Motorists' attitudes to the law and driving practices in South Africa

In Africa, 40% of traffic fatalities are pedestrians – the highest proportion globally. Yet little is known about driver characteristics that are associated with unsafe driving in African countries. We aimed to explore associations between driving practices that endanger pedestrian safety and motorists' attitudes to the law (i.e. lawlessness and normlessness), controlling for sociodemographic and personality factors. We used the Response Time Method, based on Russell Fazio's attitude paradigm, to collect information about driver behaviours, attitudes, and personality traits among a sample of 440 motorists. Male gender was associated with unsafe driving, even when controlling for the effects of personality and attitudes to the law. Unsafe driving was also associated with four dimensions of motorists' personality, namely aggression, impulsivity, risk tolerance, and altruism, even when controlling for sociodemographic factors. Lawlessness (defined as a general disregard for the law) is also an important determinant of unsafe driving, even when controlling for sociodemographic and personality factors. These findings suggest that efforts to improve pedestrian safety in South Africa should focus on changing motorists' attitudes to the law.

Significance:

This study addresses pedestrian safety in the context of South Africa. The fact that 40% of traffic fatalities are pedestrians highlights the urgent need to understand the factors contributing to unsafe driving practices. The study delves into uncharted territory by examining driver characteristics associated with unsafe driving. Through exploring associations between driving practices that endanger pedestrian safety and motorists' attitudes to the law, this article provides valuable insights that can inform targeted interventions.

Introduction

Globally, road traffic crashes account for approximately 1.35 million deaths annually, and up to 93% of these occur in low- and middle-income countries (LMICs).[1] Furthermore, 23% of all traffic fatalities are pedestrians, with the figure increasing to 40% in Africa – the highest proportion globally.[2] The high rates of pedestrian deaths in Africa are in part a function of unsafe roads and pedestrian walkways and the high reliance on walking as a primary means of mobility in the region. Pedestrians are vulnerable to mortality and morbidity because they are directly exposed to traffic with little to protect them in the event of a collision. The World Health Organization (WHO)[2] estimates that, globally, 88% of pedestrians traverse unsafe roads, which contributes directly to pedestrian injury and death. Aside from being a significant public health issue, improving the safety of pedestrians is also a social justice issue, given that social status and income are major determinants of pedestrian injuries and deaths.[2] Promoting pedestrian safety is also essential from an environmental perspective, because campaigns to slow climate change by reducing vehicle emissions and encouraging walking will be feasible only if roads are more hospitable to pedestrians.[3,4]

Road safety is a complex, multi-faceted issue; nonetheless, road user behaviours, such as violating traffic laws, remain one of the most significant determinants of pedestrian safety.[5,6] Despite the clear contribution of driver behaviour to pedestrian safety, this aspect of transportation safety remains under-researched, especially in LMICs.[5,6] Personality factors and personal attitudes are key determinants of driver behaviour and are an important focus of study, especially given that driver attitudes are potentially modifiable factors. Motorists' attitudes towards the rule of law and the officials tasked with enforcing these laws may also be an important determinant of driver safety, particularly in countries like South Africa, where many motorists act with impunity and disregard for the law.[7] Within this context, we investigated associations between driving practices that endanger pedestrian safety and motorists' attitudes to the law in South Africa, controlling for sociodemographic and personality factors.

South Africa, a middle-income country, has some of the poorest road safety outcomes in Africa despite having a comprehensive set of traffic laws by global standards.[8] Jaywalking and speeding are the primary cause of death on South African roads, collectively accounting for 56.44% of traffic-associated fatalities.[9] Pedestrian fatalities, which account for 38% of all traffic-related deaths in South Africa, are related to the lack of pedestrian-orientated infrastructure, inadequate public transport, and a large number of informal settlements adjacent to busy roads in urban areas.[10] Enforcing traffic laws in South Africa is an ongoing challenge, partly because of widespread disregard for the rule of law and many citizens' perception of the illegitimacy of law enforcement officials.[7,11] Understanding how attitudes towards the law influence safe driving could have important implications for improving pedestrian safety in South Africa.

It is well established that road safety behaviours are a function of sociodemographic and personal factors such as income, age, gender, inattentiveness, lack of concentration, driving under the influence of either alcohol or drugs, risk-taking behaviour, and not having a licence.[12] The risk of a driver colliding with a pedestrian increases for both younger and much older drivers.[13] In high-income countries, young drivers from wealthy households are more likely to speed and violate the speed limit compared to drivers of the same age from low-income homes[12], suggesting an interaction between speeding and sociodemographic factors. Although there is a large body of evidence suggesting that men are more prone to unsafe driving than women[14,15], for instance, the WHO argues that about three-quarters of all road traffic deaths occur among young male individuals under the age of 25 years, and in Africa in particular, nearly twice as many male individuals aged 15–59 die from injury-related causes (road traffic accidents, violence, and others) than those who die from tuberculosis (20% vs 10%). Some scholars have contested this figure, as their findings

indicate no significant gender differences, especially when controlling for other confounding factors.[12,13] Frequent exposure to media representations of unsafe driving, such as playing computer games that entail speeding or watching high-speed driving in action movies, has also been shown to have an adverse effect on driving behaviour.[12,16] As might be expected, drivers who do not have valid licences are more likely to exhibit unsafe driving than licensed drivers. Indeed, unlicensed drivers are 173% more likely to cause a collision with a pedestrian than a driver with a licence.[13]

Driving styles are partly determined by personality traits, namely aggression, risk-taking, impulsivity and altruism. Drivers who score high on personality tests of aggression are more likely to have a history of traffic crashes than those who score lower[17], and are less inclined to yield for pedestrians.[15] Motorists also exhibit a pronounced tendency to justify their own unsafe driving behaviours by asserting that they were provoked by other aggressive motorists.[17,18] Drivers who report high levels of trait aggression are more inclined than other drivers to feel irritated by pedestrian jaywalkers and fail to give way to them.[19] Risk tolerance (i.e. the opposite of risk aversion) is also strongly associated with unsafe driving behaviours and low levels of empathy for other road users.[14,20,21]

Unsafe driving is also positively correlated with trait impulsivity (i.e. the dimension of personality that measures an individual's tendency to act impulsively and have difficulty inhibiting behaviour).[22,23] One systematic review of associations between impulsivity and driving styles concluded that trait impulsivity is also associated with motorists' tendency to express anger, drive aggressively, speed, and violate traffic laws.[23] Indeed, promoting motorists' impulse control and their ability to regulate anger are important strategies for curbing unsafe driving, particularly among young drivers.[22–24] Impulsivity is also associated with poor self-control and an inability to anticipate consequences, which increase motorists' propensity to drink and drive, a leading cause of vehicle collisions with pedestrians.[13,23] Altruistic traits (i.e. a person's tendency to show concern for others' well-being, being cooperative and kind-hearted) are negatively correlated with deviant and dangerous driving behaviours.[12,25,26]

Some scholars have noted that personality traits are generally weak predictors of unsafe driving, suggesting that adherence to social norms and morality play a more significant role in shaping driving styles.[6,27] Indeed, motorists appear more willing to yield for children and elderly pedestrians, suggesting that driving behaviour is related to motorists' perceptions of morality and social norms. Hashemiparast et al.[28] found that normative beliefs of non-compliance contributed to lawlessness. Acemoglu and Jackson[29] suggest that this occurs because numerous individuals are disobeying the law, leading each individual to expect minimal whistleblowing, resulting in lawlessness. Robinson[30] postulates that social forces (norms of justice) are internalised by group members, essentially regulating their behaviour and thus promoting social cooperation (irrespective of whether the cooperation payoff is known or unknown). Sinclair[31] refers to this phenomenon in South Africa, where people who believe themselves to be law-abiding become part of the majority group that does not adhere to formal rules. Acemoglu and Jackson raise a related point by showing that unreasonably severe laws may, in some instances, encourage society to transgress.[29] Jovanović et al.[32] compared drivers in Serbia to those in Northern Kosovo. Their results show that a lack of enforcement affects almost every other type of behaviour (speeding, the wearing of seat belts, driving under the influence, as well as ordinary violations that involve risky behaviour).

Furthermore, unsafe driving behaviour has consistently been associated with both normlessness (i.e. a state in which social norms have been eroded)[6,33,34] and lawlessness (i.e. disregard for the law)[12,34]. Understanding how attitudes towards the law influence driving styles could have policy implications for improving pedestrian safety, especially in countries like South Africa, where high rates of criminality may create permissive social norms.[31]

Methods

Our aim was to explore associations between driving practices that endanger pedestrian safety and motorists' attitudes to the law (i.e. lawlessness and normlessness), controlling for sociodemographic and personality factors among a sample of postgraduate students and staff

members (including academic, support and administrative staff) from a university in the Western Cape, South Africa. We used the Response Time Method (RTM), based on Russell Fazio's attitude paradigm[35], to collect information about driver behaviours, attitudes, and personality traits. This method helps overcome the limitations of commonly used self-report survey methods that are prone to social desirability bias. In effect, the RTM measures the instinctive reactions of respondents while minimising the influence of cognitive biases and distortions that typically occur in traditional explicit measures such as self-report surveys. The method has mostly been used in the field of marketing, where the measurement of attitudes is important to predict future sales.[36] The RTM minimises the biases typical of declarative questionnaires, such as social pressure or the so-called 'sponsor effect', because it assesses the level of hesitation when providing a given answer. The strong values are expressed with high confidence and indicated by faster response times, whereas slower response times indicate weaker, less accessible attitudes expressed with hesitation. This approach assesses both the response time and the declarative answer, thus supplying both implicit and explicit data. These two scores are merged into a single number by multiplying the declarative with the response time score, which results in a confidence index. This confidence index provides a measure of how honest a participant has been in their response.

Ethical consideration

The study received ethical clearance from the Research Ethics Committee for Social, Behavioural and Education Research (Stellenbosch University). We also obtained institutional permission from Stellenbosch University (Project number: 16765). Participants provided informed consent electronically prior to data collection. All data were collected anonymously and were securely stored on a password-protected web-based server.

Sample

Emails were sent to all postgraduate students and all administrative, academic and support staff of Stellenbosch University. Half of the emails were linked to the test for drivers and the other half to the test for pedestrians. These were distributed evenly between staff and postgraduate students. The demographic questions we asked related to age, gender, level of education, employment status, and annual family income. The self-selected sample of participants ($n = 440$) consisted predominantly of individuals self-identifying as female (58.3%), employed individuals (54.5%), and individuals with a tertiary level of education (92.0%). Only a few individuals did not possess a university degree. From this, we can infer that slightly more than half of the sample were staff, and the rest were students. Participation was incentivised through a lucky draw that offered five participants each a ZAR500 gift voucher to an online shop.

Procedure

Participants were invited via email to complete an online response time survey using the iCode platform. Informed consent was obtained electronically prior to data collection. The online assessment took 10 minutes to complete and consisted of three parts: (1) collection of sociodemographic information (age, gender, level of education, employment status and annual family income); (2) a calibration phase in which respondents were familiarised with how the survey worked and asked to complete a series of tasks that assessed whether they could read and follow instructions and understand which buttons to press to respond to survey items; and (3) the survey phase in which a set of 50 statements were presented one-by-one in a randomised order.

Measures

We used the iCode platform to present participants with statements and asked them to indicate whether they agreed or disagreed with each statement. We measured participants' response time and the resultant confidence index for each of the statements. At the start of the survey, participants were presented with a series of 'warm-up' statements that were used for individual calibration (i.e. to create a personal baseline which took into account individual differences, such as familiarity with computers, age, fatigue, length of statements, speed, and stability of Internet connection and (in)voluntary carelessness of respondents). The survey statements are indicated in the Appendix.

Unsafe driving practices

Participants were asked to agree or disagree with seven statements that assessed unsafe driving and practices that endanger pedestrians. Participants' responses to these items were aggregated to yield an 'unsafe driving index' (a value from 0 to 7, with higher scores indicating unsafe driving that endangers pedestrians).

Personality

Participants were asked to agree or disagree with 35 statements assessing four dimensions of personality, namely aggression, altruism, impulsivity, and risk avoidance. We assessed impulsivity with the brief eight-item version of the Barratt Impulsiveness Scale.[37]. To assess aggression we used the eight-item Aggressive Beliefs and Attitudes Scale.[38] To assess normlessness we used statements derived from Kohn and Schooler's normlessness scale[39], which has been used in previous studies of driver behaviour[40,41]. We assessed risk aversion and risk perception with nine items adapted from the Perceived Risk of Risky Driving Behaviours Scale, which has been used in other studies of road safety and driver behaviour.[42,43] Finally, to assess altruism, we used the items from the Altruism Subscales of the International Personality Item Pool (IPIP).[44]

Attitudes to the law

Participants were asked to agree or disagree with eight statements assessing lawlessness (i.e. a general disregard for the law) and five statements assessing normlessness (i.e. a breakdown of social norms regulating individual conduct). We derived these statements ourselves, based on our previous qualitative and quantitative research in this area.

Data analysis

Data were cleaned and checked, and the statistical program STATA was used for the analysis. Descriptive statistics were used to provide a description of the sample characteristics. In the first step of the analysis, we calculated Cronbach's alpha to assess the internal consistency of the items included for each dimension of personality and for the subset of statements assessing attitudes to the law. Subsequently, we used simple and multiple regression analyses to identify associations between unsafe driving and sociodemographic variables, personality variables and attitudes to the law. Second, we used multivariate regression to explore associations between lawlessness and unsafe driving, controlling for sociodemographic and personality factors identified as significant in the preceding analysis. In the final step of the analysis, we estimated a multivariate regression model of factors associated with unsafe driving, including only the variables that were significant in the preceding model. The results of all regression analyses are presented as beta coefficients (with associated 95% confidence intervals). For all analyses, the level of significance (alpha) was set as 0.05.

Results

The sample consisted predominantly of individuals self-identifying as female (58.3%), employed individuals (54.5%), and individuals with a tertiary level of education (92.0%). The estimated mean score of the unsafe driving index is 3.1 ± 0.07 (s.d.=1.5, range=0–7).

Sociodemographic factors associated with unsafe driving

In the first step of the analysis, we used simple and multiple regression analysis to identify sociodemographic factors associated with unsafe driving practices (Table 1). The sociodemographic factors are categorical variables (which include education, gender, income, age, and employment status). In the bivariate analysis, unsafe driving was positively associated with male gender ($\beta=0.5$, $p<0.001$) but not with any of the other sociodemographic variables. In the multivariate analysis, controlling for the effects of all sociodemographic variables, unsafe driving was positively associated with male gender ($\beta=0.6$, $p<0.001$) and inversely associated with higher levels of education ($\beta=-0.6$, $p<0.005$).

Personality factors

We calculated Cronbach's alpha for the set of statements in each dimension of personality assessed (i.e. aggression, altruism, impulsivity, and risk aversion) to assess the internal consistency of each construct. The poor internal consistency of the items in each dimension indicated that the included questions were not measuring the same thing, suggesting that each item assessed a different component of the underlying dimension of personality. Consequently, rather than considering all statements for each dimension collectively, we instead considered each statement individually within each personality dimension. Simple and multiple regression analyses were used to explore associations between driving practices that endanger pedestrians and each statement within each of the four dimensions of motorists' personality (Table 2). For each personality dimension, we identified the best-fitting combination of statements associated with unsafe driving. In terms of aggression, unsafe driving was associated with "I sound my horn to indicate my annoyance to another road user" ($\beta=0.4$, $p<0.05$) and "Sometimes passengers tell me to calm down because I am angry at other drivers" ($\beta=0.4$, $p<0.05$), and "Some people are just bad people" ($\beta=0.5$, $p<0.05$). For altruism, unsafe driving was associated with the statement "I look down on others" ($\beta=0.5$, $p<0.01$) and for impulsivity, unsafe driving was inversely associated with "I am a careful thinker" ($\beta=-0.3$, $p<0.05$) and "I concentrate easily" ($\beta=-0.5$, $p<0.01$). Unsafe driving was also associated with risk-taking ("I drive a vehicle while feeling tired or fatigued"; $\beta=0.5$, $p<0.01$).

Attitudes to the law

Simple and multiple regression analyses were used to explore associations between driving practices that endanger pedestrians and attitudes to the law, namely lawlessness (i.e. a general disregard for the law) and normlessness (i.e. a breakdown of social norms regulating individual conduct). The link between social norms and people's attitudes to the law (including institutions of the state responsible for enforcing these laws), implies that simply changing the laws and/or increasing the penalties for transgressions is unlikely to make roads safer if people's attitudes to the laws are not also addressed.

As seen in Table 3, unsafe driving was associated with endorsing the following statements reflecting lawlessness: "I claim government benefits to which I am not entitled" ($\beta=0.7$, $p<0.01$); "I make illegal U-turns at intersections" ($\beta=0.4$, $p<0.01$); "I sometimes accept bribes in the course of my duties" ($\beta=0.4$, $p<0.05$); "I sometimes

Table 1: Sociodemographic factors associated with unsafe driving practices

	Univariate analysis		Multivariate analysis	
	Beta	**95% Confidence interval**	**Beta**	**95% Confidence interval**
Education	-0.4	-1.0–0.2	-0.6**	-1.2–0.1
Age	0.0	-0.1–0.2	0.1	-0.0–0.3
Male gender	0.5***	0.2–0.8	0.6***	0.3–0.9
Income	-0.0	-0.0–0.0	-0.0	-0.0–0.0
Employment status	-0.1	-0.4–0.2	-0.2	-0.6–0.2

***p<0.01, **p<0.05

Table 2: Motorists' personality factors associated with unsafe driving practices

		Univariate analysis		Multivariate analysis	
		Beta	95% Confidence interval	Beta	95% Confidence interval
Aggression	"Getting back at others makes me feel better"	0.1	-0.2–0.4	0.1	-0.2–0.4
	"I feel the need to get even if someone disrespects me"	-0.0	-0.3–0.2	-0.2	-0.5–0.2
	"I have the right to retaliate if I am betrayed"	0.2	-0.1–0.5	0.1	-0.2–0.4
	"I sound my horn to indicate my annoyance to another road user"	0.4**	0.1–0.7	0.4**	0.0–0.7
	"Large corporations exploit their employees"	-0.0	-0.4–0.3	-0.2	-0.5–0.2
	"Some people are just bad people"	0.6***	-1.0–0.3	0.5**	-1.0–0.1
	"Some people are simply horrible human beings"	-0.3	-0.6–0.1	-0.2	-0.6–0.2
	"Sometimes passengers tell me to calm down because I am angry at other drivers"	0.5***	0.2–0.8	0.4**	0.0–0.7
	"The rich get richer by taking advantage of the poor"	0.2	-0.2–0.5	0.1	-0.3–0.5
	"The wealthy capitalise on those who are less fortunate"	0.1	-0.3–0.4	0.2	-0.3–0.6
Altruism	"I am concerned about others"	-0.2	-0.5–0.2	-0.1	-0.4–0.3
	"I am indifferent to the feelings of others"	0.0	-0.3–0.3	0.0	-0.3–0.3
	"I anticipate the needs of others"	-0.0	-0.3–0.3	-0.0	-0.3–0.3
	"I have a good word for everyone"	-0.1	-0.4–0.3	0.1	-0.2–0.4
	"I look down on others"	0.6***	-0.9–0.4	0.5***	-0.8–0.2
	"I love to help others"	-0.4***	-0.7–0.1	-0.2	-0.5–0.1
	"I make people feel uncomfortable"	-0.1	-0.4–0.2	-0.0	-0.3–0.3
	"I make people feel welcome"	-0.1	-0.4–0.2	-0.0	-0.4–0.3
	"I make time for others"	-0.4***	-0.7–-0.2	-0.3	-0.5–0.0
	"I turn my back on others"	-0.3**	-0.6–0.1	-0.1	-0.4–0.2
Impulsivity	"I act on the spur of the moment"	0.1	-0.2–0.4	0.1	-0.2–0.3
	"I am a careful thinker"	-0.4***	-0.7–0.2	-0.3**	-0.6–0.0
	"I am self-controlled"	-0.2	-0.5–0.1	0.0	-0.3–0.3
	"I concentrate easily"	-0.6***	-0.9–0.3	-0.5***	-0.8–0.2
	"I do things without thinking"	0.1	-0.1–0.4	0.1	-0.2–0.4
	"I plan tasks carefully"	-0.3**	-0.6–0.1	-0.2	-0.5–0.1
	"I say things without thinking"	0.1	-0.2–0.4	0.1	-0.2–0.4
	"I tend to be absent-minded"	0.2	-0.1–0.5	0.2	-0.1–0.5
Risk aversion	"I sometimes follow another vehicle closely"	0.3**	0.1–0.6	0.2	-0.1–0.5
	"I drive a vehicle soon after drinking alcohol, but within the legal limit"	0.4**	0.1–0.7	0.2	-0.2–0.5
	"I drive a vehicle while distracted"	0.0	-0.3–0.3	0.1	-0.2–0.4
	"I drive a vehicle while feeling tired or fatigued"	0.6***	0.3–0.9	0.5***	0.1–0.8
	"I drive a vehicle while using a hands-free mobile phone"	0.4*	-0.0–0.7	0.3	-0.1–0.7
	"I often break speed limits"	0.1	-0.2–0.4	0.1	-0.2–0.4
	"I perform illegal driving manoeuvres"	-0.2	-0.4–0.1	-0.2	-0.5–0.1
	"Sometimes I overtake a vehicle by crossing double white lines"	0.4**	0.0–0.7	0.2	-0.2–0.6

***p<0.01, **p<0.05, *p<0.1

Table 3: Attitudes to the law associated with unsafe driving practices

		Univariate analysis		Multivariate analysis	
		Beta	95% Confidence interval	Beta	95% Confidence interval
Lawlessness	"I avoid paying fares on public transport"	0.1	-0.2–0.4	-0.0	-0.3–0.3
	"I cheat on my taxes if I have a chance"	0.0	-0.3–0.3	0.1	-0.2–0.4
	"I claim government benefits to which I am not entitled"	0.7***	0.4–1.0	0.7***	0.4–0.9
	"I make illegal U-turns at intersections"	0.3**	0.0–0.6	0.4***	0.2–0.7
	"I signal when I make lane changes"	0.1	-0.2–0.4	0.1	-0.2–0.4
	"I sometimes accept bribes in the course of my duties"	0.5***	0.2–0.8	0.4**	0.1–0.7
	"I sometimes take pencils from work for private use"	0.4***	0.1–0.6	0.3**	0.1–0.6
	"I speed up in order to make it through yellow lights"	0.6***	0.3–0.9	0.6***	0.3–0.9
Normlessness	"If something works, it is less important whether it is right or wrong"	0.2	-0.6–0.9	0.1	-0.6–0.9
	"It is OK to get around laws and rules as long as nobody is aware of it"	0.2	-0.2–0.7	0.2	-0.3–0.6
	"Some things can be wrong to do even though they are legal"	0.3	-0.6–1.2	0.2	-1.1–1.5

*** $p < 0.01$, ** $p < 0.05$

take pencils from work for private use" ($\beta = 0.3$, $p < 0.05$); and "I speed up in order to make it through yellow lights" ($\beta = 0.6$, $p < 0.01$). None of the statements about normlessness was associated with unsafe driving.

In the next step of the analysis, we used multivariate regression to explore associations between lawlessness and unsafe driving, controlling for sociodemographic and personality factors identified as significant in the preceding analysis (Table 4). In this model, unsafe driving was associated with: male gender ($\beta = 0.4$, $p < 0.01$); aggression ("Some people are just bad people", $\beta = 0.4$, $p < 0.05$); altruism ("I look down on others", $\beta = 0.5$, $p < 0.01$); and five statements asserting lawlessness ("I claim government benefits to which I am not entitled", $\beta = 0.5$, $p < 0.01$; "I make illegal U-turns at intersections", $\beta = 0.5$, $p < 0.01$; "I sometimes accept bribes in the course of my duties", $\beta = 0.3$, $p < 0.05$; "I sometimes take pencils at work for private use", $\beta = 0.3$, $p < 0.05$; and "I speed up in order to make it through yellow lights", $\beta = 0.5$, $p < 0.01$). No associations were observed between unsafe driving and impulsivity or risk aversion in this multivariate model.

In the final step of the analysis, we estimated a multivariate regression model of factors associated with unsafe driving, including only the variables that were significant in the preceding model (Table 5). In this model, unsafe driving was associated with male gender ($\beta = 0.5$, $p < 0.01$), aggression ("Some people are just bad people", $\beta = 0.4$, $p < 0.05$), altruism ("I look down on others", $\beta = 0.5$, $p < 0.01$), and all five statements affirming lawlessness ("I claim government benefits to which I am not entitled", $\beta = 0.6$, $p < 0.01$; "I make illegal U-turns at intersections", $\beta = 0.5$, $p < 0.01$; "I sometimes accept bribes in the course of my duties", $\beta = 0.3$, $p < 0.05$; "I sometimes take pencils at work for private use", $\beta = 0.4$, $p < 0.01$; "I speed up in order to make it through yellow lights", $\beta = 0.6$, $p < 0.01$). The model was significant ($p < 0.05$) and accounted for 18.4% of the variance in unsafe driving (adjusted $R^2 = 0.184$).

Discussion

This study is, to our knowledge, the first paper of its kind to use the Response Time Method to analyse driver behaviours that endanger pedestrians. Our findings are congruent with prior empirical research on the topic, strengthening arguments in favour of applying this technique in future road safety research, particularly in studies in which honest responses from participants cannot be reliably solicited using conventional survey methods. Our findings support previous research that unsafe driving behaviours are a function of sociodemographic and personality factors. Crucially, our data show that lawlessness (defined as a general disregard for the law) is also an important determinant of unsafe driving, even when controlling for sociodemographic and personality factors. As this research only focused on the attitudes of drivers, it cannot say anything, for instance, about the role of infrastructure in crashes.

Our data show that male drivers in South Africa are more likely than female drivers to engage in unsafe driving that endangers pedestrians, as previously reported in other studies on driving behaviour.[14,15,45] We found that male gender was associated with unsafe driving, even when controlling for the effects of personality and attitudes to the law, which appears to be a novel finding. Although level of education was associated with unsafe driving in multivariate regression analysis, after controlling for the effects of other sociodemographic variables, this association was not significant when considering personality factors and attitudes to the law. Furthermore, our data indicate that driving behaviours that endanger pedestrians are associated with four dimensions of motorists' personality, namely aggression, impulsivity, risk tolerance, and altruism, even when controlling for sociodemographic factors. These findings support the assertion that personality is a determinant of driving style in South Africa, as has been found in high-income Western countries.[15,19,20,22,25]

The main result from our analysis, however, points to lawlessness as a significant contributor to driving behaviours that endanger pedestrians, even when controlling for sociodemographic and personality factors. Indeed, our findings suggest that a culture of lawlessness may represent a significant risk for the safety of pedestrians in South Africa. Furthermore, our data indicate that lawlessness may be more important than normlessness, even though it is well established that social norms exert a strong influence on driver behaviours.[46] If one considers the relationship between norms and laws, three possibilities may arise: social norms and formal laws support one another; the two are in conflict; or there is some level of absence of formal laws that results in social norms governing behaviour. Du Plessis et al.[47] have shown that South Africa has a good set of traffic laws compared to other countries, but the incidences of traffic-related deaths are still relatively high. The authors argue that this divergence may be explained by social norms that counter formal laws. Similarly, Acemoglu and Jackson[29] suggest

Table 4: Multivariate logistic regression analysis of associations between lawlessness and unsafe driving, controlling for sociodemographic and personality factors

		Beta	95% Confidence interval
Sociodemographic factors	Education	-0.3	-0.7–0.2
	Male gender	0.4[1]	0.2–0.7
Aggression	"I sound my horn to indicate my annoyance to another road user"	0.1	-0.2–0.4
	"Some people are just bad people"	0.4**	0.0–0.7
	"Sometimes passengers tell me to calm down because I am angry at other drivers"	0.2	-0.1–0.5
Altruism	"I look down on others"	0.5[1]	0.2–0.7
Impulsivity	"I am a careful thinker"	-0.2	-0.5–0.1
	"I concentrate easily"	-0.3	-0.6–0.0
Risk aversion	"I drive a vehicle while feeling tired or fatigued"	0.3	-0.0–0.6
Lawlessness	"I claim government benefits to which I am not entitled"	0.5[1]	0.2–0.7
	"I make illegal U-turns at intersections"	0.5[1]	0.2–0.7
	"I sometimes accept bribes in the course of my duties"	0.3**	0.0–0.6
	"I sometimes take pencils from work for private use"	0.3**	0.0–0.6
	"I speed up in order to make it through yellow lights"	0.5[1]	0.2–0.8

$aR^2=0.211$, $p<0.05$

***$p<0.01$, **$p<0.05$

Table 5: Multivariate logistic regression analysis of associations between lawlessness and unsafe driving, controlling for sociodemographic and personality factors

		Beta	95% Confidence interval
Sociodemographic factors	Male gender	0.5***	0.2–0.8
Aggression	"Some people are just bad people"	0.4**	0.0–0.8
Altruism	"I look down on others"	0.5***	0.3–0.8
Lawlessness	"I claim government benefits to which I am not entitled"	0.6***	0.3–0.8
	"I make illegal U-turns at intersections"	0.5***	0.2–0.7
	"I sometimes accept bribes in the course of my duties"	0.3**	0.0–0.6
	"I sometimes take pencils from work for private use"	0.4***	0.1–0.6
	"I speed up in order to make it through yellow lights"	0.6***	0.3–0.9

***$p<0.01$, **$p<0.05$

lawlessness may increase when dominant societal norms are in conflict with laws, resulting in a converse effect.

Hashemiparast et al.[28] found participants' pessimism and distrust of authorities contributed to their proclivity to defy the rules and regulations of the road, disobey the law and engage in unsafe, unlawful driving practices. For example, Sinclair attributes normlessness in South Africa to widespread corruption in government.[31] More recently, Murphy et al.[48] and de Bruijn[49] found that compliance with COVID regulations was not significantly affected by the risk of legal sanction. This suggests that compliance cannot simply be enforced by authorities. Instead, normative concerns were shown to be central to promoting individuals' sense of duty to support and obey laws.[48,49] Murphy et al.[48] recommend persuading people of their sense of duty and responsibility so that they engage in actions that protect others (especially the most vulnerable) and remind them of their duty. De Bruijn et al.[49] suggest three factors play a significant role in COVID regulation compliance: morality, perception and the situation. First, they found a positive correlation between compliance and moral support of the law. Further,

when the obligation to comply was higher, so was lawfulness (moral support compliance). Second, when laws and their enforcement are perceived as just and reasonable, the probability of compliance increases. Finally, compliance increased when the environment favoured it and decreased when the environment favoured violation. Our findings in this study support the argument that the link between social norms and laws must be considered if traffic laws are to be successfully implemented, despite the possible influence of other factors (such as inadequate law enforcement and accountability of law enforcement to uphold the law).

The study has several limitations, including the small sample size, the use of a self-selected (non-probability) sample from one geographical region, and the use of survey items to assess aspects of personality that have not been validated for use in this population. The sample also consisted of university students and staff, which limits the generalisability of findings. Nonetheless, the study is a first step towards conducting similar studies in samples that are more representative of South African motorists and pedestrians. Replicating the study with more representative

probability samples could have important implications for public health measures to improve pedestrian safety in South Africa, the most pertinent of which is that simply imposing stricter rules and/or higher penalties for unsafe driving is unlikely to have any positive impact unless attitudes of lawlessness are addressed. Importantly, this study shows how novel research methods could be used to improve our understanding of factors that influence motorists' and pedestrians' safety behaviours.

Conclusion

Our study is the first of its kind in South Africa to investigate associations between driving behaviours that endanger pedestrians, personality factors, and attitudes to the law. Our data indicate that while some sociodemographic and personality factors may influence driving style, lawlessness is an important determinant. The significance of this finding is that lawlessness, unlike personality and sociodemographic factors, is a potentially modifiable risk factor for unsafe driving. To the extent that our findings are generalisable, they suggest that current traffic laws are not going to achieve improved outcomes without further attempts to improve enforcement or change motorists' attitudes to the law. Future research in South Africa could employ behavioural experiments and controlled trials to give policy directives on how to change motorists' attitudes to the law to promote pedestrian safety. Our study should also be replicated in larger, more representative probability samples.

Acknowledgements

Competing interests

We have no competing interests to declare.

Authors' contributions

J.B.: Conceptualisation; methodology; formal analysis; writing – review and editing; funding acquisition. S.d.P.: Conceptualisation; methodology; software; writing – review and editing; project administration; funding acquisition. A.J.: Conceptualisation; methodology; formal analysis; writing – review and editing. P.S.: Writing – original draft preparation; writing – review and editing; project administration.

References

1. World Health Organization. Road traffic injuries [webpage on the Internet]. No date [cited 2021 Apr 23]. Available from: https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries

2. World Health Organization (WHO). Global status report on road safety 2018. Geneva: WHO; 2018. Available from: https://www.who.int/publications/i/item/9789241565684

3. Chapman L. Transport and climate change: A review. J Transp Geogr. 2007;15:354–367. https://doi.org/10.1016/j.jtrangeo.2006.11.008

4. Stanley JK, Hensher DA, Loader C. Road transport and climate change: Stepping off the greenhouse gas. Transp Res Part A Policy Pract. 2011;45:1020–1030. https://doi.org/10.1016/j.tra.2009.04.005

5. Cinnamon J, Schuurman N, Hameed SM. Pedestrian injury and human behaviour: Observing road-rule violations at high-incident intersections. PLoS ONE. 2011;6, e21063. https://doi.org/10.1371/journal.pone.0021063

6. Disassa A, Kebu H. Psychosocial factors as predictors of risky driving behavior and accident involvement among drivers in Oromia region, Ethiopia. Heliyon. 2019;5:1–11. https://doi.org/10.1016/j.heliyon.2019.e01876

7. Adeniji AA, Mabuza LH, Titus E. Magnitude, trends and prevention of road traffic accidents in the Republic of South Africa. S Afr Fam Pract. 2020;62, Art. #5032. https://doi.org/10.4102/safp.v62i1.5032

8. Du Plessis S, Jansen A, Siebrits K. The limits of laws: Traffic law enforcement in South Africa. S Afr J Econ Manag Sci. 2019;23, Art. #3430. https://doi.org/10.4102/sajems.v23i1.3430

9. Road Traffic Management Corporation. TOS – Traffic Offence Survey [document on the Internet]. c2015 [cited 2021 Sep 17]. Available from: https://www.rtmc.co.za/images/rtmc/docs/research_dev_rep/31%20March%202015.%20TOS%20%E2%80%93%20Traffic%20Offence%20Survey.pdf

10. Road Traffic Management Corporation. National Road Safety Strategy 2016–2030 [document on the Internet]. c2016 [cited 2021 Sep 17]. Available from: https://www.westerncape.gov.za/assets/departments/transport-public-works/Documents/2017_05_18_strategic_plans_national_road_safety_strategy_2016_to_2030_approved.pdf

11. Bantjes J, Du Plessis S, Jansen A, Siebrits K, Slabbert P. Motorists' perceptions of factors that influence payment of speeding fines in Cape Town, South Africa: Application of the theory of planned behaviour. S Afr J Psychol. 2021;52(1):48–60. https://doi.org/10.1177/00812463211011746

12. Kunnawee K, Ketphat M, Jiwattanakulpaisarn P. Application of the theory of planned behaviour to predict young drivers' speeding behaviour. J East Asia Soc Transp Stud. 2013;10:2031–2048. https://doi.org/10.1136/injuryprev-2012-040590t.5

13. Jiménez-Mejías E, Martínez-Ruiz V, Amezcua-Prieto C, Olmedo-Requena R, Luna-Del-Castillo JDD, Lardelli-Claret P. Pedestrian- and driver-related factors associated with the risk of causing collisions involving pedestrians in Spain. Accid Anal Prev. 2016;92:211–218. https://doi.org/10.1016/j.aap.2016.03.021

14. Gaymard S, Tiplica T. Conditional respect towards the pedestrian: Difference between men and women and risk modeling by the Bayesian approach. Qual Quant. 2014;48:91–110. https://doi.org/10.1007/s11135-012-9751-y

15. Mphele SBM, Selemogwe MM, Kote M, Balogun SK. Who owns the road? Exploring driver and pedestrian behaviour at zebra / pedestrian crossings in Gaborone Botswana. Br J Arts Soc Sci. 2013;13:121–130.

16. Beullens K, Van den Bulck J. News, music videos and action movie exposure and adolescents' intentions to take risks in traffic. Accid Anal Prev. 2008;40:349–356. https://doi.org/10.1016/j.aap.2007.07.002

17. Kaiser S, Furian G, Schlembach C. Aggressive behaviour in road traffic: Findings from Austria. Transp Res Procedia. 2016;14:4384–4392. https://doi.org/10.1016/j.trpro.2016.05.360

18. Bowen L, Budden SL, Smith AP. Factors underpinning unsafe driving: A systematic literature review of car drivers. Transp Res F: Traffic Psychol. 2020;72:184–210. https://doi.org/10.1016/j.trf.2020.04.008

19. Chai J, Zhao G. Effect of exposure to aggressive stimuli on aggressive driving behavior at pedestrian crossings at unmarked roadways. Accid Anal Prev. 2016;88:159–168. https://doi.org/10.1016/j.aap.2015.12.026

20. Şimşekoğlu Ö. How do attitudes, personality traits, and driver behaviors relate to pedestrian behaviors? A Turkish case. Traffic Inj Prev. 2015;16:84–89. https://doi.org/10.1080/15389588.2014.880785

21. Taubman-Ben-Ari O, Shay E. The association between risky driver and pedestrian behaviors: The case of ultra-orthodox Jewish road users. Transp Res F: Traffic Psychol. 2012;15:188–195. https://doi.org/10.1016/j.trf.2011.12.005

22. Bıçaksız P, Öztürk İ, Özkan T. The differential associations of functional and dysfunctional impulsivity with driving style: A simulator study. Transp Res F: Traffic Psychol. 2019;63:1–11. https://doi.org/10.1016/j.trf.2019.02.011

23. Bıçaksız P, Özkan T. Impulsivity and driver behaviors, offences and accident involvement: A systematic review. Transp Res F: Traffic Psychol. 2016;38:194–223. https://doi.org/10.1016/j.trf.2015.06.001

24. Mirón-Juárez CA, García-Hernández C, Ochoa-Ávila E, Díaz-Grijalva GR. Approaching to a structural model of impulsivity and driving anger as predictors of risk behaviors in young drivers. Transp Res F: Traffic Psychol. 2020;72:71–80. https://doi.org/10.1016/j.trf.2020.05.006

25. Al-Tit AA. The impact of drivers' personality traits on their risky driving behaviors. J Hum Behav Soc Environ. 2020;30:498–509. https://doi.org/10.1080/10911359.2019.1700866

26. Ge Y, Qu W, Jiang C, Du F, Sun X, Zhang K. The effect of stress and personality on dangerous driving behavior among Chinese drivers. Accid Anal Prev. 2014;73:34–40. https://doi.org/10.1016/j.aap.2014.07.024

27. Elliott MA, Armitage CJ, Baughan CJ. Drivers' compliance with speed limits: An application of the theory of planned behavior. J Appl Psychol. 2003;88:964–972. https://doi.org/10.1037/0021-9010.88.5.964

28. Hashemiparast M, Negarandeh R, Montazeri A. How young pedestrians do explain their risky road crossing behaviors? A qualitative study in Iran. Health Promot Perspect. 2017;7:140–144. https://doi.org/10.15171/hpp.2017.26

29. Acemoglu D, Jackson MO. Social norms and the enforcement of laws. J Eur Econ Assoc. 2017;15:245–295. https://doi.org/10.1093/jeea/jvw006

30. Robinson PH. Natural law and lawlessness: Modern lessons from pirates, lepers, Eskimos, and survivors. Univ Ill Law Rev. 2013;2013:433–506. https://doi.org/10.2139/ssrn.1990498

31. Sinclair M. Attitudes, norms and driving behaviour: A comparison of young drivers in South Africa and Sweden. Transp Res F: Traffic Psychol. 2013;20:170–181. https://doi.org/10.1016/j.trf.2013.07.001

32. Jovanović D, Stanojecić P, Stanojević D. Motives for, and attitudes about, driving-related anger and aggressive driving. Soc Behav Pers: Int J. 2011;39(6):755–764. https://doi.org/10.2224/sbp.2011.39.6.755

33. Yang J, Du F, Qu W, Gong Z, Sun X. Effects of personality on risky driving behavior and accident involvement for Chinese drivers. Traffic Inj Prev. 2013;14:565–571. https://doi.org/10.1080/15389588.2012.748903

34. Nandavar S, Lewis I, White KM. Understanding drivers' altruistic driving decisions: A theoretically guided investigation. Transp Res F: Traffic Psychol. 2019;62:212–227. https://doi.org/10.1016/j.trf.2018.12.017

35. Fazio RH. A practical guide to the use of response latency in social psychological research. In: Hendrick C, Clark MS, editors. Research methods in personality and social psychology. London: Sage; 1990. p. 74–97.

36. Ohme R, Matukin M, Wicher P. Merging explicit declarations with implicit response time to better predict behavior. In: Chkoniya V, Madsen AO, Bukhrashvili P, editors. Anthropological approaches to understanding consumption patterns and consumer behavior. Hershey, PA: IGI Global; 2020. p. 427–448. https://doi.org/10.4018/978-1-7998-3115-0.ch023

37. Steinberg L, Sharp C, Stanford MS, Tharp AT. New tricks for an old measure: The development of the Barratt Impulsiveness Scale–Brief (BIS-Brief). Psychol Assess. 2013;25:216–226. https://doi.org/10.1037/a0030550

38. Michel JS, Hartman P, Gitter S. Development and validation of a short form aggressive beliefs and attitudes scale. Pers Individ Differ. 2015;87:130–135.

39. Kohn ML, Schooler C. Job conditions and personality: A longitudinal assessment of their reciprocal effects. Am J Sociol. 1982;87:1257–1286. https://doi.org/10.1086/227593

40. Chen CF. Personality, safety attitudes and risky driving behaviors. Evidence from young Taiwanese motorcyclists. Accid Anal Prev. 2009;41:963–968.

41. Lazuras L, Rowe R, Poulter DR, Powell PA, Ypsilanti A. Impulsive and self-regulatory processes in risky driving among young people: A dual process model. Front Psychol. 2019;10:1170. https://doi.org/10.3389/fpsyg.2019.01170

42. Ivers R, Senserrick T, Boufous S, Stevenson M, Chen H-Y, Woodward M, et al. Novice drivers' risky driving behavior, risk perception, and crash risk: Findings from the DRIVE study. Am J Public Health. 2009;99:1638–1644. https://doi.org/10.2105/AJPH.2008.150367

43. Machin MA, Sankey KS. Relationships between young drivers' personality characteristics, risk perceptions, and driving behaviour. Accid Anal Prev. 2008;40:541–547. https://doi.org/10.1016/j.aap.2007.08.010

44. Goldberg LR, Johnson JA, Eber HW, Hogan R, Ashton MC, Cloninger CR, et al. The international personality item pool and the future of public-domain personality measures. J Res Pers. 2006;40(1):84–96.

45. Matzopolous R, Myers JE, Jobanputra R. Road traffic injury: Prioritising interventions. S Afr Med J. 2008;98:692–696.

46. Jiang X, Wang W, Bengler K, Guo H, Li C. Analysis of drivers' performance in response to potential collision with pedestrians at Urban crosswalks. IET Intell Transp Syst. 2017;11:546–552. https://doi.org/10.1049/iet-its.2016.0344

47. Du Plessis S, Hartig B, Jansen A, Siebrits K. Improving payment of traffic fines with financial incentives: Discounts vs. penalties. Transp Res F: Traffic Psychol. 2020;74:298–306. https://doi.org/10.1016/j.trf.2020.08.019

48. Murphy K, Williamson H, Sargeant E, McCarthy M. Why people comply with COVID-19 social distancing restrictions: Self-interest or duty? Aust N Z J Criminol. 2020;53:477–496. https://doi.org/10.1177/0004865820954484

49. De Bruijn AL, Feldman Y, Kuiper ME, Brownlee M, Reinders Folmer C, Kooistra EB, et al. Why did Israelis comply with COVID-19 mitigation measures during the initial first wave lockdown? [preprint]. PsyArXiv. 2020 August 28. https://doi.org/10.31234/osf.io/vm8x9

Appendix: Survey statements (agree/disagree)

Unsafe driving practices:

I stop at marked pedestrian crossings when people want to cross

I stop for pedestrians who want to cross a street where there are no traffic crossings

I have been involved in an accident involving a pedestrian

I slow down when I see a marked pedestrian crossing

I think jaywalking is a problem

I eat or drink while I drive

When approaching a pedestrian cross, I look for someone who would like to cross

I fail to notice when a traffic light turns green

I avoid driving at night

I talk on my hand-held cell phone when driving

I forget where I have parked my car

Aggression:

The wealthy capitalise on those who are less fortunate

Some people are just bad people

The rich get richer by taking advantage of the poor

Appendix continued...

Getting back at others makes me feel better

Large corporations exploit their employees

I have the right to retaliate if I am betrayed

I feel the need to get even if someone disrespects me

Some people are simply horrible human beings

A passenger told me to calm down because I am angry at other drivers

I sound my horn to indicate my annoyance to another road user

Altruism:

I make people feel welcome

I anticipate the needs of others

I love to help others

I am concerned about others

I have a good word for everyone

I look down on others

I am indifferent to the feelings of others

I make people feel uncomfortable

I turn my back on others

I make time for others

Impulsivity:

I plan tasks carefully

I do things without thinking

I don't 'pay attention'

I am self-controlled

I concentrate easily

I am a careful thinker

I say things without thinking

I act on the spur of the moment

Risk aversion and risk perception:

I drive a vehicle while feeling tired or fatigued

I drive a vehicle while using a hands-free mobile phone

I drive a vehicle while distracted (e.g. due to drinking, eating, smoking, changing a CD)

I drive at 70 km/h in a designated 60 km/h speed zone

I perform illegal driving manoeuvres (e.g. doughnuts, drifting)

I don't mind driving while closely following another vehicle (at a less than 2 s following distance)

Sometimes I overtake a vehicle by crossing double white lines

I drive a vehicle soon after drinking alcohol but within the legal BAC limit of 0.05

I drive at 120 km/h in a designated 100 km/h speed zone

Lawlessness:

I claim government benefits to which I am not entitled

I avoid paying fares on public transport

Appendix continued...

I take pencils from work for private use
I cheat on my taxes if I have a chance
I accept bribes in the course of my duties
I signal when I make lane changes
I make illegal U-turns at intersections
I speed up in order to make it through yellow lights
Normlessness:
It is all right to do anything you want as long as you stay out of trouble
It is OK to get around laws and rules as long as you don't break them directly
If something works, it is less important whether it is right or wrong
Some things can be wrong to do even though it is legal to do them