South African Journal *of* Science

volume 119
number 7/8

Future climate change and South African rainfall

Anticancer properties of a Mauritian sponge

3D printed phantom for planar X-ray imaging

ASSAf
ACADEMY OF SCIENCE OF SOUTH AFRICA

Volume 119
Number 7/8
July/August 2023

South African Journal of Science

eISSN: 1996-7489

Published by
the Academy of Science of South Africa (www.assaf.org.za) with financial assistance from the Department of Science & Innovation.

Design and layout
diacriTech Technologies

Correspondence and enquiries
sajs@assaf.org.za

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

Research Article

Cover caption

The Republic of Mauritius is a maritime country with a vast potential for marine organisms that are not yet fully utilised as sources of bioactive substances. The exclusive economic zone of Mauritius harbours a diverse assemblage of *Neopetrosia* species and thus opportunity for the discovery of new bioactive agents. In their article, Beesoo et al. report on the promising anticancer activity of extracts derived from the marine sponge *Neopetrosia exigua*, collected from Mauritian waters.

Sustaining excellent science

On 13 July 2023, the National Science and Technology Forum (NSTF) announced the winners in various categories for scientists in South Africa. On show, amongst both the winners and the finalists, was an impressive array of South African scholarship and innovation, with work focussed on solving pressing problems in our country, our continent, and, indeed, our planet. This event is just one opportunity for science in South Africa to be recognised and celebrated. We at the *South African Journal of Science* congratulate all the nominees and winners, and we look forward to receiving submissions from them all for consideration for publication in our journal.

Less than two months before this date, on 16 May 2023, the Progress in International Reading Literacy Study (PIRLS) Report of 2021 data was published.[1] As the South African Preliminary Highlights Report[2] demonstrates, the results for South African children are cause for very serious concern. The economist Nic Spaull[3] refers to the findings for South Africa as a 'generational catastrophe', and amongst the key summary points made by Spaull on the findings are the following (quoted directly from Spaull[3]):

- In 2021 81% of Grade 4 learners cannot read for meaning in any language, up from 78% in 2016.
- SA came last of all 57 countries [studied], with the largest decline between 2016 & 2021.
- The average Gr4 child in SA in 2021 was 80% of a year behind their counterpart in 2016.
- Northern rural provinces experienced the largest declines in reading.
- English and Afrikaans schools did not experience a decline between 2016 and 2021.
- Brazilian Grade 4s are 3 years ahead of South African Grade 4s.
- SA does not currently have a credible or budgeted plan to catch up learning losses, despite experiencing the largest decline between 2016 and 2021.

The story these data, and others in the report, tell, is alarming. South African Grade 4 learners perform appallingly in reading for meaning when compared with children in 56 other countries, and lag three years behind those in a comparable BRICS country (Brazil). Things are getting worse, if one compares with data from previous waves. And within the pattern of general poor performance, existing inequalities, including those of language and place, are reproduced in the reading comprehension skills of Grade 4 learners in South Africa.

There are many excellent researchers working very hard to improve a serious and deteriorating situation, but these data have implications, not just for those working in fields related directly to literacy and school education, but also for all scientists. If skills deficits are so serious and getting worse, it may well have implications for the future of science in our country – the pipeline, as others have said, is broken.

The contrast between the great achievements showcased by the NSTF and the PIRLS data may lead many South African scientists to feel as though we are living simultaneously in two very different worlds – on the one hand, South Africa is a world leader in a number of scientific fields, and on the other, ours is a country of vast, and probably growing, functional illiteracy (including scientific illiteracy). The PIRLS study confirms the latter. At the same time, many South African academics face the everyday stress of doing their best to educate on a mass scale, and with few resources, poorly prepared students without the skills in basic literacy and numeracy one would expect university students to have. The effort this involves impacts on academic and research productivity from these academics, and also creates a situation in which many academics, and others, question the value of some university qualifications. Through all of this runs the river of historical and ongoing inequality: we have excellent, well-trained and equipped scientists, but we also have those excluded from having a fair chance at developing their talents so that they can aspire to the privilege of a productive and meaningful scientific career. As a scientific community, across all disciplines, we have a collective responsibility to be mindful of and to contribute to working towards the eradication of dysfunction and inequality in our educational systems, and in our society as a whole. This is an issue not just of social justice but also of sustainability of the good quality of science in our country.

In this regard, it is especially pleasing and encouraging for us that this year's winner of the Lifetime Award at the NSTF ceremony was Professor Jonathan Jansen. The commendation given for this award by the NSTF reads as follows:

> *For a distinguished contribution to the advancement of education scholarship through advanced research and publication, scholarly teaching, innovative university management especially in times of racial disharmony, science leadership, school improvement, educational innovation, capacity development, public engagement, and science advocacy.*

At the *South African Journal of Science*, we experience Jansen's support and help regularly, way beyond what could reasonably be expected from his role as President of the Academy of Science of South Africa (ASSAf), our publisher. More importantly for this discussion, though, is Jansen's ongoing work to improve the quality and sustainability of scholarship and research in South Africa. And it is no coincidence that one of Jansen's most recent publications, *Corrupted: A Study of Chronic Dysfunction in South African Universities*[4], reviewed in this issue, deals with the question of corruption in higher education in South Africa. Though it is a daunting, and probably impossible, task fully to understand the depth of dysfunction in some of our higher education institutions, a contributing or exacerbating factor must be the lack of equal access to the tools of rigorous scientific enquiry. Science is not about the production of what in today's knowledge economy are termed 'outputs', but about the use of skills to articulate and solve problems. And currently, as in the past, these skills are bound up with historical and current privilege.

To bridge the gap between the two worlds – science excellence on the one side and lack of access to functional literacy on the other – is no easy task. But we need more concerted work across disciplines to stop the gap from widening, and ultimately towards bridging it.

References

1. Mullis IVS, Von Davier M, Foy P, Fishbein B, Reynolds KA, Wry E. PIRLS 2021 International Results in Reading. Chestnut Hill, MA: Boston College, TIMSS & PIRLS International Study Center; 2023. https://doi.org/10.6017/lse.tpisc.tr2103.kb5342
2. South African Department of Basic Education (DBE). PIRLS 2021: South African Preliminary highlights report. Pretoria: DBE; 2023. Available from: https://www.up.ac.za/media/shared/164/ZP_Files/2023/piirls-2021_highlights-report.zp235559.pdf
3. Spaull N. 10 main findings from PIRLS 2021 South Africa [webpage on the Internet]. c2023 [cited 2023 Jul 13]. Available from: https://nicspaull.com/
4. Jansen JD. Corrupted: A study of chronic dysfunction in South African universities. Johannesburg: Wits University Press; 2023. https://doi.org/10.18772/12023037946

HOW TO CITE:
Swartz L. Sustaining excellent science. S Afr J Sci. 2023;119(7/8), Art. #16469. https://doi.org/10.17159/sajs.2023/16469

AUTHOR:
George Ellis[1]

AFFILIATION:
[1]Department of Mathematics and Applied Maths, University of Cape Town, Cape Town, South Africa

CORRESPONDENCE TO:
George Ellis

EMAIL:
george.ellis@uct.ac.za

HOW TO CITE:
Ellis G. Brian Warner (1939–2023): Astronomer, historian and academic leader. S Afr J Sci. 2023;119(7/8), Art. #16307. https://doi.org/10.17159/sajs.2023/16307

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

PUBLISHED:
8 August 2023

Brian Warner (1939–2023): Astronomer, historian and academic leader

Brian Warner was a great asset to the University of Cape Town (UCT) and to science in South Africa in general. He set the Department of Astronomy at UCT on the road to being the world-class institution that it now is, and helped to build up science at UCT and in South Africa to an international level.

Prior to 1970, astronomy at UCT was taught by R. H. Stoy, who was Director of the Observatory, and by Don Fernie and Pat Wild of the Physics Department. UCT's Department of Astronomy was formally founded in 1970 with Tony Fairall as Lecturer and Pat Wild as Senior Lecturer. Brian was recruited as Head of Department in 1972. He had been a Fellow of Balliol College, Oxford (1965–1967), where he met mathematics tutor Jack de Wet, who became Dean of Science at UCT in 1971 and then recruited Brian to come to UCT. De Wet wrote to Brian when he was in Texas, telling him that he was due to retire from Oxford and was returning to UCT as Dean of Science. Brian said later "That changed everything. I knew De Wet, I admired him. I wanted to go where he went."[1] They formed a close working relationship as Jack de Wet transformed science first at UCT, and then transformed the way the then Foundation for Research Development worked (it later became the National Research Foundation).

Brian attended East Grinstead County Grammar School, and as a schoolboy befriended Patrick Moore who lived in East Grinstead. Brian and his friends used Moore's telescope. He wrote in Moore's obituary in *The Guardian*[2]:

> *There are many individuals in successive generations of professional astronomers who owe a great deal to the books and personal support of Patrick Moore. He introduced children of all ages to astronomy, and some of them became prominent professionals in astrophysics and planetary sciences. Having, as a schoolboy, lived within bicycling distance of Patrick's house in East Grinstead, I benefited from his encouragement and generosity of time and, indeed, from his introduction to those who were later to become my teachers and mentors. And I still value the inscribed books he gave me during those years. The scope of his books went well beyond introductory and popular texts – his early lunar and planetary publications were well-rounded reviews that included much from the professional literature. Some books, such as the one on Neptune (1989), were useful contributions to the history of astronomy.*

Brian studied astronomy at University College London, already publishing two papers[3,4] before completing his doctoral thesis[5] in 1964. For his thesis research he travelled to the Radcliffe Observatory in Pretoria to use the observatory's 1.9-m telescope. Here he met David Thackeray, who specialised in astronomical spectroscopy. At the University of Texas at Austin he helped develop the new field of high-speed photometry for studying variable stars and measuring stellar radii by observing lunar occultations. He became a major figure in astronomy worldwide through his books and papers on cataclysmic variable stars. He published two key books[6,7] and many papers (for example[8,9]) on the topic, including theoretical as well as observational papers[10]. He was Head of the Department of Astronomy until December 2004; Renee Kraan-Korteweg took over as Head in January 2005. But he continued to do research.

He was, however, much more than an astronomer: he was a meticulous researcher of broad interests who wrote books on flora, and on the history of astronomy in South Africa. Particularly he wrote about Sir John Herschel in regard to both his astronomy and his botanical studies, for example co-authoring *Flora Herscheliana: Sir John and Lady Herschel at the Cape 1834 to 1838*[11] and authoring *Cape Landscapes: Sir John Herschel's Sketches 1834–1838*[12].

An important paper which he published in the *South African Journal of Science* was on the meeting between Charles Darwin and John Herschel at the Cape of Good Hope in 1836.[13] Brian wrote: "with Herschel himself having speculated on evolution just before he met Darwin, it is probable that he stimulated at least the beginnings of the latter's lifelong work on the subject" – a key observation regarding the history of the theory of evolution. Brian wrote poetry[14,15], and built a harpsichord and a clavichord, both requiring great woodworking virtuosity.

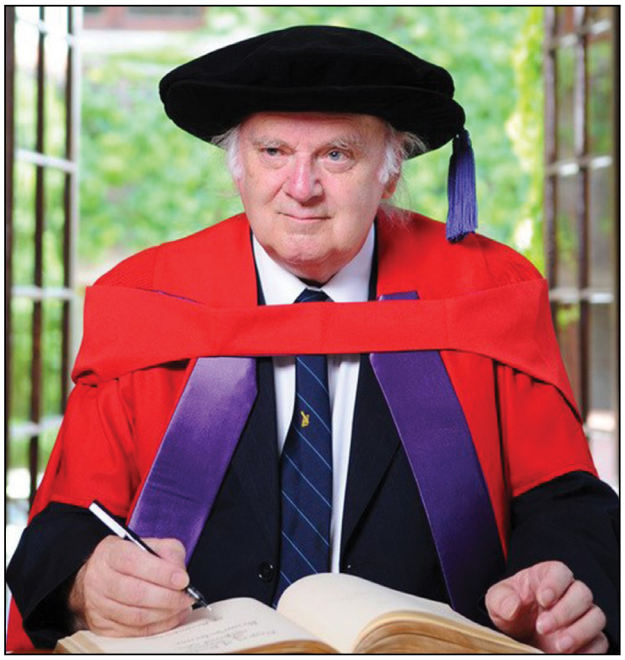
He was a strong supporter of science in general at UCT and in South Africa. From 1981 to 1983, he served as President of the Royal Society of South Africa, and he was a founding member of the Academy of Science of South Africa. But his reach was global: he was Vice-President of the International Astronomical Union from 2003 to 2009 (other South African astronomers who have held this position are Richard Stoy, Michael Feast, and Renee Kraan-Korteweg).

His work was recognised when he was made a Fellow of UCT, and an Honorary Fellow of the Royal Astronomical Society. He received the McIntyre Award from the Astronomical Society of Southern Africa, the John F.W. Herschel Medal from the Royal Society of South Africa, and the Gill Medal from the Astronomical Society of Southern Africa. He was made an Honorary Fellow of the Royal Society of South Africa, an Honorary Member of the Royal Astronomical Society of New Zealand, and won the Science-for-Society Gold Medal of the Academy of Science of South Africa. He was awarded an Honorary Doctorate from the University of Cape Town in 2009, made an Honorary Fellow of University College London the same year, as well as being elected a Fellow of The World Academy of Sciences. He was on the board of the South African National Library for 10 years.

The Academy of Science of South Africa's book *Legends of South African Science* states[1]:

> *Warner considers building up the astronomy department at UCT as his crowning achievement. When he took charge in the beginning, it consisted of only two people. Today there are so many staff and students that he can't remember the names of everyone. "It's a big department and I'm proud to have laid the foundation for it," he says. Without it and the people it has trained, South Africa would not have been in the position to build and run the ten-metre mirror SALT, which helped to win the bid to host the SKA.*

He made a great contribution to UCT and to South Africa in many ways. I personally was grateful to be able to talk to him on many topics, including astronomy.



Brian Warner (photo CC-BY-ND: www.news.uct.ac.za)

Acknowledgement

I thank Patrick Woudt for helpful comments on this obituary.

References

1. Academy of Science of South Africa (ASSAf). Legends of South African science: Brian Warner. Pretoria: ASSAf; 2017. Available from: https://research.assaf.org.za/bitstream/handle/20.500.11911/74/BRIAN%20WARNER.PDF?sequence=19&isAllowed=y
2. Barker D. Sir Patrick Moore obituary. The Guardian (London). 2012 Dec 09. Available from: https://www.theguardian.com/science/2012/dec/09/sir-patrick-moore
3. Warner B. Rilles near the Lunar Crater Plato. J Br Astron Assoc. 1960;70:299–300.
4. Warner B. The emission spectrum of the night side of Venus. Mon Notices Royal Astron Soc. 1960;121(3):279–283. https://doi.org/10.1093/mnras/121.3.279
5. Warner B. Abundances in late-type stars [PhD thesis]. London: University College London; 1964.
6. Warner BJ. Cataclysmic variable stars. Cambridge: Cambridge University Press; 1996. https://doi.org/10.1017/CBO9780511586491
7. Warner BJ. High speed astronomical photometry. Cambridge: Cambridge University Press; 1988.
8. Warner BJ. Observations of dwarf novae. In: Symposium – International Astronomical Union Vol. 73. Cambridge: Cambridge University Press; 1976. p. 85–140. https://doi.org/10.1017/S007418090001189X
9. Warner B. Absolute magnitudes of cataclysmic variables. Mon Notices Royal Astron Soc. 1987;227(1):23–73. https://doi.org/10.1093/mnras/227.1.23
10. Warner B. Atomic oscillator strengths – III. Alkali-like spectra. Mon Notices Royal Astron Soc. 1968;139(1):115–128. https://doi.org/10.1093/mnras/139.1.115
11. Warner B. Rourke J. Flora Herscheliana: Sir John and Lady Herschel at the Cape 1834 to 1838. Johannesburg: Brenthurst Press; 1998.
12. Warner B. Cape landscapes: Sir John Herschel's sketches 1834–1838. Cape Town: University of Cape Town Press; 2006.
13. Warner B. Charles Darwin and John Herschel. S Afr J Sci. 2009; 105(11/12):432–439. https://doi.org/10.4102/sajs.v105i11/12.147
14. Warner BJ. Dinosaurs' end: Scientific poems. Cape Town: Firfield Pamphlet Press; 1996.
15. Warner B. Scatological verse. Illustrated by Tony Grogan. Cape Town: Snailpress; 2007.

AUTHOR:
David J.N. Limebeer[1]

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

CORRESPONDENCE TO:
David Limebeer

EMAIL:
david.limebeer@eng.ox.ac.uk

HOW TO CITE:
Limebeer DJ. Barry Dwolatzky (1952–2023): In memoriam. S Afr J Sci. 2023;119(7/8), Art. #16358. https://doi.org/10.17159/sajs.2023/16358

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

PUBLISHED:
8 August 2023

Barry Dwolatzky (1952–2023): In memoriam

On the morning of 16 May 2023, I received an email message from Estelle Trengove with the sad news of Barry Dwolatzky's passing. Estelle is the current Head of the School of Electrical and Information Engineering at the University of the Witwatersrand (Wits) as well as one of Barry's past graduate students. Barry was an active participant in the affairs of Wits University until the very end of his life. I met him for the last time at the opening of the Wits Innovation Centre that he helped to create; this was only four weeks before his death.

Barry and I enrolled to study Electrical Engineering at Wits in 1971. I recall in retrospect that many of the students of the day looked ridiculous. In our case this was probably a reaction to our recent release from national service. Distinctive features of our attire included long unkempt hair, garish bell-bottomed jeans, and tie-dyed T-shirts that were topped off with socks that didn't match anything. Our first meeting was probably on the Wits library lawn, in Pop's café, or in the student canteen. I still remember his smile, his softly spoken dry humour and a slight stammer that never left him. The class of 1971 included some very talented students, but we all found it a demanding course and down time away from our studies had to be rationed. A fond memory was Friday evenings with Barry at Jean Bodart's house. As engineering students go, Jean was a cultured man with an impressive collection of tape recordings and long-playing records that included several folk luminaries such as Pete Seeger, Bob Dylan, and Joan Baez. In these settings Barry made me feel uncomfortably inferior – even at this young age he was widely read with a rapidly evolving interest in South African politics.

Another fond undergraduate recollection was our Friday afternoon technical drawing classes that were held on the top floor of what came to be the Richard Ward building. An off-curriculum objective of these classes was to make planes out of A0 drawing paper and attempt to hit targets on the far side of Juta Street; despite our best efforts, the railway lines remained stubbornly out of range. Nobody would have guessed that in later life Barry would turn into a property developer who would be instrumental in the refurbishment of some of the buildings we flew over as students. Barry and I were separated in 1973, when he had to repeat a couple of courses; his multiple outside interests had spread him too thinly.

My 50-year friendship with Barry reminds me of two oscillators operating at slightly different frequencies that would slowly drift in and out of phase. Sometimes we were in contact and other times less so. After graduating in 1975, Barry stayed on at Wits to do a PhD in digital control systems. Motivated in part by his strong distaste for the military, and the prospect of further 'camps', he went to the University of Manchester's Institute of Science and Technology as a postdoc and then to Imperial College London. At roughly the same time I left South Africa to go to Cambridge. As fortune would have it, the head of the control group at Imperial was another Witsie, David Mayne, whose PhD supervisor had been the well-connected British elder statesman, John Westcott. John was apparently able to 'fix things' effortlessly within the UK civil service.

Barry's move to Imperial College required a new work permit. This involved a six-week wait and multiple trips from Manchester to the Home Office in London for a string of encounters with obstructive bureaucrats. Frustrated by all this, Barry explained his difficulties to John Westcott, who immediately picked up the phone and called a man named 'Bunter'. "His name is Dwolatzky, and the application number is xxx .... Thank you, old chap. Regards to your wife." John Westcott then explained: "If you go to Lunar House now your work permit will be ready." The full story can be found in Barry's memoir – *Coded History: My Life of New Beginnings*, which tells a more complete story of Barry's life as both an activist and a scientist. For what I am sure were sound reasons, John Westcott appointed Barry, while David Mayne hired me – we were united again.

Ever since his undergraduate days Barry had been interested in computers and computer programming. These interests were put to good use at Imperial where he joined John Westcott's group that was working on the computer modelling of the British economy. Their focus was to find ways of employing control-theoretic methods to develop the UK's economic policies. Barry worked on the 12th floor of the Electrical Engineering building at Imperial — we regularly got to meet for lunch. Barry's group included a fiery Turk and an equally volatile Greek. Barry's role was to develop software and mediate in all-to-frequent arguments. Barry did not like conflict and did not believe in the idea that the strength of an argument is proportional to the volume with which it is delivered.

As was routinely the case, fixed-term research grants would end, and the associated staff had to seek employment elsewhere. The conclusion of one such grant triggered my move from Cambridge to Imperial College, while the end of another resulted in Barry's move to the GEC-Marconi Research Centre on the outskirts of Chelmsford in Essex. This turned out to be a good move for Barry because he got involved in the use of artificial intelligence in robotic manufacturing. Barry was particularly proud of his contributions to flexible robotic manufacturing using object-oriented programming. As with many technological advances, artificial intelligence, computers, and high-speed digital communications can produce socially repressive consequences such as social credit scores, vaccine passports and manipulable digital currencies. Despite his enthusiastic support for these new technologies, he was less keen on the idea of his bathroom scales talking to the fridge, which would them conspire jointly with the local supermarket to 'improve' his diet.

When Barry was still living in London, and commuting to GEC, he received the news that he was seriously ill and would be immediately transferred to St Bartholomew's Hospital in London (a world-renowned cancer hospital). Barry had been diagnosed with hairy cell leukaemia, which at the time had a poor prognosis. I was amazed at the apparent calmness with which he took this terrible news – but that was Barry.

In 1989 it appeared that Barry's leukaemia was in remission and for various reasons he decided to return to South Africa. Another point of convergence was Professor Hu Hanrahan's offer of a senior lectureship for Barry at Wits.

Barry Dwolatzky

In 1972 Barry and I were in our second year, and Hu Hanrahan was then a young professor who struggled to teach us some electronics. His return to South Africa was the start of a good period in Barry's life: his cancer was in remission, he married Rina, he became a marathon runner, and an embryonic property developer.

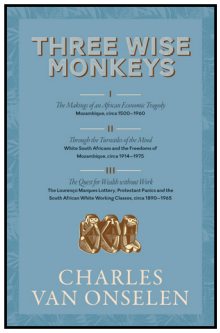
Barry's many contributions to Wits have been covered by others, and so I will not deal with this aspect of his life in any detail. His Wits-related achievements included setting up several new undergraduate programming courses; he was founder and one-time director of the Joburg Centre for Software Engineering; he had the vision to spearhead the fundraising effort that transformed several rundown buildings in Braamfontein into the Tshimologong Digital Innovation Precinct and was its first director. He also persuaded IBM to establish their 13th research laboratory that is attached to Tshimologong. In 2016 he was honoured with a Vice-Chancellor's Award for Research and Teaching. He also worked alongside Wits University's Deputy Vice-Chancellor, Professor Lynn Morris, to establish the Wits Innovation Centre, which was launched only a month before his untimely death.

From a human perspective Barry was a simple man, but he was no simpleton. As part of Wits' centenary celebrations, we wrote a paper on some South African contributions to engineering that was recently published in this journal. While drafting the paper, I was amazed by his breadth of knowledge of the history of South African engineering and especially Wits' contribution to World War II radar; a topic I thought I knew well. He also had a high social IQ and was remarkably good in his dealings with people; I never knew him to become impatient or irritable, but he was also able to quietly stand his ground when necessary.

In closing, I would also like to mention that Barry was instrumental in arranging my current appointment as a Distinguished Professor at Wits, a gesture for which I am most grateful.

God's speed my friend, you will be missed by all those who had the privilege of knowing you.

BOOK TITLE:
Three wise monkeys



AUTHOR:
Charles van Onselen

ISBN:
9781776192502 (paperback, 330 pp)

PUBLISHER:
Jonathan Ball Publishers, Cape Town; ZAR1500

PUBLISHED:
2023

REVIEWER:
Bárbara Direito[1]

AFFILIATION:
[1]Research Fellow: CIUHCT, NOVA School of Science and Technology, Caparica, Portugal

CORRESPONDING EMAIL:
b.direito@fct.unl.pt

HOW TO CITE:
Direito B. Exploring the dynamics of the relationship between South Africa and southern Mozambique. S Afr J Sci. 2023;119(7/8), Art. #15840. https://doi.org/10.17159/sajs.2023/15840

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

PUBLISHED:
8 August 2023

Exploring the dynamics of the relationship between South Africa and southern Mozambique

Charles van Onselen is a leading South African historian whose career has spanned decades. With this trilogy, he has made yet another important contribution to the historiography of southern Africa. Standing firmly in the field of social history, but foraying into economic, political and cultural history, van Onselen uses a broad approach, both in terms of time and space, to look at some of the different dynamics of the historical relationship between South Africa and southern Mozambique, to which he has been devoted in the past few years.

This collection is made up of three volumes:

Volume 1: Mizaru – 'See No Evil' – The makings of an African economic tragedy: Mozambique, circa 1500–1960

Volume 2: Kikazaru – 'Hear No Evil' – Through the turnstiles of the mind: White South Africans and the freedoms of Mozambique, circa 1914–1975

Volume 3: Iwazaru – 'Speak No Evil' – The quest for wealth without work: The Lourenço Marques lottery, Protestant panics and the South African white working classes, circa 1890–1965

The title of the collection – and of each of the three volumes – is in itself particularly meaningful: by referencing the old Japanese proverb of the Wise Monkeys, van Onselen is evoking what in these historical exchanges between the two neighbouring territories has been silenced, especially in the context of colonialism and imperialism. That is precisely what he hopes to uncover: "Only by seeing, hearing and speaking honestly about the past can we hope to understand a troubled present" (Vol. 1, p. VIII).

The particular dynamics of the relationship that the author proposes to examine in detail, as well as the methodologies chosen to do so, are laid out in the general introduction to the collection, entitled "Intersections of Church, Nation and State: South Africa and Mozambique, circa 1650–1970" (Vol. 1, p. 1–67). Therein van Onselen makes a strong case for bringing together the teachings of Braudel and Hobsbawm and examining the connections between South Africa and southern Mozambique since the very earliest times. By setting aside the traditional basis of analysis in historical research, the nation-state, favouring instead a transnational, regional approach, and by looking at the deeper history of southern Africa from the 1650s, but mainly after the Union of South Africa was formed in 1910, van Onselen hopes to follow more closely how the relationship between South Africa and southern Mozambique unfolded, emphasising instances of tension, competition, dependence or cooperation.

Importantly, the main argument goes, the political, cultural and economic dynamics of the relationship to be examined in the three volumes were profoundly shaped by asymmetries that are at the root of Mozambique's current underdevelopment. Mozambique's predicament, van Onselen posits, is that it was doubly colonised – first by "the Portuguese and commerce" since the 16th century, and then by "South Africa and industry." Although officially under Portuguese colonial rule, in the late 19th century, southern Mozambique (the region below the Sabi River, known in Portuguese as the Sul do Save) became "in effect, South Africa's fifth province, the subject of a form of sub-imperialism both bedevilling and retarding 'state' formation in a colony" (Vol. 1, p. 14). Religion is another key variable in this argument: van Onselen attributes a central role to Protestantism and Catholicism and the ways in which they helped shape different societies on either side of the border and informed official policies, as well as relationships between the two neighbours.

Van Onselen addresses the root causes and the consequences of the historical asymmetries between South Africa and Mozambique in Volume 1. Following in the footsteps of his recent *The Night Trains: Moving Mozambican Miners to and from the Witwatersrand Mines, 1902-1955*[1], where he examined in detail the harrowing experience that led thousands of Mozambican men to work in South African mines after the mining revolution, in Volume 1 of the trilogy he explores the long path that led Mozambique from "global slavery to one of regional industrial servitude" (Vol. 1, p. 67). Looking at a glance through centuries of Mozambique's history, van Onselen focuses especially on the 19th-century shift from a territory politically and economically centred in the north and around relations in the Indian Ocean World, to a territory centred in the Sul do Save and around relations with South Africa, as seen by the growing importance of the Lourenço Marques port. South Africa's "predations" on the Sul do Save since the 1870s, van Onselen tells us, were also financial and commercial, and its industrialisation was made on the back of thousands of low-paid Mozambican workers.

In Volumes 2 and 3, the 20th-century culture wars being fought within Protestant South Africa, and their reflection in neighbouring Catholic Mozambique, take centre stage. Highlighting the close entanglement between religion and politics in an increasingly conservative South Africa, van Onselen examines first how Mozambique, and especially the then capital Lourenço Marques (now Maputo), praised for its buoyant nightlife and pristine beaches, came to occupy an important role in the lives and the imagination of white South Africans looking for a refuge from the strictures of Calvinist morals. Appealing to different tourist populations, from adventurous young women to working- and lower-middle-class men, Lourenço Marques came to signify the possibility of living a freer life, even if just for a few days. And while the Portuguese, van Onselen shows us, were intent on stimulating these exchanges and even dreamed of turning Lourenço Marques into a French Riviera of sorts, namely through the construction of the high-end Polana Hotel, the South African government was less pleased. Fearing the subversive potential of a city where classes, sexes and races mixed in ways that could complicate the stricter hierarchies imposed at home, South African authorities put in place systems of surveillance and control of cross-border movements, but ultimately failed to curb Lourenço Marques' touristic appeal.

A second cultural war, this time fought in the airwaves, is the focus of the second part of Volume 2. Van Onselen centres on the fascinating history of an English-language commercial radio station, Lourenço Marques Radio (LM Radio), based in the Mozambican capital, from its beginnings in the 1930s to its demise in the 1970s under pressure from Afrikaner nationalists, revealing in the process a Portuguese government pliable to South African interests. Much like travelling to Mozambique, listening to LM Radio became an escape for the elusive South African youth. Popular shows like the 'LM Hit Parade', aired on Sunday nights and broadcast to South Africa, led to fears of the potentially dangerous effects of popular culture, while in the 1960s, in the context of decolonisation wars, South African authorities feared rather politically subversive messages.

Finally, in Volume 3, van Onselen tells the story of "a religious war lost" (Vol. 3, p. 18). What was at stake in this war was controlling the hearts, the minds and especially the wallets of South Africans. In this rich and detailed transnational history of gambling in South Africa, van Onselen focuses especially on the lives and dealings of an Australian businessman and a Portuguese ex-priest working in Johannesburg as a "Native Affairs Curator," each with capital and influence on either side of the border, who came together to obtain a concession to run the Lourenço Marques Lottery. Aimed especially at white working classes in the Rand, where gambling was essentially illegal and frowned upon by the church, the lottery, thriving between the wars, also faced opposition in Mozambique. Importantly, van Onselen shows us how activities like dog racing, pinballs or lotteries came to occupy different roles for different sectors of the population in South African society during the 20th century, in the context of changing economic and political circumstances, but also how Afrikaans and English Protestants – at odds on other issues – and the state aligned to control the future of the disposable income of first working-class men, and later women and families.

Three Wise Monkeys will appeal to scholars of southern Africa and wider audiences alike. It remarkably shows how Mozambique is part of South African history, not just through the hardship of thousands of Mozambican workers, so fundamental to sustaining the mining industry, but also through different relationships and cultural exchanges. From an interdisciplinary perspective, beyond the study of history as narrowly understood, van Onselen's work throws light on a number of intersecting fields, including the studies of extractive labour practices, the entanglement between politics and religion and international politics. By limiting his analysis to mostly English-language primary sources, however, van Onselen tells this story essentially from a South African perspective. In particular, the points made about the inner workings of the Portuguese colonial government, about the ways in which the Sul do Save was captured by South African interests, or about the social and economic life of the region could be nuanced by referencing other important scholarship drawing mostly on Portuguese-language sources.[2-4] Furthermore, the cultural exchanges discussed in this trilogy appear as one-sided, when more varied sources could perhaps have helped to understand how different expressions of popular culture, namely coming from South Africa, circulated, were received and were appropriated in Mozambique, by settler populations and especially by African populations. For a broader picture of the dynamics and relationships forming on both sides of the border, van Onselen's *Three Wise Monkeys* could usefully be read alongside works like Todd Cleveland's recent exploration about tourism in Africa, several chapters of a recent edited volume on popular culture in Africa under former Portuguese colonial rule, as well as with a monograph on football in Lourenço Marques.[5-7] In spite of these shortcomings, these three volumes greatly enhance our knowledge about the history of southern Africa and specifically about the relationship between South Africa and Mozambique, and they should be read by scholars from a wide range of fields.

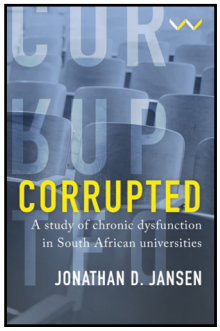
Acknowledgement

Research for this Book Review was supported by FCT – Fundação para a Ciência e a Tecnologia, I.P. (SFRH/BPD/114805/2016 and CEECIND/01948/2017), as well as by CIUHCT (UIDB/00286/2020 and UIDP/00286/2020).

References

1. Van Onselen C. The night trains: Moving Mozambican miners to and from South Africa, 1902–1955. Johannesburg: Jonathan Ball; 2019.
2. Havik PJ, Keese A, Santos M, editors. Administration and taxation in former Portuguese Africa 1900–1945. Newcastle upon Tyne: Cambridge Scholars Publishing; 2015.
3. Covane LA. Migrant labour and agriculture in southern Mozambique with special reference to lower Limpopo valley, 1920–1992 [PhD thesis]. London: SOAS; 1996.
4. Pélissier R. Naissance du Mozambique: Résistance et révoltes anticoloniales 1854–1918 [Birth of Mozambique: Anticolonial resistance and revolts 1854–1918]. Orgeval: Pélissier; 1984. French.
5. Cleveland T. A history of tourism in Africa: Exoticization, exploitation, and enrichment. Athens, OH: Ohio University Press; 2021. https://doi.org/10.2307/j.ctv224txwk
6. Domingos N, editor. Cultura popular e Império. As lutas pela conquista do consumo cultural em Portugal e nas suas colónias [Popular Culture and Empire. The Struggles for the Conquest of Cultural Consumption in Portugal and its Colonies]. Lisbon: Imprensa de Ciências Sociais; 2021. Portuguese.
7. Domingos N. Football and colonialism: Body and popular culture in urban Mozambique. Athens, OH: Ohio University Press; 2017. https://doi.org/10.2307/j.ctv224tx04

BOOK TITLE:
Corrupted: A study of chronic dysfunction in South African universities



AUTHOR:
Jonathan D. Jansen

ISBN:
9781776147946 (paperback, 320 pp)

PUBLISHER:
Wits University Press, Johannesburg; ZAR350

PUBLISHED:
2023

REVIEWER:
Nkosinathi E. Madondo[1]

AFFILIATION:
[1]Academic Literacy and Language Unit, Mangosuthu University of Technology, Durban, South Africa

EMAIL:
madondone@mut.ac.za

HOW TO CITE:
Madondo NE. A timely, deeper, and complex analysis of dysfunctionality in South African public universities. S Afr J Sci. 2023;119(7/8), Art. #16073. https://doi.org/10.17159/sajs.2023/16073

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

PUBLISHED:
8 August 2023

A timely, deeper, and complex analysis of dysfunctionality in South African public universities

In his own words, the author of the book under review, Professor Jonathan Jansen, points out that the existing accounts of malfunction in dysfunctional universities in South Africa cannot give us the complex and deep analysis we need to understand the persistent instabilities in South African public universities. Jansen cites publications on "corruption and mismanagement in these dysfunctional universities" (p. 2), but these are limited and often lack empirical data and theoretical rigour. The purpose of this book is, thus, to go beyond immediate experiences and observations of dysfunction in public universities to excavate more deeply into the root cause/s of persistent dysfunction. In what is a profound contribution of Jansen's book, he provides a deeper theoretical and layered understanding or explanation of persistent dysfunction.

A key issue in the book is an awareness and critical analysis of mechanisms in the structural domain (social structures, finance system, education system, corruption, resources, power and so on); cultures (discursively constituted sets of ideas in language and other sign systems that constrain and enable an emergence of events, observations, and experiences); and agency (peoples' thoughts and actions enabling or constraining them to do or say what they can or cannot).[1] Like Archer[1] and Bhaskar[2] Jansen is of the belief that, to come as close as possible to the 'truth', a layered understanding or explanation is necessary, where structures, culture and agency are analysed separately. In other words, Jansen is asking a question and is providing an alarming answer to the role that structures, cultures or agency play in contributing to persistent dysfunction in public universities. Based on this understanding, Jansen argues that a deterministic view is too simplistic to provide us with a nuanced explanation of persistent dysfunction. In his analysis, he does not dismiss history, race or class. He acknowledges that the legacy of apartheid neglect and inequity may well be a key contributor to dysfunction in historically disadvantaged universities. Historically, white universities were institutionally supported, in material and symbolic terms, by the apartheid regime, with far less support given to historically black universities. But a key contribution of Jansen's analysis is that he goes well beyond a narrow, simple and deterministic view of history – he shows that the real and root causes of persistent dysfunction in public, particularly black, universities are far more complex.

Using political economy as a theoretical framework, Jansen is able to provide a deeper analysis, explanation and understanding of persistent dysfunction in public historically black universities in South Africa. His central thesis is based on the idea that structures like the financial system, in which 26 public universities receive billions of rands annually in state funds, could be one of the reasons such universities end up not paying needed attention to their mandate, which is the academic project. Instead, vice-chancellors and deputy vice-chancellors, chairs of council and council members, students' representatives, and unions compete for scarce state resources – a scramble which is compounded by political interference. The result is an industrial scale of corruption in these institutions. In relation to this point, Ngcamu and Mantzaris[3(p.7)] argue that, in these institutions, "The university systems that detect corruption have been deliberately weakened by university leaders to enable fraudsters to access funds through corrupt means." This is the central thesis of the book and, in my view, accurate.

Jansen shows that these events and patterns of corruption in universities are hardly obvious or visible in the eyes of the public, given that most of us in the public domain, at most, view universities only as places for education. We do not think of them as places to be associated with corruption. This perception itself may provide some scaffolding for the compromise of the academic project in these institutions.

Jansen does not dispute that there is corruption in historically white universities, but the extent of it cannot be equalled to that occurring in historically black and disadvantaged universities. In historically black and disadvantaged universities, there is blatant stealing and looting of resources in the form of state funding, in its various forms.

The credibility of Jansen's analysis and thus argument is based on empirical evidence of "more than a hundred interviews with senior people involved in, and knowledgeable about, South African universities", having been involved as "assessors and administrators in various universities" (p. 13), as well as his direct involvement in some of these institutions, either as the deputy vice-chancellor or an administrator. Such evidence forms the basis to understand the problem. This book is unique in that Jansen puts a human face and biographical experience to a problem of political economy.

In the final chapter of the book, Jansen proposes that, to address the current chronic dysfunction, among other things, universities need to go back to their mandate, to pursue an academic project. This is no more and no less than the production of knowledge (through research), which can be achieved only if universities genuinely subscribe to the ideals of integrity, argues Jansen – doing the right and the correct thing at all times.

Jansen further notes that, for this idea to be realised, universities need to reduce the numbers and depoliticise university councils, the highest governing structures in universities, and appoint professionals with integrity. Such persons are likely not to conflate their roles of governance with the roles of management. The same thinking, in terms of depoliticisation, extends to students and unions. Jansen argues that students are at universities to learn. He does not dispute that students should be involved in politics. If, however, their involvement is geared towards securing tenders for constituencies outside university as they sit on procurement committees, this can have dire consequences. For example, where procurement decisions do not favour their preferred candidates (and by implication their own interests), students may make the institutions ungovernable. Jansen suggests further that

where trade unions narrowly pursue their own interests in universities, such involvement will derail and compromise the academic project.

A major and important contribution of this book is the core notion that the academic project must be revived. I think therefore that the book would have benefitted from a clearer indication by Jansen of how people involved in universities may come to understand what an academic project is. What do academic leaders, academic teachers, students, and so on, need to do to realise, truthfully, the academic project, especially in the midst of violence that has engulfed these universities? Whilst this book has many strengths, it is less strong at taking full account of the fact that the contexts of our universities and the espoused academic project have a long history of being dominated by "white and neoliberal world views"[4]. Jansen, very importantly, redresses an imbalance in our understanding of corruption in our universities. Building on this scholarship, we now need a more complete and synthesised analysis of the core issues.

References

1. Archer M. Culture and agency: The place of culture in social theory. Cambridge: Cambridge University Press; 1996.
2. Bhaskar R. A realist theory of science. Sussex: The Harvester Press; 1978.
3. Ngcamu BS, Mantzaris E. Anatomy and the detection of corruption in 'previously disadvantaged' South African universities. 2023;20(1):323–349. https://doi.org/10.35683/jcman1001.197
4. Radcliffe SA. Decolonising geographical knowledges. Trans Inst Br Geogr. 2017;42(3):329–333. https://doi.org/10.1111/tran.12195

AUTHORS:
Lenore Manderson[1]
Susan Levine[2]

AFFILIATIONS:
[1]School of Public Health, University of the Witwatersrand, Johannesburg, South Africa
[2]Department of Anthropology, University of Cape Town, Cape Town, South Africa

CORRESPONDENCE TO:
Lenore Manderson

EMAIL:
lenore.manderson@wits.ac.za

HOW TO CITE:
Manderson L, Levine S. COVID-19 research and science infrastructure in South Africa. S Afr J Sci. 2023;119(7/8), Art. #16294. https://doi.org/10.17159/sajs.2023/16294

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
COVID-19, pandemic preparedness, science and technology innovation, vaccine development

PUBLISHED:
8 August 2023

COVID-19 research and science infrastructure in South Africa

Significance:

Biomedical laboratory and field scientists, as well as social scientists, in South Africa and elsewhere on the continent, responded to the challenges of COVID-19 with speed. African-wide experience with infectious disease, and the networks and infrastructure to conduct new research and implement field trials, were part of the global effort to contain the pandemic. But in order to contribute, scientists necessarily set aside ongoing research, including on some of the most persistent infections – HIV, TB, malaria. This situation highlights the precarity of science research programmes and the challenges of sustaining research capacity when agendas, funds and acknowledgements reinforce global inequalities.

It has been more than 3 years since the first cases of COVID-19 in South Africa, and over a third of the population has been fully vaccinated and transmission has slowed. It continues to do so, offering us time to draw breath and to reflect on how, as a country, we have navigated the pandemic and stalled the worst economic impacts at household level for many South Africans. Since then, there has been growing awareness of the costs of preventive interventions to contain the pandemic for those most vulnerable to infections, and the economic repercussions of these interventions.[1-4] At the same time, we now have the opportunity to reflect on how, by utilising existing strengths in science infrastructure, resources, networks and skills, South Africans played a major role in curbing the spread of infection. In this Commentary, we discuss policies and early public health interventions across Africa in response to the pandemic, then turn to the role of scientists in developing biomedical interventions. In doing so, we draw attention to the importance of science capacity, and the distribution of power in research as well as policy.

The epidemiological picture of the pandemic, here and elsewhere on the continent, remains somewhat unclear, the consequence of varied capacity within and between countries to case find, maintain records, and report administrative data. The limits of information on cases reflect people's reluctance to present for diagnosis among populations with poor access to medical care, due to distance from services and lack of right to receive care, the latter particularly for people without IDs or marginalised from society for structural reasons. Counter directives to shelter in place and avoid social interactions, such as queuing, that might increase transmission of COVID-19, also discouraged people from presenting to clinics. In addition, the relative youth and low density of populations in some countries and local areas likely reduced infection rates, severe morbidity and mortality.

The rapid response of governments across the continent to implement non-pharmaceutical interventions to slow transmission and to introduce fiscal measures to minimise the economic costs of the pandemic was less predictable. South Africa was one of many countries that used existing grant mechanisms and additional cash transfers to support people who might be harshly affected otherwise, as social interventions forced people out of work and the economy contracted. In Kenya, Uganda, Nigeria, and elsewhere, governments likewise supplemented existing social protection funds and introduced cash transfer programmes, despite challenges in administrating the programmes, and facilitating registration and access.[5] In Uganda, various monetary policies were introduced to mitigate the harsh social and economic constraints designed to limit transmission, although these policies did not include funds to help support households.[6] Elsewhere, some households received food or financial aid.[7] The swift implementation of mechanisms of control by African governments illustrates country capacity to respond to health emergencies, at least as well as other countries across the globe, and capacity to plan for future pandemics and other threats to personal, community and national security, lives and livelihoods.

These apparent successes contradict conventional accounts of pandemics, such as Ebola in 2014–2016, which suggest a lack of preparedness within health systems, and little capacity to manage exponential infection without input from outside – from WHO, other multilateral agencies, and scientists from the Global North.[8] But while there were problems in responding to the Ebola pandemic in Sierra Leone, Liberia and Guinea, Senegal's success in averting infection spread suggests that this was not inevitable.[9] Moreover, considerable work had been conducted in the preceding decade on pandemic preparedness. In particular, responses to the H1N1 pandemic in 2009 led to continent-wide summits and the subsequent development of pandemic preparedness plans in 39 countries.[10] Whether and how this anticipatory work informed national action as COVID-19 spread has yet to be documented. While COVID-19 might have captured imagination in the Global North as the first major pandemic since the Spanish Flu, South Africans were already long familiar with pandemics and their chronicity and endemicity, HIV/AIDS and TB among others. While devastating, people were prepared for a measure of pandemic compliance, unseen in countries like the USA.

At the same time as urgent pragmatic actions were implemented by governments to contain COVID-19, scientists across disciplines also responded with speed. Many were able to draw on the social and intellectual capital and infrastructures of nearly four decades of science research and policy responses to HIV, and the long history of research across disciplines on TB and neglected diseases of poverty. This corps of highly trained researchers, and its resources and experiences, was quickly deployed to work on the virology, immunology and vaccinology of SARS-CoV-2.[11,12] Those of us with an eye on scientific developments on the continent, sensitive to the ways in which science agendas are made elsewhere, must celebrate the role of African scientists in conducting critical research on COVID-19, and appreciate how it came about.

In a series of manuscripts on which we are working[13], we illustrate how the J&J trial, which was administered from the Masiphumelele Research Site in Cape Town from September 2020, was possible precisely because of the sustained work of its scientists on TB and HIV. Because of their day-to-day, year-to-year work on these pernicious endemic diseases, these scientists were able to quickly mount a clinical trial of the candidate vaccine for SARS-CoV-2. Money flowed to facilitate the collaboration of scientists across the USA, Latin America and India and to ensure the infrastructure, technical networks, and systems of patient recruitment and monitoring, so enabling the rapid development and roll out of the trial and the verification of results. The ready availability of resources in part reflected US presidential hubris; there was limited acknowledgement that this work was possible only because of decades of laboratory and population research on other diseases. In contrast to the work on COVID-19, this earlier research was often poorly funded, and research teams would wait to year end for advice on new funding and to learn if staff could be retained.

Other population-based research infrastructures in South Africa, including the Africa Research Institute in KwaZulu-Natal and the MRC/Wits Agincourt Research Unit in Mpumalanga, likewise drew on systems and structures for population data collection, community engagement, and data management systems, developed over several decades, to initiate epidemiological and other studies on COVID-19.[14,15] This was also true elsewhere on the continent (e.g. for Kenya[16] and Senegal[17]), with numerous institutes and networks of scientists pivoting from other research to COVID-19. Intellectual and material infrastructure across Africa explains how innovations and basic and applied science projects were up and running so early in the pandemic, even as scientists grappled with shortcomings.

In speaking at a web-based Inkundla, held by the international network Future Earth prior to the Sustainability Research and Innovation Congress (SRI) (online and in Pretoria, 20–24 June 2022), Uzma Alam (Africa Institute for Health Policy, Kenya) argued the value of using COVID-19 as a case study of the flexibility and adaptability that characterise science on the continent. Others speaking in this webinar, such as Tom Achoki (University of Pretoria, South Africa) and Richard Wamai (Northeastern University, USA), made the same point, reiterating the value of governance frameworks such as the African Union's Science, Technology and Innovation Strategy for Africa 2024 to ensure scientific competitiveness and economic transformation. They noted too the deftness of African-based biomedical and social scientists, in laboratories and in the field, to enable focused responses to the pandemic.

But despite South African participation in this work, agendas are still set largely in the Global North, and funds flow accordingly. The tilt of agenda setting impacts profoundly on the capacity of scientists in South Africa and across the continent to maintain research programmes in areas in which they were trained, to maintain scientific and institutional competitiveness, and to contribute to reducing economic costs and social suffering of a wide range of diseases. Wider disparities in funding severely limit the competitiveness of African-based scientists, as Julie Livingston[18] and Marissa Mika[19] have illustrated for oncology. In the excitement of new funding calls and increased resources, scientists across disciplines run the risk that important work, including basic, translational, implementation and social science research, is diverted when funds follow the shifting agendas set elsewhere.

Recognition of the politics of research, including in relation to funding, the flow of data, the appropriation of knowledge, and the acknowledgement of the work, varies across disciplines. While social scientists have argued the need for a 'theory of the south', even much of this discourse has been generated from scholars in the Global North, with limited acknowledgement of the empirical and theoretical work produced in 'the South'. As scientists on the continent worked to temper the transmission of COVID-19, so the ongoing challenges of effective vaccines, improved treatments, and adherence to the management of other infections largely slipped from sight. But the continued transmission of HIV and other endemic diseases requires sustained and well-supported research, including the training of new researchers and maintenance of laboratories and community study settings. The risk with the diversion of effort, and consequent slides in funding to and distraction from core research, is that, as scientists, we might lose our intellectual edge as we struggle to have it recognised.

Competing interests

We have no competing interests to declare.

References

1. Manderson L, Levine S. COVID-19, risk, fear, and fall-out. Med Anthropol. 2020;39(5):367–370. https://doi.org/10.1080/01459740.2020.1746301
2. Chavarro D, Kagaha A, Kaunda-Khangamwa BN, Zakumumpa H, Manderson L. COVID-19 policies: Human rights approaches to protecting vulnerable groups in Africa [document on the Internet]. c2021 [cited 2022 September 22]. http://dx.doi.org/10.13140/RG.2.2.19451.85283
3. Manderson L, Chavarro D, Kaunda-Khangamwa B, Kagaha A, Zakumumpa H. Containing COVID-19 and the social costs on human rights in African countries. Humanit Soc Sci Commun. 2022;9(1):347. https://doi.org/10.1057/s41599-022-01357-4
4. Presidency of South Africa. South Africa Covid-19 Country Report First Edition June 2021. Pretoria: DPME (Department of Planning, Monitoring and Evaluation), GTAC (Government Technical Advisory Centre) & NRF (National Research Foundation); 2022. Available from: https://www.gov.za/documents/south-africa-covid-19-country-report-first-edition-june-2021-30-jun-2022-0000
5. Human Rights Watch. "We are all vulnerable here": Kenya's pandemic cash transfer program riddled with irregularities. New York: Human Rights Watch; 2021.
6. Okumu IM, Kavuma SM, Bogere G. Uganda and COVID-19: Macroeconomic policy responses to the pandemic. Johannesburg: South African Institute of International Affairs; 2021. Available from: https://saiia.org.za/research/uganda-and-covid-19-macroeconomic-policy-responses-to-the-pandemic/
7. Ogunji J, Iheanacho S, Ogunji CV, Olaolu M, Oleforuh-Okoleh V, Amaechi N, et al. Counting the cost: The effect of COVID-19 lockdown on households in South East Nigeria. Sustainability. 2021;13(10), Art. #12417. https://doi.org/10.3390/su132212417
8. Martineau F, Wilkinson A, Parker M. Epistemologies of Ebola: Reflections on the experience of the Ebola Response Anthropology Platform. Anthropol Q. 2017;90(2):475–494. https://doi.org/10.1353/anq.2017.0027
9. Desclaux A. Ebola imaginaries and the Senegalese outbreak: Anticipated nightmare and remembered victory. Africa. 2020;90(1):148–166. https://doi.org/10.1017/S0001972019000986
10. Sambala EZ, Kanyenda T, Iwu CJ, Iwu CD, Jaca A, Wiysonge CS. Pandemic influenza preparedness in the WHO African region: Are we ready yet? BMC Infect Dis. 2018;18, Art. #567. https://doi.org/10.1186/s12879-018-3466-1
11. Bhiman JN, Moore PL. Leveraging South African HIV research to define SARS-CoV-2 immunity triggered by sequential variants of concern. Immunol Rev. 2022;310(1):61–75. https://doi.org/10.1111/imr.13086
12. Moyo-Gwete T, Moore PL. Leveraging on past investment in understanding the immunology of COVID-19-the South African experience. S Afr J Sci. 2022;118(5–6), Art. #3171. https://doi.org/10.17159/sajs.2022/13171
13. Levine S, Manderson L. Unblinding: Politics, care, and the J&J vaccine trial in South Africa. In: Auerbach J, Jansen J, editors. Racial logics and the politics of the biomedical science. Dordrecht and New York: Springer. In press.
14. Harling G, Gómez-Olivé FX, Tlouyamma J, Mutevedzi T, Kabudula C, Mahlako R, et al. Protective behaviors and secondary harms resulting from nonpharmaceutical interventions during the COVID-19 epidemic in South Africa: Multisite, prospective longitudinal study. JIMR Public Health Surveill. 2021;7(5), e26073. https://doi.org/10.2196/26073
15. Cohen C, Kleynhans J, Von Gottberg A, McMorrow M, Wolter N, Bhiman JN, et al. SARS-CoV-2 incidence, transmission, and reinfection in a rural and an urban setting: Results of the PHIRST-C cohort study, South Africa, 2020-21. Lancet Infect Dis. 2022;22(6):821–834. https://doi.org/10.1016/S1473-3099(22)00069-X

16. Agoti CN, Ochola-Oyier LI, Dellicour S, Mohammed KS, Lambisia AW, De Laurent ZR, et al. Transmission networks of SARS-CoV-2 in coastal Kenya during the first two waves: A retrospective genomic study. eLife. 2022;11, e71703. https://doi.org/10.7554/eLife.71703

17. Diallo AI, Faye A, Tine JAD, Ba MF, Gaye I, Bonnet E, et al. Factors associated with the acceptability of government measures to address COVID-19 in Senegal. Revue Epidemiol Sante Publique. 2022;70(3):109–116. https://10.1016/j.respe.2022.03.123

18. Livingston J. Improvising medicine: An African oncology ward in an emerging cancer epidemic. Durham, NC: Duke University Press; 2012. https://doi.org/10.1515/9780822395768

19. Mika M. Africanizing oncology: Creativity, crisis and cancer in Uganda. Athens, OH: Ohio University Press; 2021.

AUTHORS:
Sibusiso Biyela[1,2]
Ntokozo N. Msomi[3]
Anina Mumm[1,2]

AFFILIATIONS:
[1]SciBraai, Johannesburg, South Africa
[2]ScienceLink, Johannesburg, South Africa
[3]Language and Communication Department, Faculty of Arts and Design, Durban University of Technology

CORRESPONDENCE TO:
Sibusiso Biyela

EMAIL:
sibusiso@sciencelink.co.za

HOW TO CITE:
Biyela S, Msomi NN, Mumm A. South African iLukuluku podcast shows we can talk about science in African languages. S Afr J Sci. 2023;119(7/8), Art. #15648. https://doi.org/10.17159/sajs.2023/15648

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
science engagement, podcast, African languages, translation, decolonisation

PUBLISHED:
8 August 2023

South African iLukuluku podcast shows we can talk about science in African languages

Significance:

ILukuluku is the name of our African-language science communication podcast that includes curious Zulu-speaking communities in scientific discourse. The podcast entertains and educates Zulu listeners of all ages, by breaking down complex scientific topics using facts, linguistic quirks and humour. We contend that the iLukuluku podcast is a practical, proudly South African example of how ordinary citizens, science communicators, translators, scientists, and black people in particular, can work together to communicate science beyond English.

Ilukuluku is a Zulu word for 'curiosity'.

It is also now the name of one of Mzansi's first African-language science communication podcasts that specifically includes curious Zulu-speaking communities in scientific discourse. Each episode is an entertaining, light-hearted dialogue between Zulu-speaking hosts Ntokozo Nomasiko Msomi, a professional language practitioner, and Sibusiso Biyela, a professional bilingual science communicator. Msomi and Biyela worked on several other science translation projects prior to starting their podcast. During this time, they would often have lively debates with one another in a mixture of English and Zulu, in just the same way many other South Africans do. On many occasions, their conversations would boil down to how certain technical or scientific words and concepts, for instance 'dinosaur', 'climate change' or 'planet', just do not exist in Zulu. They would argue about how, or even if, the language could be developed to better express scientific concepts. They recorded some of their debates, and their 'presenter chemistry' was palpable to friends and colleagues who listened.

We were thus encouraged to share these often hilarious and gripping conversations with Zulu-speaking listeners from all walks of life. And so the iLukuluku podcast was born, as a collaboration between local science communication non-profit organisation SciBraai (https://scibraai.co.za/) and People Of Colour Podcasts (https://www.pocpods.com/). The podcast immediately gained media attention in publications like the *Weekend Argus*[1], and began trending the moment it was published on platforms like *Apple iTunes, Spotify,* and *Google Podcasts*. From a science communication and translation perspective, we believe that the podcast entertains, educates and includes Zulu listeners of all ages, because it breaks down complex scientific topics by blending facts, linguistic quirks and humour.

But we also believe that this podcast is part of something bigger; something we support with both 'vigour' and 'passion', the other meanings of *iLukuluku*; and something for which there is strong evidence: by coining new terms to help ordinary folks talk about science in their own African languages, we can begin to undo the distrust of science, and the exclusion from science, that remain entrenched by the legacy of colonisation on the African continent.[2-4] In fact, whether or not a person cares about the relevance of our colonial past, science communication research shows that talking to people about science in their own language does build trust in science.[4] We therefore believe that translating science in African languages will include more people in the public understanding of science, science discourse and science education, and it can even improve health care.[2]

In this Invited Commentary, we contend that the iLukuluku podcast is a practical, proudly South African example of how ordinary citizens, science communicators, translators, scientists, and black people in particular, can work together to communicate science beyond English. We also argue that our podcast supports the broader African project of decolonising science and science communication, and so we are calling on potential funders who want to support our work to email us at info@sciencelink.co.za.

Is it even possible to talk about science in indigenous languages?

Yes, it is, because South Africa has done it before. In 1948, Afrikaans became the dominant language of the state after the Afrikaner nationalists secured a majority stake in parliament, which was the culmination of a decades-long effort to develop the Afrikaans language.[5] This era also ushered in the policy of apartheid, and the Afrikaans language was used to assert dominance of the new white elite over the black majority of the country.[5] This new dominance of a language that had been considered a 'kitchen' language not long before the 1950s, took decades of effort by Afrikaners.[5] They preserved their language through literary works such as poetry, short story collections, and by developing lexicons. And they did all this with the understanding that language is a crucible of culture and ethnicity.

By the 1970s, the Afrikaner-led government had also long recognised the importance of science and technology in legitimising Afrikaans as a global language. They poured considerable state resources into developing scientific literature and other materials in Afrikaans, and they heavily promoted innovators in Afrikaans, such as Christiaan Barnard, the Afrikaner responsible for the first human heart transplant.[3]

To us, this is such a powerful example of inclusive science communication. It likely enabled Afrikaans-speakers to join the global scientific discourse, and to talk about science in their own language, in their own communities, and in their own homes. Indeed, the way Afrikaans was promoted in South Africa demonstrates that with the right political will, resources and pride in one's mother tongue, it is possible to include an indigenous South African language in science communication. Just knowing it can be done is one of the main reasons we started the iLukuluku podcast. And yet, perhaps more importantly, the apartheid government's Afrikaans language policies also reveal exactly

how enforcing non-native language discourse excluded indigenous languages from scientific discourse. This legacy, left across the African continent by oppressive regimes and colonialism, is why we know we need to continue working on projects like iLukuluku.

Language inclusivity builds public trust in science

The well-researched link between language exclusion and a lack of trust in science, particularly in South Africa, is worth expanding on. We hope a deeper understanding of this key issue will motivate science communicators, funders, academics and language practitioners to pour resources, like training and funding, into making science more accessible in indigenous languages.

By 1976, the apartheid government had decided that the medium of instruction in black schools also had to be Afrikaans. This of course led to the Soweto Uprising of 1976, during which young people had recognised that Afrikaans was a tool of the oppressor. They fought back against an ethnically based language policy that was forced onto the black majority, most of whom had already been subjected to the Bantu Education system that was designed to disadvantage them socio-economically.[6] Even today, Afrikaans language policies still spark controversy in some elite South African universities that use it (completely legally!) to exclude black students.[7]

In the past, even scientific research itself had been used to exploit indigenous people around the world. Professor Linda Tuhiwai Smith put it best in the opening lines of her *Decolonizing Methodologies*[8] book:

> *From the vantage point of the colonized, a position from which I write, and choose to privilege, the term 'research' is inextricably linked to European imperialism and colonialism… The ways in which scientific research is implicated in the worst excesses of colonialism remains a powerful remembered history for many of the world's colonized peoples.*[8]

This distrust in science is understandable. Anecdotally, many Zulu speakers use the colloquial saying, *izinto zabelungu*, which means 'things for white people', when referring to science and technology.

Even in our own limited experience as school pupils in public schools in Gauteng and KwaZulu-Natal in the early 2000s, many of our peers feared maths and science, and considered these the most difficult of the subjects. Likely, many South Africans of all ages still relate to this perception today. We suspect that science and maths may be even more intimidating in South Africa than in other parts of the world, partly because these subjects were, and still are, mainly taught in English. This despite English being the second, third or even fourth language of millions in South Africa.[9]

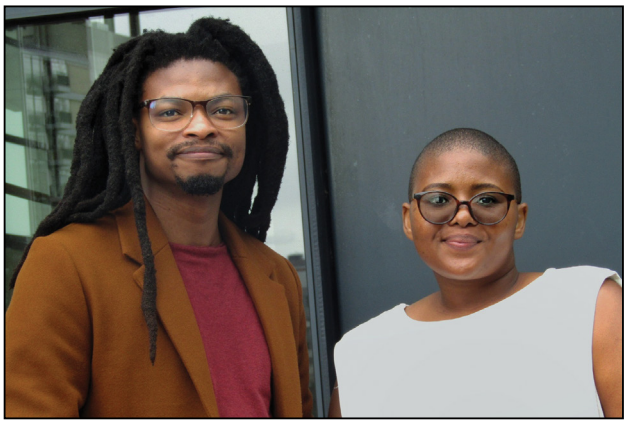
To their credit, the post-apartheid South African government did recognise that the damage of colonialism and apartheid to indigenous languages had to be undone. In 1996, the South African government set out to preserve the heritage of South Africa's rich diversity by introducing policies to develop the country's nine indigenous languages. They intended to develop them all into languages of science too, so that knowledge could be created in those languages.[3] Yet today, nearly 30 years later, we still have very little science and maths materials in indigenous languages in schools or in public libraries. So, while the equality of indigenous African languages has been guaranteed by Section 6 of the South African constitution[3], in our view it has clearly not been a political priority.

How iLukuluku helps develop Zulu as a language of science

One of the ways in which we want to cultivate Zulu in particular as a language of science, is to discuss science-related topics in a conversational, colloquial way in Zulu.

We felt that a podcast is the perfect platform because it is a popular and accessible digital medium. SciBraai, the non-profit science communication organisation with which we are affiliated, approached POC Podcasts for help because they too actively create spaces in popular media for voices in Zulu and other indigenous languages. Unlike how science is usually discussed at school or in popular media aimed at Zulu speakers, where a topic is presented as a (boring) lecture primarily in English, or peppered with anglicisms, Msomi and Biyela's goal in each episode is to find or invent the Zulu words for a topic through conversation. And, although they draw on their professional experience and preparations to make informed, factual statements, their lively quips are unscripted. This gives the show a chat-around-the-*braai*-between-friends vibe.

Biyela's main role as co-host is to explain topics, like the Big Bang, how the immune system works and even dreams, in Zulu. Msomi then uses this new knowledge to translate or deduce new Zulu words around the topic. The nature of the Zulu language is that all sounds and words are phonetically consistent, which means that any new term a listener hears will be easy to spell. They would also understand the meaning behind it because Zulu terms tend to be self-explanatory. The listener goes along on the word-search journey, so that they can follow the reasoning behind any new terms that Msomi and Biyela coin. We think this approach also helps Zulu speakers gain an intimate understanding of the science, as compared to merely memorising borrowed English terms.[10]



Sibusiso Biyela (left) and Ntokozo Nomasiko Msomi (right), hosts of the iLukuluku podcast (photo credit: Sibusiso Biyela).

Biyela and Msomi also engage listeners on social media by inviting them to agree or disagree with the terms they coin, and listeners can even offer their own translations to enrich the communal lexicon. In the first episode, for example, we asked, "what is 'dinosaur' in Zulu? Is it *idayinaso* or is it *igonqonqo*?". The idea for that episode came from a previous science communication project in which Biyela wrote about a dinosaur fossil discovery in Zulu, but struggled to find the words.[11]

In several other episodes, we noted clear differences in how certain natural phenomena are understood by English speakers versus Zulu speakers. For example, one listener did not know that the planet known as Venus in English was the same night sky object known as *Ikhwezi* in Zulu.

Table 1 gives an excerpt from Episode 2[12] which was about the Periodic Table of the Elements in which we broke down what an atom is.

In another discussion, we learned for the first time that sulfur, a common term heard in English chemistry classes, already has a name in Zulu: *isibabuli*. Interestingly, many Zulu-speaking learners would say *iSulphur* when talking about chemistry, not making the link to the medicinal yellow powder they were already familiar with outside of the lab, known in their own language as *isibabuli.*

These instances support the idea that there are asymmetries between English and many other African languages, as reflected in the history of colonisation in Africa.[13] We are hopeful that our podcast can help to better align science discourse in our multilingual society. Our podcast is a 'pilot project' that experiments with the idea of bringing together

Table 1: Excerpt from Episode 2 – Periodic Table of the Elements

Speaker	Zulu	English (translated)
Ntokozo	Yini i- ***atom*?**	What is an **atom?**
Sibusiso	Uma ubuka leli tafula, uyakwazi ukulisika libe uhhafu, uphinde ulisike futhi uhhafu, nalowo hhafu kanjalo kanjalo kugcine sekuyizingcucu.	Well, take a look at this table here, if you cut it in half, and cut that half in half again, and cut it in half again and so on, you end up with a lot of tiny pieces.
	Ososayensi bakudala bazibuza ukuthi ungakwazi yini ulokhu uqhubekile *forever,* uthathe into enkulu uyenze into encane.	Scientists a long time ago asked themselves what would happen if you cut something like this forever, taking large things and turning them into smaller things.
	Labo sosayenzi bathi uma ikhona into enjalo, into engakwazi uku-*divide*-ka uma usufikile kuyona—	Those scientists said that if there is such a thing that would be the smallest thing you could not divide—
Ntokozo	I- ***atom.*** *Ok,* i- ***atom*.** Singasho yini ukuthi *it is where all things originate?*	That's an **atom.** Ok, that's an **atom**. Could we then say that it's where all things originate?
Sibusiso	Singasho kanjalo, *that's what we used to think,* kodwa akusa sebenzi kanjalo.	Yeah, we could say that, that's what we used to think, but it doesn't work like that anymore.
	Abakaze bayithola leyo-*atom* ababeyifuna, *so* manje, basebethola ukuthi ayiyodwa i-*atom* ekhona, kunama-*atom* amaningi.	Those scientists never found such an atom they were looking for, so now, they actually found that there isn't just one atom but there are many.
Ntokozo	Kanjani futhi manje? Akukwazi phela!	Wait, how is that possible? That can't be!
Sibusiso	Ingoba phela izinto zenziwe ngezinto eziningi. Izinto ezenza amanzi akuzona izinto ezenza ipulangwe.	This is because things are made up of many other things. Things that make up water are not the same things that make up a plank.
Ntokozo	*OK*, sengiyakuzwa ukuthi uthini. Bengithi uthi uma ufika phansi kwi-*atom* uthola ukuthi nawo maningi ama-*atom.*	Ok, I understand what you're saying. I thought you meant that when you get down to the level of the atom, it is made up of many other atoms.
Sibusiso	Sisazofika kuyo i-*atom* ngoba ososayensi sebathola ukuthi nayo i-*atom* iya-*divide*-ka. Kodwa sizokhuluma nge-*quantum physics.*	We'll get to that topic on the atom one day as scientists long since discovered that the atom itself still is divisible. But we will later talk about quantum physics.

an experienced English-to-Zulu translator and an experienced science communicator, supported by people who believe that science should engage and include all members of society.

In this Commentary, we have tried to show that South Africans need more words to talk about climate change, black holes, energy, evolution, technology, health and other scientific topics in their mother tongue, and that it is in fact possible to translate science into indigenous languages. In sharing the rationale behind translating science, and how we try to engage people who speak indigenous languages in science, we hope it is clear how much of our 'passion', 'vigour' and 'curiosity' goes into science communication projects like iLukuluku.

If you also care about communicating science beyond English and want to know more, or if you are able to sponsor a season of iLukuluku, please email us at info@sciencelink.co.za.

Competing interests

We have no competing interests to declare.

References

1. Africa K. Zulu podcast about science breaks barriers. Independent Online. 24 April 2022. Available from: https://www.iol.co.za/weekend-argus/lifestyle/zulu-podcast-about-science-breaks-barriers-3352c27b-f568-498b-a5d4-dbb05f10965e
2. Levin M. Language as a barrier to care for Xhosa-speaking patients at a South African paediatric teaching hospital. S Afr Med J. 2006;96(10):1076–1079. https://pubmed.ncbi.nlm.nih.gov/17164939/
3. Prah KK. Challenges to the promotion of indigenous languages in South Africa. Cape Town: Centre for Advanced Studies of African Society; 2007.
4. Kago G, Cissé M. Using African indigenous languages in science engagement to increase science trust. Front Commun. 2021;6:40–46. https://doi.org/10.3389/fcomm.2021.759069
5. De Kadt J. Language development in South Africa – Past and present. The Politics of Language in South Africa. Princeton, NJ: Woodrow Wilson School of Public and International Affairs, Princeton University; 2005.
6. McKeever M. Educational inequality in apartheid South Africa. Am Behav Sci. 2017;61(1):114–131. https://doi.org/10.1177/0002764216682988
7. Du Toit N. Open Stellenbosch: A university education in exclusion. Daily Maverick. 25 August 2015. Available from: https://www.dailymaverick.co.za/opinionista/2015-08-25-open-stellenbosch-a-university-education-in-exclusion/
8. Smith LT. Decolonizing methodologies: Research and indigenous peoples. London: Bloomsbury Academic; 2012.
9. Fish D, Allie S, Pelaez N, Anderson TA. Cross-cultural comparison of high school students' responses to a science centre show on the physics of sound in South Africa. Public Underst Sci. 2017;26(7):806–814. https://doi.org/10.1177/0963662516642725

10. Dlodlo TS. Science nomenclature in Africa: Physics in Nguni. J Res Sci Teach. 1999;36(3):321–331. https://doi.org/10.1002/(SICI)1098-2736(199903)36:3%3C321::AID-TEA6%3E3.0.CO;2-8

11. Biyela S. Ososayensi Bathole Isilwane Sasemandulo Esandisa Umlando Wohlu Lwezilwane [Scientists have discovered an ancient animal that extends the history of the animal list]. SciBraai. 14 November 2018. Zulu. Available from: https://scibraai.co.za/ososayensi-bathole-isilwane-sasemandulo-esandisa-umlando-wohlu-lwezilwane/

12. Biyela S, Msomi N. The Periodic Table of the Elements [podcast on the Internet]. iLukuluku; 2022. Available from: https://podcasts.apple.com/us/podcast/periodic-table-of-elements/id1617276271?i=1000555949609

13. Alexander N. The potential role of translation as social practice for the intellectualisation of African languages. Cape Town: University of Cape Town; 2005.

AUTHORS:
Mduduzi Mbiza[1]
Saurabh Sinha[2]

AFFILIATIONS:
[1]Research Associate, University of Johannesburg, Johannesburg, South Africa
[2]Professor and Deputy Vice-Chancellor: Research and Internationalisation, University of Johannesburg, Johannesburg, South Africa

CORRESPONDENCE TO:
Saurabh Sinha

EMAIL:
ssinha@uj.ac.za

HOW TO CITE:
Mbiza M, Sinha S. The Fourth Industrial Revolution: Conceptual paradox or catalyst for achieving the Sustainable Development Goals? S Afr J Sci. 2023;119(7/8), Art. #16090. https://doi.org/10.17159/sajs.2023/16090

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
Fourth Industrial Revolution, Sustainable Development Goals, artificial intelligence, South Africa, machine learning

FUNDING:
US Fulbright Grant (#PS00332353), South African National Research Foundation

PUBLISHED:
11 May 2023

The Fourth Industrial Revolution: Conceptual paradox or catalyst for achieving the Sustainable Development Goals?

Significance:

The topic of the Fourth Industrial Revolution (4IR) became significant in South Africa from 2017, through advocacy, amongst others, by the University of Johannesburg and subsequently through the appointment of the Presidential Commission on 4IR. Preceding industrial revolutions each focused on a single technology; 4IR, however, speaks to a confluence of technologies and a synergy of computing, data, and communications technology, with artificial intelligence rapidly redefining the world of work. Conceptual and geopolitical challenges and potential negative societal implications notwithstanding, we argue that the 4IR paradigm shift is critical to South Africa and to realising the Sustainable Development Goals.

Introduction

Around 2017, while not new in nomenclature, a strong pitch for the Fourth Industrial Revolution (4IR) developed in South Africa. The University of Johannesburg (UJ) was a key advocate, and 4IR soon became a significant undertaking through various national and regional processes, such as the appointment of the Presidential Commission on 4IR and the role of UJ as a Pan-African epicentre for critical intellectual inquiry. UJ, supported by its Council, committed to over ZAR500 million (UJ Council resolution, March 2019) in 4IR and associated change management processes over a 5-year period. UJ's systemic and systematic approach offered both breadth (a transdisciplinary approach) and depth (progressing the technical discipline of artificial intelligence (AI) specifically). In terms of the latter, UJ hosts the Institute for Intelligent Systems, now a significant partner for both government and industry.

The conceptual debate

Around the world, and in South Africa, 4IR as a concept is much debated. At the centre of this debate is the fact that previous industrial revolutions had at their core a single revolutionary breakthrough or innovation. The First Industrial Revolution occurred during the late 1700s and early 1800s, and was characterised by the rise of steam-powered machinery and widespread utilisation of the power loom. The Second Industrial Revolution, from the late 1800s to the early 1900s, introduced mass production techniques and the widespread adoption of electricity. The Third Industrial Revolution during the latter half of the 1900s was propelled by advancements in computer and digital technologies.

4IR, on the other hand, is characterised by the convergence of multiple technological breakthroughs, particularly in AI and machine learning. It enables physical, digital, and biological systems to converge, leading to smart factories, integrated workflows and new lifestyles. 4IR is significantly different from past industrial transformations due to the combination of technologies involved, and is expected to have a much wider range and a stronger impact on people's lives than previous industrial revolutions.[1]

AI integration into several industries is already in progress and is anticipated to bring considerable benefits. The technology is expected to drive productivity and efficiency, by augmenting repetitive tasks and processing massive amounts of data. Moreover, new jobs are being created in fields such as data analysis and software engineering for those skilled with AI-based products, services and the product–service continuum. By incorporating AI into specific areas, novel products and services will be developed.

4IR is facilitating rapid technological evolution, which presents a range of open questions in terms of conceptualisation and societal implications. These current developments will surely provide novel solutions to existing problems, or cause disruption – or both. Stakeholders must be prepared to navigate this complex dynamic going forward.

The term 'industrial revolution' is coming under scrutiny, given that the current landscape of technology has shifted, leading to a need for a more nuanced and comprehensive term. The increasing complexity and interconnectedness of technologies requires a broader interpretation than is traditionally offered by the phrase 'industrial revolution'. A more accurate phrase is needed to appropriately reflect the current state of technological advancement.

We argue that there are many indications that the period of significant transformation we are currently experiencing can indeed be termed the 4IR. Emerging technologies like AI, robotics, the Internet of Things, blockchain and quantum computing are revolutionising many aspects of our lives, such as workflows, communication systems, and access to information. Everywhere we look, there are indications that we are experiencing an enormous transformation – ranging from how we interact with one another to how we gain access to information.

Exponential growth in AI technology is largely attributed to advancing computing abilities, bolstered data storage and processing capabilities, and the widespread utilisation of high-speed Internet access. This acceleration in AI development and deployment has been heavily facilitated by technological advances made over the last few decades.

The Internet of Things, connected devices, and cloud computing have democratised AI solutions. This means that businesses, of any size, can now access and deploy AI solutions to automate processes across different

spheres. As a result of these developments, a new wave of innovators is arriving on the scene – and leveraging AI to tackle various economic and societal issues. All these developments do indeed suggest that the term 'revolution' is appropriate.

From a geopolitical perspective, 4IR is being driven, primarily by the West, especially the World Economic Forum. This raises the important question of whether this agenda accurately reflects the needs of other regions. However, when one sees increasing technological and AI developments occurring, independently, in China, East Asia, and parts of India, it is possible to conclude that 4IR has a larger global reach than initially thought. This emphasises the significance Africa should play from its unique perspective to keep up with global trends and to redefine 4IR to promote socio-economic inclusivity.

As 4IR has emerged as a major driving force of innovation, technology, and economic growth on the global stage, South Africa has taken note. It is imperative that South Africa capitalises on the advances of 4IR to further its development objectives. Therefore, this topic deserves ongoing discussion as it continues to be of utmost importance in the current landscape.

South Africa is currently confronted with an energy crisis, which requires multiple approaches to be tackled successfully. In this context, 4IR technologies and thought paradigms might well be useful when searching for viable solutions. However, this situation is too complex to be adequately explored in this article.[2]

With escalating global competition, South Africa has strategic opportunities to take the steps needed to bring its development up to date. The implementation of 4IR technologies and advancements can be used to propel the nation to overcome roadblocks that have previously held it back; these include poverty, inequality, and unemployment. The adoption of technological innovation could be key towards accelerating the development needed to improve South Africa's future. However, automation and AI technologies have undoubtedly created job losses and increased income inequality. The implementation of these systems has also sparked debates about privacy and security concerns as well as raised questions regarding ethics and bias in AI.

In this article, we argue that 4IR is crucial for South Africa. 4IR has instigated a necessary change in perspective and, by fostering innovation, is revitalising optimism and presenting opportunities for rapid advancement. Additionally, we contend that 4IR or similar thinking will be vital in fulfilling the United Nations' Sustainable Development Goals (SDGs).[3] These 17 goals were established by the United Nations in 2015 to advance sustainable development and tackle pressing global issues, including education, poverty, and health and well-being. Given that 4IR is at the forefront of worldwide innovation, South Africa has the opportunity to leverage technological innovations to attain the SDGs.

At the same time, South Africa has the opportunity to contribute in alternative ways to 4IR's advancement and to the SDGs. For example, by incorporating African value systems into developing technologies, such as the creation of the operating system Ubuntu (Linux) to ensure that 'Ubuntu' principles were incorporated. Ubuntu is a societal value system from the African continent and stands for 'I am because you are'. Ubuntu, the operating system was founded by Mark Shuttleworth, a South African entrepreneur. The Ubuntu value system and the associated operating system have prolifically progressed open access globally. This brings a unique edge to several SDGs – including SDG 17 (Sustainable Partnerships).[4] The worldwide goal of achieving the 17 SDGs by 2030 seeks to ensure that no one is left behind and there is a brighter and more sustainable future for all. All nations are called upon to advance sustainability and preserve the earth, regardless of wealth and disadvantage.

We recognise that the implementation of 4IR in South Africa has challenges. The country must address the digital divide, the unequal distribution of technology and Internet access. This is a significant challenge for South Africa, as many of its citizens still lack resources such as technology and Internet connectivity, which makes it difficult for them to participate in 4IR. Compounding all of these, South Africa's digital divide is also a function of extreme social inequality.

Some argue that South Africa, owing to the current energy crisis, should instead pursue 3IR. While we recognise this approach, the 4IR thought paradigm allows individuals to leapfrog in thinking – for example, instead of thinking of only an electricity grid, one can think about a smart grid. The latter, with AI enablement, creates an array of opportunities – innovative business models, policies, technological advancements, a sustainable environment, and others. Furthermore, energy infrastructure is a long-term investment, and if these issues are not considered at the outset, the costs of addressing these at a later stage will be significantly higher.

South Africa must also enhance its education and training to ensure that its citizens have the skills necessary to advance 4IR and reap the benefits thereof. Like the rest of Africa, South Africa has the demographic advantage of a young population; this group is technology-savvy and could be rapidly deployed to form a '4IR movement'. Universities are a microcosm of society. Considering this analogy in terms of 4IR, UJ is an exceptional example of what can be achieved; the university's 2021/22 ranking as number 1 in South Africa by the Times Higher Education (THE) Impact Ranking further demonstrates international acknowledgement.

Fighting poverty through 4IR

The global community continues to grapple with poverty – an issue that has been further exacerbated by the COVID-19 pandemic. With millions pushed into poverty and economic hardship, solutions are urgently needed. Fortunately, 4IR technologies have tremendous potential for alleviating poverty and improving living standards. It is up to policymakers and stakeholders around the world to identify ways in which 4IR can be harnessed to make a real difference in the fight against poverty.

One of the critical mechanisms of 4IR in helping to end poverty is the gathering and analysis of data. With the increased use of technologies such as big data and AI, vast amounts of information can be collected about poverty levels, living standards, and the distribution of resources. This information can inform policies and initiatives to reduce poverty, making them more effective and targeted. Data analytics must, however, be done at an aggregate level to ensure the highest ethical standards.

Another way that 4IR can contribute to ending poverty is by creating new economic opportunities. The digital economy and e-commerce enabled by 4IR technologies provide new avenues for economic growth and job creation, especially in developing countries where traditional employment opportunities may be limited. The digital economy is particularly well suited to the needs of developing countries, as it is less dependent on physical infrastructure and can provide people with access to global markets and customers. For instance, in 2021, the South African government introduced the ICT special economic (tax) zone incentivising the establishment of businesses in the virtual space.

Improving education through 4IR

SDG 4 requires that we ensure equitable quality education for all and create lifelong learning opportunities. AI and machine learning can be leveraged to optimise and personalise educational content, customising it according to individual learner needs and parameters.

On the other hand, by using gamification, virtual and augmented reality can make education more engaging, interactive, and enjoyable for learners, helping to foster a lifelong love of learning.[5] Technologies can also provide a more inclusive learning platform or bridge for students who are differently enabled.

Data collection and analysis using educational technologies is an effective way to gain insight into the learning process. This information can then be used to help teachers and administrators make decisions about how to enhance student learning outcomes.

Solving water-related crises through 4IR

SDG 6 aims to provide universal clean water and access to sanitation. Despite the current efforts of the South African government, work towards achieving this goal needs to accelerate due to the rapidly increasing demand for reliable water sources, and especially in view of climate change.

The water industry is gradually beginning to adopt 4IR technologies. Countries worldwide are exploring ways to introduce these innovations to improve public health and simultaneously increase economic opportunities for their citizens. Implementing such measures will bring many benefits over time.

Improving health and well-being through 4IR

The health sector has already drawn extensively on 4IR. Early evidence indicates that this has had an immense impact.[6]

4IR can improve access to health care and support the prevention and management of disease. For example, telemedicine and mobile health applications can provide remote consultations and support to people who live in distant areas or have limited access to healthcare facilities. Wearable technology and sensors can also monitor health metrics in real time, allowing for early detection of health issues and prompt intervention. Furthermore, AI advances traditional methods of diagnostics; this complementary approach extends the limited resources available in developing parts of the world.

Recommendations

Governments, businesses, and civil society organisations must work together to ensure all communities can access affordable, reliable, high-speed Internet connectivity. This includes investing in digital infrastructure and providing digital skills training with particular emphasis on youth and historically disadvantaged communities. By advancing the ICT economic zone initiative, governments can also provide subsidies or other incentives to promote the adoption of new technologies and entrepreneurship.

Additionally, governments can establish regulations and standards that safeguard personal information and prevent cyberattacks. Zero-rating of education opportunities, including around cybersecurity, can be another effective approach.

Governments can also establish partnerships with the private sector, civil society organisations, and academia to drive innovation and promote the development of new and appropriate technologies that further advance the SDGs.

In an overarching manner, and across the three tiers of government, recommendations such as those made in the Presidential Commission's report[7] must also be implemented.

Conclusion

While we acknowledge the conceptual definitional 'mismatch' of 4IR in the context of industrial revolutions, the confluence of technologies, and particularly AI, is significant today, and requires focused and continued attention. For this reason, we agree with the World Economic Forum and others that this AI-associated confluence deserves 'revolutionary' recognition.

It is crucial that the benefits of 4IR are shared equitably across society and that countries address any negative implications. Going forward, proactive steps must be taken by governments, businesses, civil society, and other stakeholders to fully understand the impact of technology on society.

We must take a multifaceted approach to technology that considers its implications in the broader societal and political context. A more integrated method of technological progress, which draws upon the insights of various disciplines and stakeholders, will foster a sense of collective ownership. What is needed is an innovative framework for developing technology with an ethical underpinning which emphasises fairness, sustainability, and justice.

Acknowledgements

S.S. is supported by a US Fulbright Grant (#PS00332353). In the South African context, US Fulbright is jointly supported by the South African National Research Foundation. During compilation of this manuscript, S.S. was based, as part of the US Fulbright undertaking, at Princeton University.

Competing interests

We have no competing interests to declare.

References

1. Morgan J. Will we work in twenty-first century capitalism? A critique of the fourth industrial revolution literature. Econ Soc. 2019;48(3):371–398. https://doi.org/10.1080/03085147.2019.1620027
2. David LO, Nwulu NI, Aigbavboa CO, Adepoju OO. Integrating fourth industrial revolution (4IR) technologies into the water, energy & food nexus for sustainable security: A bibliometric analysis. J Clean Prod. 2022;363, Art. #132522. https://doi.org/10.1016/j.jclepro.2022.132522
3. Mbiza M, Sinha S. Technology and sustainable development: A hamlet in rural South Africa shows how one can power the other. The Conversation. 2023 January 17 [cited 2023 Mar 16]. Available from: https://theconversation.com/technology-and-sustainable-development-a-hamlet-in-rural-south-africa-shows-how-one-can-power-the-other-197355
4. Sinha S, Lutchman V. Transdisciplinary education: Enabling the Sustainable Development Goals using the Fourth Industrial Revolution. In: Leal Filho W, Pretorius R, de Sousa LO, editors. Sustainable development in Africa. World Sustainability Series. Cham: Springer; 2021. https://doi.org/10.1007/978-3-030-74693-3_9
5. Du Preez J, Sinha S. A paradigm shift in higher education in the context of the Fourth Industrial Revolution. IEEE Potentials. 2021;40(2):13–18. https://doi.org/10.1109/MPOT.2020.3044279
6. Castro e Melo JAG de M e, Faria Araújo NM. Impact of the Fourth Industrial Revolution on the health sector: A qualitative study. Healthc Inform Res. 2020;26(4):328–334. https://doi.org/10.4258/hir.2020.26.4.328
7. South African Government. Report of the Presidential Commission on the 4th Industrial Revolution [document on the Internet]. c2020 [cited 2023 Apr 22]. Available from: https://www.gov.za/documents/report-presidential-commission-4th-industrial-revolution-23-oct-2020-0000

AUTHOR:
Chris Callaghan[1]

AFFILIATION:
[1]School of Management, Anglia Ruskin University, Chelmsford, UK

CORRESPONDENCE TO:
Chris Callaghan

EMAIL:
chris.callaghan@aru.ac.uk

HOW TO CITE:
Callaghan C. Revisiting how scientific research drives technological change: The Fifth Industrial Revolution. S Afr J Sci. 2023;119(7/8), Art. #15556. https://doi.org/10.17159/sajs.2023/15556

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
Fourth Industrial Revolution, Fifth Industrial Revolution, global productivity growth, industrial revolutions

PUBLISHED:
11 May 2023

Revisiting how scientific research drives technological change: The Fifth Industrial Revolution

Significance:

Moll, Marwala, and Ntlatlapa highlight salient criticisms of terminologies and definitional uncertainties associated with the term 'Fourth Industrial Revolution' (4IR). Scientific research on technological change seems to suggest a need for theoretical synthesis to address a failure of 4IR notions to consider the central role of a *revolution in the scientific/knowledge creation process itself* – that is seemingly a causal driver of current technological and societal changes. The term 'Fifth Industrial Revolution' might helpfully be used to differentiate 4IR debates from those deriving from revolutionary changes in science itself that may underlie our current trajectory of technological change.

Moll[1] argues there has been "no technological revolution, let alone a 'Fourth Industrial Revolution'" (4IR). He refers to Schwab's[2] World Economic Forum discussions of the 4IR:

> *We have yet to grasp fully the speed and breadth of this new revolution. Consider the unlimited possibilities of having billions of people connected by mobile devices, giving rise to unprecedented processing power, storage capabilities and knowledge access. Or think about the staggering confluence of emerging technology breakthroughs, covering wide-ranging fields such as artificial intelligence, robotics, the Internet of Things, autonomous vehicles, 3D printing, nanotechnology, biotechnology, materials science, energy storage and quantum computing, to name a few. Many of these innovations are in their infancy, but they are already reaching an inflection point in their development as they build on and amplify each other in a fusion of technologies across the physical, digital and biological worlds.*

For Moll[1]: "What I want to question is the way we use the word 'revolution' to describe and understand our activities. On that score, I argue that we have not witnessed a 'grand', overall technological revolution in recent times. It is important that scientists and technologists understand this." In response, referencing Kuhn, Marwala[3] argues the 4IR represents "a scientific paradigm shift" and that Moll's "argument that the 4IR does not constitute a revolution is thus unfounded". Ntlatlapa[4] also challenges Moll's conclusion that changes associated with the 4IR fail to meet "the criteria for an industrial revolution".

These debates reflect contested definitional understandings of technological change in scientific literature. Scholarly scientific perspectives of contemporary technological change, and its societal implications, abound, typically couched in metaphors such as that of 4IR[2], Industry 4.0[5], the First, Second, and Third Industrial Revolutions[6], and the 'second machine age'[7], amongst others.

Research on the topic has sought to discover core theoretical mechanisms, or fundamental causal drivers of contemporary technological changes. Recent studies of contributions of technological changes in scientific research and biomedicine[8,9] and disaster response[10] suggest some patterns in how roots of societal technological change might derive from innovations in the scientific or knowledge creation process itself. These patterns are relevant to the debates Moll, Marwala, and Ntlatlapa engage with.

These patterns suggest the central role of innovations in the scientific process itself are not sufficiently considered in popular debates about industrial revolutions. Revisiting the key notion of 'productivity', and relating productivity revolutions to some scientific ideas, one might describe the current 'revolution' we are about to fully experience, as a 'Fifth Industrial Revolution' (5IR).[11,12]

In light of Moll's, Marwala's, and Ntlatlapa's discussions, one might criticise the introduction of yet another term for the confluence of technological change we face. Indeed, many technological developments, including ChatGPT, have been associated with a dramatic change in technology use, and, irrespective of terms like 4IR and 5IR, these impacts are tangible and quantifiable. However, the 5IR stream of literature[11,12] seems concerned with a 'revolution' in productivity that becomes evident or measurable at the aggregate level (which is arguably yet to materialise – see Robert Gordon's extensive work on the topic).

It is entirely possible that AI might now 'show up' in the aggregate productivity statistics and drive a global productivity revolution. In that the 5IR literature focuses on the aggregate level and is concerned with identifying a 'revolutionary' increase in global productivity growth, we might therefore expect this to inevitably occur at some time in the future.

The 5IR literature has built on the 4IR literature, and is complementary to previous literature. As such, it does not seek to 'replace' the 4IR in current terminology because 5IR is concerned with extending scientific work on the aggregate relationships between technological change and global productivity growth.

Why is the aggregate level important? Schumpeterian theory predicts that, as new technologies develop, they can disrupt pre-existing products, processes, businesses, and other previous knowledge recipes – making them obsolete.[13] Thus, technological innovation can also create obsolescence, and it is only at the aggregate level that we can see the net results of overall systemic change. Technological change is typically path dependent[14,15] and there is no guarantee that it proceeds in a societally optimal way. An example of this is the QWERTY keyboard,

which was built to slow down typing to accommodate limitations of mechanical typewriters. 5IR literature therefore seems interested in overall, or aggregate, productivity growth.

Therefore, 5IR might be a useful term to describe a synthesis of the proliferation of some of the definitions and descriptive popular debates associated with the 4IR, to build a separate literature that extends these debates with a focus on scientific work on global productivity growth. Two rationales summarise some patterns and trends in the scientific literature on technological change.

First, the 5IR concept acknowledges widespread popular use of the term 4IR, *but the new term 5IR* may be necessary to re-anchor debates about productivity revolutions in scientific work often ignored in popular discussions.[11]

The primary industrial revolution has been discussed in terms of its radical productivity enhancements that caused radical societal change.[16] As discussed in the following sections, anchoring debates to scientific work on productivity may offer us a more scientific (historical) definition, as few would disagree that radical societal change has historically occurred due to productivity revolutions.

Second, and relatedly, building on Moll's[1] criticism of "fanciful, rhetorical, science-fiction like evocations", setting aside popular (or populist) notions of these phenomena, and other descriptions of the 4IR that are largely descriptive and atheoretical, one might ask, *what does scientific study of technological change tell us*? Is there a pattern that is grounded in scientific work that can unite popular debates?

Many popular debates seemingly do not sufficiently draw from seminal scientific notions of technological change[17] and the evolution of this literature that explicitly models it theoretically[13,18,19]. Similarly, these debates do not seem to engage sufficiently with current ongoing scientific conversations on the topic.[20-22]

What seemingly unites the scientific literature is a simple notion – that many historical advances, or revolutionary improvements in quality of life, in human history were *driven by the same phenomenon*: a step change in the way we generated knowledge – the effectiveness of our scientific system itself[23]. In Nielsen's[23] words:

> *Revolutions are sometimes marked by a single, spectacular event…But often the most important revolutions aren't announced with the blare of trumpets…We are in the midst of a great change in how knowledge is constructed…A change of similar magnitude is going on today: we are reinventing discovery…To historians looking back a hundred years from now, there will be two eras of science: pre-network science, and networked science. We are living in the time of transition to the second era of science.*

Nielsen gives examples of very large-scale projects, such as the Polymath Project (solving difficult mathematical problems using crowdsourcing), GenBank (collecting global genetic information), Wikipedia (the online encyclopaedia), and Galaxy Zoo (mapping the galaxy using global participants), that seem to demonstrate revolutionary productivity in scientific knowledge production.

Nielsen's examples seem to highlight the workings of some causal mechanisms underlying these radical productivity improvements, and some of the 'why' of current technological change. Building on Nielsen's work, the term 5IR might therefore usefully provide a synthesis and a clear logic (differentiated from fragmented 4IR thinking) to argue that a productivity revolution is currently underway due to the causal influence of *radical innovations in the scientific research or knowledge creation system itself*.

Thus, instead of a 4IR focus largely on the outputs of this revolution, or its 'contents', 5IR thinking should focus more precisely on the fundamental cause of this change. This change seems to be caused by a revolution in how we generate knowledge – not simply regarding (big) data and information – but in the scientific/knowledge creation process itself[9] and a coming paradigm change in scientific research itself[24].

Nielsen's[23] explanations of the causal mechanisms underlying this revolution also seem to reconcile an important paradox in the science/technology literature – of the failed predictions of Romer's[25] and other endogenous growth models, as well as recombinant growth theory[26], to reflect in evidence of constant or increasing returns to scale of idea creation. Romer's Nobel Prize winning work suggests that productivity results from non-rivalrous 'knowledge recipes' that, once created, can often be used by others at little cost. In short, this should ultimately improve productivity over time, with increasing returns to scale in knowledge creation. This body of work is based on rigorous theoretical modelling, and this failure seems to be at the crux of debates about the failure of technological change to deliver radical aggregate productivity enhancements. It recalls the longstanding Solow Paradox, whereby, according to Solow himself in the late 1980s, you "can see the computer age everywhere but in the productivity statistics".[20,27]

This failure seems well documented[28-33], and also seems to be related to what Jones[28] describes as a 'burden of knowledge' effect, analogous to Romer's own description of a 'fishing out' effect in research. The argument here is that science – *the research process itself*, and theoretical models featured in discussions deemed worthy enough to earn their theorists Nobel awards – seem to underlie predictions of the fundamental trajectory of technological change.

In short, Romer's work seemingly has a radical implication – that returns to the creation of ideas could possibly be exponential, demonstrating increasing returns over time. Nielsen's work may offer evidence of how these returns can be achieved, and reasons for why social structures of science are slowing these shifts. As suggested by Neo-Schumpeterian theory, economic and technological advances are typically held back by socio-institutional forces of human societies.[34]

Moll, Marwala, and Ntlatlapa's insightful discussions provide a useful reference point to highlight the broad and fragmented nature of broader 4IR debates. These 4IR debates could be supplemented to include reference to an extensive body of theoretical and empirical work on how changes in the knowledge creation/scientific process itself may also explain the current trajectory of technological change. The term 5IR might helpfully differentiate 4IR work to include a primary focus on science itself and a revolution within science. This revolution in the processes of science may fundamentally be driving a mode transition, from declining returns to idea creation to increasing returns. As Nielsen[23] predicts, there will ultimately be no doubt that we are in the midst of a revolution, if the revolution entails the reinvention of discovery itself.

In conclusion, some implications derive from discussions above that extend previous work on predicting outcomes of technological change.[35]

First, Amara's Law, that 'we tend to overestimate the effect of a technology in the short run and underestimate the effect in the long run'[36] implies that if the radical changes predicted by Romer and Nielsen do arise from idea proliferation then they will be profound. Although some have argued that the most impactful technological changes have largely been limited to Internet and communication technologies and entertainment[7] – which seemingly supports some of Moll's observations that large-scale change is yet to be experienced – 5IR logics suggest that large-scale societal change will now accelerate.

Second, and finally, probabilistic innovation theory[10] predicts (perhaps fancifully) not only increasing returns to scale in research, but ultimately that real-time research productivity will be possible, and that research will ultimately be able to solve all 'solvable' problems – problems that are inherently solvable. We might all agree that not even 'fanciful, rhetorical, or science-fiction-like evocations' can describe what such a future would look like. More critical engagement in these conversations is urgently needed.

Competing interests

There are no competing interests to declare.

References

1. Moll I. Why there is no technological revolution, let alone a 'Fourth Industrial Revolution'. S Afr J Sci. 2023;119(1/2), Art. #12916. https://doi.org/10.17159/sajs.2023/12916
2. Schwab K. The Fourth Industrial Revolution. Geneva: World Economic Forum; 2016.
3. Marwala T. The Fourth Industrial Revolution has arrived. Comments on Moll (S Afr J Sci. 2023;119(1/2), Art. #12916). S Afr J Sci. 2023;119(1/2), Art. #15429. https://doi.org/10.17159/sajs.2023/15429
4. Ntlatlapa N. Defining the Fourth Industrial Revolution. Comments on Moll (S Afr J Sci. 2023;119(1/2), Art. #12916). S Afr J Sci. 2023;119(1/2), Art. #15436. https://doi.org/10.17159/sajs.2023/15436
5. Agostini L, Nosella A. Industry 4.0 and business models: A bibliometric literature review. Bus Process Manag J. 2021;27(5):1633–1655. https://doi.org/10.1108/BPMJ-03-2021-0133
6. Rifkin J. The Third Industrial Revolution: How lateral power is transforming energy, the economy, and the world. London: Macmillan; 2011.
7. Brynjolfsson E, McAfee A. The second machine age: Work, progress, and prosperity in a time of brilliant technologies. New York: WW Norton & Company; 2014.
8. Callaghan CW. Crowdsourced 'R&D' and medical research. Br Med Bull. 2015;115(1):67–76. https://doi.org/10.1093/bmb/ldv035
9. Callaghan CW. Developing the transdisciplinary aging research agenda: New developments in big data. Curr Aging Sci. 2018;11(1):33–44. https://doi.org/10.2174/1874609810666170719100122
10. Callaghan CW. Disaster management, crowdsourced R&D and probabilistic innovation theory: Toward real time disaster response capability. Int J Disaster Risk Reduct. 2016;17:238–250. https://doi.org/10.1016/j.ijdrr.2016.05.004
11. Callaghan CW. Transcending the threshold limitation: A fifth industrial revolution? Manag Res Rev. 2019;43(4):447–461. https://doi.org/10.1108/MRR-03-2019-0102
12. Callaghan CW. The global productivity growth and research productivity declines: The (urgent) need for a 'fifth industrial revolution' imperative. Int J Bus Innov Res. 2021;24(2):197–217. https://doi.org/10.1504/IJBIR.2021.112816
13. Aghion P, Howitt P. A model of growth through creative destruction. Econometrica. 1992;60(2):323–351. https://doi.org/10.2307/2951599
14. Arthur WB. Comment on Neil Kay's paper – 'Rerun the tape of history and QWERTY always wins'. Res Policy. 2013;42(6):1186–1187. https://doi.org/10.1016/j.respol.2013.01.012
15. Kay NM. Rerun the tape of history and QWERTY always wins. Res Policy. 2013;42(6):1175–1185. https://doi.org/10.1016/j.respol.2013.03.007
16. Mokyr J. The contribution of economic history to the study of innovation and technical change: 1750-1914. In: Hall BH, Rosenberg N, editors. Handbook of the economics of innovation: Volume 1. Amsterdam: Elsevier; 2010. p. 11–50. https://doi.org/10.1016/S0169-7218(10)01002-6
17. Solow RM. A contribution to the theory of economic growth. Q J Econ. 1956;70(1):65–94. https://doi.org/10.2307/1884513
18. Romer PM. Endogenous technological change. J Political Econ. 1990;98(5Pt2):S71–S102. https://doi.org/10.1086/261725
19. Grossman GM, Helpman E. Quality ladders in the theory of growth. Rev Econ Stud. 1991;58(1):43. https://doi.org/10.2307/2298044
20. Acemoglu D, Dorn D, Hanson GH, Price B. Return of the Solow paradox? IT, productivity, and employment in US manufacturing. Am Econ Rev. 2014;104(5):394–399. https://doi.org/10.1257/aer.104.5.394
21. Acemoglu D, Restrepo P. Tasks, automation, and the rise in US wage inequality. Econometrica. 2022;90(5):1973–2016. https://doi.org/10.3982/ECTA19815
22. Acemoglu D, Restrepo P. Automation and new tasks: How technology displaces and reinstates labor. J Econ Perspect. 2019;33(2):3–30. https://doi.org/10.1257/jep.33.2.3
23. Nielsen M. Reinventing discovery: The new era of networked science. Princeton, NJ: Princeton University Press; 2012. https://doi.org/10.1515/9781400839452
24. Hey T, Tansley S, Tolle KM. The Fourth Paradigm: Data-intensive scientific discovery. Redmond, WA: Microsoft Research; 2009.
25. Romer PM. The origins of endogenous growth. J Econ Perspect. 1994;8(1):3–22. https://doi.org/10.1257/jep.8.1.3
26. Weitzman ML. Recombinant growth. Q J Econ. 1998;113(2):331–360. https://doi.org/10.1162/003355398555595
27. Solow RM. You can see the computer age everywhere but in the productivity statistics. New York Review of Books. 1987:36.
28. Jones BF. The burden of knowledge and the "death of the renaissance man": Is innovation getting harder? Rev Econ Stud. 2009;76(1):283–317. https://doi.org/10.1111/j.1467-937X.2008.00531.x
29. Jones CI. Growth: With or without scale effects? Am Econ Rev. 1999;89(2):139–144. https://doi.org/10.1257/aer.89.2.139
30. Jones CI. R & D-based models of economic growth. J Political Econ. 1995;103(4):759–784. https://doi.org/10.1086/262002
31. Gordon RJ. U.S. productivity growth: The slowdown has returned after a temporary revival. International Productivity Monitor. 2013(25):13–19.
32. Cowen T. The great stagnation: How America ate all the low-hanging fruit of modern history, got sick, and will (eventually) feel better: A Penguin eSpecial from Dutton. New York: Penguin; 2011.
33. Callaghan CW. Lessons and insights from the global productivity slowdown: A research management agenda. Afr J Sci Technol Innov Dev. 2022;14(5):1265–1273. https://doi.org/10.1080/20421338.2021.1945775
34. Castellacci F. Evolutionary and new growth theories. Are they converging? J Econ Surv. 2007;21(3):585–627. https://doi.org/10.1111/j.1467-6419.2007.00515.x
35. Callaghan CW. Surviving a technological future: Technological proliferation and modes of discovery. Futures. 2018;104:100–116. https://doi.org/10.1016/j.futures.2018.08.001
36. Ratcliffe S. Oxford essential quotations. 6th ed. Oxford: Oxford University Press; 2018. https://doi.org/10.1093/acref/9780191866692.001.0001

Are we sinking African cheetahs in India?

AUTHORS:
Kelly A. Marnewick[1]
Michael J. Somers[2]
Jan A. Venter[3]
Graham I.H. Kerley[4]

AFFILIATIONS:
[1]Department of Nature Conservation, Tshwane University of Technology, Pretoria, South Africa
[2]Eugène Marais Chair of Wildlife Management, Mammal Research Institute, Centre for Invasion Biology, Department of Zoology and Entomology, University of Pretoria, Pretoria, South Africa
[3]Department of Conservation Management, Nelson Mandela University, George, South Africa
[4]Centre for African Conservation Ecology, Nelson Mandela University, Gqeberha, South Africa

CORRESPONDENCE TO:
Kelly Marnewick

EMAIL:
MarnewickKA@tut.ac.za

HOW TO CITE:
Marnewick KA, Somers MJ, Venter JA, Kerley GIH. Are we sinking African cheetahs in India? S Afr J Sci. 2023;119(7/8), Art. #15617. https://doi.org/10.17159/sajs.2023/15617

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
cheetahs, India, sink, restoration, reintroduction

PUBLISHED:
15 May 2023

Significance:
The current initiative to export African cheetahs to India has a limited scientific basis, placing the Asian subspecies and the translocated animals at risk. There is no evidence that this will benefit African cheetah conservation. We call for a globally coordinated approach to cheetah conservation, based on sound science.

Failed conservation actions waste money and can result in sink populations where species decline and ultimately may go extinct. We, therefore, cannot afford any failed conservation efforts, and poor conservation decisions should be avoided. The recent translocation of 20 African cheetahs, *Acinonyx jubatus jubatus* – 8 from Namibia and 12 from South Africa – to a reserve in India, with the aim to establish a free-ranging population of cheetahs in and around the release site, and with further multiple translocations planned from South Africa, raises concerns regarding the scientific basis of these translocations and their contribution to conservation. Recent statements by project scientific advisory members that it is actually an "experimental reintroduction of cheetahs into India"[1], suggesting that the outcome is uncertain, raises additional ethical concerns.

We clarify these concerns and suggest future actions in the context that this initiative may serve as a sink for African cheetahs.

1. Is it good science?

The Asiatic cheetah, *A. j. venaticus,* is extinct in India and should not be replaced, without appropriate scientific consideration, by the African cheetah. The relict population of the Asiatic cheetah in Iran holds the only extant members of the subspecies with an estimate of 50 mature individuals.[2] Genetic evidence points to historical translocations of African cheetahs to India[3], and others even suggest that the cheetah was never indigenous to India[4]. This makes the Iranian population even more valuable, as it does not appear to have been hybridised through anthropogenic contact with the African cheetah.[3] The genetic integrity of the African and Asian cheetah lineages should therefore be maintained.[5] The current translocation to India carries the risk of future outbreeding, ultimately swamping the Asian lineage if African cheetahs were to come into contact with Asiatic cheetahs through uncoordinated translocations or even natural dispersal from future introduced populations.

Numerically, the founder African cheetah population in India (20 individuals[6]) is smaller than the relict *A. j. venaticus* population in Iran. Therefore, we argue that the recovery of the Iranian population could, in the long term, with concerted conservation and political investment, provide more appropriate animals for the Indian restoration project, with direct conservation value to the subspecies and a low risk of compromising genetic integrity.

As the founder population (20 cheetahs) and the maximum number of cheetahs (36 cheetahs)[6] that can be supported at the Indian introduction site are too small to avoid breeding between related individuals, they will be highly vulnerable to stochastic demographic risks[7]. Related cheetahs will need to be removed from the Indian population on an ongoing basis to ensure that inbreeding does not occur, and to allow for demographic management of the population. Clearly, this introduced population is, therefore, not viable and will need recurring supplementation, representing an ongoing drain on African cheetah populations and scarce conservation resources.

The estimated 21–36 cheetahs that can be sustained at the release site was determined through prey-base calculations.[6] This is a well-tested and commonly utilised method for evaluating reintroduction viability. However, in a contradictory statement, the lead author of the Action Plan for Introduction of Cheetah in India has claimed that the release area can probably only sustain 15 cheetahs.[8] This raises concerns about the quality of the data and the analyses thereof on which the Action Plan is based, and supports concerns around the prey and space requirements of cheetahs in India.[9] With 20 cheetahs already introduced, two subsequent deaths[10], the birth of four cubs and another possible pregnancy[11], the population of more than 23 cheetahs could likely be exceeding the capacity of the prey base.

The viability analysis, risk assessment and data on which these are based for this project are not available for scientific scrutiny. As such, it is impossible to evaluate what risks were identified and how they were mitigated against. This is particularly important in this scenario in which the animals are likely to come into contact with novel pathogens, unknown and unpredictable ecological interactions (including predator–prey and interspecific carnivore guild interactions), a high poaching threat[12] and undefined conflict with humans. Many of these uncertainties could be reduced and proactively mitigated against by using published science on the biology of the species, e.g. cheetah spatial ecology has been used to predict the long-distance post-release movements exhibited by particular cheetahs in India.[13]

No clear exit strategy, as required by the IUCN Guidelines for Reintroductions, has been defined should the project not succeed or cause unpredicted harm to other wildlife or humans. Cheetahs are expected to be returned to Africa for demographic management[14] or through experimental failure and the African cheetah population will be exposed to unknown risks from these returned animals.

In South Africa, there is no formal, peer-reviewed metapopulation management plan for cheetahs in fenced reserves. Management of these cheetahs is fragmented and not formally goal-driven at a national level. Data for the metapopulation, upon which the sustainability of the India project is assessed, remain unpublished[6], including population size, growth rates, sex ratios, demand for cheetahs in South Africa and the number of cheetahs needed

to maintain the metapopulation. No modelling or sensitivity analysis has been published using recent data on the possible long-term impact of removals on the South African population.

Claims of South Africa having 'excess cheetahs' in the metapopulation are used to support the Indian reintroduction[6], yet there are no data to support this. Simultaneously, captive cheetahs are being used to supplement the metapopulation within South Africa.[15] About 9% of the total metapopulation has been supplemented with 39 captive animals (~66% of these since January 2019).[14] This indicates the opposite scenario to an excess in the metapopulation.

Despite claims to the contrary[1], we hold that this project has not made the best use of available science or used best scientific practice in its planning.

2. Does it support cheetah conservation?

Conservation projects should not compete by consuming funds or redirecting the limited funding available. The Indian cheetah project will cost approximately USD50–60 million to create three small populations.[6] This funding could be used for *in situ* projects that directly benefit the conservation of extant Indian wildlife, e.g. tiger *Panthera tigris* and Gir lion *Panthera leo leo* conservation or cheetahs in Africa.

Scarce conservation habitat should be allocated to the conservation of indigenous species and not be occupied by exotics, as per the tiger 'conservation' projects in South Africa (e.g. Tiger Canyon and Save China Tigers) and the proposal to take rhinos to Australia.[16] There is also evidence that progress on a long-standing critical Gir lion conservation project has been disrupted due to the cheetah reintroduction at the same site and may even be halted by legal action as a result.[17] Clearly, these translocated African cheetahs will be displacing indigenous conservation opportunities.

Sending African cheetahs to India on an experimental basis generates the perception of an excess of African cheetahs and, by extension, that African populations no longer need conservation. Yet, at a continental scale, the cheetah is in decline.[18] Furthermore, there are areas in Mozambique, Malawi, Zambia, Angola, and several West African countries of sufficient size, with sufficient prey, that could be restored for cheetahs within 10 years, either naturally or through assisted recovery if effective conservation action is taken.[19] There have been several successful reintroductions into such areas, where cheetahs sourced from South Africa were reintroduced into well-protected parks.[1,20] Restoration projects in these countries directly benefit cheetah conservation and local ecosystems in the African subspecies' historical range. They indicate a growing need for African cheetahs within Africa, rather than an excess.

An unpublished Population Viability Assessment using data from 2017 was used to guide an export quota of 13 cheetahs from South Africa for reintroduction into range states per year.[21] The plan is to send 12 cheetahs per year for 10 years from South Africa to India and 12 have already left for India.[22] This is the maximum amount that is potentially available from the metapopulation, leaving no cheetahs for restoration into African range states. Clearly, these cheetahs should remain part of Africa's natural heritage[16] and be used to maximise the conservation of African species and benefit African people.

The 50% mortality rate of cheetahs predicted for the Indian reintroduction[6] is much higher than the observed survival rate of 85% for reintroductions in the metapopulation in South Africa[23]. The predicted mortality rate in India is aligned with the relocation of damage-causing animals in the large areas of unfenced Namibian farmland[24], where cheetahs roam outside protected areas and are vulnerable to persecution and other anthropogenic threats. The cause of this predicted high mortality rate in India is not identified, and suggests a lack of resources to properly maintain the population, a lack of understanding of the reintroduction process or general unsuitability of the area for cheetahs.

Ten of the current tranche of 20 cheetahs are expected to die as a consequence of this translocation (with more to follow); one female cheetah has died due to what appears to be a pre-existing renal condition[9], another due to what appears to be a stress-related condition[6] and two cheetahs escaped causing conflict with villagers, with one cheetah being stoned and possibly injured[25] and multiple captures needed to return them to the park and one returned to captivity. It appears that the welfare of the animals is being compromised through a lack of mitigation of threats to their post-release survival and inadequate fencing, and the conservation potential of the animals is being squandered.

The South African fenced metapopulation model for cheetah conservation is unique. These cheetahs are maintained in fenced reserves, away from human populations and livestock and are generally habituated to vehicles and often also to people on foot. The cheetahs sourced from Namibia were wild born, but have been maintained in captivity, also making them habituated to people and unsuitable for release in the larger reserve.[11] There is evidence that the female cheetahs taken to India from Namibia have not been 'rewilded', will not be released into the larger reserve, and have been allowed to breed in fenced enclosures as part of a captive-type breeding project.[11] Therefore, these animals may not be ideal candidates for creating a founder population of wild cheetahs in an area where people and livestock are not excluded, as planned in India.

The potential for human–wildlife conflict is elevated when habituated animals come into contact with humans and their livestock. Large-scale sensitisation programmes are required to prevent conflict, as cheetahs can roam far beyond park boundaries. Conflict-related killing is a leading threat to cheetahs across their range[18], with few conflict mitigation projects being effective in the long term[26], and good legislative frameworks as preventative tools[1] do not prevent the killing. Two cheetahs have already left the reintroduction site. One roamed more than 20 km into a village where villagers, fearful for the lives of their children and livestock, threw stones at the cheetah and it was possibly injured. The cheetah has been returned to the park.[25] The other cheetah is still at large and has reportedly 'triggered panic among the surrounding villagers'[27]. If the prey base becomes depleted due to unsustainable predation, as suggested, then breakouts can be expected to increase. This human element amplifies the risk of creating a sink for African cheetahs in India.

Conclusion and the way forward

The recent global commitment to increase protected area targets at the United Nations' Convention on Biological Diversity (CoP15) should lead to an increase in protected areas, creating more opportunities for African cheetahs in Africa.

A more coordinated and science-based approach to cheetah conservation globally is needed, particularly in light of the recent warning that cheetahs are in a more precarious position than indicated by their current IUCN "Vulnerable" status.[18] *In situ* conservation action for Asian cheetahs should be prioritised and supported. Within Africa, a range-wide science-based metapopulation plan must be developed for cheetahs in fragmented populations. A collaborative peer-reviewed exercise must identify and prioritise suitable areas for indigenous cheetah restoration through reintroductions of respective individuals of the appropriate subspecies in both Asia and Africa. This should include complete, transparent risk assessments and clear references to best scientific and management practices for reintroductions. The principles of evidence-based management and transparency should be applied in future decisions, and can provide a framework for other taxa in Africa.

In South Africa, a managed metapopulation plan for cheetahs is required to guide the science-based management of these cheetahs to maximise their contribution to cheetah conservation. National-level management must be done collaboratively, and actions must be taken to maximise the management objectives of individual reserves and cheetah conservation across the range, and ensure that best practices are followed.

Establishing a cheetah sink out of Africa will threaten African (and Asian) cheetahs and undermine South Africa's reputation as a science-based leader in the conservation management of large mammal populations.

Note in Proof

A third cheetah death among the cheetahs introduced from Africa to India was reported on 9 May 2023 (https://pib.gov.in/PressReleasePage.aspx?PRID=1922858). This highlights the need for urgent interventions to prevent any further mortalities.

Competing interests

We have no competing interests to declare.

References

1. Tordiffe AS, Jhala YV, Boitani L, Cristescu B, Kock RA, Meyer LRC, et al. The case for the reintroduction of cheetahs to India. Nat Ecol Evol. 2023;7:480–481. https://doi.org/10.1038/s41559-023-02002-2
2. Khalatbari L, Jowkar H, Yusefi GH, Brito JC, Ostrowski S. The current status of Asiatic cheetah in Iran. Cat News. 2017;66:10–13.
3. Rai N, Verma SK, Gaur A, Iliescu FM, Thakur M, Golla TR, et al. Ancient mtDNA from the extinct Indian cheetah supports unexpectedly deep divergence from African cheetahs. Sci Rep. 2020;10(1), Art. #4618. https://doi.org/10.1038/s41598-020-60751-7
4. Thapar V, Thapar R, Ansari Y. Exotic aliens: The lion and the cheetah in India. New Delhi: Aleph Book Company; 2013.
5. Prost S, Machado AP, Zumbroich J, Preier L, Mahtani-Williams S, Meissner R, et al. Genomic analyses show extremely perilous conservation status of African and Asiatic cheetahs (*Acinonyx jubatus*). Mol Ecol. 2022;31:4208–4223. https://doi.org/10.1111/mec.16577
6. Jhala YV, Ranjitsinh MK, Bipin CM, Yadav SP, Kumar Alok, Mallick Amit, et al. Action plan for introduction of cheetah in India. New Delhi: Wildlife Institute of India, National Tiger Conservation Authority and Madhya Pradesh Forest Department; 2021.
7. Caughley G. Directions in conservation biology. J Anim Ecol. 1994;63:215–244. https://doi.org/10.2307/5542
8. Koshy J. Kuno National Park unsuitable to host all 20 cheetahs, not enough prey, says scientist. The Hindu. 03 March 2023. Available from: https://www.thehindu.com/sci-tech/energy-and-environment/kuno-park-unsuitable-to-host-all-20-cheetahs-says-scientist-behind-reintroduction-plan/article66577021.ece
9. Gopalaswamy AM, Khalatbari L, Chellam R, Mills MG, Vanak AT, Thuo D, et al. Introducing African cheetahs to India is an ill-advised conservation attempt. Nat Ecol Evol. 2022;6(12):1794–1795. https://doi.org/10.1038/s41559-022-01922-9
10. Mudur GS. Kuno National Park: Tragic death of the magnificent Uday. The Telegraph Online. 25 April 2023. Available from: https://www.telegraphindia.com/india/kuno-national-park-tragic-death-of-the-magnificent-uday/cid/1932140
11. Mudur GS. Kuno National Park awaits cheetah births. Telegraph India. 29 March 2023. Available from: https://www.telegraphindia.com/india/kuno-national-park-awaits-cheetah-births/cid/1925767
12. Special dog appointed to prevent illegal poaching at Kuno forest area The Free Press Journal. 04 April 2023. Available from: https://www.freepressjournal.in/bhopal/mp-kuno-gets-illu-a-german-shephard-dog-to-guard-cheetahs-and-other-wild-animals
13. Wachter B, Portas R, Melzheimer J. The introduction of African cheetahs to India was planned without considering their spatial ecology. Conserv Sci Pract. 2023;e12943. https://doi.org/10.1111/csp2.12943
14. University of Pretoria. University of Pretoria vets lead revival of India's extinct cheetah. 20 July 2022. Available from: https://www.up.ac.za/news/post_3089400-university-of-pretoria-vets-lead-revival-of-indias-extinct-cheetah#:~:text=University%20of%20Pretoria%20(UP)%20academics,the%20survival%20of%20the%20species
15. Endangered Wildlife Trust [unpublished data]. In: Buk KG, Van der Merwe VC, Smit M, Naude VN, editors. Biodiversity management plan for cheetah in South Africa. Draft report version 1.6. 2021.
16. Hayward MW, Ripple W, Kerley GIH, Landman M, Plotz R, Garnett S. Neocolonial conservation: Is moving rhinos to Australia conservation or intellectual property loss? Conserv Lett. 2018;11(1), Art. #12354. https://doi.org/10.1111/conl.12354
17. Mazoomdar J. Cheetahs in, govt to re-examine plan to shift Gir lions to Kuno. Indian Express. 28 March 2023. Available from: https://indianexpress.com/article/india/cheetahs-in-govt-to-re-examine-plan-to-shift-gir-lions-to-kuno-8522539/2023
18. Durant S. IUCN Red List assessment for cheetah – a species to watch. IUCN SSC Cat News. 2022;76:14–16.
19. Cheetah Conservation Initiative (CCI). Africa Range-Wide Cheetah Conservation Initiative. Cheetah distribution and status: Africa 2022 update [webpage on the Internet]. c2022 [cited 2023 May 05]. Available from: https://cheetahconservationinitiative.com/cheetah-maps/
20. Sievert O, Fattebert J, Marnewick K, Leslie A. Assessing the success of the first cheetah reintroduction in Malawi. Oryx. 2022;56(4):505–513. https://doi.org/10.1017/S0030605321000788
21. The Scientific Authority of South Africa. Non-detriment finding for *Acinonyx jubatus* (cheetah). 24 September 2020. Government Gazette number 45552. Available from: https://www.dffe.gov.za/sites/default/files/gazetted_notices/nemba_cheetahnondetrimentfindings_g45552gon1532.pdf
22. Omarjee L. SA to send 12 cheetahs to India to boost conservation efforts. News 24. 27 January 2023. Available from: https://www.news24.com/fin24/climate_future/environment/sa-to-send-12-cheetahs-to-india-to-boost-conservation-efforts-20230127
23. Marnewick K, Hayward MW, Cilliers D, Somers MJ. Survival of cheetahs relocated from ranchlands to fenced protected areas. In: Hayward MW, Somers MJ, editors. The re-introduction of top order predators: Chapter 13. Oxford: Blackwell Publishing; 2009. p. 282–306. https://doi.org/10.1002/9781444312034.ch13
24. Boast LK, Chelysheva EV, Van der Merwe V, Schmidt-Küntzel A, Walker EH, Cilliers D, et al. Cheetah translocation and reintroduction programs: Past, present, and future. In: Marker L, Boast LK, Schmidt-Küntzel A, editors. Cheetahs: Biology and conservation. London: Elsevier; 2018. p. 275–289. https://doi.org/10.1016/B978-0-12-804088-1.00020-4
25. Cheetah escapes from Kuno, scares villagers. The Wildlife India. 02 April 2023. Available from: https://www.thewildlifeindia.com/2023/04/Cheetah-Escapes-From-Kuno-Scares-Villagers.html
26. Webber AD, Hill CM, Reynolds V. Assessing the failure of a community-based human-wildlife conflict mitigation project in Budongo Forest Reserve, Uganda. Oryx. 2007;41(2):177–184. https://doi.org/10.1017/S0030605307001792
27. Another cheetah escaped from Kuno National Park. AP7am. 6 April 2023. Available from: https://www.ap7am.com/lv-372662-another-cheetah-escaped-from-kuno-national-park

AUTHOR:
Keyan G. Tomaselli[1,2]

AFFILIATIONS:
[1]Dean's Office, Faculty of Humanities, University of Johannesburg, Johannesburg, South Africa
[2]Professor Emeritus and Fellow, School of Applied Human Sciences, University of KwaZulu-Natal, Durban, South Africa

CORRESPONDENCE TO:
Keyan Tomaselli

EMAIL:
keyant@uj.ac.za

HOW TO CITE:
Tomaselli KG. Academic xenophobia in South Africa – issues, challenges and solutions: Reflections on an ASSAf Roundtable. S Afr J Sci. 2023;119(7/8), Art. #16278. https://doi.org/10.17159/sajs.2023/16278

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
xenophobia, internationalisation, glocalisation, South Africa, foreign African nationals

PUBLISHED:
8 August 2023

Academic xenophobia in South Africa – issues, challenges and solutions: Reflections on an ASSAf Roundtable

Significance:

Anti-black African xenophobia is a pressing issue in South Africa, from the streets to our universities. Physical threats against honest university administrators and non-nationals have become legion. Yet foreign academics and teachers contribute directly to local education and employment which would be much poorer without them. This Commentary, which reports on a Roundtable discussion of this issue, explores the form of this prejudice, the consequences for the institutions and individuals concerned, and the implications for South African glocalisation.

In vogue during the dark days of late apartheid were terms like 'international hotel', or more idiotically, 'international toilet'. 'International' in the apartheid lexicon meant multiracial and multinational. These categories were introduced after the Transkei's 'independence' in 1976 to facilitate black Bantustan and foreign African diplomats who were visiting 'white' South Africa. Anyone could now legally use facilities officially designated as 'international'. After apartheid, 'international' took on some new negative connotations when associated with the Afrophobic term, '*makwerekere*', black African in-migrants ('foreigners') whose presence in a liberated South Africa was often resented.

What does 'international' mean for any South African city hosting one or more universities? I actively assisted in the recruitment of international students and interinstitutional collaborations in the 1990s under the auspices of the University of KwaZulu-Natal (UKZN) International Office. That Office, at the time, was looking to 'decolonise' UKZN by attracting fee-paying African students onto the university's diversifying campuses.

Glocalisation, until the international supply chain was disrupted by COVID-19 in 2020, was the new buzz word. Until then, 'being local' meant that universities needed to go global. Still, most universities are now operating in an internationally competitive intercultural knowledge market. If our universities remain parochial, so will our students, our research topics and our educational priorities. National policymaking will reflect this parochialism, as echoed by the panellists who contributed to the 9th Academy of Science of South Africa (ASSAf) Presidential Roundtable Discussion held on 24 November 2022. The topic was 'The Threat of Academic Xenophobia to the Future of South African Universities'.

If UKZN was indeed the 'Premier University for African Scholarship' as it claimed from 2004, then we must understand that Africa is not just the KwaZulu-Natal Province, South Africa or even Africa. As the African diaspora is everywhere now across the Western world, that's the identity that we need to embrace, and that was the advice offered by the Roundtable panellists. But more, our institutional and employment and recruitment policies must reflect an international remit. It is internationalisation at every level that vests excellence in the world's top universities because these institutions attract the best students and scholars from everywhere, as Jonathan Jansen, President of ASSAf, observed while chairing the panel.

Universities, although they are places of open learning, are not immune to 'academic xenophobia', a term invented for the Roundtable discussion. In this Commentary on the Roundtable invited by the *South African Journal of Science*, I attempt to provide a frame of reference and some affirmative thoughts deriving from the discussion.

The scourge of xenophobia

The visible signs of anti-immigrant violence that play out on the streets and poorer areas of South Africa are economically disruptive. This behaviour transgresses human rights and sometimes even involves assassination.[1] In the academic arena, prejudice directed against 'foreigners' (i.e. students and academic staff recruited predominantly from the former frontline states of Zimbabwe and Zambia), occurred from the mid-1990s when the first wave of excited, exciting and exhilarating international graduate scholars and newly appointed lecturers arrived expectantly on many South African campuses. For some universities this influx positively reshaped these institutions in terms of diversity – intellectually, conceptually and culturally – repositioning them as regional players. The in-migrants' own earlier historical experiences of arrested liberation had matured them to benefit and be benefitted by the brave new world that South Africa had promised in the mid-1990s.

From pervasive intellectual, cultural and historical parochialism, reinforced by the cultural and academic boycotts during the apartheid era, South Africa instantly from 2 February 1990, with the release of Nelson Mandela and the unbanning of the liberation movements, became the focus of intrigued global attention. One of the first to set up an International Office in 1992 was UKZN, which in short time had signed up well over 200 collaboration agreements, with its campuses swamped with visiting delegations. South African academics and students were thereby invited to become global academic citizens. They were recipients of international grants, engaging in collaborative projects and securing previously unimaginable teaching, sabbatical, research, exchange and resourcing opportunities.

Instead of capitalising on this once-off and extraordinary Mandela effect for the national benefit, the reaction from many students and staff, often the first generation of their cohort, was anxiety. It seems that little had changed by the end of 2022, as was indicated by everyone who spoke on the ASSAf panel.

The Roundtable panellists discussed their respective experiences and what can be done to develop a progressive employment regime that is built on the academic values of the modern cosmopolitan university. The discussion started rather starkly with Nicole Fritz's future-alert. Director of the Helen Suzman Foundation, Fritz drew attention to the impending sub-Saharan intra-continental migration that is being caused by climate change. This mass movement is anticipated to become an "existential issue" in the near term.

Simultaneously, staying the threatened termination of Zimbabwean exemption work permits was one of the Foundation's projects. The South African Department of Home Affairs' announcement in 2022 that permits would be revoked was made without consultation with the affected constituencies, including schools, colleges and universities that employ Zimbabwean expatriates. These kinds of ad hoc decisions indicate a lack of foresight by the much-criticised department with regard to migration planning and burden shearing, not to mention academic productivity (see for example Grant[2]). Fritz asked whether universities were educating their graduates to address these kinds of concerns.

The panellists who followed were Evance Kalula (Zambia), Emeritus Professor of Law, University of Cape Town; Precious Simba (Zimbabwe), Lecturer, Department of Education Policy Studies, Stellenbosch University; and South African Sakhela Buhlungu (Vice-Chancellor, University of Fort Hare [UFH]). The chair was Jonathan Jansen, ASSAf President.

The theme of the Roundtable

My approach here is to précis the issues common to the panel and then, using my own experience deriving from my closeness to the early UKZN International Office, to draw out some measures to close the discussion on an affirmative note.

Jansen[3] has argued that one of the most serious threats facing higher education and the scientific enterprise in South Africa is the rising tide of academic xenophobia. The contradictions unleashed by internationalisation during the 1990s have intensified, it seems, given the examples offered by the panellists. While 14 of the 15 universities surveyed by Simba proclaim themselves as African training and research institutions, a "disconnect" often occurs within them in the form of xenophobia targeted at "foreign African nationals." Yet, as Fritz pointed out in comparison, the USA-quartered enabler of the criminal capture of the South African Revenue Services (SARS), Bain and Co, was spared the same kind of popular opprobrium, even though exposed in the Nugent and Zondo Commissions and in the national news media. The Nugent Commission, led by Judge Robert Nugent, investigated the failure of governance at SARS, which enabled state capture by international criminals detailed in no less than eight volumes by the Commission headed by Judge Raymond Zondo (available here: https://www.statecapture.org.za/site/information/reports)

Resentment by South African hosts usually emanates from a sense of insecurity and entitlement, triggered by competition for resources and opportunities. Nevertheless, Kalula observed that while he has "encountered hostility from colleagues and line managers," that his students "from across all races, have been very receptive [to foreigners]." The panel spoke candidly, largely from the heart, illustrated with examples of discrimination against foreigners. They talked about themselves being in the "trenches," and indeed just six weeks after the Roundtable, Buhlungu survived a second attempt on his life, although sadly his bodyguard was killed. Reports of other senior UFH managers having been previously (and since) attacked and assassinated by those engaged in "mischief" now made the national press. "Fort Hare under Siege" screamed the *Sunday Times* (8 January 2023, p. 1).

Previously, one of my former UKZN colleagues, Malawian Gregory Kamwendo, who had moved to University of Zululand (UZ), was executed for exposing degree fraud. Vice-chancellors, deans and legal officers often require 24-hour protection. This mafia type behaviour emanates from organised criminals internal to universities who are looting public resources and state-owned enterprises, and assassinating honest staff and whistleblowers who resist corruption. Kamwendo's killers and those of his South African UZ colleagues involved, however, considered the university merely a resource to be milked at whatever price. Dysfunctional universities riddled with fraud and corruption were largely inherited from the now dis-established Bantustans, although some such universities, as Jansen reminded, have succeeded in reinventing themselves. These, and many more equally alarming instances, are documented in Jansen's book *Corrupted*[4], launched within a few months of the panel discussion.

The international contribution

In justifying his appointment of a foreign African national as an acting registrar in the face of such threats during a period of UFH meltdown, Buhlungu observed that the institution "pulled through, not because of his 'foreignness', but for his integrity, firmness and his principled approach to administration."

As UKZN Vice-Chancellor Brenda Gourley observed during the 1990s, tactically leveraging a 1980s' union slogan, 'an injury to one university is an injury to all'. Donors, international partners and funders lose faith and cancel collaborations, withdraw resources and terminate staff–student exchanges. Fearful local students who can afford the fees divert to expensive but safe private colleges. Everyone is affected, no matter their locations and no matter the security arrangements at the safer campuses.

The issue of personal safety backgrounded each panellist's concern. Notwithstanding such anxiety, Kalula revealed that he had personally benefitted from being taught by exiled lecturers from South Africa, Uganda and Nigeria in his native Zambia. But he acknowledged that a university in which its nationals were in the minority "has a problem." However, simply "blaming foreign academics, and especially African ones, for the absence of local ones is wrong," he cautioned. The fear of victimisation on the basis of their countries of origin has deterred many African intellectuals now teaching in the global diaspora from applying to South African universities. The 1990s outmigration from universities of native South Africans into very well-paid government jobs, the civil service and business, resulted in a shortage of local academics in many rural institutions, creating a vacuum that has been filled by foreign expatriates, predominantly Zimbabweans and Nigerians. Indeed, the early loss included UFH's first black post-apartheid vice-chancellor who was appointed as the national unity government's first minister of education.

'Foreigners', a term used pejoratively in some selection committees on which I myself served during the transition, were sometimes accused during the first decade of liberation of being opportunistic interlopers exploiting unfair advantage of gaps in the academic sector. The result is that internal discrimination with the ending of apartheid was supplanted by antagonism against African foreigners who it is claimed: (1) usurp jobs from South Africans; (2) compete with South Africans for scarce resources; (3) exhibit a culture of entitlement in that South Africa owes them for taking an anti-apartheid stance; and (4) engage in criminality, although there is little evidence to support this allegation (Fritz). South Africans like to play at being the "victim" – being at the mercy of the Other, being stereotyped and accused "of everything" (Buhlungu).

Underpinning these assumptions is that wealth is understood by the criminals and some political ideologues to be a technical redistributive process rather than also a productive procedure that is created by entrepreneurial activity, capacity building and intellectual investment. Initiative is what counts, not 'entitlement', that has so often resulted in institutional failure. The "killing fields" to which Buhlungu referred in the *Sunday Times* front page occurred, ironically, because of UFH's first clean audit in 30 years.

Backgrounding the panel's deliberations was that South Africa has been engaged in nation-building during a conjuncture when some other nations and empires have been fragmenting. Identity creation requires the forging of a single South Africanism in the face of extensive domestic

diversity. Othering the Other (other Africans) is one way of forging an 'us'–'them' dichotomy towards fostering a national identity. Yet, as Simba cautioned about ringfencing nationality in the global academic marketplace, "Universities should always see themselves as part of a larger community of knowledge and should always hold themselves to the high standards that the society holds us to." This claim to African identity in the 14 local universities' mission statements, she concluded, raises questions on the nature of the disconnect between institutional PR, their "imagined" identity and the actual "performance" in practice. It was becoming almost impossible, said Simba, for non-South Africans to join South African academia due to ongoing "ring-fencing of academic space": "… what is happening … is a hierarchisation of the 'better black'" and even within that inclusion, "there is a black person that the labour policies preferred, and by doing so we start seeing the pushing out and exclusion of foreign nationals."

South Africans are thus still ill-prepared for globalisation, competitive job markets and professional mobility required by transnational academia. Buhlungu observed that:

> *We need to expand and diversify the range of non-locals to go beyond the two dominant countries (Zimbabwe and Nigeria) and bring in scholars from across Africa, and to also bring in more Indians and whites to help end black to black xenophobia.*

The global opportunities now available to the educated professional classes, including South Africans, need to be better understood.

Internationalisation policies are key to growth, as Kalula pointed out in offering the example of Malaysia and Singapore. On dividing into two separate states in 1965, the internationalisation policy of Singapore's National University of Singapore led to the university thriving through the recruitment of expatriate lecturers and becoming one of the highest-ranked institutions in Asia. In contrast, the University of Malaya stagnated by rejecting non-local faculty and becoming inward-looking. In Africa, many of the top universities became shadows of their former selves once they severed their links with their parent institutions located in the colonial metropoles. "No university in the world has ever become a global centre of academic and research excellence through nativist thinking in its academic appointments policy," said Jansen. But if nativism is pursued, then "We might as well be a church, mosque or synagogue based on faith or allegiance." As Jansen points out, some South African universities

> *would quite literally fall apart were it not for other African academics willing to work in rural areas and uphold their academic programmes in everything from undergraduate teaching to postgraduate supervision and, of course, senior administration.*

The building of a strong local pool in a fair and non-prejudicial manner is not easy because this requires the remedying of the reluctance of South Africans to work at geographically remote campuses.

Concluded Simba: "Our [South African] colleagues should create space for African foreign nationals, to allow for different experiences, different ideas, different accents, different skin tones, different kinds of black, different kinds of academics…." To achieve this, African foreign nationals, suggested Kalula, should be encouraged "to apply their trade in the interests of South Africa":

> *You South Africans are very angry people. You are angry against each other. You are angry against foreigners and even one politician in his lucid moment characterised this xenophobia as self-hate. This history is self-destructive, a lost dream, so foreign academics must come here to help in the South African mission, rather than 'taking over', or thinking of themselves as substitutes.*

One potential benefit of travelling academics pollinating our shores is recovery of the 1990s 'going global' dream. Let's now identify some affirmative possibilities of internationalisation towards this goal.

Internationalisation

Having discussed attraction factors over the past 30 years with my own international African students – who populated about half of my Centre's graduate student complement at UKZN – they explained that South Africa is a Europe next door for Southern African Development Community countries, in that its universities offer affordable access to top quality academic programmes, technology, resources and libraries. For me, and as reported at a UKZN corporate relations conference over a decade ago, there are many benefits that international students and academics bring to South African institutions, including:

- Wider understandings of their respective conditions to South Africans who tend to be ignorant of Africa.
- Foreign African graduates return home and often establish similar programmes in their own universities. They then draw on South Africans as external examiners, collaborate with them on pan-African research projects, and send on their own students to do graduate work at the South African institution that hosted them.
- International students bring a political stability to class discussions, and a maturity of purpose to their studies.
- International students show high initiative, take responsibility for their own learning, time and finances, and they are self-motivated. They become our teachers and add significant value to the classroom.
- International collaboration is best driven by academics themselves. International offices are the facilitators, the ambassadors; the cooperating academics are the drivers, the international lecturers are the advisors, and the benefit is institutional.
- International offices are not just administrative ventures gobbling up scarce resources. They are income generators, the global recruitment arms of a university. A university's business plans should indicate how return on investment will occur over specific periods.

An international university needs more than international hotels and toilets. It needs international students and staff, international visitors and international research collaborations. But Marina Waruru's[5] take-away of the Roundtable concluded that:

> *The presence of xenophobic practices in South Africa's higher learning institutions can be blamed on a lack of strong leadership in the universities, which has turned a blind eye on the vice by allowing expediency to prevail over merit.*

Leadership is key, but it has to get buy-in from all sectors of the academy. Fortunately, observes Simba, "From my experience I can, however, say that students are open, want to learn and do not care about your origin." In contrast, she suggests: "Our colleagues, though, are not as accommodating, and there are some who are pushing for a policy change so that universities do not employ foreign academics."

We are living in a global world and global job markets and we local South Africans can also become global citizens pursuing international career opportunities. We do not have to be angry; we certainly should not be killing each other over finite resources, and we should be growing our wealth rather than squandering it through corruption.

Why is this difficult to understand? Positive intercultural values indeed are what we should be teaching in the academy, in schools and in kindergartens. That was the underlying message of the Roundtable.

The recording of the Roundtable at which these issues were discussed is available here: http://hdl.handle.net/20.500.11911/260

Competing interests

I have no competing interests to declare.

References

1. Kunda JKL. Xenophobia in South Africa: Revisiting Tutu's handwriting on the wall? Crit Arts. 2009;23(1):120–123. https://doi.org/10.1080/02560040902739022
2. Grant J. A harrowing tale from Home Affairs. Visas: An unwelcome and time-consuming interlude to academic publishing. ANFASA Magazine. 2016;6(2):1–3. Available from: https://www.anfasa.org.za/wp-content/uploads/2022/08/ANFASA-Volume-6-2-2022Final-.pdf
3. Jansen J. Xenophobia is threatening the future of the SA university. University World News. 05 October 2022. Available from: https://www.universityworldnews.com/post.php?story=20221004102540205
4. Jansen JD. Corrupted: A study of chronic dysfunction in South African universities. Johannesburg: Wits University Press; 2023. https://doi.org/10.18772/12023037946
5. Wararu M. Strong leadership needed to confront academic xenophobia. University World News. 30 November 2022. Available from: https://www.universityworldnews.com/post.php?story=20221129083657870

AUTHORS:
Donrich Thaldar[1,2]
Aliki Edgcumbe[1]
Dusty-Lee Donnelly[1]

AFFILIATIONS:
[1]School of Law, University of KwaZulu-Natal, Durban, South Africa
[2]Petrie-Flom Center for Health Policy, Biotechnology, and Bioethics, Harvard Law School, Cambridge, Massachusetts, USA

CORRESPONDENCE TO:
Donrich Thaldar

EMAIL:
thaldard@ukzn.ac.za

HOW TO CITE:
Thaldar D, Edgcumbe A, Donnelly D-L. How to interpret core concepts in POPIA? Recommendations on the draft Code of Conduct for Research. S Afr J Sci. 2023;119(7/8), Art. #15062. https://doi.org/10.17159/sajs.2023/15062

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
research code of conduct, consent, special personal information, responsible party, de-identification

FUNDING:
US National Institutes of Health (U01MH127690)

PUBLISHED:
8 August 2023

How to interpret core concepts in POPIA? Recommendations on the draft Code of Conduct for Research

Significance:

The draft Code of Conduct for Research is a welcome development, but there is room for improvement in the way that it interprets core concepts in POPIA. In particular, it should: remove the provision regarding *consent* to future research; clarify that *special personal information* is a subclass of personal information that must comply with an extra layer of requirements; not exclude an individual researcher employed by a research institution from qualifying as a *responsible party*; and clearly differentiate *de-identification* in POPIA from corresponding terms used in other jurisdictions.

Introduction

The publication of the draft Code of Conduct for Research[1] (draft CCR) by the Academy of Science of South Africa (ASSAf) in September 2022 is a welcome development. Overall, the draft CCR promises to be a useful document for the South African research community. It is written in plain language, contains useful diagrams and user-friendly hyperlinked cross-references, examples in the research context, and references to additional resources.

In terms of the draft CCR's substance, the most striking positive element is how it deals with the concept of *public interest*. Public interest is a central concept in the *Protection of Personal Information Act 4 of 2013* (POPIA), and is important for research, as it is relevant when considering a research exception for allowing the processing of special personal information (section 27(1)(*d*)(*i*)), and a research exemption from the conditions for processing personal information (section 37(1)(*a*)). The Information Regulator has proffered a 'basic formulation' of public interest in a *Guidance Note*[2], but this 'basic formulation' has been critiqued in the literature as misaligned with South African case law on public interest[3]. To ASSAf's credit, the draft CCR (Table 5, page 24) does not simply follow the *Guidance Note*'s embattled 'basic formulation', but adopts a more pragmatic – and legally justifiable – meaning of public interest.

However, there is still a need to improve the way that the draft CCR deals with four other core concepts in POPIA, namely (1) consent; (2) special personal information; (3) responsible party; and (4) de-identification. In this article, we explain why there is a need for improvement of the draft CCR in the case of each of these concepts, and we make recommendations on how to improve the draft CCR in this regard.

Consent

The interpretation of the meaning of *consent* in POPIA has been the subject of academic debate; members of our research group contended that consent, for purposes of POPIA, must be *specific*, as clearly provided for in section 1 of POPIA.[4-6] However, there have been other scholars who have argued for an interpretation of consent in POPIA as meaning *broad* consent.[7] Also, we have contended that POPIA constitutes a new layer of legal rules, and that POPIA compliance should therefore be differentiated from ethics compliance, as these are two distinct sets of rules, and the one does not subsume or replace the other.[4-6] In this light, we commend ASSAf on confirming this law/ethics distinction in the draft CCR, and on making it clear that consent in POPIA must be *specific* – at least in the context of an *initial* research project.

However, we must raise concern about the following statement in the draft CCR that relates to *future* research (Table 4, page 20)[1]:

> *POPIA Consent for future use is allowed as long as the future uses of the Personal Information are not speculative, are described as fully as possible, and further use of the Personal Information is restricted.*

Our concern has two parts. First, this statement in the draft CCR departs from POPIA. There is a difference between having a specific purpose and having a purpose that is merely 'not speculative'. Accordingly, this risks watering down the requirement of *specific* consent, and calls for serious reflection. Second, given POPIA's research exceptions (in sections 15(3)(*e*) and 27(1)(*d*)), consent for *future* research projects is *not* necessarily a POPIA requirement. In other words, from the perspective of POPIA, researchers may rely on consent for further processing, but they do not have to. Accordingly, we recommend that the statement in the draft CCR pertaining to consent for *future* research should be removed. From the perspective of POPIA, it is both problematic and unnecessary.

However, depending on the circumstances, consent to a further research project may be required by the institutional research ethics committee. This raises the question: how should researchers integrate the POPIA and ethics requirements regarding consent at the stage of collecting information from research participants? In brief, for the *initial* research project, consent must be for a *specific* purpose, as required by section 13(1) of POPIA. Researchers may include additional provisions if required by institutional research ethics committees. As an example, consider the question: 'May we contact you again for a future research project or follow up on your responses?' Such a provision is self-evidently only an ethical consideration. From a legal perspective it is not consent to future

research, much less *specific* consent. At the same time, it may also be wise to request research participants to provide consent now for *future* research. The mode of consent that is appropriate – e.g. specific, tiered, or broad consent – is determined by the relevant institutional research ethics committee, and is an ethics requirement, not a POPIA requirement. It is essential to approach this enquiry from a pragmatic perspective. One can readily imagine the futility of trying to predict all the uses of data that may become possible in the light of advancing data science techniques. It is precisely for these kinds of situations that we recommend researchers familiarise themselves with POPIA's research exemptions. ASSAf may consider including this as a consolidated legal-ethical note for clarification.[8]

Special personal information

POPIA broadly regulates two types of information: personal information and special personal information, which are each defined in POPIA. Importantly, special personal information is defined with reference to personal information (in section 1 of POPIA) as follows: '"special personal information' means personal information as referred to in section 26'. Ergo, special personal information is a *subclass* of personal information. Logic dictates that provisions in POPIA that apply to personal information would *include* special personal information *qua* subclass of personal information, unless special personal information is specifically excluded.

Now consider POPIA's processing requirements. In respect of *personal information*, section 11 provides that at least one of the legal grounds for processing personal information listed in that section must be present. In respect of *special personal information*, section 26 provides that if personal information qualifies as special personal information, it may not be processed – except if at least one legal ground for processing special personal information listed in section 27 is present. Importantly, given that special personal information is a subclass of personal information, special personal information is not exempted from compliance with section 11. Accordingly, sections 26 and 27 of POPIA apply as an *extra layer* of rules over the basic provisions of section 11.9 Stated differently, sections 26 and 27 are *not* an *alternative* compliance pathway for special personal information; they constitute an *additional* compliance pathway for special personal information.

However, the draft CCR (paragraph 4.3.3.3.5, page 19) states that 'Any of the following legal justifications [referring to the grounds listed in section 11 of POPIA] must apply when the Research does NOT include Special Personal Information.' This has the effect of exempting special personal information from compliance with section 11. We suggest that this is a mistake that should be corrected. Moreover, it would assist the research community if a *general* provision could be included in the draft CCR that clarifies that special personal information is a subclass of personal information that must comply with an *extra layer* of requirements. This is important, as the issue is not only relevant to the processing limitations in POPIA, but also to further processing and the research exceptions (section 15(3)(*e*) for personal information and section 27(1)(*d*)(*i*) for special personal information).

Responsible party

Section 1 of POPIA defines a responsible party as 'a public or private body or any other person which, alone or in conjunction with others, determines the purpose of and means for processing personal information'. Accordingly, who qualifies as a responsible party is a question of fact: who determines the purpose of, and means for, processing personal information? In the research context, a responsible party is likely to include both (1) individual researchers; and (2) research institutions – i.e. juristic persons that conduct research.[10] Importantly, no juristic person can ever determine the purpose of, and means for, processing personal information on its own – there must always be at least one individual who makes this determination.[10] We acknowledge that it may be possible, pragmatic even, to proceed from the standpoint that in a typical employment relationship the employee merely acts as an agent of their employer. But we warn that it is dangerous to assume that this will always be the case.

In research it is the principal investigator who will exert primary control over determining the purpose and means of processing. In our view, researchers and their employing research institutions would all be 'responsible parties' as defined in section 1 of POPIA. Importantly, this statutory definition cannot be narrowed by a code of conduct if that would amount to amending the provisions of the statute.[10] However, that is exactly what the draft CCR purports to do. It proposes that a responsible party does *not include* an individual researcher if such a researcher is in the employ of a research institution. In other words, irrespective of the *facts* of who actually determines the purpose of, and means for, processing personal information, the draft CCR proposes that only the juristic person *qua* employer should be legally liable for POPIA compliance. This is a narrowing of the statutory definition provided for in POPIA and is clearly not permissible.

The draft CCR states (Table 11, page 61): 'An entity acting through its employees is vicariously liable for the actions of those employees, provided that the employee is acting within the course and scope of their employment.' This is true, but it does not mean that the employee is excluded from liability. Vicarious liability is a legal concept adopted in the branch of law that imposes liability for causing harm to another (known in South Africa as *delict*). Where the person who caused the harm was an employee acting within the course and scope of their employment, then the law holds *both* the employee and the employer jointly and severally liable. Importantly, vicarious liability thus does not exclude the employee from liability. The injured party is free to choose whether they will sue the employee, the employer or both. Vicarious liability is thus a tool in the hand of a *plaintiff* to choose who to sue; it is *not a defence* in the hand of an employee. In the context of POPIA and research, the plaintiff will be a *research participant* (or to use POPIA terminology, a *data subject*) who intends to sue for damages in terms of section 99. As a general rule in our law, a plaintiff has a procedural *right to choose who to sue* among potential defendants.[10,11] The plaintiff therefore has the right to choose to sue: (1) the individual researcher who determined the purpose of, and means for, processing personal information; or (2) the juristic person that employed the individual researcher (based on vicarious liability); or both (1) and (2).

Accordingly, if vicarious liability is properly understood, it is clear that it cannot serve as a rationale for narrowing POPIA's definition of responsible party. Quite the opposite – it highlights that the draft CCR's proposed exclusion of individual researchers from qualifying as responsible parties, if such researchers are in the employ of research institutions, will infringe on the legal procedural rights of data subjects, and hence compromise their right to access the courts, protected by section 34 of the South African Constitution.[12]

Consider the following hypothetical facts: Professor X is the principal investigator of a research project on HIV. Professor X writes a research protocol that provides for the collection of biospecimens and health information from people living in a local community. The protocol sets out the purpose of, and means for, processing the data generated from studying the biospecimens and the health information collected directly from the research participants. Professor X is in the employ of University Y. The research protocol is approved by University Y's health research ethics committee. However, after all the data have been generated and collected, Professor X materially fails to comply with POPIA. As a result, there is a data breach and the identities of the research participants, who are HIV positive, become public knowledge. Z, one of the research participants who has been identified as being HIV positive, is ostracised by his community and loses his work. Z's legal aid attorney writes a letter to both Professor X and University Y, requesting a meeting in an attempt to settle the matter in a non-litigious fashion. The meeting takes place and University Y's representatives offer a sincere apology to Z. They also agree to issue a public apology. Z accepts this offer from University Y. However, Professor X refuses. She states that people must learn to be open about their HIV status, and leaves the meeting. Z has the right to choose to sue Professor X alone for damages. Any attempt to deny Z this right would infringe on Z's constitutional right of access to justice.

Table A: Recommendations for deletions, alterations, and inclusions in the draft CCR

Core concept	Recommendations
Consent	*Key recommendation:* Remove the explanation relating to *consent* for future research. Table 4, page 20, delete*:* ~~'POPIA Consent for future use is allowed as long as the future uses of the Personal Information are not speculative, are described as fully as possible, and further use of the Personal Information is restricted.'~~*Further recommendation:* For clarity, ASSAf may consider providing a consolidated legal-ethical note, which explains that the mode of consent for *future* research is determined by the relevant institutional research ethics committee.
Special personal information	*Key recommendation:* Clarify that *special personal information* is a subclass of *personal information* that must comply with an *extra layer* of requirements. We recommend that ASSAf include a *general* provision in the draft CCR that clarifies this. Paragraph 4.3.3.3.5 and Table 4, page 19, amend: 'Any of the following legal justifications [referring to the grounds listed in section 11] must apply ~~when the Research does NOT include Special Personal Information~~ to the processing of Personal Information'.Paragraph 4.3.3.3.5, page 19, recommended wording of the suggested inclusion: Note: Special Personal Information is a subclass of Personal Information. When Research includes Special Personal Information, it must comply with an extra layer of requirements. [See Table 5].'
Responsible party	*Key recommendation:* Do not exclude individual researchers employed by research institutions and public bodies from qualifying as *responsible parties*. Page 7, paragraph 2.2.2, amend*:* '~~In most instances,~~ The private or Public Body that employs, or directly controls the researchers will be the Responsible Party , along with the researcher(s) who determine(s) how and why personal information is processed.Page 7, paragraph 2.2.2, delete*:* ~~'The researchers will only be Responsible Parties in their individual capacity if a Responsible Party (private or Public Body) does not employ or control them; in other words if they are Independent Researchers. The definition of 'employee' is in the Labour Relations Act 66 of 1995 or the Basic Conditions of Employment Act 75 of 1997. Even if a researcher is not an employee, they may be still under the control of a Research Institution.'~~Table 11, page 61, delete: ~~'In most instances, the private or Public Body that employs, or directly controls the researchers will be the Responsible Party. The researchers will only be Responsible Parties in their individual capacity if the Responsible Party (private or Public Body) does not employ or control them; in other words if they are Independent Researchers.'~~*Further recommendation:* The draft CCR may include a recommendation that public bodies and research institutions indemnify their own employees against litigation.
De-identification	*Key recommendations:* • Throughout the draft CCR, clearly differentiate *de-identification* in the South African context from corresponding terms used in other jurisdictions. • Delete the terms *anonymise/anonymised/anonymisation* when used as a synonym for de-identification. (For example, see paragraph 1.1.2.2.1, page 5, and paragraph 1.3, page 64.) • Explain references to foreign terms and documents. (For example, see Table 9, page 40.) • Remove references to foreign tests (used in the draft CCR to determine whether data is de-identified). The test within POPIA is unique. (For example, see paragraph 1.7, page 65, and Table 1, page 72.) *Further recommendations:* We recommend that the use of foreign terms should be avoided, as these clearly have different definitions, standards, and tests to what is required by POPIA. Where they are employed, it should be made clear that *de-identification* in POPIA is not equivalent to corresponding terms used in other jurisdictions and that the test for de-identification is unique to South Africa.

Clearly, the draft CCR's proposed exclusion of individual researchers from qualifying as responsible parties, if such researchers are in the employ of research institutions, is not only misaligned with POPIA, but is also on shaky constitutional ground. We suggest that ASSAf rethink this issue. Instead of attempting to extract individual researchers from litigation, which is not legally possible, a better way to ameliorate the position of individual researchers in the face of strict liability litigation is for research institutions to indemnify their own employees against such litigation.[10] This approach can be *recommended* in the draft CCR, and each research institution can decide whether, and how, to implement it. While indemnification will not extract individual researchers from the litigation process, it will ensure that they are not personally bankrupted by their legal costs and a potential damages award.

De-identification

The use of foreign terminology in the draft CCR, especially the use of *anonymisation* as a synonym for POPIA's reference to *de-identification*, is problematic. While *anonymisation* and *de-identification* are used in other jurisdictions and frequently appear in biomedical literature, the concepts are distinct and often defined in conflicting ways, both within statues and

the literature.[13] Because there is no global consensus on the meaning and use of the terms *de-identification* and *anonymisation*, particular care is needed when employing them. Therefore, we recommend that these terms should not be used synonymously in the draft CCR; instead, they should be explained to clarify how POPIA's *de-identification* differs from corresponding terms found in other jurisdictions. The correct term in South Africa is *de-identification* – and it has been given a distinct definition in section 1 of POPIA, where it means

> *to delete any information that— (a) identifies the data subject; (b) can be used or manipulated by a reasonably foreseeable method to identify the data subject; or (c) can be linked by a reasonably foreseeable method to other information that identifies the data subject.*

For information to be de-identified, and excluded from the scope of POPIA, section 6(1)(*b*) provides that the information must be de-identified 'to the extent that it cannot be re-identified'. The test employed is unique to South Africa – and quite stringent: if any information can be 'used', 'manipulated' or 'linked' by 'a reasonably foreseeable method' to re-identify the data subject, it is not de-identified information in terms of POPIA (section 1 of POPIA).

Although the term *de-identification* is also used in the USA, the concept (and its meaning) differs from that in POPIA. Within the *Health Insurance Portability and Accountability Act of 1996* (HIPAA), de-identification relies on applying one of two methods: (1) the removal of 18 personal identifiers from a data set, with the important proviso that the researcher has no 'actual knowledge' that the residual information may identify an individual (the *Safe Harbor* method, §164.514(b)(2)); or (2) a determination made by a suitably qualified expert who must establish that the risk of re-identification 'is very small' (the *Expert Determination* method, §164.514(b)(1)).[14] Both methods speak to the overarching principle that, to be de-identified, the information 'does not identify an individual and… there is no reasonable basis to believe that the information can be used to identify an individual' (§164.514(a)). Notably, a covered entity may still assign a code or other means of record identification to allow de-identified information to be re-identified at a later stage (§164.514(a)), whereas this would not be permitted under POPIA's definition of de-identification.

The term *de-identification* is also used in the United Kingdom's *Data Protection Act 2018* (DPA). Section 171(1) of the DPA provides that it is an offence to 'knowingly and recklessly re-identify information that is de-identified personal data'. It is in this context that section 171(2)(a) of the DPA provides that 'personal data is "de-identified" if it has been processed in such a manner that it can no longer be attributed, without more, to a specific data subject'. Therefore, de-identification in the context of HIPAA and the DPA is clearly not equivalent to de-identification in POPIA.

Furthermore, although *pseudonymisation* is used within the draft CCR, it is not a concept found in POPIA. Rather, it appears in other international instruments such as the European Union's General Data Protection Regulation 2016/679 (GDPR) where, in Article 4(5), pseudonymisation is described as the process of storing additional information separately that can be used to re-identify the data subject.[15] As re-identification is possible in the case of pseudonymised data, the data continue to be considered personal data and must be treated as such. Thus, data that can be re-linked using a code, algorithm or pseudonym remain personal data under the GDPR, and the position is the same under POPIA. Although POPIA does not expressly refer to the term, it remains an important data privacy safeguard, and section 19 is broad enough to encompass pseudonymisation of data in its requirement that 'appropriate, reasonable technical and organisational measures' are used to secure the privacy of data subjects.

For this reason, the draft CCR should avoid the use of terms that clearly have different standards and definitions to what is required by POPIA. Instead, it should be made clear that *de-identification* in South Africa is not equivalent to corresponding terms employed in other jurisdictions and that the test for de-identification is unique to South Africa. We recommend that ASSAf clarifies the South African position to resolve any confusion around the use of these terms.

Conclusion

The draft CCR is breaking new ground in a relatively new field of the law. If approved by the Information Regulator, the draft CCR will – for the duration of its 5-year lifecycle – become an important document for the South African research community. Accordingly, before submitting the draft CCR to the Information Regulator, it is essential that ASSAf irons out the issues that we have identified in this article – especially because these issues relate to four core concepts in POPIA. Our recommendations are summarised in Table A.

Acknowledgements

We thank Beverley Townsend for her useful comments on previous drafts of this article. Any remaining errors are our own. We acknowledge the support by the US National Institute of Mental Health of the US National Institutes of Health (award number U01MH127690). The content of this article is solely our responsibility and does not necessarily represent the official views of the US National Institute of Mental Health or the US National Institutes of Health.

Competing interests

We have no competing interests to declare.

References

1. Academy of Science of South Africa (ASSAf). Code of Conduct for Research [document on the Internet]. c2022 [cited 2022 Oct 28]. Available from: https://www.assaf.org.za/wp-content/uploads/2022/09/20220923_ASSAF_Draft-Code_V8.7.pdf
2. Information Regulator. Guidance note on [the] processing of special personal information [document on the Internet]. c2021 [cited 2022 Oct 28]. Available from: https://www.justice.gov.za/inforeg/docs/InfoRegSA-GuidanceNote-Processing-SpecialPersonalIn formation-20210628.pdf
3. Thaldar D. Research and the meaning of 'public interest' in POPIA. S Afr J Sci. 2022;118(3/4), Art. #13206. https://doi.org/10.17159/sajs.2022/13206
4. Townsend BA, Thaldar DW. Navigating uncharted waters: Biobanks and informational privacy in South Africa. S Afr J Hum Rights. 2019;35(4):329–350. https://doi.org/10.1080/02587203.2020.1717366
5. Thaldar DW, Townsend B. Genomic research and privacy: A response to Staunton et al. S Afr Med J. 2020;110(3):172. http://doi.org/10.7196/SAMJ.2020.v110i3.14431
6. Swales L. The Protection of Personal Information Act 4 of 2013 in the context of health research: Enabler of privacy rights or roadblock? Potchefstroom Electron Law J. 2022;25:1–32, Art. #9490. https://doi.org/10.17159/1727-3781/2022/v25i0a11180
7. Staunton C, Adams R, Botes M, Dove ES, Horn L, Labuschaigne M, et al. Safeguarding the future of genomic research in South Africa: Broad consent and the Protection of Personal Information Act No. 4 of 2013. S Afr Med J. 2019;109(7):468. https://doi.org/10.7196/SAMJ.2019.v109i7.14148
8. Thaldar D, Townsend B. Protecting personal information in research: Is a code of conduct the solution? S Afr J Sci. 2021;117(3/4), Art. #9490. https://doi.org/10.17159/sajs.2021/9490
9. Thaldar DW, Townsend BA. Exempting health research from the consent provisions of POPIA. Potchefstroom Electron Law J. 2021;24:1–32. https://doi.org/10.17159/1727-3781/2021/v24i0a10420
10. Swales L, Thaldar D, Donnelly DL. Why research institutions should indemnify researchers against POPIA civil liability. S Afr J Sci. 2022;118(3/4), Art. #13205. https://doi.org/10.17159/sajs.2022/13205
11. Parekh v Shah Jehan Cinemas (Pty) Ltd 1982 (3) SA 618 (D).
12. Constitution of the Republic of South Africa, 1996.
13. Chevrier R, Foufi V, Gaudet-Blavignac C, Robert A, Lovis C. Use and understanding of anonymization and de-identification in the biomedical literature: Scoping review. J Med Internet Res. 2019;21(5), e13484. http://www.jmir.org/2019/5/e13484/
14. Health Insurance Portability and Accountability Act, 1996, United States of America.
15. General Data Protection Regulation (EU) 2016/679, European Union.

AUTHORS:
Beverley Townsend[1,2]
Amy Gooden[2]
Marietjie Botes[2,3*]
Donrich Thaldar[2,4]

AFFILIATIONS:
[1]York Law School, University of York, York, UK
[2]School of Law, University of KwaZulu-Natal, Durban, South Africa
[3]SnT Interdisciplinary Centre for Security, Reliability, and Trust, University of Luxembourg, Luxembourg, Luxembourg
[4]Petrie-Flom Center for Health Policy, Biotechnology, and Bioethics, Harvard Law School, Cambridge, Massachusetts, USA
*Current: Department of Medicine, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa

CORRESPONDENCE TO:
Amy Gooden

EMAIL:
215031300@stu.ukzn.ac.za

HOW TO CITE:
Townsend B, Gooden A, Botes M, Thaldar D. Repurposing research data for commercial use: POPIA, a foil or a facilitator? S Afr J Sci. 2023;119(7/8), Art. #15075. https://doi.org/10.17159/sajs.2023/15075

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
research, commercialisation, repurposing, POPIA, Draft Code of Conduct for Research

FUNDING:
US National Institutes of Health (U01MH127690)

PUBLISHED:
8 August 2023

Repurposing research data for commercial use: POPIA, a foil or a facilitator?

Significance:

For considerations of public policy, research in South Africa is not (merely) a purpose in itself, but a fundamental contributor to the South African economy. Because research data – which may consist of personal information – often has commercial value, it is important that we consider legal pathways for repurposing research data for commercial use. The POPIA draft Code of Conduct for Research should provide guidance in this regard.

Introduction

It is common practice for institutions around the world to conduct research using donated biological samples and related data. But what if a research institution wants to use the data generated from these samples for commercial applications, or license the use of the data, or sell the data to a commercial entity? Allegations of the commercialisation of data initially collected for research are not new, with companies, universities, and state departments purportedly selling and commercialising biological samples and/or the data originally collected for research purposes.[1-7] It is not only a matter of the repurposing of data initially collected for research for commercial use, but crucially that these activities may have been done without the knowledge or consent of the participants to whom the data relate. Individuals, it would seem, are less likely to donate altruistically if, unbeknown to them and without their consent, an organisation is to benefit financially from their contribution.[7]

Sound data protection practices form the backbone of lawfully and successfully securing public confidence in data use.[8] This is particularly true in light of the dramatic increase in the volume of digitally available research information coupled with the ability to share such data widely. To this end, the South African *Protection of Personal Information Act 4 of 2013* (POPIA) sets out measures to safeguard personal information. While much of the literature has focused on data collected for a commercial purpose – such as data stored in a commercial biobank – being further processed (or 'repurposed') for research use[9-18], we consider here the implications when the position is reversed: that is, when data collected for research are repurposed for commercial use. This would typically be the case if research yields data that are commercially valuable. Examples would be data sets containing the genomic sequences and health information of large groups of people.

Our intention in this Commentary is threefold. First, we suggest that the commercialisation of research data is a significant issue for the South African research community, as it is not only a public policy imperative, but, in certain crucial circumstances, a statutory duty.

Second, and in light of this, we analyse how POPIA might act both as a foil and a facilitator of secondary data use for purposes of commercialisation. This is to say that we consider, first and foremost, the legal position regarding the repurposing and commercialisation of research data, as stipulated in POPIA, rather than taking any ethical position one way or the other. We acknowledge that compelling ethical reasons underlie the requirement for, and availability of, high-quality, curated, unbiased, and representative data, and that there is enormous value and utility in granting wider access to discrete, aggregated data sets. Equally, however, and precipitated on important value judgements involving autonomy and self-determination, personal data may be used beyond the context for which consent was initially granted or 'repurposed' for other uses where the downstream commercial applications cannot always be anticipated, often contrary to user expectations. For example, the unsupervised and unregulated mass collection of biometric data was recently reported in 23 African countries by various third-party actors, the use of which is for unknown purposes.[19] These are, unfortunately, not isolated anomalies, but are symptomatic of troubling precedent-setting trends in countries with inadequate privacy legislation and enforcement, and where privacy protection is not taken seriously. Guided by values, virtues, and considerations of human well-being, we need to rethink not just how data sets are used and re-used and for what purpose, and how innovative applications are built, but for whom and at what cost. Here, we thus demonstrate only, and narrowly, how POPIA helps to set guardrails for the repurposing of data for commercial use through the deployment of lawful data practices.

Finally, we make recommendations on how the proposed POPIA draft Code of Conduct for Research (for ease, we refer to it simply as the 'draft CCR'), recently published by the Academy of Science of South Africa (ASSAf)[20], can support the legislative framework by guiding the South African research community in this regard.

The importance of commercialising research data

Research and research data are invaluable to innovation and contribute significantly to the advancement of society. Research, while primarily and traditionally the domain of university research groups and university hospitals, is now increasingly being conducted by various actors and is frequently used in commercial applications.[21] Universities remain one of the primary institutions driven to commercialise research. Investment in emerging research areas leads to better funding, new jobs, industry development, and enhanced collaboration between universities and industry, which is often required in translating research into valuable products and therapies for the public.[22,23] Examples of such applications are in producing vaccines[24,25], in clinical trials[26,27], and in the development of medications, treatments, and products by pharmaceutical companies, and others[22,26-29]. Innovation and commercialisation promote economic growth and, in turn, advance human well-being and benefit society.[23]

South African science and innovation policy supports the commercialisation of research results.[30-33] This public policy position of support for commercialisation has even been enacted as a *statutory duty* in certain circumstances. The *Intellectual Property Rights from Publicly Financed Research and Development Act 51 of 2008* (IPR Act) provides that where publicly financed research results in intellectual property (IP), such IP *must* be protected and commercialised in a way that benefits South Africa (section 2 of the IPR Act). In other words, from a South African public policy viewpoint, research is not (merely) a purpose in itself, but has the instrumental goal of building the South African economy. This public policy perspective of research underpins our analysis in this article.

Analysis

Preliminary observations

Concerning commercialisation, there are two scenarios at play: first, consideration of the repurposing of data, initially collected for a commercial purpose, for a subsequent research purpose; and second, consideration of the inverse position, that is, where data initially collected for research are commercialised. The tangential question of whether research participants are entitled to the benefits of such commercialisation is beyond the scope of this article. We provide two points of clarification before the position regarding purpose specification and further processing is described.

First, POPIA does not apply to the processing of de-identified personal information that is incapable of being re-identified (section 6(1)(*b*)), hence research data that have been de-identified and then used for commercial purposes fall beyond the remit of POPIA. True de-identification of data (and because of their nature, genomic data) is questionable.[9,34-37] Such data fall within POPIA, unless de-identified or otherwise excluded in terms of section 6.

Second, the *Regulations Relating to Research with Human Participants* (Human Research Participant Regulations)[38], published in terms of the *National Health Act 61 of 2003* (NHA), provide that research participants must be informed of, inter alia, 'the expected benefits of the research' (regulation 5(*e*)) and 'the availability of beneficial products or interventions post-research (regulation 5(*n*)) – the implications of which are that any possible commercialisation must be communicated to research participants.

Purpose specification and further processing limitation

At the outset, section 13(1) of POPIA requires that personal information 'be *collected for a specific, explicitly defined and lawful purpose*' that is '*related to a function or activity of the responsible party*' and that (except for cases that fall within section 18(4) of POPIA) the data subject (the individual 'to whom personal information relates' (section 1 of POPIA) which, in this case, is the research participant) must *be made aware* of the purpose of the collection of the personal information (section 13(2)). Therefore, where data that are initially collected for the specific and explicitly defined purpose of research are subsequently re-used for a different purpose, such as selling the data, and reasonably practicable steps are not taken to make the data subject aware of this, this contravenes section 13(2) of POPIA.

However, following this, and taking into account the specific purpose for which the data were collected (referred to in section 13 of POPIA), any further processing of personal information must be in accordance with (or compatible with) this specific purpose (section 15(1) of POPIA). So, in the event that the data were collected for a research purpose, any subsequent research would be allowable as it is in accordance with its initial purpose, that is, of research. However, on a reading of sections 13 and 15 of POPIA, *prima facie*, this means that where the activities of the responsible party do not include commercial purposes, and the data were not collected for a commercial purpose (that is, the purpose was solely for research), then the responsible party is not permitted to further process this information for a commercial purpose as it falls outside the scope of their activities and is not compatible with that initial purpose – namely research.

POPIA: 'Compatible' purpose

POPIA sets out what is meant by 'compatible' and what are deemed 'not incompatible' purposes in section 15. Others in the literature have posed antithetical meanings of 'compatibility' as situated within research: whether an initial collection of personal information for medical research may be more broadly interpreted to mean that *any* other medically related research is compatible, or whether compatibility should be attributed a far narrower meaning – that is to say, data collected for medical research on a specific medical condition such as, diabetes, for example, would exclude medical research on other conditions, say asthma.[39] We do not argue the particularities of these positions here, save to say that we examine the position regarding the 'compatibility' of data collected for research being repurposed for 'commercial use' – a use that we interpret in an extended sense as *any* commercial use and *any further* commercial use.

As a point of departure, section 15(2)(*a*)–(*e*) of POPIA implores us to assess the following criteria when determining compatibility: the relationship between the purpose of the further processing and collection (section 15(2)(*a*)), the nature of the information (section 15(2)(*b*)), the effect that the further processing may have on the data subject (section 15(2)(*c*)), the manner in which the information was collected (section 15(2)(*d*)), and any contractual duties that exist between the parties (section 15(2)(*e*)). Thus, to determine whether further processing is permitted in a given situation will depend upon the circumstances of the case, the conditions and contracts in place, as well as the information's nature and effect, and its intended use.

Where data were collected for a non-commercial purpose, such as for research, the presupposition is that the data cannot be commercialised, as it would fall outside of the scope of the original purpose of collection, thereby precluding further processing for commercial purposes (in terms of section 15(2)(*a*) of POPIA). But all is not lost. Notwithstanding the section 15(2) compatibility criteria, POPIA lists instances where an otherwise incompatible purpose may be deemed compatible.

POPIA: 'Not incompatible' purpose

POPIA clarifies specific instances, in section 15(3), that are deemed *not in*compatible (that is, they *are* compatible) with the purpose of collection. These instances are when: *consent* has been granted for the further processing (section 15(3)(*a*)), the information is contained in a *public record* or the data subject has *deliberately made it public* (section 15(3)(*b*)), the information is needed for *legal purposes or national security interests* (section 15(3)(*c*)), further processing is required in order to *prevent a threat to health or safety* (section 15(3)(*d*)), the information is used for '*historical, statistical or research purposes*' and is further processed only for those purposes and will not be published in an identifiable manner (section 15(3)(*e*)), or an *exemption has been granted in terms of section 37* (section 15(3)(*f*)).

In light of the above, consider a situation in which data that were initially collected for a commercial purpose are then used for a research purpose. In this instance, further processing is allowed in terms of section 15(3)(*e*) of POPIA as the '*historical, statistical and research purpose*' provision can be relied upon, provided that the 'responsible party ensures that the further processing is carried out solely for such purposes and will not be published in an identifiable form' (section 15(3)(*e*)). However, the inverse is not true: data collected initially for a research purpose, save for relying on one of the provisions described in section 15(3) of POPIA in the paragraph above, cannot simply be re-used for commercial purposes. So, *unless* the commercial purpose falls within one of the section 15(3)(*a*)–(*f*) provisions, it will be of an *incompatible purpose*. Save for these particular instances, POPIA restricts the repurposing of research data for commercial purposes.

Accordingly, in situations where the information is in the public domain or is deliberately made public by the data subject, in terms of section 15(3)(*b*) of POPIA, the information may be used for the further commercialised purpose. Moreover, and importantly, in terms of section 15(3)(*a*) of POPIA, where the consent of the data subject has been obtained, a researcher may use their information for a commercial purpose because

POPIA deems it to then be compatible with the purpose of collection (in this case, for research). But this purpose (that is, commercialisation) would still need to be specific and explicitly defined – in line with section 13(1) of POPIA. Although Staunton et al.[40] contended that POPIA can somehow be 'purposively interpreted' to allow for *broad* consent, we suggest that such an interpretation is incorrect.[41] The language used in POPIA is unambiguous. Unless an exemption has been granted in terms of POPIA, the consent requirement in POPIA is *specific* consent – both for the initial collection (section 13(1)) and for the further processing (section 15(1) and (3)(*b*)).[42,43]

What then of commercial applications which have a wide (and undetermined) scope at the time of collection and repurposing? In this regard, a contract between the responsible party and the data subject may shed some light on whether, and under what conditions, further processing is allowable. If a research participant were to agree to donate their biological material and/or data for the purpose of *research*, and the contract was *silent or prevented* the research institution from subsequently using these data for commercial purposes, then the further processing of research data for commercial reasons would not be allowable (in terms of section 15(2)(*e*) of POPIA). Much will, however, depend on the initial and subsequent purpose (for research or for commercial use), what has been specifically consented to, and the nature and extent of such consent. As a solution of last resort, data subjects may need to re-consent if the stipulated consents are not in place.

Enabling repurposed research data commercialisation

POPIA offers a further possible enabling solution – one that is an alternative to obtaining consent – in section 15(3)(*b*). This section deems information contained in a 'public record' or that which the data subject has 'deliberately made public' not incompatible with the purpose of collection. In addition to the research-related provisions already available to researchers, section 15(3)(*b*) of POPIA may prove helpful in supporting and growing science and technology innovations as envisaged in the South African National Development Plan.[44] Such considerations inform economic growth and improve health systems, education, and infrastructure, by creating a public record of personal data and collaborating with participants to voluntarily make their information public.

By way of illustration, the Finnish biobank, Auria, demonstrates how biobank and personal research data can be used for commercial purposes to benefit participants.[45] In this context, Finish citizens place trust in researchers and public institutions by willingly donating, and agreeing to make public, personal research data for commercial use.[46] A model based on commercialisation – with reciprocal benefit and data made publicly available by the participant – that strengthens collaboration and participation with data subjects is advantageous in that it not only fosters long-term relationships between the parties (be they researchers or otherwise), but remains in compliance with data protection laws.[47] Accordingly, and following POPIA, data collected for research purposes may be repurposed for commercial use if such data are contained in a public record or deliberately made public by the data subject. A model, not dissimilar to that of the one adopted in Finland, could be extended to South African research repositories that want to commercialise data.

Conclusion and recommendations regarding the draft CCR

The draft CCR was developed, inter alia, to create legal certainty regarding the interpretation of POPIA's provisions, to ensure that personal information used in research is protected, and to assist research institutions and independent researchers in complying with POPIA.[20] POPIA does not mention the commercialisation of personal information, and therefore it is left up to the draft CCR to address this issue and guide researchers and research institutions, while protecting participants. But how does the draft CCR recognise, and manage, situations in which researchers want to use personal information, including special personal information, either (1) from a previous research project; or (2) for a different purpose (which could include commercial purposes)?[20] Where researchers seek to use personal information collected from a previous research project for further research, or where researchers intend to use personal information for a new research-related purpose, the draft CCR requires researchers to provide certain information in a new research protocol, including: (1) the conditions under which the personal information was originally collected (including disclosures to research participants and information about consent); (2) how researchers will ensure that the personal information is used only for research and will not be published in an identifiable form; (3) how the notification requirement in section 18 of POPIA will be complied with; and (4) whether permission has been granted by the responsible party who initially processed the personal information.[20] The draft CCR acknowledges the possibility of a changing purpose – which, although not explicitly stated in the draft CCR, may imply repurposing research data for commercial use. If this is indeed the case, this must be done in accordance with section 15(1) of POPIA.

However, the draft CCR concerns itself only with situations where personal information collected for another purpose is re-used for *research*.[20] If the purpose changes to another type of research, it may be possible for this to fall within the scope of section 15(1) of POPIA as it remains within the boundaries of research and may be what the draft CCR envisioned when drafting this section. The draft CCR approaches research as a one-directional, inward flow of personal information collected for other purposes *into research*, but fails to provide for an outward flow of personal information initially collected for purposes of research *into commercialisation*. As we highlighted above, the commercialisation of research results is central to South African science and innovation policy – and a statutory duty to commercialise exists where public funds are used and commercially viable IP is involved.

Accordingly, given the importance of the commercialisation of research results to the South African research community, we suggest that the draft CCR address this important issue and lay out a *clear roadmap* for the South African research community on how to commercialise personal information initially collected for research. It would also be helpful for the draft CCR to include practical examples of consent statements for consent to the repurposing of personal data (initially collected for research) for commercialisation.

Against this recommendation, it can be argued that the topic of the commercialisation of research results is not research, and should therefore be excluded from the draft CCR. The draft CCR is, after all, intended to deal only with research. Although this argument would have technical merit, such a position is, in our estimation, myopic. Considerations about the commercialisation of personal data initially collected for research are often – and should be – a part of the design of research projects, and are a reality facing researchers and research institutions – whether initially planned or not. In other words, practically, decision-making regarding the commercialisation of research results is intrinsically integrated with research decision-making. Accordingly, the South African research community would benefit from a code of conduct for research that acknowledges this and considers how the commercialisation of research data might be realised.

Acknowledgements

We acknowledge support from the US National Institute of Mental Health of the US National Institutes of Health (award number U01MH127690). The content of this article is solely our responsibility and does not necessarily represent the official views of the US National Institute of Mental Health or the US National Institutes of Health.

Competing interests

We have no competing interests to declare.

References

1. Stokstad E. Major U.K. genetics lab accused of misusing African DNA. Science News. 2019 Oct 30. https://doi.org/10.1126/science.aba0343
2. Njilo N. Stellenbosch University demands return of DNA samples – But UK lab hits back. Times Live. 2019 Oct 16. Available from: https://www.timeslive.co.za/news/south-africa/2019-10-16-stellenbosch-university-demands-return-of-dna-samples-but-uk-lab-hits-back/

3. Blanchard S, Randall I. South African scientists demand the return of hundreds of tribal DNA samples after a British institute was accused of trying to use them to make money. Daily Mail. 2019 Oct 14. Available from: https://www.dailymail.co.uk/sciencetech/article-7570501/UK-lab-told-return-DNA-African-tribes-accused-trying-commercialise-them.html

4. The Associated Press. Texas sued over sale of baby blood samples. CBS News. 2010 Dec 09. Available from: https://www.cbsnews.com/news/texas-sued-over-sale-of-baby-blood-samples/

5. Aaronson B. Baby blood battle continues. The Texas Tribune. 2010 Dec 08. Available from: https://www.texastribune.org/2010/12/08/lawsuit-alleges-dshs-sold-baby-dna-samples/

6. Moore v Regents of the University of California 793 P.2d 479 (Cal. 1990).

7. Pike ER. Securing sequences: Ensuring adequate protections for genetic samples in the age of big data. Cardozo L Rev. 2015;37:1977–2033. https://ssrn.com/abstract=2658306

8. Organisation for Economic Co-operation and Development (OECD). The role of data in building trust. The path to becoming a data-driven public sector. Paris: OECD Publishing; 2019. https://doi.org/10.1787/059814a7-en

9. Varga EA. You want to do what? My mother's choice to have direct-to-consumer genetic testing. J Genet Counsel. 2012;21(3):382–385. https://doi.org/10.1007/s10897-012-9482-1

10. Gutmann Koch V. PGTandMe: Social networking-based genetic testing and the evolving research model. Health Matrix. 2012;22(1):33–74.

11. Stoeklé HC, Mamzer-Bruneel MF, Vogt, G, Hervé C. 23andMe: A new two-sided data-banking market model. BMC Med Ethics. 2016;17(19):1–11. https://doi.org/10.1186/s12910-016-0101-9

12. Bathe OF, McGuire AL. The ethical use of existing samples for genome research. Genet Med. 2009;11:712–715. https://doi.org/10.1097/GIM.0b013e3181b2e168

13. Caenazzo L, Tozzo P. The future of biobanking: What is next? BioTech. 2020;9(23):1–6. https://doi.org/10.3390/biotech9040023

14. Anderlik MR. Commercial biobanks and genetic research. Am J PharmacoGenomics. 2003;3(3):203–215. https://doi.org/10.2165/00129785-200303030-00006

15. O'Doherty KC, Christofides E, Yen J, Bentzen HB, Burke W, Hallowell N, et al. If you build it, they will come: Unintended future uses of organised health data collections. BMC Med Ethics. 2016;17:1–16. https://doi.org/10.1186/s12910-016-0137-x

16. Swede H, Stone CL, Norwood AR. National population-based biobanks for genetic research. Genet Med. 2007;9:141–149. https://doi.org/10.1097/GIM.0b013e3180330039

17. Martin-Sanchez FJ, Aguiar-Pulido V, Lopez-Campos GH, Peek N, Sacchi L. Secondary use and analysis of big data collected for patient care. Yearb Med Inform. 2017:28–37. https://doi.org/10.15265/IY-2017-008

18. Andreotta AJ, Kirkham N, Rizzi M. AI, big data, and the future of consent. AI Soc. 2021:1–14. https://doi.org/10.1007/s00146-021-01262-5

19. Unsafe national biometric data collection in 23 African countries [webpage on the Internet]. c2022 [cited 2023 Feb 09]. Available from: https://www.ictworks.org/national-biometric-data-collection/#.Y-Tihi2l1R4

20. Academy of Science of South Africa (ASSAf). POPIA Code of Conduct for Research. Pretoria: ASSAf; 2023. https://www.assaf.org.za/wp-content/uploads/2023/04/ASSAf-POPIA-Code-of-Conduct-for-Research.pdf

21. Quinn P. Research under the GDPR – A level playing field for public and private sector research? Life Sci Soc Policy. 2021;17(4):1–33. https://doi.org/10.1186/s40504-021-00111-z

22. Caulfield T, Ogbogu U. The commercialization of university-based research: Balancing risks and benefits. BMC Med Ethics. 2015;16:1–7. https://doi.org/10.1186/s12910-015-0064-2

23. Levy HV. The transformation of basic research into commercial value: Economics aspects and practical issues. JEMI. 2011;7:1–15.

24. Moore JP, Wilson IA. Decades of basic research paved the way for today's 'warp speed' Covid-19 vaccines. STAT. 2021 Jan 05. Available from: https://www.statnews.com/2021/01/05/basic-research-paved-way-for-warp-speed-covid-19-vaccines/

25. Wardle JL, Baum FE, Fisher M. The research commercialisation agenda: A concerning development for public health research. Aust N Z J Public Health. 2019;43(5):407–409. https://doi.org/10.1111/1753-6405.12930

26. Institute of Medicine (US) Committee on Health Research and the Privacy of Health Information. The value, importance, and oversight of health research. In: Nass SJ, Levit LA, Gostin LO, editors. Beyond the HIPAA Privacy Rule: Enhancing privacy, improving health through research. Washington DC: National Academies Press; 2009. p. 111–152.

27. Sinclair J. Medical research needs more commercialisation support. Research Professional News. 2021 Mar 24. Available from: https://www.researchprofessionalnews.com/rr-news-australia-industry-2021-3-medical-research-needs-more-commercialisation-support/

28. Burningham S, Ollenberger A, Caulfield T. Commercialization and stem cell research: A review of emerging issues. Stem Cells Dev. 2013;22(1):80–84. https://doi.org/10.1089/scd.2013.0317

29. Samuel G, Hardcastle F, Broekstra R, Lucassen A. Exploring how biobanks communicate the possibility of commercial access and its associated benefits and risks in participant documents. BMC Med Ethics. 2022;23:1–14. https://doi.org/10.1186/s12910-022-00829-1

30. South African Department of Science and Technology (DST). The bio-economy strategy. Pretoria: DST; 2013. Available from: https://www.gov.za/sites/default/files/gcis_document/201409/bioeconomy-strategya.pdf

31. South African Department of Science and Innovation (DSI). South Africa foresight exercise for science, technology and innovation. Pretoria: DSI; 2019. Available from: https://www.naci.org.za/wp-content/uploads/2020/07/South-African-Foresight-Exercise-For-Science-Technology-and-Innovation-2019.pdf

32. South African Department of Science and Innovation (DSI). Strategic plan 2020-2025. Pretoria: DSI; 2020. Available from: https://www.dst.gov.za/images/2020/DST_2020_2025_STRAT_PLAN_FINAL_JULY.pdf

33. Higher Education, Science, Technology and Innovation Institutional Landscape (HESTIIL) Ministerial Committee. A new pathway 2030: Catalysing South Africa's NSI for urgent scaled social and economic impact. HESTIIL; 2020. https://www.dst.gov.za/images/2021/Higher%20Education,%20Science,%20Technology%20and%20Innovation%20Institutional%20Landscape%20Review%20Report.pdf

34. Academy of Science of South Africa (ASSAf), Department of Science and Technology (DST). Human genetics and genomics in South Africa: Ethical, legal and social implications. Pretoria: ASSAf/DST; 2018. http://dx.doi.org/10.17159/assaf.2018/0033

35. Cacchio J. What you don't know can hurt you: The legal risk of peering into the gene pool with direct-to-consumer genetic testing. UMKC L Rev. 2018;87:219–244.

36. Gymrek M, McGuire AL, Golan D, Halperin E, Erlich Y. Identifying personal genomes by surname inference. Science. 2013;339(6117):321–324. https://doi.org/10.1126/science.1229566

37. Erlich Y, Williams JB, Glazer D, Yocum K, Farahany N, Olson M, et al. Redefining genomic privacy: Trust and empowerment. PLoS Biol. 2014;12(11):1–5. https://doi.org/10.1371/journal.pbio.1001983

38. Regulations relating to research with human participants GN R719 GG 38000 of 19 September 2014, South Africa.

39. Moodley K, Kleinsmidt A. Allegations of misuse of African DNA in the UK: Will data protection legislation in South Africa be sufficient to prevent a recurrence? Developing World Bioeth. 2020:1–6. https://doi.org/10.1111/dewb.12277

40. Staunton C, Adams R, Botes M, Dove ES, Horn L, Labuschaigne M, et al. Safeguarding the future of genomic research in South Africa: Broad consent and the Protection of Personal Information Act No. 4 of 2013. S Afr Med J. 2019;109(7):468–470. https://doi.org/10.7196/SAMJ.2019.v109i7.14148

41. Thaldar DW, Townsend B. Genomic research and privacy: A response to Staunton et al. S Afr Med J. 2020;110(3):172–174. https://doi.org/10.7196/SAMJ.2020.v110i3.14431

42. Thaldar DW, Townsend BA. Exempting health research from the consent provisions of POPIA. Potchefstroom Electron Law J. 2021;24:1–31. https://doi.org/10.17159/1727-3781/2021/v24i0a10420

43. Townsend BA, Thaldar DW. Navigating uncharted waters: Biobanks and informational privacy in South Africa. South Afr J Hum Rights. 2019;35(4):329–350. https://doi.org/10.1080/02587203.2020.1717366

44. Republic of South Africa National Planning Commission. National Development Plan 2030: Our future – Make it work. Pretoria: NPC; 2012.

45. Lehtimäki H, Helén I, Snell K, Eriksson P, Montonen T. Sustainable value creation in the commercialisation of innovation: The case of Auria Biobank. Int J Entrep Innov Manag. 2019;23(5):451–465. https://doi.org/10.1504/IJEIM.2019.102035

46. Gaskell G, Gotweis H, Starkbaum J, Gerber MM, Broerse J, Gottweis U, et al. Publics and biobanks: Pan-European diversity and the challenge of responsible innovation. Eur J Hum Genet. 2013;21(1):14–20. https://doi.org/10.1038/ejhg.2012.104

47. Kujala J, Lehtimäki H, Myllykangas P. Value co-creation in stakeholder relationships: A case study. In: Freeman R, Kujala EJ, Sachs S, editors. Stakeholder engagement: Clinical research cases. Dordrecht: Springer; 2017. p. 15–30.

AUTHORS:
Donrich Thaldar[1,2]
Amy Gooden[1]
Dusty-Lee Donnelly[1]

AFFILIATIONS:
[1]School of Law, University of KwaZulu-Natal, Durban, South Africa
[2]Petrie-Flom Center for Health Law Policy, Biotechnology, and Bioethics, Harvard Law School, Cambridge, Massachusetts, USA

CORRESPONDENCE TO:
Donrich Thaldar

EMAIL:
ThaldarD@ukzn.ac.za

HOW TO CITE:
Thaldar D, Gooden A, Donnelly D-L. Toward an open access genomics database of South Africans: Legal considerations. S Afr J Sci. 2023;119(7/8), Art. #15069. https://doi.org/10.17159/sajs.2023/15069

ARTICLE INCLUDES:
☐ Peer review
☐ Supplementary material

KEYWORDS:
data protection, genomics, open science, Research Code of Conduct

FUNDING:
US National Institutes of Health (U01MH127690)

PUBLISHED:
8 August 2023

Toward an open access genomics database of South Africans: Legal considerations

Significance:

There may be significant benefits to an open access genomics database of South Africans. The proposed Code of Conduct for Research should be amended to provide a clear roadmap – aligned with the *Protection of Personal Information Act 4 of 2013* – for open access genomics projects.

Introduction

Why should individual-level genomic data be private? Concerns range from unintentionally discovering previously unknown family members to insurance discrimination based on health risks disclosed by the genomic data. Although many research participants express concern about the privacy of their genomic data, the picture of the nature and extent of their concerns is a complex and highly variable one.[1,2] Also, there are persons who are willing to share their genomic data in the public domain without requiring any privacy guarantees.[1-5] An iconic example is the Harvard Personal Genome Project (Harvard PGP) initiated in 2005[6], which has since become a global network of projects.[7] It publishes the whole genome sequences of its research participants online for anyone around the world to download[6] – no registration required, no paywall, and no data access committee. This is a truly open access, individual-level genomic database. Furthermore, its research participants may choose to supplement their genome sequences by also including phenotype and health information in the open access database.[6]

Imagine a South African version of the Harvard PGP, i.e. an open access, individual-level genomic database composed of the genomic data of thousands of South Africans, freely available to all. Advances in science and technology have resulted in improvements in the time, cost, and methods involved in genome sequencing[8,9], and the focus has now shifted to filling the gaps in the amount, and reliability, of population-level data about newly discovered genes and their links to disease. However, the success and effectiveness of various medicines and therapies, as well as the realisation of precision medicine, may be hampered by differences in the reference group and population on which clinical trials are conducted. This poses a real concern for countries like South Africa, whose major population groups are grossly underrepresented in existing genomic reference sets.[10-14] Establishing an *inclusive, open access genomics project* would not only align with all the benefits typically associated with open science, but may also offer a solution to the problem of underrepresentation.

But, is there a legal pathway to establishing such an *open access* genomics project in South Africa? In this article, we explore this question from the perspective of the *Protection of Personal Information Act 4 of 2013* (POPIA). Furthermore, the Academy of Science of South Africa (ASSAf) has recently submitted its long-awaited proposed Code of Conduct for Research (proposed CCR)[15], in terms of POPIA, to the Information Regulator. As such, where relevant, we make recommendations on how the proposed CCR should be amended to clarify the relevant law and provide guidance with regard to such an open access genomics project(s).

Terminology

A core concept in this article is *research*. As POPIA does not define research, it should be understood in its general meaning. A widely used definition of research is that of the Organisation for Economic Cooperation and Development (OECD), which reads as follows: 'Any creative systematic activity undertaken in order to increase the stock of knowledge, including knowledge of man, culture and society, and the use of this knowledge to devise new applications.'[16] The proposed CCR (at paragraph 1.1.2.3.1.) offers a definition of research that is very similar to the OECD definition, namely that research 'includes the range of activities that a private or Public Body conduct to extend knowledge through disciplined enquiry or systematic investigation'[15]. The only problem with the proposed CCR's definition is that it seems to exclude individual researchers who are not part of a 'private or public body', while the rest of the proposed CCR clearly contemplates the inclusion of such individual researchers. We suggest that the proposed CCR should be revised to include independent individual researchers within the ambit of its definition of research. This can be accomplished either by explicit inclusion of independent individual researchers in the definition, or by removing the reference to the entities that conduct research – in line with the OECD definition.

Analysis

Are individual-level genomic data regulated by POPIA?

The first part of the POPIA analysis is to establish whether individual-level genomic data fall under the two main types of information regulated by POPIA, namely personal information and special personal information:

- Personal information is information that relates to, inter alia, an 'identifiable, living, natural person' (section 1 of POPIA). As genomic data relate to a living natural person, and the person can be identified using the genomic data, genomic data qualify as *personal information*.
- Special personal information is a subclass of personal information that relates to, inter alia, a natural person's race, health, or biometric information (section 26 of POPIA). Genomic data relate to all three of these and therefore clearly qualify as *special personal information*.

Accordingly, individual-level genomic data are indeed regulated by POPIA at two levels: frst at the level of personal information, and second at the level of special personal information.[17]

Importantly, regulation of individual-level genomic data by POPIA commences from the moment that the genomic data are *generated* through sequencing and recorded in an electronic device (section 3(1)).[17,18] POPIA does not apply to biological samples or to DNA.[17,18] This is because POPIA applies only to personal information that is *entered in a record*, while genomic information is naturally present in DNA, rather than being entered in DNA.[17,18] Individual-level genomic data, once generated and recorded, will likely always fall within POPIA, as it seems unlikely that such genomic data can be de-identified.[19,20]

Uploading genomic data to an open access database

After individual-level genomic data are generated, the next step in the context of an open access genomics project would be for the project to upload data subjects' genomic data to the project's open access online database. This would qualify as *processing* the genomic data in terms of section 1 of POPIA (which includes 'any operation or activity or any set of operations' such as collecting, recording, organising, storing, disseminating, or making available in any other form). As a general rule, the processing of personal information and special personal information is only lawful if a legal ground for processing is present. *Consent* is a legal ground for processing in the case of both personal information (section 11(1)(*a*) of POPIA) and special personal information (section 27(1)(*a*) of POPIA). Consent is defined by POPIA as a 'voluntary, specific, and informed expression of will' (section 1). Accordingly, to be POPIA compliant, data subjects must voluntarily agree to the uploading of their genomic data to the open access online database, while understanding the consequence – that this will make their genomic data public – and the possible privacy risks thereof – that their genomic data contains information of a personal nature from which they can be identified.[21-24] For consent to be informed, data subjects would also need to understand that this will impact upon their data subject rights (section 5 of POPIA), including their rights in relation to the cross-border transfer of their data (sections 57 and 72 of POPIA), discussed below. But how does one know whether data subjects understand this, and whether their consent is therefore truly *informed*?

The Harvard PGP's solution was to develop a new consent model that they called *open consent*.[25,26] This entails, inter alia, that prospective participants are provided with resource material that explains not only the benefits to science of participating in the Harvard PGP, but also the potential risks to participants of being identified through their open genomic data.[27] In contrast with most research projects that do not assess whether prospective participants objectively understand what they are consenting to, open consent requires prospective participants to take an online entrance examination to *objectively assess their understanding*.[26,27] This examination can be taken repeatedly, but only prospective participants who achieve full marks are admitted as participants in the Harvard PGP.[26,27] In other words, the rationale behind open consent is that the heightened risk to privacy (by making one's genomic data open access) is offset by the heightened measure of objective assessment to ensure that consent is truly *informed*.

We suggest that objective assessment of understanding should be a requirement for open access genomics projects in South Africa, given the heightened risk to data subjects. Provided that this is complied with, i.e. that data subjects voluntarily agree to the uploading of their genomic data to the open access online database and pass an objective assessment showing that such consent is informed, the processing will be lawful.

Accessing the data subject's genomic data on the Internet

Once the genomic data are published on the Internet, anyone can access and use the data for research or for any other purpose, as the data are open access. This would qualify as further processing of the genomic data in terms of section 15 of POPIA. Again, as a general rule, (further) processing of personal information and special personal information is only lawful if a legal ground for such processing is present. One such legal ground for (further) processing of both personal information (section 15(3)(*b*) of POPIA) and special personal information (section 27(1)(*e*) of POPIA) is if information has *deliberately been made public by the data subject*.

This raises the question: must data subjects themselves perform every action necessary to make the relevant information public, or can other persons act as their agents? Although not covered in POPIA, the proposed CCR provides for situations where data subjects 'consented' that an 'intermediary' can intentionally make their personal information public.[15] This is a welcome provision in the proposed CCR. However, from a legal perspective, it would be better to use the stronger word 'instruct', as this would imply the nominate contract of 'mandate', which entails that one party (the mandatee) gratuitously performs a service for the other party (the mandator).[27-30] This nominate contract originates from Roman law, and automatically entails that the mandatee must exercise reasonable care when performing the mandate on instruction of the mandator.[29,30] We suggest that a mandate construction is essential to comply with sections 15(3)(*b*) and 27(1)(*e*) of POPIA.

Accordingly, it would be important in the context of open access genomics projects that data subjects not only *consent* to the uploading of their genomic data to an open access online database, and hence to making it public, but simultaneously also *instruct* the open access genomics project to perform said action. If this is done, the open access genomics project has a mandate to make the data subjects' genomic data public. Given that mandate is such a well-established part of South African law, this should suffice for compliance with sections 15(3)(*b*)) and 27(1)(*e*) of POPIA respectively. As a consequence, anyone would be able to lawfully use the genomic data, thus succeeding in the open science objective of the open access genomics project.

The cross-border aspect

An open access genomics project would make data openly accessible on the Internet – which means that the data would be available beyond physical geographical borders, and may thus bring about the provisions regarding transfers of personal information outside of South Africa in section 72 of POPIA. If the genomic data are stored on a server outside of South Africa, as is the case with many cloud services, this in itself would constitute a cross-border data flow. Further, whenever the data are downloaded outside of South Africa, there is a cross-border data flow. POPIA's provisions on cross-border data flows thus apply.

The cross-border transfer of personal information may only take place if, inter alia, the transfer is required in terms of a *contract* between the data subject and the responsible party (section 72(1)(*c*) of POPIA). As we have suggested above, data subjects should *instruct* the open access genomics project to upload their genomic data to its open access online database, as this would constitute the *contract* of mandate. Accordingly, the project giving access to anyone anywhere in the world to download the data would be in pursuance of the terms of the *contract*, and hence comply with POPIA's regime for the cross-border transfer of personal information.

However, because genomic data are not only personal information, but also special personal information, there is an additional requirement, namely that if a party who downloads the data are in a country that is not deemed to provide adequate protection for the processing of personal information, the open access genomics project must obtain prior authorisation from the Information Regulator (section 57(1)(*d*) of POPIA). Given that the South African Information Regulator has not yet issued a list of countries that it deems as providing adequate protection, no country currently qualifies as such. Also, even if the Information Regulator issues such an adequacy list, the purpose of the open access genomics project is to make its genetic data easily available to anyone in the world, regardless of whether the recipient is in a country that is deemed to provide adequate protection or not. Accordingly, the open access genomics project would need to apply for prior authorisation from the Information Regulator in order to comply with POPIA.

Importantly, if a code of conduct has come into force in the relevant sector of society, the prior authorisation requirement for the cross-border transfer of special personal information ceases to apply (section

57(3) of POPIA). Accordingly, if and when the proposed CCR comes into force, there will be respite for open access genomics projects in this regard. This highlights the importance of having sufficient provisions in the proposed CCR to properly regulate open access genomics projects.

Various rights of data subjects

Data subjects giving consent and instructing the open access genomics project to upload their genomic data to an open access online database do not exhaust the data subjects' rights from the perspective of POPIA. We briefly analyse other relevant rights of data subjects in the context of open access genomics projects and consider how these rights apply in the context of an open access genomics project.

First, data subjects would have the right to request (in terms of section 23 of POPIA) information from the open access genomics project about the identity of all third parties who have access to their genomic data. We suggest that this would place a duty on the open access genomics project: (1) to require would-be data downloaders to first register on the project website; and (2) to take reasonable measures – such as a verification email – to verify the registration information. Should data subjects exercise their right to request information about the identity of all third parties who have access to their genomic data, the open access genomics project would be in a position to provide this to them.

Second, data subjects would have the right to be notified of, inter alia, the data being collected, the identity of the responsible party, and the purpose of collection (section 18(1) of POPIA). An exception to this right is when data are collected for the purpose of research (section 18(4)(*f*)(*ii*) of POPIA). Also, data subjects can waive this right if they consent to non-compliance with the notification requirement (as provided for in terms of section 18(4)(*a*) of POPIA). Accordingly, an open access genomics project would have two options: (1) expand the registration requirement mentioned above by requiring would-be data downloaders to declare that they intend to use the data for research; and (2) incorporate a waiver of the notification right in the consent process for data subjects. Given that (1) does not provide any guarantees, we suggest that the best solution would be to implement both (1) and (2).

Third, data subjects would have the right to withdraw their consent at any time (section 11(2)(*b*) of POPIA), to object to the processing of their data on reasonable grounds (section 11(3)(*a*) of POPIA), and to request that their data be deleted (section 24 of POPIA). In the context of an open access genomics project, this would require that the project removes the data from its website. However, the project would not be under any obligation to take steps to have the data deleted by others who have already downloaded such data. In this regard, the Harvard PGP promises to take an extra step, namely that they will 'request any organizations or researchers with whom the PGP has any formal data sharing agreements to likewise delete your data and information within a reasonable time frame'[31]. We suggest that this would be a good policy to follow. Furthermore, the informed consent process should ensure that data subjects are aware of these rights, and how to exercise them.

Conclusion and recommendations

Privacy is a right of persons to be exercised as they deem fit. Persons are autonomous moral agents, and provided that they understand the risks to their privacy, they should be free to make their own genomic data public. Moreover, as there is public benefit in such open sharing of genomic data, this is, in principle, something that should be welcomed from a public policy perspective. However, reasonable protective measures – aligned with POPIA – should still be put in place.

We recommend that the proposed CCR should clearly provide a roadmap for a prospective open access genomics project to follow. Flowing from our analysis in this article, we suggest the landmarks in this roadmap are the following:

1. *General*. The regulation of genomic data by POPIA commences from the moment that the genomic data are generated through sequencing and are recorded on an electronic device. POPIA does not apply to biological samples or to DNA.[18] POPIA always applies to genomic data, as such data cannot be de-identified.
2. *Pertinent elements of consent*. The open access genomics project must provide resource material to prospective participants that explains, most pertinently in the context of POPIA: (1) what the open access genomics project entails and the possible risks to their privacy; and (2) their right to request information about persons who access their data, their right to withdraw, and the consequences of exercising these rights.
3. *Objective assessment.* The open access genomics project must require prospective participants to pass an objective assessment that assesses their understanding of the content of the resource material in order to ensure that their consent is truly *informed*.
4. *Consent plus mandate*. The open access genomics project must ensure that its participants: (1) consent to the uploading of their genomic data to an open access online database, and hence consent to making it public; and (2) instruct the open access genomics project to upload their genomic data to its open access online database.
5. *Registration*. The open access genomics project must: (1) require data downloaders to first register on the project website; (2) take reasonable measures, such as a verification email, to verify the registration information; and (3) require data downloaders to declare that they intend to use the data for research.

Note that the requirement for obtaining prior authorisation from the Information Regulator for cross-border transfers of data will fall away once a code of conduct for research is issued. As such, the suggested roadmap need not include prior authorisation. Implementing the above suggestions in the proposed CCR will provide clarity for the establishment of an open access genomics project, which in turn will benefit all South Africans.

Competing interests

We have no competing interests to declare.

Acknowledgements

Work on this article was supported by the US National Institute of Mental Health of the US National Institutes of Health (award number U01MH127690) under the Harnessing Data Science for Health Discovery and Innovation in Africa (DS-I Africa) programme. The content of this article is solely our responsibility and does not necessarily represent the official views of the US National Institute of Mental Health or the US National Institutes of Health.

References

1. Clayton EW, Evans BJ, Hazel JW, Rothstein MA. The law of genetic privacy: Applications, implications, and limitations. J Law Biosci. 2019;6(1):1–36. https://doi.org/10.1093/jlb/lsz007
2. Wan Z, Hazel JW, Clayton EW, Vorobeychik Y, Kantarcioglu M, Malin BA. Sociotechnical safeguards for genomic data privacy. Nat Rev Genet. 2022;23:429–445. https://doi.org/10.1038/s41576-022-00455-y
3. Haeusermann T, Greshake B, Blasimme A, Irdam D, Richards M, Vayena E. Open sharing of genomic data: Who does it and why? PLoS ONE. 2017;12(5):1–15. https://doi.org/10.1371/journal.pone.0177158
4. Ball MP, Thakuria JV, Zaranek AW, Clegg T, Rosenbaum AM, Wu X, et al. A public resource facilitating clinical use of genomes. Proc Natl Acad Sci USA. 2012;109(30):11920–11927. https://doi.org/10.1073/pnas.1201904109
5. Ball MP, Bobe JR, Chou MF, Clegg T, Estep PW, Lunshof JE, et al. Harvard Personal Genome Project: Lessons from participatory public research. Genome Med. 2014;6(10):1–7. https://doi.org/10.1186/gm527
6. The Harvard Personal Genome Project. About [webpage on the Internet]. No date [cited 2022 Oct 15]. Available from: https://pgp.med.harvard.edu/about
7. Personal Genome Project: Global network 'The Personal Genome Project' [homepage on the Internet]. No date [cited 2022 Oct 15]. Available from: https://www.personalgenomes.org/
8. Adams J. DNA sequencing technologies. Nat Educ. 2008;1(1):193. https://www.nature.com/scitable/topicpage/dna-sequencing-technologies-690/

9. Preston J, VanZeeland A, Peiffer DA. Innovation at Illumina: The road to the $600 human genome [document on the Internet]. c2021 [cited 2022 Oct 15]. Available from: https://media.nature.com/original/magazine-assets/d42473-021-00030-9/d42473-021-00030-9.pdf

10. Wits University. Why and how Africans need to participate in genetic studies [webpage on the Internet]. c2022 [cited 2022 Oct 27]. Available from: https://www.wits.ac.za/news/latest-news/research-news/2022/2022-02/why-and-how-africans-need-to-participate-in-genetic-studies-.html

11. Adepoju P. Tackling Africa's underrepresentation in genomics studies. Nature Africa News. 2022 April 05. Available from: https://www.nature.com/articles/d44148-022-00051-6

12. Jackson C. Africa's missing genomic data and its impact on health care. GEN. 2020 September 08. Available from: https://www.genengnews.com/insights/africas-missing-genomic-data-and-its-impact-on-health-care/

13. Bentley AR, Callier SL, Rotimi CN. Evaluating the promise of inclusion of African ancestry populations in genomics. NPJ Genom Med. 2020;5:1–9. https://doi.org/10.1038/s41525-019-0111-x

14. Bentley AR, Callier S, Rotimi CN. Diversity and inclusion in genomic research: Why the uneven progress? J Community Genet. 2017;8(4):255–266. https://doi.org/10.1007/s12687-017-0316-6

15. Academy of Science of South Africa (ASSAf). Code of Conduct for Research [document on the Internet]. c2023 [cited 2023 Jun 06]. Available from: https://www.assaf.org.za/wp-content/uploads/2023/04/ASSAf-POPIA-Code-of-Conduct-for-Research.pdf

16. Organisation for Economic Cooperation and Development (OECD). Frascati Manual: Guidelines for collecting and reporting data on research and experimental development [webpage on the Internet]. c2015 [cited 2022 Oct 15]. Available from: https://www.oecd.org/publications/frascati-manual-2015-9789264239012-en.htm

17. Thaldar DW, Townsend BA. Exempting health research from the consent provisions of POPIA. Potchefstroom Electron Law J. 2021;24:1–32. http://dx.doi.org/10.17159/1727-3781/2021/v24i0a10420

18. Thaldar D. Why POPIA does not apply to DNA. S Afr J Sci. 2021;117(7/8), Art. #11286. https://doi.org/10.17159/sajs.2021/11286

19. Pike ER. Securing sequences: Ensuring adequate protections for genetic samples in the age of big data. Cardozo L Rev. 2015;37:1977–2033. https://ssrn.com/abstract=2658306

20. Townsend BA, Thaldar DW. Navigating uncharted waters: Biobanks and informational privacy in South Africa. S Afr J Hum Rights. 2019;35(4):329–350. https://doi.org/10.1080/02587203.2020.1717366

21. Gymrek M, McGuire AL, Golan D, Halperin E, Erlich Y. Identifying personal genomes by surname inference. Science. 2013;339(6117):321–324. https://doi.org/10.1126/science.1229566

22. Erlich Y, Williams JB, Glazer D, Yocum K, Farahany N, Olson M, et al. Redefining genomic privacy: Trust and empowerment. PLoS Biol. 2014;12(11):1–5. https://doi.org/10.1371/journal.pbio.1001983

23. Sweeney L, Abu A, Winn J. Identifying participants in the Personal Genome Project by name. SSRN. 2013:1–4. http://dx.doi.org/10.2139/ssrn.2257732

24. Malin B, Sweeney L. How (not) to protect genomic data privacy in a distributed network: Using trail re-identification to evaluate and design anonymity protection systems. J Biomed Inform. 2004;37(3):179–192. https://doi.org/10.1016/j.jbi.2004.04.005

25. Lunshof JE, Chadwick R, Vorhaus DB, Church GM. From genetic privacy to open consent. Nat Rev Genet. 2008;9:406–411. https://doi.org/10.1038/nrg2360

26. Zarate OA, Green Brody J, Brown P, Ramírez-Andreotta MD, Perovich L, Matz J. Balancing benefits and risks of immortal data: Participants' views of open consent in the Personal Genome Project. Hastings Center Rep. 2016;46(1):36–45. https://doi.org/10.1002/hast.523

27. Angrist M. Eyes wide open: The Personal Genome Project, citizen science and veracity in informed consent. Per Med. 2009;6(6):1–13. https://doi.org/10.2217/pme.09.48

28. Hutchison D, Pretorius CJ, Du Plessis J, Eiselen S, Floyd T, Hawthorne L, et al. The law of contract in South Africa. 2nd ed. Cape Town: Oxford University Press; 2012.

29. Van Zyl DH, Joubert DJ. Mandate and negotiorum gestio. In: The law of South Africa. 3rd ed. Vol. 28(1). Johannesburg: LexisNexis; 2020.

30. Dendy M. Agency and representation. In: The law of South Africa. 3rd ed. Vol. 1. Johannesburg: LexisNexis; 2013.

31. Harvard Personal Genome Project. Consent form. Harvard PGP [cited 2022 Oct 26]. Available from: https://my.pgp-hms.org/static/PGP_Consent_2021-07-12_online.pdf

AUTHORS:
John B. Noonoo[1]
Edem Sosu[1,2]
Francis Hasford[1,2]

AFFILIATIONS:
[1]School of Nuclear and Allied Sciences, University of Ghana, Accra, Ghana
[2]Radiological and Medical Sciences Research Institute, Ghana Atomic Energy Commission, Accra, Ghana

CORRESPONDENCE TO:
John Noonoo

EMAIL:
jbnoonoo@gmail.com

DATES:
Received: 05 July 2022
Revised: 14 Dec. 2022
Accepted: 12 Jan. 2023
Published: 08 Aug. 2023

HOW TO CITE:
Noonoo JB, Sosu E, Hasford F. Three-dimensional image quality test phantom for planar X-ray imaging. S Afr J Sci. 2023;119(7/8), Art. #14269. https://doi.org/10.17159/sajs.2023/14269

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

DATA AVAILABILITY:
☐ Open data set
☐ All data included
☒ On request from author(s)
☐ Not available
☐ Not applicable

EDITOR:
Michael Inggs

KEYWORDS:
3D printing, phantom, image quality, contrast, resolution

FUNDING:
None

Three-dimensional image quality test phantom for planar X-ray imaging

We aimed to produce a simple, inexpensive 3D printed phantom as a prototype for image quality assessment of contrast, contrast-to-noise ratio (CNR), signal-to-noise ratio (SNR) and resolution in planar X-ray imaging systems. The test phantom was designed using SOLIDWORKS software, printed with a polylactic acid material and filled with paraffin wax. Circular aluminium sheets were used as inserts for contrast evaluation. A planar X-ray system was used for imaging and DICOM images were evaluated using ImageJ software. We evaluated spatial resolution, contrast, CNR and SNR. For resolution, full width at half maximum (FWHM) was measured on different grid sizes. For contrast, intensity of grey values and standard deviation were calculated on the different grid sizes. For CNR and SNR, difference in greyscale of investigated tissue and background per standard deviation of noise in the background was calculated. Resolution of the system was evaluated to be 1.57 and 1.80 lp/mm on grids A and B respectively. Contrast showed differential attenuation per variation in thickness. CNR increased from −13.7 for a thickness of 0.00 mm to 24.90 for a thickness of 28 mm. CNR did not change for a thickness greater than 16.0 mm. The SNR of the system fell in the acceptable range of ≥ 5. The results from the analyses performed indicate that the test phantom has great potential to be a good substitute for the commercially available phantoms on the market, especially for low-resource settings.

Significance:

This study highlights the emergence of 3D printing technology and its suitability within radiology and medical physics for the production of cost-effective phantoms which can serve as substitutes for commercial phantoms in low-resourced medical facilities in low- and middle-income countries.

Introduction

Image quality assessment as a component of quality control in radiology departments is one of the many job descriptions of the clinical physicist. A variety of modality-specific phantoms are used in quality assurance examinations. However, these phantoms are expensive and sometimes delicate, and expert technicians are needed for their use and to evaluate their data.[1,2] Typically, commercially available phantoms are in the price range of USD4000 to USD10 000, depending upon the specifications and the applicable imaging modality.

Since its introduction in the 1980s, three-dimensional (3D) printing technology has progressed from its use in research facilities to being a widely used method for the construction of phantoms for use in diagnostics and radiotherapy.[3-5] Besides the printer technology, attempts have been made to synthesise materials that can be utilised to create 3D models such as resins. The printing process and material determine the quality of the printed product. 3D printing has recently been utilised to create phantoms for multimodality and modality-specific imaging.[6]

According to Huda and Abrahams[7], image quality in radiological imaging is determined by factors such as contrast, spatial resolution, signal-to-noise ratio and noise. Contrast[5] is defined as the difference between the mean greyscale in a region of interest in a study material (S_t) and the mean greyscale in a background area of interest (S_b). This is commonly referred to as the contrast-to-noise ratio (CNR) in digital imaging. In the presence of noise, it is an object-size-independent estimate of the signal strength in the study tissue. This is represented as[8]:

$$CNR = \frac{S_t - S_b}{\sigma_b} \qquad \text{Equation 1}$$

where σ_b is defined as the greyscale standard deviation of the noise in the background.

Spatial resolution is described[9] as the capacity of an imaging modality to distinguish two neighbouring structures as distinct from one another, i.e. image detail visibility. The resolution is mostly estimated using the full-width-at-half-maximum (FWHM) measure in units of line pairs per millimetre (lp/mm).

Signal-to-noise ratio is a measure of true signal (real anatomy) to noise. A lower SNR normally results in images with a gritty appearance. In radiology, SNR is proportional to the amount of contrast in the square root of transmitted photons.[10]

Noise is an undesirable feature in images as it obstructs visualisation and comprehension of an anomaly of interest.[11] The two most prevalent sources of noise in medical images are anatomical noise and radiographic noise. Anatomical noise is the term for undesirable anatomical anomalies in an image. As a result, anatomical noise characterisation is task dependent and independent of the inherent performance of a detector. Radiographic noise describes unwanted image variations that are not produced by the image subject.

A good quality control program can be used to evaluate the clinical performance of imaging systems. The outcomes of routine image quality control are compared to those acquired during equipment acceptance testing or to predefined baseline values at regular intervals. Differences in image quality are indicated by deviations from the acceptance test or baseline values. Periodic quality control helps to discover departures from optimal efficiency and lays the groundwork for continual development by providing frequent feedback. This might be beneficial to patient diagnosis and therapy.[12]

Three-dimensional (3D) printing has gained prominence in recent times for building volumetric objects with the help of a computer-aided design application and the use of a wide range of materials such as ceramics, resins, metals and thermoplastics (e.g. acrylonitrile butadiene styrene (ABS), polyethylene terephthalate glycol-modified (PETG) and polylactic acid (PLA)).[13] 3D printing in radiology has been used predominantly in the construction of phantoms for diagnostic radiology, nuclear medicine and radiotherapy. It has also been used in the printing of breast phantoms using materials as tissue substitutes for their attenuation coefficients.[14,15] Simple symmetrical phantoms for use in computerised tomography (CT) have been manufactured through 3D printing using tissue equivalent materials like resins and thermoplastics.[16] Anthropomorphic phantoms depicting the whole body[17] as well as body parts such as the spine[18] and head[19] have all been manufactured and are playing a key role in radiology departments worldwide.

Studies have been done on the manufacture of 3D-printed image quality assessment phantoms suitable for conventional X-ray imaging, mammography and fluoroscopy. One study[6] used PLA material to print a low-contrast phantom with air holes of different radii ranging from 0.5 mm to 4.5 mm and irradiated with a fluoroscopy machine of 40 kV – 70 kV and the results were feasible. The manufacture of a universal image quality phantom for use in general X-ray, mammography, CT and fluoroscopy has been explored.[2]

Three relevant issues are prevalent in image quality assessment in resource-limited facilities: the high cost of commercial phantoms, lack of human resources, and time constraints, with cost being chief among them. In this study, we therefore aimed to use 3D printing technology to develop an in-house image quality assessment phantom for resolution, contrast, contrast-to-noise ratio and signal-to-noise ratio for general X-ray imaging systems of a low-resourced centre in a low- to middle-income country (LMIC).

Materials and methods

Materials

SOLIDWORKS® software (Version 2019, Dassault Systemes, France) was used for the design of the test phantom. The Creality Slicer software (V.4.8.2 build 177 win 64, Shenzhen Creality 3D Technology Co., Limited, China) was used for G-code conversion of the SOLIDWORKS design to STL files and a Creality CR-20 Pro 3D Printer (Shenzhen Creality 3D Technology Co., Limited, China) was used to print the PLA material (density = 1.250 g/cm^3) into the 3D test phantom.

The PLA material is composed of hydrogen (0.058), carbon (0.541), nitrogen (0.018) and oxygen (0.383). It has a specific density of 1.43 g/cm^3, tensile strength of 28.8 MPa, bending strength of 58.6 MPa and a Hounsfield unit of -530 ± 25. It has linear attenuation coefficients of 0.439, 0.286 and 0.244 at kiloelectron volts of 30, 45 and 60 keV, respectively.[13,14]

The 3D-printed phantom was filled with paraffin wax, and aluminium sheets of 99% purity were inserted into the wax. Image quality assessment, using the phantom, was done on the Philips DuraDiagnost Release 4 X-ray machine (Koninklijke Philips N.V, Netherlands).

ImageJ software (Version 1.51, US National Institutes of Health and the Laboratory for Optical and Computational Instrumentation (LOCI, University of Wisconsin), USA) was used for image analysis. A tape measure, the Ocean software, Piranha quality control meter, the collimator and beam alignment quality control test tool and a beam alignment phantom were used for quality control procedures.

Phantom design and modelling

A circular-shaped phantom was designed, with a whole-body diameter of 150 mm, radius of 75 mm, and thickness of 45 mm and with eight circular holes on the surface (Figure 1). The holes were arranged in a coordinated orientation with equal tolerance in between them for accuracy in alignment. This arrangement was relevant for the measurement of contrast. Opposite the circular holes was a group of four squares of equal dimensions (27 mm × 27 mm) with gridlines of varied spacing in decreasing order of visibility, which is relevant for resolution. The design was saved in the STL format, a universally accepted format for computer-aided designs. ISO/ASTM 52900:2015 was used for the design and fabrication process.

3D printing

The fused deposition modelling method was used for printing the thermoplastic material. Printing of the phantom was done using different parameters. The axial setup was in the order of x-50, y-130 and z-9.99.

The temperature of the printer's nozzle remained constant at 240 °C. Printing time ranged from 1 h to 18 h for the different components of the phantom. A constant bed temperature of 60 °C was maintained throughout the printing process. The fan embedded in the printer, which primarily is used for cooling, was maintained at 50% of its capacity throughout the printing process.

Extruder filling

Paraffin wax was used as the filling material for the printed 'image quality' phantom (Figure 2). Paraffin wax was chosen for its high density and similarity to human tissue by properties. Paraffin wax candles were heated to 80 °C, allowed to cool and poured into the printer extruder (phantom) while in a semi-liquid state to fill in the empty spaces evenly without air gaps. The phantom was left in the open for up to 2 h to solidify and evenly fill every space. The finished test phantom had the specifications shown in Table 1.

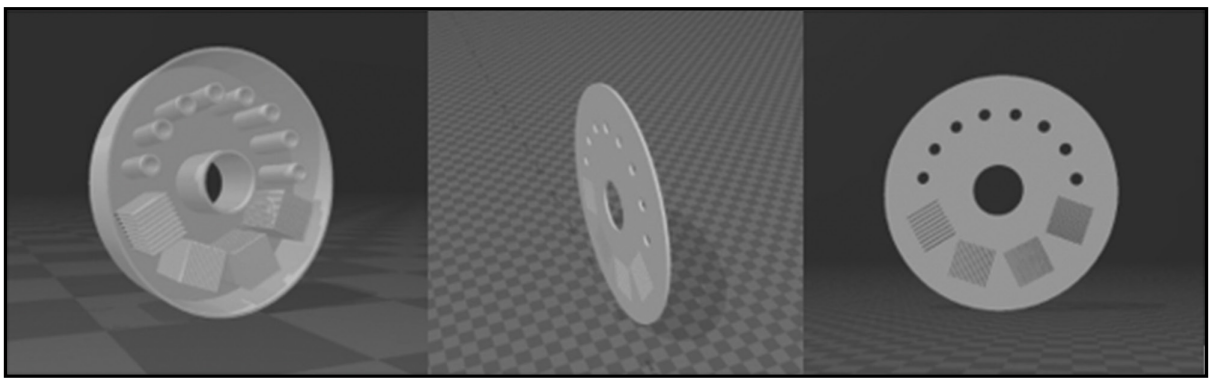


Figure 1: The eight-hole design concept and animation done with SOLIDWORKS.

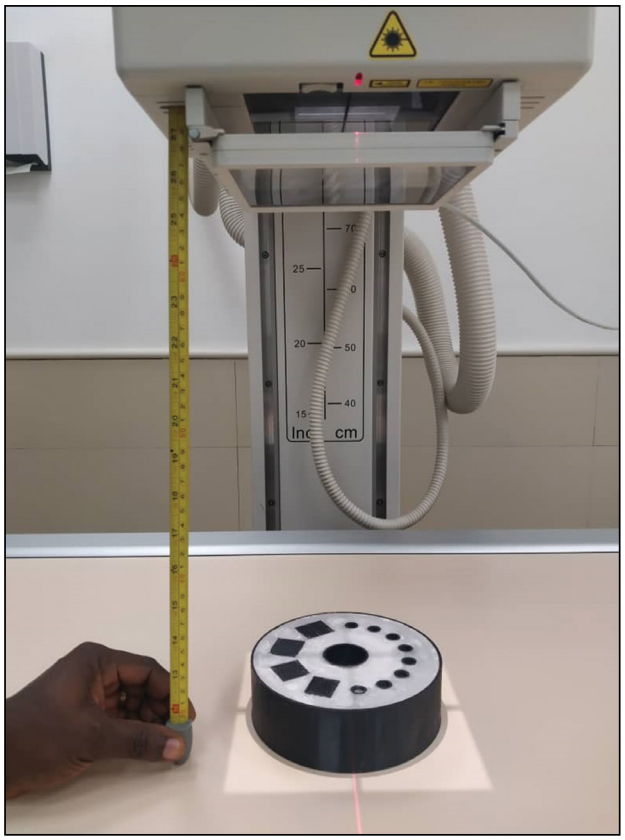


Figure 2: The 3D-printed phantom fully filled with paraffin wax and evenly set up for exposure with a conventional X-ray machine.

Table 1: Specifications and related dimensions of the test phantom

Specification	Dimensions (mm)
Phantom radius (R)	75
Phantom thickness (t)	45
Radius of circular holes (r)	5
Tolerance between circular holes (T_r)	8
Radius of aluminium discs	3.5
Thickness of aluminium discs	0.8
Spacing of resolution lines/bars	0.8/1.0/1.2/1.4
Radius of aluminium rod	4
Thickness of aluminium rod	45–10 (factor of 5)
Contrast squares	27 × 27

Exposure and acquisition of DICOM image

Images for the image quality assessment were acquired using the setup depicted in Figure 2. The circular holes on the phantom were filled with aluminium discs of diameter 7 mm. The discs, of 0.8 mm thickness, were placed on top of each other to form varying thicknesses in seven of the eight holes, with the first hole unfilled.

The phantom was set up on the couch of the X-ray machine just beneath the source. The distance between the source and the phantom was 700 mm. The X-ray source was set to 52 kVp, 63.0 ms and 32 mAs with 1.35 μGym^2.

DICOM (Digital Imaging and Communications in Medicine) is the primary file format for storing and transferring medical images.The DICOM image acquired (Figure 3) was uploaded to the ImageJ software and analysed using the various image quality assessment tools.[4] The acquired X-ray showed fine and clear details with minimal or no noise or artefacts.

Resolution

Analysis of resolution was done by drawing regions of interest across the gridlines. A polynomial fit of distance versus greyscale values was performed for each gridline. For each grid, the FWHM representing the resolution of the X-ray system was determined. The average resolution of each grid was calculated using Equation 2:

$$\text{Grid X, Line Y}$$
$$\text{Resolution peak} = \frac{\text{Highest curve peak}}{2} \qquad \text{Equation 2}$$
$$\text{Resolution} = \text{Highest value on xaxis} - \text{Lowest value on x} - \text{axis}\ldots.$$

Contrast

Contrast was analysed by drawing regions of interest in the phantom image. Each of the holes in the image represented a different thickness and contrast due to the difference in the number of aluminium discs. Each thickness corresponded to the position of the circular hole in increasing order of contrast. A polynomial graph of thickness against mean greyscale value was plotted to show the curvature of contrast.

Contrast-to-noise ratio

Noise was determined by using the contrast-to-noise ratio (CNR). The standard deviation of the mean greyscale values was calculated. For each hole and thickness, the average greyscale in the region of interest (ROI) in the hole was found as well as the average greyscale in the ROI in the surrounding background. These two parameters were used to find the CNR together with the standard deviation of the noise in the background.

Signal-to-noise ratio

The SNR was estimated using Equation 3. For each thickness, the mean grey value in the ROI was found. The standard deviation of the mean

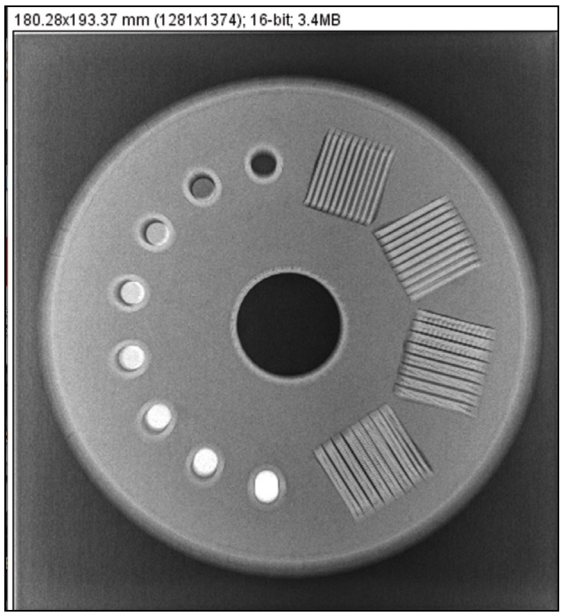


Figure 3: DICOM image of the phantom for image quality assessment.

grey values within the ROI was calculated. These two parameters were used to find the SNR.

$$\text{SNR} = \frac{\text{Mean grey value within ROI}}{\text{Standard deviation of grey value within ROI}} \quad \text{Equation 3}$$

Estimation of covariance (CoV) of the measured grey values gives a good insight into what is happening with the increase in thickness and was calculated as:

$$\text{CoV} = \frac{\text{standard deviation of grey values}}{\text{mean grey values}} \times 100\%.... \quad \text{Equation 4}$$

Results and discussion

Resolution

The spatial resolution of the planar X-ray system was determined from the gridlines of the phantom. The phantom consisted of four grids: A, B, C and D. The width of the lines/bars and their spacing for each grid were 0.8 mm, 1.0 mm, 1.2 mm and 1.4 mm, respectively. The ImageJ program evaluated the resolution by means of a data set collected in the parameters of distance and greyscale values. A polynomial graph of distance versus greyscale value was plotted for each line of every grid. Image resolution was evaluated quantitatively.

Table 2 shows the accumulation of the distance and initial and final grey values of Grids A and B. The distance and grey values for each line of the grids were generated from the ImageJ software. The final grey value is calculated by subtracting the initial grey value from the minimum value. This was done for all lines of Grids A and B.

Spatial resolution of the planar X-ray scanner was determined by calculating the full width at half maximum (FWHM) from each of the graphs in Figure 4. The FWHM was calculated using Equation 5.

$$\text{FWHM} = |z_1 - z_2| \quad \text{Equation 5}$$

where z_1 is the minimum distance (mm) value corresponding to the minimum grey value at $1/2h_{max}$ and z_2 is the maximum distance (mm) value corresponding to the maximum grey value at $1/2h_{max}$. Subtracting the minimum value of distance from the maximum value gives the resolution value for the grid and line under consideration.

Table 3 shows the spatial resolution measured from the three lines for Grids A and B. It can be observed that the grey value increases constantly with a corresponding increase in distance until it peaks and decreases steadily from the point of the highest peak as it approaches zero.

From Table 3, it can be seen that the average spatial resolution decreases with decreasing spatial frequency. The spatial frequency of Grid A (Table 2) produced an average FWHM of 1.57 mm, while the spatial frequency of Grid B (i.e. 1.0 mm width and 1.0 mm spacing) produced an average

Table 2: Cumulative table of distance and initial and final pixel grey values for Grids A and B

Distance (mm)	Grey value (initial)						Grey value (final)					
GAL1 – GBL3	GAL1	GAL2	GAL3	GBL1	GBL2	GBL3	GAL1	GAL2	GAL3	GBL1	GBL2	GBL3
0.0	1304.3	1033.2	1429.8	901.0	1028.0	2061.0	360.1	88.2	106.1	18.0	0.0	86.0
0.1	1107.5	945.0	1486.1	1117.3	1616.9	2138.6	163.3	0.0	162.4	234.3	588.9	163.6
0.3	1115.4	988.5	1436.8	2016.8	2431.4	2009.2	171.2	43.5	1131.0	1133.8	1403.4	34.2
0.4	1268.9	1146.1	1551.0	2496.5	2682.5	2147.8	324.7	201.1	227.3	1613.5	1654.5	172.8
0.6	1529.2	1498.3	1792.7	2491.9	2660.0	2470.3	585.0	553.3	469.0	1608.9	1632.0	495.3
0.7	2056.8	1959.4	2129.8	2418.5	2690.2	2677.1	1112.6	1014.4	806.1	1535.5	1662.2	702.1
0.8	2425.4	2174.7	2592.1	2543.3	2848.3	2795.0	1481.2	1229.7	1268.4	1660.3	1820.3	820.0
1.0	2604.2	2538.4	2665.9	2742.1	2881.9	2775.2	1660.0	1593.4	1342.2	1859.1	1853.9	800.2
1.1	2652.4	2534.0	2648.1	2813.6	2761.9	2832.4	1708.2	1589.0	1324.4	1930.6	1733.9	857.4
1.3	2701.8	2623.3	2660.5	2850.0	2670.5	2858.0	1757.6	1678.3	1336.8	1967.0	1642.5	883.0
1.4	2509.6	2612.4	2523.2	2802.0	2770.2	2770.7	1565.4	1667.4	1199.5	1919.0	1742.2	795.7
1.5	2552.0	2553.5	2112.5	2755.4	2694.1	2910.4	1607.8	1608.5	788.8	1872.4	1666.1	935.4
1.7	2362.0	2078.9	1731.5	2375.3	2528.0	2804.0	1417.8	1133.9	407.8	1492.3	1500.0	829.0
1.8	1940.0	1577.3	1388.4	2247.7	2386.3	2455.8	995.8	632.3	64.7	1364.7	1358.3	480.8
2.0	1376.4	1248.5	1323.7	2288.6	2265.1	2182.7	432.2	303.5	0.0	1405.6	1237.1	207.7
2.1	1024.6	1102.9	1387.6	2155.0	1977.5	1978.8	80.4	157.9	63.9	1272.0	949.5	3.8
2.3	944.2	1016.4	1449.6	1585.6	1396.9	1975.3	0.0	71.4	125.9	702.6	368.9	0.3
2.4	1044.5	1068.9	1328.5	1033.9	1035.3	2175.8	100.3	123.9	4.8	150.9	7.3	200.8
2.5	1249.4	1149.4	1570.1	883.0	1030.0	2427.0	305.2	204.4	246.4	0.0	2.0	452.0
Min value	**944.2**	**945.0**	**1323.7**	**883.0**	**1028.0**	**1975.3**						

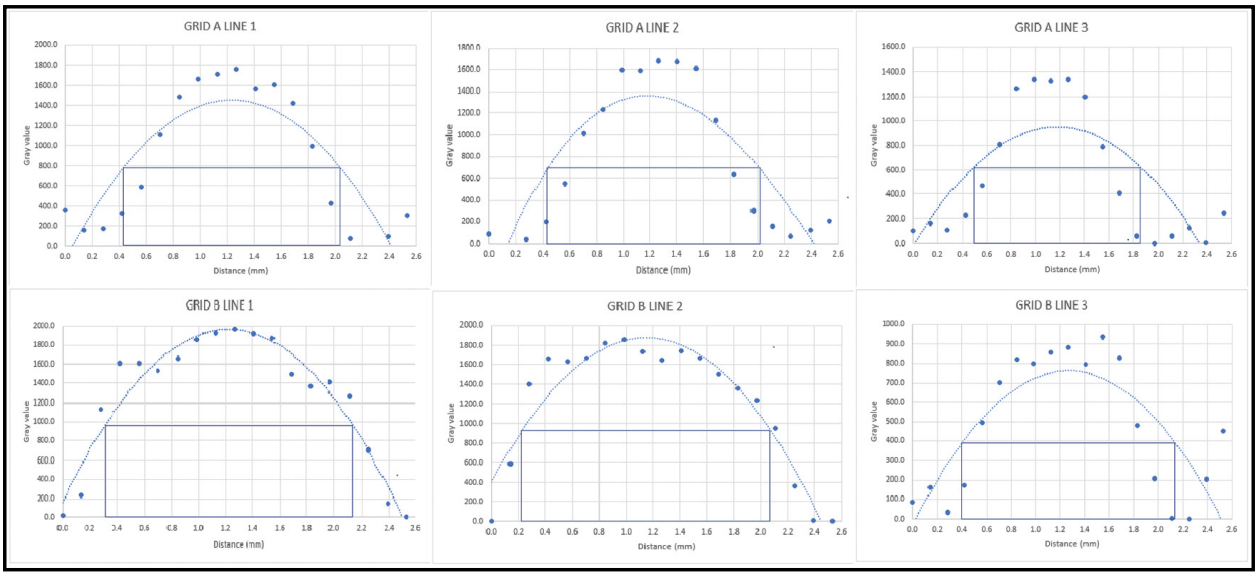


Figure 4: Resolution graphs for lines 1, 2 and 3 for Grids A and B.

Table 3: Resolution for lines of Grids A and B

GRID	Line 1			Line 2			Line 3			Average of differences
	Max	Min	Diff	Max	Min	Diff	Max	Min	Diff	
A	2.06	0.40	1.66	2.02	0.42	1.60	1.90	0.45	1.45	1.57
B	2.12	0.32	1.80	2.08	0.22	1.86	2.14	0.40	1.74	1.80

FWHM of 1.80 mm. This means that Grid A could be used to resolve lines/bars that are 1.57 mm wide with 1.57 mm spacing, while Grid B can be used to resolve lines/bars that are 1.80 mm wide with 1.80 mm spacing. The smaller the FWHM, the better the spatial resolution. Due to the difficulty in the 3D printing of Grid C (i.e. 1.2 mm width and 1.2 mm spacing) and Grid D (i.e. 1.4 mm width and 1.4 mm spacing), images of Grid C and D were not included for calculating the spatial resolution. As the resolution of Grid A is lower than those of Grid B, Grid C and Grid D, it was more efficient in resolving structures with sizes less than 1.80 mm.

Resolution of the phantom could be attributed to the type of material used for measuring the spatial resolution or exposure parameters, such as the source to image (phantom) distance, output voltage, tube current, and exposure time used in the acquisition of images. When the source to image distance increases, the X-ray beam diverges, forming a cone shape and thereby affecting the intensity of the X-ray beam and quantity of X-rays. Also, attenuation due to low kilovoltage peak may lead to the desired image not generating.

In this study, the FWHM was calculated based on the slit (i.e. grids) method on the digital detector. The resultant penumbral image provided a line spread function or Gaussian curve, from which the FWHM was estimated, due to a partial block of the radiation from the source by the grids.[20,21]

Contrast

Contrast was measured from eight holes of the same diameter and radius (Figure 3), filled with different thicknesses of aluminium inserts. Using the elliptical measurement tool in the Radiant DICOM viewer, an area of 0.1289 cm^2 was drawn in the centre of each circular image (i.e. as region of interest) to obtain the intensity of grey values and their standard deviation. The thickness of the target increased with greyscale values, but greyscale values remained fairly the same beyond a target thickness of 16 mm. There is an attainment of saturation in grey value for an aluminium thickness beyond 16 mm. Also, the CoV became better with an increase in thickness saturating from ≥ 20.0 mm. A CoV of ≤5% is normally acceptable and ≤10% is within a good range. Data from this analysis are presented in Table 4 and Figure 5.

The X-ray intensity attenuation across the material is approximately the same for thicknesses greater or equal to 16 mm. This implies that the maximum contrast that can be measured by the phantom using exposure parameters of 52 kVp, 32 mA, 63 ms and a source to phantom distance of 700 mm is 4095 at a phantom thickness greater or equal to 16 mm. This indicates the attainment of a saturation point for the mean grey values.

Contrast-to-noise ratio

The CNRs for each of the eight circular targets (holes) were also calculated using Equation 1.[8]

Table 5 shows the average greyscale in the ROI in the investigated tissues, backgrounds and standard deviations.

Negative CNR values indicate less signal than noise, and positive CNR values indicate more activation signal than noise.[22]

From Figure 5b, there is a steady increase in the CNR per aluminium thickness. This is because an increase in the depth of aluminium discs increases the relative X-ray transmittance in the phantom. The CNR increased from a negative value of −13.7 for a thickness of 0.00 mm to 24.90 for a thickness of 28.00 mm. In the exposure of the phantom for assessment, the first hole was left empty (without any aluminium insert) and this accounted for the 0.00 cm thickness. This is because without any aluminium insert present, there is no attenuation of X-rays in the medium, indicating less signal than noise. The CNR steadily increased per increase in thickness because attenuation increased as aluminium thickness increased. However, the CNR did not change considerably with an increase in thickness from 16 mm to 28 mm. The standard deviation within this range also showed a steady change, indicative of the attainment of a saturation point.

Table 4: Position of holes and their corresponding thicknesses, pixel grey values and signal-to-noise (SNR) ratios

Position	Thickness (mm)	Min grey value	Max grey value	Mean grey value	Standard deviation	SNR	Covariance (%)
1	0.0	618	976	815.44	61.93	13.17	7.6
2	4.0	1393	2019	1744.7	102.65	16.99	5.9
3	8.0	2638	3229	2892.29	99.01	100.26	3.4
4	12.0	3104	3835	3485.29	114.48	30.44	3.3
5	16.0	3431	4095	3793.71	113.39	33.46	3.0
6	20.0	3713	4095	4018.94	85.80	46.84	2.1
7	24.0	3719	4095	4009.25	86.95	46.11	2.1
8	28.0	4039	4095	4064.00	84.97	47.83	2.1

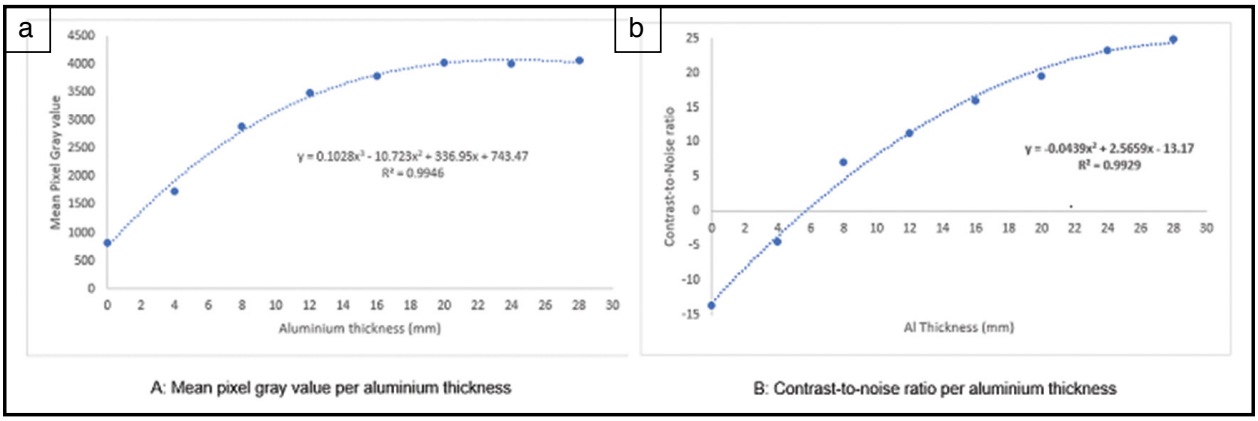


Figure 5: (a) Mean pixel grey value per aluminium thickness and (b) contrast-to-noise ratio per aluminium thickness.

Table 5: Thickness, mean grey pixel values of the region of interest (ROI) in investigated tissue and background, with associated standard deviations and contrast-to-noise ratios (CNR)

Position	Thickness (mm)	Mean grey value in ROI in investigated tissue	Mean grey value in ROI in background	Standard deviation of background	CNR
1	0.0	815.44	2084.17	92.88	–13.66
2	4.0	1744.7	2153.48	94.22	–4.34
3	8.0	2893.29	2199.22	98.48	7.05
4	12.0	3485.39	2276.26	107.63	11.23
5	16.0	3793.71	2240.47	97.51	15.93
6	20.0	4018.94	2145.87	96.05	19.50
7	24.0	4009.25	1994.34	86.96	23.17
8	28.0	4064.00	1849.41	88.91	24.91

Signal-to-noise ratio

The signal-to-noise ratio (SNR) for each of the eight circular targets (holes) was calculated using Equation 3[22] as shown in Table 4. The SNR as a measure compares a desired signal to the level of background noise. The higher the CNR is between structures, the lower the SNR needed. From Figure 5b, there is a steady increase in the CNR per aluminium thickness, hence a higher SNR is required for differentiation. The SNR changed steadily from a thickness of 20 mm to 28 mm. The standard deviation within this range also showed a steady change, indicative of the attainment of a saturation point. According to the Rose model, the image quality of a system is acceptable if the SNR is greater or equal to 5.[23]

Conclusion

This study has shown that 3D printing techniques can be used for the manufacture of test phantoms for image quality assessment in planar X-ray imaging. We successfully designed and printed a test phantom for in-house use in a low-resourced medical imaging facility in a LMIC that assessed image quality successfully. The phantom demonstrated the capability of being used for analysing image quality parameters, including resolution, contrast and CNR on general X-ray imaging systems. Subsequent plans include acceptance testing and commissioning tests for clinical use. This in-house quality control equipment, at a unit price of USD150, could be a good substitute for relatively expensive commercially available phantoms.

Acknowledgements

This research was carried out at the Medical Physics Department of the School of Nuclear and Allied Sciences in the University of Ghana as part of the master's degree of J.B.N. The authors are grateful to the department for its support.

Competing interests

We have no competing interests to declare.

Authors' contributions

J.B.N.: Conceptualisation, methodology, data collection, sample analysis, data analysis, validation, data curation, writing – the initial draft, writing – revisions, project management. F.H.: Conceptualisation, methodology, sample analysis; data analysis, validation, data curation, writing – revisions, student supervision, project leadership, project management. E.S.: Conceptualisation, methodology, sample analysis, data analysis, validation, data curation, writing – revisions, student supervision, project leadership, project management.

References

1. Madamesila J, McGeachy P, Villarreal Barajas JE, Khan R. Characterizing 3D printing in the fabrication of variable density phantoms for quality assurance of radiotherapy. Phys Medica. 2016;32(1):242–247. https://doi.org/10.1016/j.ejmp.2015.09.013
2. Groenewald A, Groenewald WA. Development of a universal medical X-ray imaging phantom prototype. J Appl Clin Med Phys. 2016;17(6):356–365. https://doi.org/10.1120/jacmp.v17i6.6356
3. Ehler ED, Barney BM, Higgins PD, Dusenbery KE. Patient specific 3D printed phantom for IMRT quality assurance. Phys Med Biol. 2014;59(19):5763–5773. https://doi.org/10.1088/0031-9155/59/19/5763
4. Leng S, McGee K, Morris J, Alexander A, Kuhlmann J, Vrieze T, et al. Anatomic modeling using 3D printing: Quality assurance and optimization. 3D Print Med. 2017;3(1), Art. #6. https://doi.org/10.1186/s41205-017-0014-3
5. Chan HK, Griffin J, Lim JJ, Zeng F, Chiu ASF. The impact of 3D printing technology on the supply chain: Manufacturing and legal perspectives. Int J Prod Econ. 2018;205:156–162. https://doi.org/10.1016/j.ijpe.2018.09.009
6. Kapetanakis I, Fountos G, Michail C, Valais I, Kalyvas N. 3D printing X-ray quality control phantoms: A low contrast paradigm. J Phys Conf Ser. 2017;931(1), Art. #012026. https://doi.org/10.1088/1742-6596/931/1/012026
7. Huda W, Abrahams RB. X-ray-based medical imaging and resolution. Am J Roentgenol. 2015;204(4):W393–W397. https://doi.org/10.2214/AJR.14.13126
8. De Crop A. Image quality evaluation in X-ray medical imaging based on Thiel embalmed human cadavers [PhD thesis]. Ghent: Ghent University; 2015. http://hdl.handle.net/1854/LU-8514957
9. Vajuhudeen Z, Jones J. Spatial resolution. Radiopaedia.org. 2009 Jun 05. https://doi.org/10.53347/rID-6318
10. Cohen-Adad J, Wald LL. Chapter 2.1 – Array coils. In: Cohen-Adad J, Wheeler-Kingshott CAM, editors. Quantitative MRI of the spinal cord. Amsterdam: Academic Press; 2014. p. 59–67. https://doi.org/10.1016/B978-0-12-396973-6.00005-8
11. Samei E. Performance of digital radiographic detectors: Quantification and assessment methods. Advances in digital radiography: RSNA Categorical Course in Diagnostic Radiology Physics. Durham, NC: Duke University Medical Center; 2003. p. 37–47.
12. Reis C, Pascoal A, Sakellaris T, Koutalonis M. Quality assurance and quality control in mammography: A review of available guidance worldwide. Insights Imaging. 2013;4(5):539–553. https://doi.org/10.1007/s13244-013-0269-1
13. ISO/PRF 17296-1 – Additive manufacturing – General principles [Internet]. [cited 2021 Dec 03]. Available from: https://standards.iteh.ai/catalog/standards/iso/a31c7a1c-bb3c-4b56-acb9-33e9602e9a2b/iso-prf-17296-1
14. He Y, Liu Y, Dyer BA, Boone JM, Liu S, Chen T, et al. 3D-printed breast phantom for multi-purpose and multi-modality imaging. Quant Imaging Med Surg. 2019;9(1):63–74. https://doi.org/10.21037/qims.2019.01.05
15. Schopphoven S, Cavael P, Bock K, Fiebich M, Mäder U. Breast phantoms for 2D digital mammography with realistic anatomical structures and attenuation characteristics based on clinical images using 3D printing. Phys Med Biol. 2019;64(21):215005. https://doi.org/10.1088/1361-6560/ab3f6a
16. Solomon J, Ba A, Bochud F, Samei E. Comparison of low-contrast detectability between two CT reconstruction algorithms using voxel-based 3D printed textured phantoms. Med Phys. 2016;43(12):6497–6506. https://doi.org/10.1118/1.4967478
17. Lee MY, Han B, Jenkins C, Xing L, Suh TS. A depth-sensing technique on 3D-printed compensator for total body irradiation patient measurement and treatment planning. Med Phys. 2016;43(11):6137. https://doi.org/10.1118/1.4964452
18. Javan R, Bansal M, Tangestanipoor A. A prototype hybrid gypsum-based 3-dimensional printed training model for computed tomography-guided spinal pain management. J Comput Assist Tomogr. 2016;40(4):626–631. https://doi.org/10.1097/RCT.0000000000000415
19. Kamomae T, Shimizu H, Nakaya T, Okudaira K, Aoyama T, Oguchi H, et al. Three-dimensional printer-generated patient-specific phantom for artificial in vivo dosimetry in radiotherapy quality assurance. Phys Med. 2017;44:205–211. https://doi.org/10.1016/j.ejmp.2017.10.005
20. Malliori A, Daskalaki A, Dermitzakis A, Pallikarakis N. Development of physical breast Phantoms for X-ray imaging employing 3D printing techniques. Open Med Imaging J. 2020;12(1):1–10. https://doi.org/10.2174/1874347102012010001
21. Wang S, Pavlicek W, Roberts CC, Langer SG, Zhang M, Hu M, et al. An automated DICOM database capable of arbitrary data mining (including radiation dose indicators) for quality monitoring. J Digit Imaging. 2011;24(2):223–233. https://doi.org/10.1007/s10278-010-9329-y
22. Welvaert M, Rosseel Y. On the definition of signal-to-noise ratio and contrast-to-noise ratio for fMRI data. PLoS One. 2013;8(11):e77089. https://doi.org/10.1371/journal.pone.0077089
23. Bath M. Evaluating imaging systems: Practical applications. Radiat Prot Dosim. 2010;139:26–36. https://doi.org/10.1093/rpd/ncq007

AUTHORS:
Raphael B.H. Gameli[1]
Elliot H. Alhassan[2]
Abudu B. Duwiejuah[3]
Emmanuel D. Abarike[2]
Abdul-Aziz Bawa[4]

AFFILIATIONS:
[1]Department of Environment and Sustainability Sciences, Faculty of Natural Resources and Environment, University for Development Studies, Nyankpala Campus, Tamale, Ghana
[2]Department of Fisheries and Aquatic Resources Management, Faculty of Biosciences, University for Development Studies, Nyankpala Campus, Nyankpala, Ghana
[3]Department of Biotechnology and Molecular Biology, Faculty of Biosciences, University for Development Studies, Nyankpala Campus, Tamale, Ghana
[4]Spanish Laboratory Complex, University for Development Studies, Nyankpala Campus,Tamale, Ghana

CORRESPONDENCE TO:
Abudu Duwiejuah

EMAIL:
abalu096@gmail.com

DATES:
Received: 01 Mar. 2022
Revised: 11 May 2023
Accepted: 12 May 2023
Published: 08 Aug. 2023

HOW TO CITE:
Gameli RB, Alhassan EH, Duwiejuah AB, Abarike ED, Bawa A-A. Binary and ternary metals adsorption from greywater using spent green tea as a novel adsorbent. S Afr J Sci. 2023;119(7/8), Art. #13352. https://doi.org/10.17159/sajs.2023/13352

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

DATA AVAILABILITY:
☐ Open data set
☒ All data included
☐ On request from author(s)
☐ Not available
☐ Not applicable

EDITOR:
Michael Inggs

KEYWORDS:
adsorption, binary system, greywater, mercury, green tea

FUNDING:
None

Binary and ternary metals adsorption from greywater using spent green tea as a novel adsorbent

Adsorption is one of the most easy-to-operate, less costly, efficient and, most importantly, environmentally friendly methods of removing toxic metals from aqueous environments. We used spent Impra Green Tea Ginseng Flavoured to recover mercury (Hg^{2+}), lead (Pb^{2+}) and cadmium (Cd^{2+}) in binary and ternary systems from greywater. We undertook this study in binary and ternary systems at adsorbent dosages with a corresponding 100 mL varied initial metal concentrations of the greywater. The adsorption efficiency at varied concentrations and dosages in the binary systems by the spent tea waste ranged from 38.5% to 100% for lead, 11.50% to 100% for cadmium and was 100% for mercury. In the ternary system, the adsorption efficiency of toxic metals ranged from 28.91% to 72.85% for cadmium and was 100% for mercury and lead. The maximum adsorption capacity (Q_e) for toxic metals in the binary system ranged from 38.46 to 81.97 mg/g for Pb^{2+} and 12.64 to 56.82 mg/g for Cd^{2+}. The Langmuir adsorption isotherm model was the best fit for the adsorption of toxic metals by Impra Green Tea Ginseng Flavoured. The pH under which the experiments were conducted showed very high removal efficiency for lead and mercury but lower removal efficiencies for cadmium. Spent Impra Green Tea Ginseng Flavoured can be used as an effective and low-cost adsorbent of toxic metals from greywater or wastewater. Based on our findings, further studies should be conducted to determine the effects of varying the contact time, temperature and elevated metal concentrations in the greywater or other wastewater.

Significance:

- This study provides useful information on how spent Impra Green Tea Ginseng Flavoured can be used as an effective and low-cost adsorbent of toxic metals from greywater or wastewater.

Introduction

Water pollution is currently an issue of great concern as it is very significant to all living organisms.[1] Wastewater from industries, landfills and smaller firms serves as a significant source of environmental pollution. This pollution results from the high amount of toxic metal ions present in the wastewater discharged.[2] These metals are non-biodegradable and toxic, and have a high tendency to be incorporated into the food chain or food web and the ability to accumulate in the body of living organisms.[3] The most common metals found in the effluents of most industrial wastewater include cadmium, mercury, lead, copper, zinc and chromium.[2]

Most applicable studies have focused on developing and designing efficient and low-cost methods and strategies to remove these heavy metals from water bodies and the environment as a whole. Some of these conventional strategies for recovering toxic metals from water bodies and the environment include reverse osmosis, evaporation, ion-exchange, solvent extraction, coagulation, and sorption.[4] Although these approaches are commended to some extent, they are not without certain limitations which include cost, technical limitations, ineffectiveness when metal concentrations are above 1 mg/L, high energy consumption, production of a large amount of residual sludge and, finally, intervention technologies show high sensitivity to the operational conditions.[2]

The adsorption process is considered to be one of the most studied metal recovery approaches using low-cost adsorbents. Tea waste can be used as an adsorbent for toxic metal removal from wastewater[5] because of its efficiency and ideal suitability as a vital material in removing toxic metal species such as cadmium, lead and mercury.[6] Tea waste contains lignin, hemicellulose, cellulose and hydroxyl groups in its cell walls and ion exchangeability which gives it the ability to remove toxic metals by means of adsorption processes.[6,7]

Treated greywater has various beneficial purposes such as car washing, lawn irrigation, landscaping, garden watering and can contribute significantly to an increase in agricultural productivity.[8] Waste green tea leaves can remove toxic contaminants as they are natural adsorbents.[9] This also serves as a possible solution to the disposal of wastewater and its management problems. Similar studies have reported brewed tea waste removal of 99.01%, 84.23% and 83.45% for Pb, Cd and Zn, respectively[10], spent Chinese green tea removal that ranged from 98.18% to 99.89% for Cd, 98.79–99.99% for Cr, 98.18–99.98% for Hg and 86.20–99.99% for Pb[11], spent green tea removal that ranged from 99.99% to 100% for Hg, 99.99% to 100% for Pb and 11.11% to 18.28% for Cd in mono systems[12] and tea waste biochar removal of 68.2% for Cr.[13] Additionally, employing waste tea residue, removal efficiencies of 100% for Cu and 99.99% for Ni ions were attained.[14]

Tea waste has recently gained popularity as an effective adsorbent for removing metal ions from waste streams due to its capacity to defeat these pollutants. Tea leaves' insoluble cell walls are primarily composed of cellulose, tannins, lignin and structural proteins. Because these components contain functional groups, particularly carboxylate, phenolic hydroxyl and oxyl groups[15], they have good potential as metal scavengers from solutions and wastewaters. Toxic metals co-exist and constitute contaminants that pose a serious threat to the aquatic ecosystem due to their toxicity. Toxic metal removal from wastewater, particularly greywater, would significantly help improve public and environmental health. The use of tea waste as an adsorbent for the removal of toxic metals in wastewater via the process of adsorption is very ecofriendly, economic and highly efficient. Potable drinking water can be further conserved when treated greywater is used in place of potable drinking water for backyard gardens, flushing toilets and other basic activities. Most developing countries are incapable of adopting the use of activated carbon for wastewater treatment due to the technical expertise and cost. Finding a low-cost, efficient technology to remove heavy metals from water becomes essential. Adsorption is a technique that is quite effective for this purpose, but cost is a key factor, and the sorts of adsorbents that are typically utilised are pricey. Our study therefore explored the use of a low-cost tea waste as an adsorbent to simultaneously remove toxic metals from greywater.

Materials and methods

Greywater sampling and technique

Greywater from bathtubs, hairdressers and laundry was collected in the Nyankpala Campus of the University for Development Studies, Ghana. The greywater was collected into polypropylene bottles and kept in an ice chest and conveyed to the Spanish Laboratory Complex in the Nyankpala Campus. The sampled greywater was then mixed to have a uniform composition. This was done because the quantity of toxic metals in greywater can be variable depending on the source and it is dependent on the source.[16] All other equipment (hand gloves, ice chest and voltic plastic bottles) for greywater sampling were properly cleaned. The sampling technique employed was purposive sampling. Samples were taken over two days in January 2020 at 6:30 and 19:30 local time.

Preparation of Impra Green Tea Ginseng Flavoured adsorbent

The spent tea bags of Impra Green Tea Ginseng Flavoured acquired were emptied and soaked in hot distilled water for 2 h and then washed continuously. This washing was to ensure that the remaining solution became colourless. The Impra Green Tea Ginseng Flavoured was used due to its properties such as large surface area and functional groups. Spent tea is likely to have some coloured components, tannins, proteins and polysaccharides that are hydrolysable.[10] Washing them was meant to remove any substance that might cause contamination. After continuous rinsing with hot distilled water to remove the colour, the tea was air dried for 24 h. In the adsorbent preparation, deionised water was used in the washing of the tea waste adsorbent until the solution became colourless. This current experiment was based on methods and techniques used in previous studies for the preparation of raw adsorbent materials such as tea waste.[17,18] The particle size of the tea waste used was between 150 μm and 250 μm.

Preparation of stock solution used for spiking of the greywater

To achieve 1 mg of each of the toxic metal compounds, the molecular weights of $HgCl_2$ (271.50), $Pb(NO_3)_2$ (331.21) and $Cd(NO_3)_2$ (236.42) were ascertained and divided by the atomic weights of Hg (200.60), Pb (207.20) and Cd (122.41), respectively. The stock solutions were set up by dissolving precisely weighed 1.35 g of mercury chloride ($HgCl_2$), 1.60 g of lead nitrate ($Pb(NO_3)_2$) and 1.93 g of cadmium nitrate $Cd(NO_3)_2$ in distilled water to prepare solutions of 1000 mg/L concentration each of the toxic metal in a 1000 mL volumetric flask.

Analysis of greywater and adsorption experiment

The greywater sampled was filtered using a glass filtering funnel and Whatman's qualitative filter paper (Ashless, circle, Cat No. 1442 125, 125 mm Ø) and was stored in a 35 mL sampling bottle as control and analysed to determine the initial concentrations of mercury, lead and cadmium. The results obtained after the analysis of the control sample for mercury, lead and cadmium were 0.00 mg/L, 0.00 mg/L and 0.86 mg/L, respectively. The experiment was then carried out by spiking the raw greywater at maximum contamination limits of 0.10 mg/L, 0.10 mg/L and 0.04 mg/L for mercury, lead and cadmium, respectively. The respective binary or ternary stock solutions were each pipetted into 1000 mL of the greywater to obtain the resulting concentrations in binary and ternary systems (Table 1).

For the binary system, the concentrations in the spiked greywater were 0.10 mg/L for mercury and 0.10 mg/L for lead; 100 mL of greywater was used and a dosage of 1 g of the tea waste adsorbent was taken into a conical flask (Table 1). The concentrations in the spiked greywater were 0.10 mg/L for mercury, 0.10 mg/L for lead, and 0.90 mg/L for cadmium in the ternary system and 100 mL of greywater and a dosage of 1 g of the tea waste adsorbent was taken into a conical flask (Table 1). Again, similar experiments were carried out for mercury, lead and cadmium in the binary and ternary systems at increasing concentrations of 1.00 mg/L, 1.00 mg/L and 1.86 mg/L, respectively, at 100 mL of greywater and at 3 g of tea waste adsorbent dosage (Table 1). At 5 g of the tea waste adsorbent dosage and at a volume of 100 mL of greywater, the concentrations of mercury, lead and cadmium in the binary and ternary systems were 5.00 mg/L, 5.00 mg/L and 5.86 mg/L, respectively (Table 1), with the pH ranging from 6.05 to 7.74 at a temperature of 25 °C. All experiments were carried out at the Spanish Laboratory Complex of the University for Development Studies, Nyankpala Campus. The elutes obtained were transported to the Ecological Laboratory of the University of Ghana for analysis using a Perkin Elmer PIN Accle 900T GRAPHITE Atomic Absorption Spectrophotometer (AAS) (Waltham, USA). The detection limits of the metal ions were 0.0001 mg/L for Hg, 0.003 mg/L for Pb and 0.0008 mg/L for Cd. The percentage recovery of standard for each metal was calculated as recovery (%) $= \frac{mean\ value}{Added\ amount} \times 100$ and the recovery (%) obtained was 99.70%.

Table 1: Contamination limits of toxic metals in the greywater at varied dosages in binary and ternary systems

Batch	Metal	Concentration (mg/L) at 1 g of dosage	Concentration (mg/L) at 3 g of dosage	Concentration (mg/L) at 5 g of dosage
Binary	Hg vrs Pb	0.10 vrs 0.10	1.00 vrs1.00	5.00 vrs 5.00
	Hg vrs Cd	0.10 vrs 0.90	1.00 vrs 1.00	5.00 vrs 5.86
	Pb vrs Cd	0.10 vrs 0.90	1.00 vrs 1.86	5.00 vrs 5.86
Ternary	Hg vrs Pb vrs Cd	0.10 vrs 0.10 vrs 0.90	1.00 vrs 1.00 vrs 1.86	5.00 vrs 5.00 vrs 5.86

Calculation for adsorption efficiency of mercury, lead and cadmium

The equilibrium concentration of the adsorbent and the uptake of the toxic metal which is denoted by the symbol Q_e for each toxic metal at each adsorbent dosage was calculated using Equation 1. The removal efficiency (Q_e) was calculated as:

$$Q_e = \frac{c_i - C_f}{M} \times V \quad \text{Equation 1}$$

In percentage, adsorption capacity was calculated as:

$$Q_e = \frac{(C_i - C_f)}{C_i} \times 100 \quad \text{Equation 2}$$

where $\mathbf{Q_e}$ is the adsorption capacity, $\mathbf{C_i}$ is the initial concentration of the toxic metal, $\mathbf{C_f}$ is the final concentration of the toxic metal after adsorption, **M** is the amount or dosage of adsorbent, and **V** is the volume of the solution.

Adsorption isotherms

Langmuir and Freundlich isotherm models were mathematically expressed based on their assumptions. The Langmuir isotherm model was first described in the removal of the gas molecule onto an analogous solid surface.[19] This isotherm is often used to determine the maximum removal or adsorption capacity including the type of interaction between the metals and the adsorbent.[20] This can be known from a linear mathematical equation provided the reaction is linear. Below is the Langmuir isotherm formula in the linear form:

$$\frac{C_e}{Q_e} = \frac{1}{K_L \cdot Q_m} + \frac{C_e}{Q_m} \quad \text{Equation 3}$$

$$\frac{C_e}{Q_m} = \frac{1}{K_L Q_{max}} + \frac{C_e}{Q_{max}} \quad \text{Equation 4}$$

where C_e is the concentration of the adsorbate at equilibrium (mg/g), K_L is the Langmuir constant (L/mg) and Q_{max} (mg/g) is the number of adsorbed molecules on the adsorbent surface at any time.[21]

This model goes with R_L as the separation factor. This enables us to better ascertain the important characteristics of the Langmuir adsorption isotherm model. Also, it is a dimensionless constant. R_L is expressed as:

$$R_L = \frac{1}{1 + K_L C_o} \quad \text{Equation 5}$$

where K_L is the Langmuir constant (mg/g), and C_o is the adsorbate initial concentration.

When $R_L>1$, the adsorption is considered to be unfavourable; when $R_L=1$ it is linear; when $R_L=0$, it is irreversible, and finally when $0<R_L<1$, it is favourable.[22]

The Freundlich model delineates the reversible and imperfect less ideal adsorption process. This isotherm often fits adsorption processes which occur on heterogeneous surfaces in the gas phase.[23] In the Freundlich adsorption isotherm, high sufficient pressure results in the infinite limit, which means it does not fit best to a wide range of data from adsorption experiments.[23] The standard Freundlich adsorption isotherm model is:

$$Q_e = K_F C_e^{1/n} \quad \text{Equation 6}$$

where Q_e is the quantity of toxic metal removed at equilibrium, per gram of the adsorbent (mg/g); K_F is the Freundlich isotherm constant (mg/g); C_e is the concentration of the adsorbate at equilibrium (mg/L); n is the empirical constant; and $\frac{1}{n}$ is the adsorption intensity. The linear form of the Freundlich model is:

$$\text{Log } Q_e = \text{Log} K_F + \frac{1}{n} \log C_e \quad \text{Equation 7}$$

The $\frac{1}{n}$ shows how energy is relatively distributed and how heterogeneous the adsorption sites are.[22,23] If $\frac{1}{n}$ <1, it means the adsorption is normal. If $\frac{1}{n}>1$, it indicates that there is co-operative adsorption. If $n=1$, it means two-phase partition that does not rely on concentration has occurred.[22]

Results and discussion

Adsorption of toxic metals by tea in the binary system

The adsorption efficiency of mercury and lead at concentrations of 0.10, 1.00 and 5.00 mg/L by the tea waste adsorbent was 100% for mercury and ranged from 38.52% to 100% for lead (Table 2). In the binary systems, we observed a good removal efficiency for both Hg^{2+} and Pb^{2+} with Hg^{2+} having complete removal efficiency throughout at all dosages. Pb^{2+} had complete removal efficiency at 1 g and 3 g but showed a lower removal efficiency at an adsorbent dosage of 5 g. As both metals in the system were bivalent, it is most likely that the active or sorption sites were occupied or reacted with Hg^{2+} before Pb^{2+}. Another reason for the fall of removal efficiency can be accrued to the formation of clusters and complexes of ions at the active sites by Hg^{2+} which prevented or obstructed effective utilisation of the active sites.[24]

Table 2: Removal percentage of toxic metals by tea waste adsorbent in the binary system

Metal	Dosage (g)	Initial concentration (mg/L)	Final concentration (mg/L)	Percentage (%)
Hg: Pb	1	0.10: 0.10	0.00: 0.00	100: 100
	3	1.00: 1.00	0.00: 0.00	100: 100
	5	5.00: 5.00	0.00: 3.07	100: 38.52
Hg: Cd	1	0.10: 0.90	0.00: 0.00	100: 100
	3	1.00: 1.86	0.00: 1.17	100: 37.36
	5	5.00: 5.86	0.00: 3.26	100: 44.36
Pb: Cd	1	0.10: 0.90	0.00: 0.24	100: 73.10
	3	1.00: 1.86	0.00: 1.26	100: 37.36
	5	5.00: 5.86	0.91: 5.19	81.84: 11.50

The adsorption efficiencies of the tea waste adsorbent in the binary systems was 100% for mercury and ranged from 37.36% to 100% for cadmium (Table 2). In the binary system of Hg^{2+} and Cd^{2+}, Hg^{2+} had complete removal efficiency at all dosages and concentrations. Cd^{2+} had complete removal efficiency at 1 g of adsorbent dosage and lower removal at 3 g and 5 g. The overall performance shows that Hg^{2+} had a higher removal efficiency than Cd^{2+} ($Hg^{2+} > Cd^{2+}$). This is as a result of the nature of the surface and the interaction with the competing metal ions. Hg^{2+} has an ionic radius of 1.02 Å and Cd^{2+} has an ionic radius of 0.97 Å, which depicts the size of the metal ions.[25] Based on the size of the metal ions, Hg^{2+} has a larger size and hence is capable of occupying the active binding sites of the adsorbent more quickly which explains why it had a greater affinity for the tea waste. The strength of adsorption of metals onto agriculture waste adsorbent (biochar) is dependent on the ionic radii of the metal ions, electronegativity, charge, active site affinity, and the type of metal binding.[25]

The adsorption efficiency of tea waste for lead and cadmium in binary systems ranged from 81.84% to 100% for lead and 11.50% to 73.10% for cadmium (Table 2). In the binary system of Pb^{2+} and Cd^{2+}, Pb^{2+} had complete removal efficiency at all dosages. Cd^{2+} showed a promising removal efficiency at 1 g of adsorbent dosage by achieving higher removal efficiency, and subsequent dosages recorded a lower removal efficiency. Adsorption efficiency of Pb^{2+} by Impra Green Tea Ginseng Flavoured was greater than that for Cd^{2+} ($Pb^{2+} > Cd^{2+}$) because of the strong and higher electronegativity of Pb^{2+} (2.33) than Cd^{2+} (1.69).[25] Due to this characteristic, Pb^{2+} would occupy active sites faster than Cd^{2+}. Toxic metals such as Pb^{2+} have greater affinity with adsorbents that have cellulose and lignin surface sites.[26] The binding sites of tea waste constitute carboxyl groups and a -OH binding group which has a greater affinity for Pb^{2+}. Lead having greater removal efficiency than cadmium can also be attributed to the fact that the affinity of Pb^{2+} for most functional groups in organic matter is higher than that of Cd^{2+}. This is due to the variations in the metal ion chemical makeup and attributes or characteristics between Pb^{2+} and Cd^{2+}.[27]

Adsorption of toxic metals by tea in the ternary system

The adsorption efficiency for mercury, lead and cadmium in the ternary system at varied concentrations was 100% for both mercury and lead and ranged from 20.88% to 72.85% for cadmium (Table 3). In the ternary system, Hg^{2+}, Pb^{2+} and Cd^{2+} had effective removal. There was complete removal of Hg^{2+} and Pb^{2+} at all dosages. The trend for this experiment was Pb^{2+}>Hg^{2+}>Cd^{2+}. Pb^{2+} had a greater affinity to the binding or active sites of the tea waste adsorbent than did Hg^{2+} and Cd^{2+}. This is due to the electronegativity of the metal ions (Pb^{2+}=2.33, Hg^{2+}=2.00, Cd^{2+}=1.69) which gives Pb^{2+} the upper hand in the surface electrostatic attraction and the complexation of the inner sphere surfaces of the adsorbent.[25] Also, Pb^{2+} has a better affinity for functional groups that exist in organic matter such as the carboxyl groups and -OH binding sites than Hg^{2+} and Cd^{2+}.[26] In a similar study done by Wan et al.[26], Pb^{2+} had greater affinity to the tea waste functional groups and higher maximum adsorption capacity than Cd^{2+} and Cu^{2+} in their batch experiment. Their experiment showed the trend Pb^{2+}>Cu^{2+}>Cd^{2+}, indicating that Pb^{2+} had the greatest affinity to the tea wastes' functional groups.

Langmuir and Freundlich adsorption isotherms

The Langmuir adsorption isotherm was employed to explain the nature and estimate the adsorption capacity of the toxic metals onto the tea waste adsorbent in the greywater in two batches (binary and ternary systems). The maximum adsorption capacity (Q_o) for the three metals in the batch experiment ranged from 38.46 to 81.97 mg/g for Pb^{2+} and 12.64 to 56.82 mg/g for Cd^{2+}. In the binary systems, Q_{max} ranged from 12.64 mg/g to 81.97 mg/g and 38. 61 mg/g in the ternary system. R_L ranged from −0.58 to 4.30 for metals in the binary systems and 8.83 for metals in the ternary system. The correlation coefficient (R^2) for the Langmuir adsorption isotherm could not be computed for toxic metals that had 100% complete adsorption (Figures 1–5). R^2 values for the Langmuir adsorption isotherm for mercury, lead and cadmium metals in the binary systems ranged from 0.7710 to 1.00 (Table 4). In the ternary systems, the R^2 obtained was 0.6109 for cadmium (Table 4).

The findings show that the Langmuir adsorption isotherm model was the best fit for the adsorption of toxic metals by Impra Green Tea Ginseng Flavoured adsorbent. R^2 values depict the order of the batch experiment as binary system > ternary system. The adsorption of mercury, cadmium and lead ions onto the tea adsorbent was monolayer with equal affinity in binding sites for adsorption. The Langmuir isotherm is used when the adsorption process is monolayered with an identical and finite number of sites.[22,23] This model was used to ascertain the nature and type of adsorption that took place on the surface of the adsorbent. A similar study done by Wan et al.[26] obtained an R^2>0.97 for all R^2 values for both Langmuir and Freundlich isotherms. Hg^{2+} had the highest maximum adsorption capacity throughout the entire experiment. A similar study recorded maximum adsorption capacities for Pb^{2+}, Cd^{2+} and Cu^{2+} of 33.49, 16.87 and 21.02 mg/g, respectively, at optimum conditions using tea waste as an adsorbent.[26] From the results, lead and cadmium negative values for K_L indicate a weak interaction. Whilst higher K_L values suggested a strong affinity between the metal ions and the tea adsorbent. R_L for Cd^{2+} in the binary and ternary system was greater than 1, implying the sorption process was unfavourable. This result indicates

Table 3: Removal percentage of toxic metals by tea waste adsorbent in the ternary system

Metal	Dosage (g)	Initial concentration (mg/L)	Final concentration (mg/L)	Percentage (%)
Hg: Pb: Cd	1	0.10: 0.10: 0.90	0.00: 0.00: 0.24	100: 100: 72.85
	3	1.00: 1.00: 1.86	0.00: 0.00: 1.47	100: 100: 20.88
	5	5.00: 5.00: 5.86	0.00: 0.00: 4.17	100: 100: 28.91

Table 4: Modelling of Langmuir and Freundlich isotherms

Metal	Langmuir parameter					Freundlich parameter			
	Q_{max} (mg/g)	K_L (l/ mg)	R_L	R^2		$1/n$	N	K_F (mg/g)	R^2
Pb: Hg	38.46	-2.30×10^{-16}	2.00	1.00		1.50	0.67	18.26	0.3390
Cd: Hg	56.82	0.66	5.30	0.7710		−78.13	−0.01	48.05	0.0001
Cd: Pb	12.64	−0.32	0.42	0.9972		−1.90	−0.53	28.26	0.9478
Pb: Cd	81.97	8.20×10^{-17}	2.00	1.00		−0.06	−15.54	18.26	0.6743
Cd: Pb: Hg	38.61	1.14	7.71	0.6109		−3.26	−0.31	31.86	0.2886

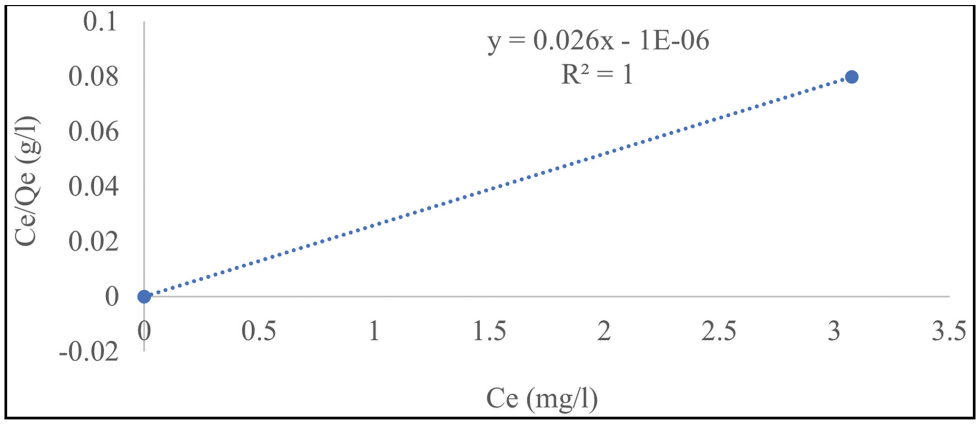


Figure 1: Langmuir isotherm graph for the adsorption of lead in greywater by spent Impra Green Tea Ginseng Flavoured in a binary system (Pb: Hg).

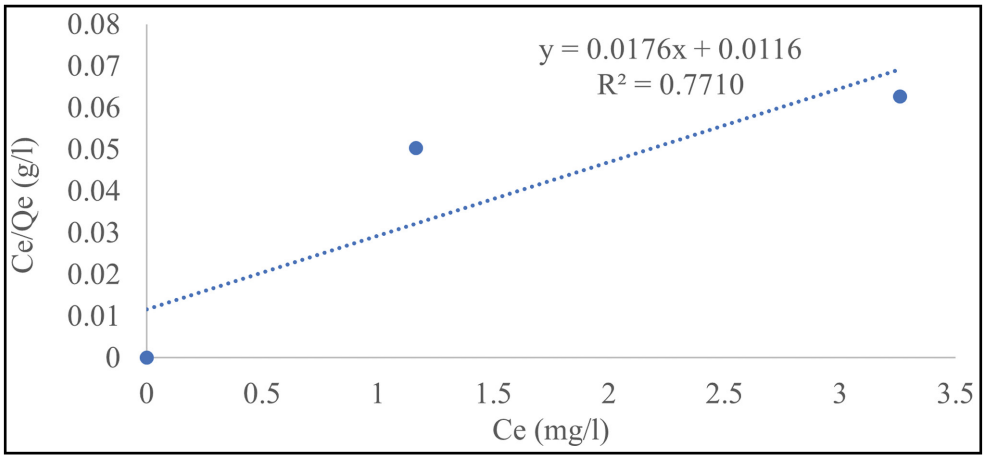


Figure 2: Langmuir isotherm graph for the adsorption of cadmium in greywater by spent Impra Green Tea Ginseng Flavoured in a binary system (Cd: Hg).

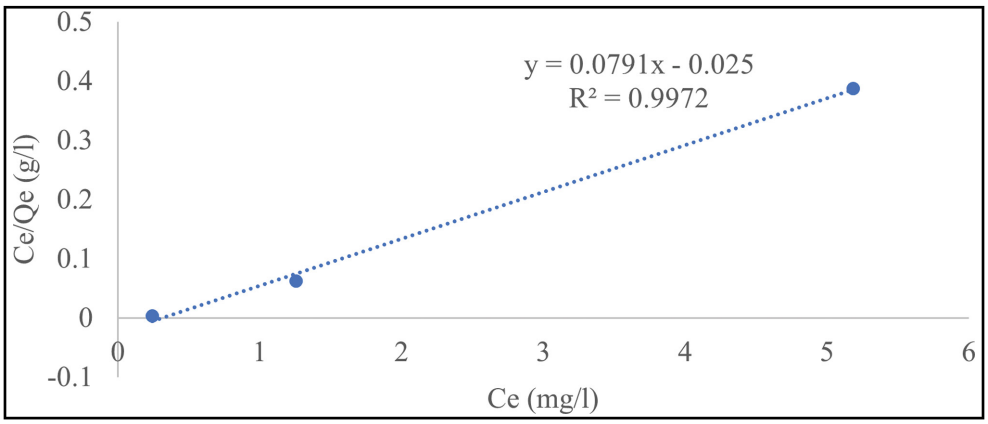


Figure 3: Langmuir isotherm graph for the adsorption of cadmium in greywater by spent Impra Green Tea Ginseng Flavoured in a binary system (Cd: Pb).

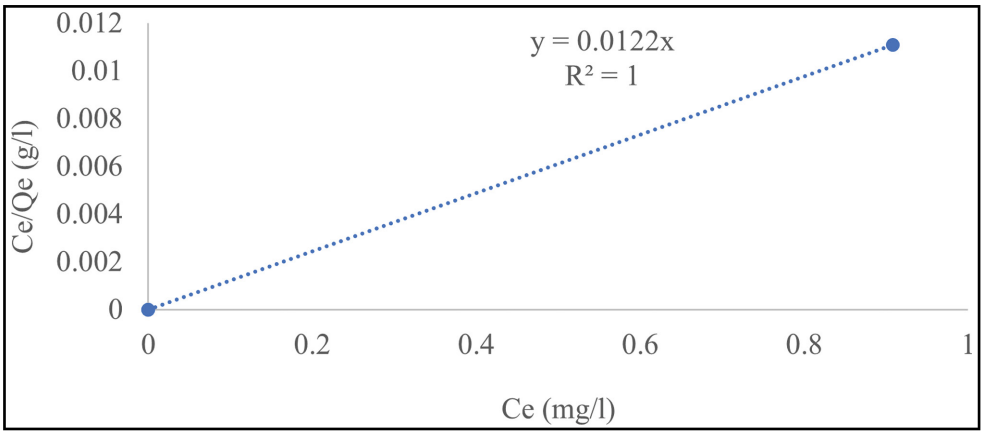


Figure 4: Langmuir isotherm graph for the adsorption of lead in greywater by spent Impra Green Tea Ginseng Flavoured in a binary system (Pb: Cd).

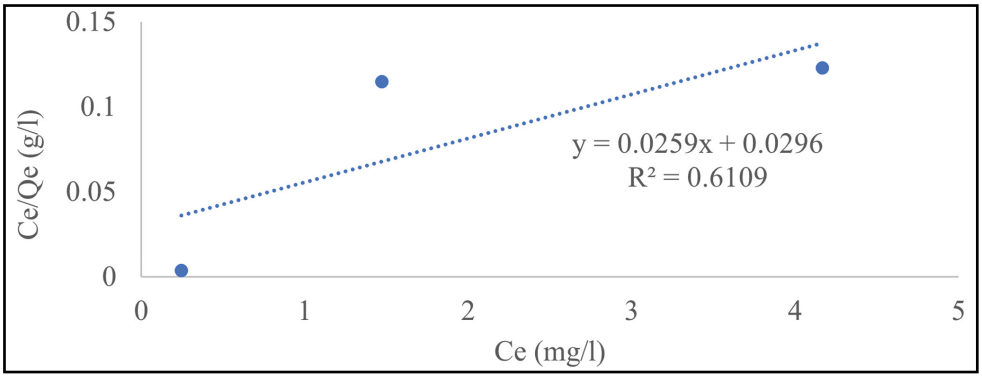


Figure 5: Langmuir isotherm graph for the adsorption of cadmium in greywater by spent Impra Green Tea Ginseng Flavoured in a ternary system (Hg: Pb: Cd).

that tea waste can be used as a good sorbent for the removal of toxic metals such as mercury and lead.

The Freundlich adsorption isotherm was utilised to consider the fitness of the batch experiments with the heterogeneity of surface sites of the tea waste adsorbent and the adsorption of toxic metals. For toxic metals in the binary systems, $1/n$ ranged from –78.13 to 1.50 and –3.26 for the ternary system (Table 4). N values for binary and ternary systems ranged from 15.54 to 0.67 (Table 4). K_F (mg/g) values for the batch experiments ranged from 18.26 to 48.05 mg/g (Table 4). Freundlich adsorption isotherm correlation coefficient values for toxic metals in the binary systems ranged from 0.0001 to 0.9478, and 0.2886 for the ternary system (Table 4). The Freundlich adsorption isotherm was employed to describe the adsorption process between the toxic metals under study (Hg^{2+}, Pb^{2+} and Cd^{2+}) and the spent Impra Green Tea Ginseng Flavoured adsorbent (Figures 6–10). The experimental data obtained fit the model well with Cd^{2+} fitting the model best. The Freundlich isotherm offers an expression that enables the description of heterogeneous surfaces of adsorbent and the exponentially distributed actives sites on them and their energies.[22] In the binary and ternary systems, only some experimental data obtained fitted this model. The $1/n$ values for metals in the binary and ternary systems were mostly less than unity.

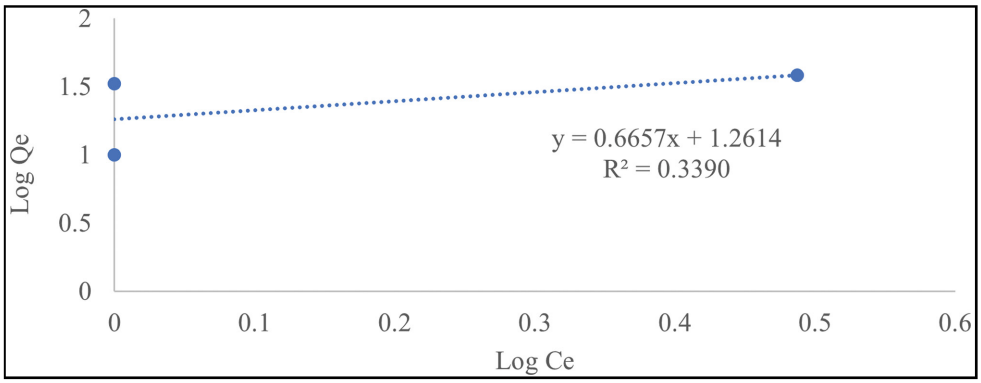


Figure 6: Freundlich isotherm graph for the adsorption of lead in greywater by spent Impra Green Tea Ginseng Flavoured in a binary system (Pb: Hg).

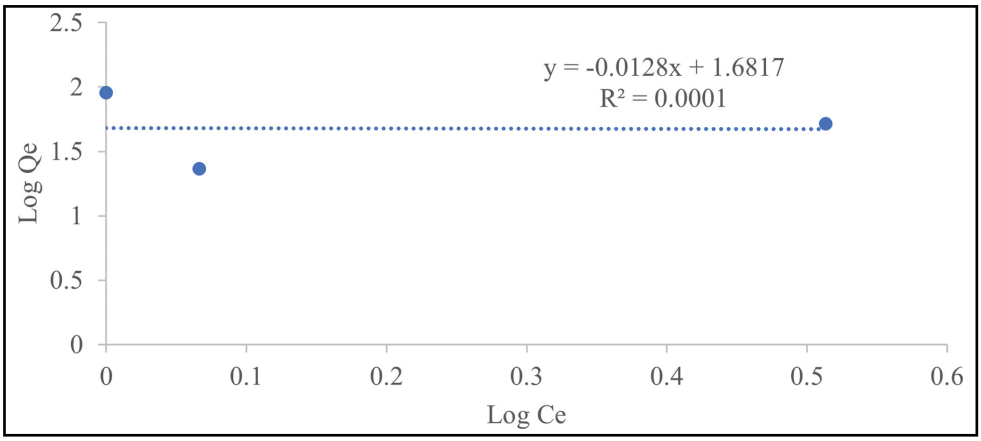


Figure 7: Freundlich isotherm graph for the adsorption of cadmium in greywater by spent Impra Green Tea Ginseng Flavoured in a binary system (Cd: Hg).

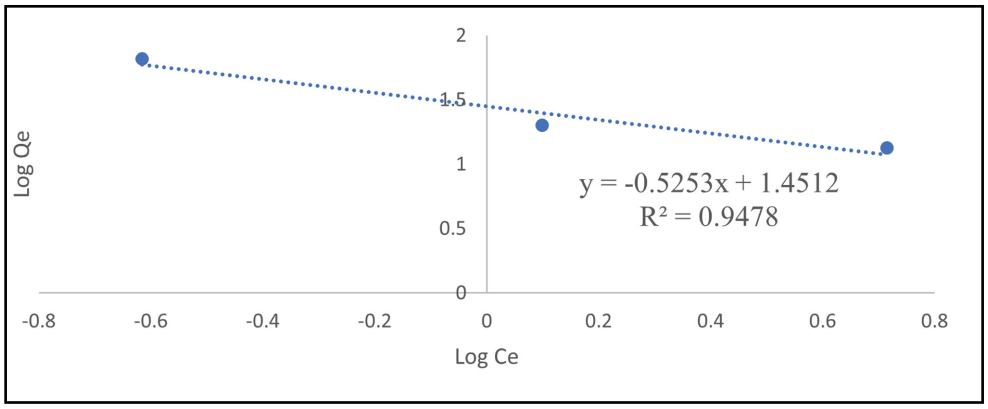


Figure 8: Freundlich isotherm graph for the adsorption of cadmium in greywater by spent Impra Green Tea Ginseng Flavoured in a binary system (Cd: Pb).

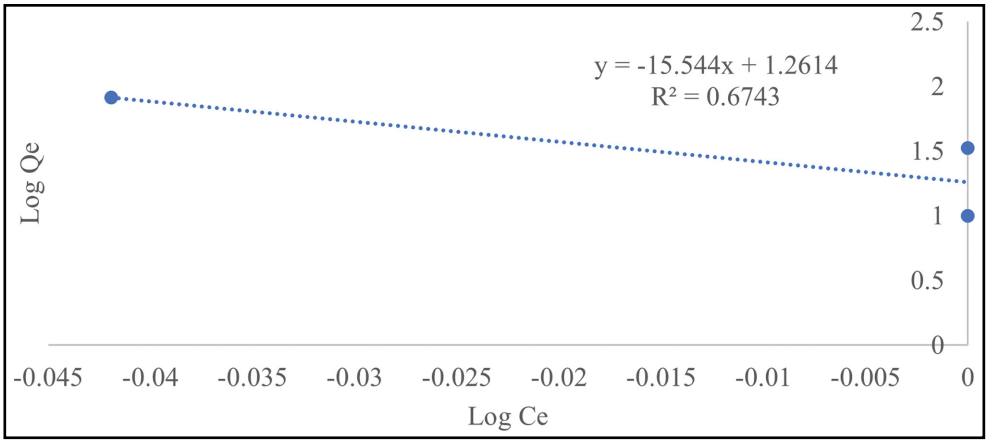


Figure 9: Freundlich isotherm graph for the adsorption of lead in greywater by spent Impra Green Tea Ginseng Flavoured in a binary system (Pb: Cd).

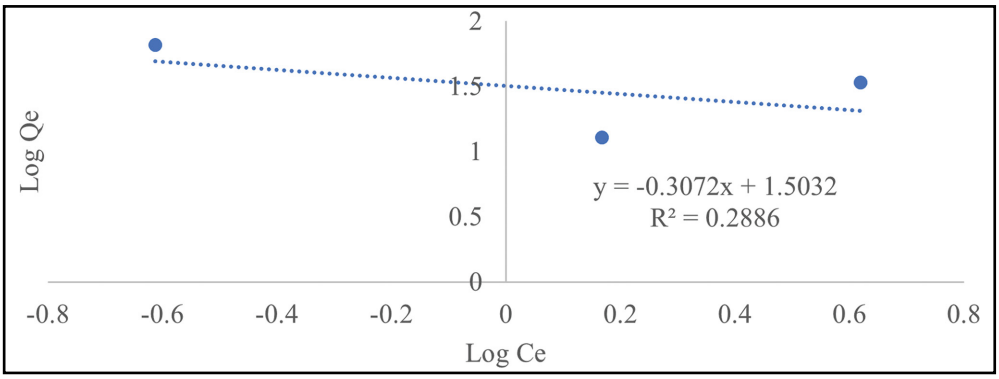


Figure 10: Freundlich isotherm graph for the adsorption of cadmium in greywater by spent Impra Green Tea Ginseng Flavoured in a ternary system (Cd: Pb: Hg).

Conclusion

Spent Impra Green Tea Ginseng Flavoured was used as an adsorbent to remove mercury, lead and cadmium from greywater. The removal of toxic metals from greywater by a low-cost tea waste adsorbent was highly efficient and effective throughout the experiment except for cadmium which yielded lower removal efficiency. The experimental data fitted better in the Langmuir adsorption isotherm than in the Freundlich isotherm. The adsorption capacity and rates are dependent on the spent Impra Green Tea Ginseng Flavoured dosage, initial concentration, pH solution, particle size, resident time and some other experimental conditions. Spent Impra Green Tea Ginseng Flavoured can be used as an alternative, effective and cheap adsorbent of toxic metals from greywater and wastewater. There is a need for further studies on using tea waste adsorbent to remediate other wastewater to have a better understanding and establish a wider range of applicability of the adsorbent. We did not vary the contact time, temperature or toxic metal concentrations. Further study should therefore be undertaken to vary the contact time, temperature and metal concentrations in the greywater.

Acknowledgements

We thank Mr Saeed Abdullah for his assistance during the experiment and Mr Prince Owusu (a Technician in the Ecological Laboratory, University of Ghana) for the sample analysis.

Competing interests

We have no competing interests to declare.

Authors' contributions

R.B.H.G.: Conceptualisation, writing – initial draft, sample analysis, data analysis, validation. E.H.A., A.B.D.: Conceptualisation, writing – initial draft. E.D.A., A.-A.B.: Methodology, data collection.

References

1. Gerenfes D. Assessment of heavy metals removal by binary and ternary mixed oxide nanocomposites. Int J Novel Res Life Sci. 2019;6(2):22–32.
2. Reynel-Avila HE, Mendoza-Castillo ID, Hernández-Montoya V, Bonilla-Petriciolet A. Multicomponent removal of heavy metals from aqueous solution using low-cost sorbents. In: Water production and wastewater treatment. New York: Editorial Nova Science Publishers; 2011. p. 69–99.
3. Wan-Ngah SW, Hanafiah MAKM. Removal of heavy metal ions from wastewater by chemically modified plant wastes as adsorbents: A review. Bioresour Technol. 2008;99:3935–3948. https://doi.org/10.1016/j.biortech.2007.06.011
4. Wang J, Chen C. Biosorbents for heavy metals removal and their future. Biotechnol Adv. 2009;27:195–226. https://doi.org/10.1016/j.biotechadv.2008.11.002
5. Amarasinghe B, Williams RA. Tea waste as a low cost adsorbent for the removal of Cu and Pb from wastewater. J Chem Eng. 2007;132 (1–3):299–309. https://doi.org/10.1016/j.cej.2007.01.016
6. Nandal M, Dhania G. Tea wastes as a sorbent for removal of toxic metals from wastewater. Int J Curr Eng Technol. 2019;4(1):243–247.
7. Hussain S, Anjali KP, Hassan ST, Dwivedi PB. Waste tea as a novel adsorbent. Appl Water Sci. 2018;8(6):165. https://doi.org/10.1007/s13201-018-0824-5
8. Rodríguez C, Sánchez R, Lozano-Parra J, Rebolledo J, Schneider N, Serrano J, et al. Water balance assessment in schools and households of rural areas of Coquimbo region, North-Central Chile: Potential for greywater reuse. Water. 2020;12:2915. https://doi.org/10.3390/w12102915
9. Jeyaseelan C, Gupta A. Jeyaseelan C, Gupta A. Green tea leaves as a natural adsorbent for the removal of Cr (VI) from aqueous solutions. Air Soil Water Res. 2016;9:ASWR-S35227. https://doi.org/10.4137/ASWR.S35227
10. Çelebi H, Gök G, Gök O. Adsorption capability of brewed tea waste in waters containing toxic lead(II), cadmium (II), nickel (II), and zinc(II) heavy metal ions. Sci Rep. 2020;10:1–12. https://doi.org/10.1038/s41598-020-74553-4
11. Duwiejuah AB, Amadu Y, Gameli BHR, Bawa A-A, Imoro ZA, Alidu SM, et al. Spent Chinese green tea as an adsorbent for simultaneous removal of toxic metals from aqueous solution. Chem Afr. 2022;5:2107–2114. https://doi.org/10.1007/s42250-022-00459-5
12. Gameli BHR, Duwiejuah AB, Bawa A-A. Adsorption of toxic metals from greywater using low-cost spent green tea as a novel adsorbent. Sci Afr. 2022;17:e01296. https://doi.org/10.1016/j.sciaf.2022.e01296
13. Katha PS, Ahmed Z, Alam R, Saha B, Acharjee A, Rahman M. Efficiency analysis of eggshell and tea waste as low cost adsorbents for Cr removal from wastewater sample. S Afr J Chem Eng. 2021;37:186–195. https://doi.org/10.1016/j.sajce.2021.06.001
14. Claude NJ, Shanshan L, Khan J, Yifeng W, Dongxu H, Xiangru L. Waste tea residue adsorption coupled with electrocoagulation for improvement of copper and nickel ions removal from simulated wastewater. Sci Rep. 2022;12:3519. https://doi.org/10.1038/s41598-022-07475-y
15. Singh SR, Singh AP. Adsorption of heavy metals from waste waters on tea waste. Glob J Eng Res. 2012;12(1):19–22.
16. Eriksson EÃ. Donner E. Metals in greywater: Sources, presence and removal efficiencies. Desalination. 2009;248 (1–3):271–278. https://doi.org/10.1016/j.desal.2008.05.065
17. Hussain S, Saima KPA, Hassan T, Brat P. Waste tea as a novel adsorbent: A review. Appl Water Sci. 2018;8(6):1–16. https://doi.org/10.1007/s13201-018-0824-5
18. Thakur LS, Parmar M. Adsorption of heavy metals (Cu^{2+}, Ni^{2+}, Zn^{2+}) from synthetic waste water by tea waste adsorpbent. Int J Chem Phys. 2013;2(6):6–19.
19. Khayyun ST, Mseer AH. Comparison of the experiment results with the Langmuir and Freundlich models for copper removal on limestone adsorbent. Appl Water Sci. 2019;9:170. https://doi.org/10.1007/s13201-019-1061-2
20. Boukhlifi F, Ouchabi M, Amar A, Jabri M, Kali A, Chraibi S, et al. Adsorption of crystal violet onto an agricultural waste residue: Kinetics, isotherm, thermodynamis and mechanism of adsoprtion. Sci World J. 2020;2020, Art. #5873521. https://doi.org/10.1155/2020/5873521
21. Kecili R, Hussain CM. Nanomaterials in chromatography. Amsterdam: Elsevier; 2018.

22. Ayawei N, Ebelegi AN, Wankasi D. Modelling and interpretation of adsorption isotherms. J Chem. 2017; 2017, Art. #3039817. https://doi.org/10.1155/2017/3039817

23. Al-Ghouti A, Da'ana AD. Guidelines for the use and interpretation of adsorption isotherm models: A review. J Hazard Mater. 2020;393, Art. #122383z. https://doi.org/10.1016/j.jhazmat.2020.122383

24. Naeem AS, Imran M, Amjad M, Ghulam A, Tahir M, Murtaza B, et al. Batch and column scale removal of cadmium from water using raw and acid activated wheat straw biochar. Water. 2019;11:17. https://doi.org/10.3390/w11071438

25. Duwiejuah AB. Eco-friendly biochars for the adsorption of heavy metals from aqueous phase [MPhil thesis]. Tamale: University for Development Studies; 2017.

26. Wan S, Ma Z, Xue Y, Ma M, Xu S, Qian L, et al. Sorption of lead (II), cadmium (II) and copper (II) ions from aqueous solution using tea waste. Ind Eng Chem Res. 2014;53:3629–3635. https://doi.org/10.1021/ie402510s

27. Cay S, Uyanik A, Ozasik A. Single and binary component adsorption on copper (II) and cadmium (II) from aqueous solution using tea industry waste. Sep Purif Technol. 2004;38:273–280. https://doi.org/10.1016/j.seppur.2003.12.003

AUTHORS:
Bongiwe P. Mshengu[1]
Cordelia Z. Buthelezi[2]
Roshila Moodley[2,3]

AFFILIATIONS:
[1]Department of Chemistry, University of Pretoria, Pretoria, South Africa
[2]School of Chemistry and Physics, University of KwaZulu-Natal, Durban, South Africa
[3]Department of Chemistry, University of Manchester, Manchester, UK

CORRESPONDENCE TO:
Bongiwe Mshengu

EMAIL:
bongiwe.mshengu@up.ac.za

DATES:
Received: 31 Jan. 2022
Revised: 02 Feb. 2023
Accepted: 03 Feb. 2023
Published: 08 Aug. 2023

HOW TO CITE:
Mshengu BP, Buthelezi CZ, Moodley R. Elemental, phytochemical, and toxicological assessment of *Cissus rotundifolia* (Forssk.) Vahl. S Afr J Sci. 2023;119(7/8), Art. #13160. https://doi.org/10.17159/sajs.2023/13160

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

DATA AVAILABILITY:
☐ Open data set
☒ All data included
☐ On request from author(s)
☐ Not available
☐ Not applicable

EDITOR:
Teresa Coutinho

KEYWORDS:
Cissus rotundifolia, heavy metals, nutritional value, phytosterol, pheophytin

FUNDING:
University of KwaZulu-Natal, South African National Research Foundation

Elemental, phytochemical, and toxicological assessment of *Cissus rotundifolia* (Forssk.) Vahl

Cissus rotundifolia (Forssk.) Vahl. (Vitaceae) is a wild plant that is commonly used by communities from rural areas as a food and medicine. There are limited studies on the phytochemical composition and the impact of soil quality on the elemental distribution in this plant. In this study, we report a phytochemical analysis to identify the phytocompounds responsible for the reported biological activities of *C. rotundifolia*. We also examined the impact of soil quality on elemental uptake by the edible parts of *C. rotundifolia* collected from eight geographical locations in KwaZulu-Natal (South Africa) to assess the nutritional benefits and potential heavy metal toxicities. Three secondary metabolites (stigmasterol, β-sitosterol, and pheophytin *a*) were isolated, and their structures were characterised by high-resolution mass spectrometry and nuclear magnetic resonance data. The plant was found to contribute adequately to the recommended dietary allowances for essential nutrients without exceeding tolerable upper intake limits and with low concentrations of toxic heavy metals. The average concentrations of microelements in the edible parts were found to be in decreasing order of Fe>Mn>Se>Zn>Cu>Cr>Ni>Co. The bioaccumulation factors indicate that the plant controls the uptake of metals from the soil and would make a good indicator and biological monitor for cadmium toxicity. However, a health risk assessment exposed carcinogenic risks on regular consumption of the plant obtained from sites close to pollution sources, such as roads and landfills. The findings from this study show the synergies when consuming medicinal plants and provide evidence for *C. rotundifolia* as a nutraceutical.

Significance:

This study provides additional scientific knowledge on the phytochemical composition of *C. rotundifolia*. Three phytocompounds (stigmasterol, β-sitosterol, and pheophytin *a*) were isolated, and their presence may be correlated to this plant's antidiabetic, anti-inflammatory, and antibacterial properties. This study shows that *C. rotundifolia* contributes adequately to the recommended dietary allowances for essential elements, and the plant is safe for human consumption if collected from non-polluted sites. The carcinogenic and non-carcinogenic estimates for the toxic metals due to consumption of the plant signify the possibility of developing cancer over time if the plant is consumed frequently from polluted sites.

Introduction

The World Health Organization (WHO) estimated that 80% of the world's population depends on plants for their primary needs, such as food and medicine.[1] Medicinal plants are a rich source of bioactive phytochemicals, which can be used directly as new drugs or developed into lead compounds for optimisation into novel drugs against various diseases.[1] Nonetheless, information on most medicinal plants' phytochemistry still needs enhancement.[1] In addition, local people from disadvantaged communities consume wild-growing edible plants to meet their nutritional needs and to diversify their diet.[2,3] Essential elements obtained from these plants are necessary for the metabolic processes and functioning of the human body, and their deficiencies or excessiveness can have detrimental health effects.[4,5] Therefore, it is essential to study the nutritional value of wild edible plants.

Wild edible plants and medicinal plants are often considered safe for human consumption; however, many studies have reported on the contamination of these plants by toxic elements from their growing environments.[6-8] As a result, studies focused on evaluating edible plants for heavy metal toxicity are of great importance to human health. *Cissus rotundifolia* (Forssk.) Vahl. (Vitaceae) is a wild, edible medicinal plant that grows in various regions of East and Central Africa, Zimbabwe, Mozambique, Egypt, Yemen, Saudi Arabia, and South Africa.[9,10] The plant has several biological activities, including antioxidant, antibacterial, anti-diabetic, anti-inflammatory, anti-fertility, anti-ulcerative, anti-parasitic, antimalarial, and cytotoxicity.[9,11-13]

Previous phytochemical studies on *C. rotundifolia* resulted in the characterisation of a triterpene, flavonoids, steroidal saponin, cissuxinoside, and cissoic acid.[9,14] Another study on the phytochemistry of *C. rotundifolia* identified 1-O-(4-coumaroyl)-β-D-glucopyranose as an active principle against diabetes.[9] However, compounds responsible for the other biological activities of this plant have not been identified. The cooked leaves of *C. rotundifolia* are widely consumed by people in rural areas of South Africa and other parts of the world for their nutritional and therapeutic benefits.[9,15] The plant is also used as a cheap thickening agent.[10] The leaves and fruits of *C. rotundifolia* collected from different regions of Saudi Arabia, and Yemen were reported to contain enough essential nutrients, including carbohydrates, proteins, fats, vitamins, minerals, and amino acids.[15-18] The mineral composition of the stem of *C. rotundifolia* and the impact of soil quality on elemental distribution in this plant have not been reported.

In this study, we report on the phytochemical analysis of *C. rotundifolia*, intending to find additional bioactive compounds responsible for the plant's purported biological activities. We also report on the elemental composition of the leaves and stems collected from different geographical sites in KwaZulu-Natal (South Africa) to assess the plant's nutritional benefits and potential heavy metal toxicities.

Materials and methods

Sample preparation and analysis

C. rotundifolia plant and soil samples were collected from eight sites in KwaZulu-Natal, South Africa (Figure 1): Margate (S1), Ballito (S2), Richards Bay (S3), Shongweni (S4), Cato Ridge (S5), St Lucia (S6), Amanzimtoti (S7), and Umgababa (S8). Except for sampling sites S4 and S5, the sites are along the coast; it was assumed that they are non-polluted as they are far from roads, landfills, and other pollution sources. S4 and S5 are inland and very close to pollution sources (S4 is close to a landfill site, and S5 is close to a railway line and a truck stop).

Collected plant material was placed in plastic bags, sealed, and stored. Plant specimens from each sampling site were submitted for authentication by the taxonomist, Mr Edward Khathi, in the School of Life Sciences (University of KwaZulu-Natal, South Africa), where a voucher specimen (Buthelezi C 01-18274) was deposited. The soil was collected from the eight sites at 10–15 cm below the plant's roots using a plastic spade. Plant samples were washed with double-distilled water, dried in the oven at 50 °C for 24 h, and then crushed using a pestle and mortar. Plant material from site 7 (Amanzimtoti) was used for phytochemical analysis due to the large quantity available at this site. Composite soil samples from each site were obtained by coning and quartering. Soil samples (10 g) were sieved through a 2 mm mesh sieve to get uniform particle size, air dried, and crushed using a mortar and pestle. All dried samples were stored in sealed plastic bags and kept in a refrigerator at 4 °C.

Nuclear magnetic resonance (NMR) analysis was performed using the Bruker Avance III spectrometers (400 MHz and 600 MHz). The isolated compounds were dissolved in deuterated chloroform ($CDCl_3$) or methanol (MeOD), and tetramethylsilane was used as the internal standard. Chemical shifts (δ) and coupling constants are reported in ppm and Hertz (Hz), respectively. Column chromatography was packed using Merck silica gel 75–230 μm. Thin-layer chromatography (TLC) plates were pre-coated with Merck silica gel 60 F254. Agilent GC-MSD was used to obtain the mass of the compounds.

The elemental analysis of the plant and soil was performed using analytical grade solvents from Sigma Aldrich (St. Louis, MO, USA). The laboratory glassware and the glass bottles were soaked in 3 M nitric acid (HNO_3), rinsed with double distilled water, and air dried. A CEM Microwave Accelerated Reaction System with MARSXpress™ vessels and infrared temperature sensors (CEM Corporation, NC, USA) was used for acid digestion of the plant samples. Inductively coupled plasma-optical emission spectrometry (ICP-OES) (PerkinElmer, Optima 5300Dual View, Billerica, MA, USA) was used for elemental analysis.

Extraction and purification

The dried powdered roots (572 g), stems (461 g), and leaves (562 g) were extracted with dichloromethane (DCM) and MeOH for 48 h consecutively. The extracts were filtered and concentrated using a rotary evaporator to give the crude extracts, which were stored in the fridge until further use. The extracts from the stems and leaves had similar TLC profiles. The DCM crude extract from the roots (11.65 g), stems (17.65 g), and leaves (50.44 g) were fractioned on a silica gel column using different proportions of hexane and ethyl acetate (starting from 100% hexane and increasing the polarity by 10% to 100% ethyl acetate). The collected fractions were profiled using TLC to give six main fractions. Compound **1** (110 mg, 0.94% of the DCM extract) and compound **2** (35 mg, 0.3% of the DCM extract) were isolated from fractions 1 and 2 of the DCM root extract, respectively. The DCM extract of the leaves was fractionated to give three main fractions, and compound **3** (6.4 mg, 0.013% of the DCM extract) was isolated from fraction 2.

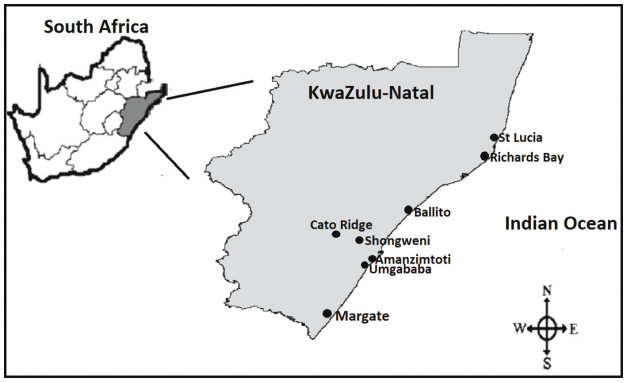


Figure 1: Map showing selected sampling sites in KwaZulu-Natal, South Africa.

Elemental analysis

Plant material (0.25 g) and the certified reference material (CRM) were weighed into microwave vessels with 10 mL HNO_3 in three replicates. The samples were pre-digested for 30 min before microwave digestion. The power was set to 100% (1600 W at 180 °C) for 20 min. The holding time and cooling times were set to 15 min each. After digestion, the samples were filtered through 0.45 μm nylon membrane filters with a 25 mm syringe into 50 mL volumetric flasks topped up to the mark with double-distilled water, transferred into 15 mL polypropylene vials, and refrigerated at 4 °C before analysis. The elements measured were: arsenic (As), cadmium (Cd), cobalt (Co), chromium (Cr), copper (Cu), iron (Fe), manganese (Mn), nickel (Ni), palladium (Pd), selenium (Se), and zinc (Zn); the elements were measured using ICP-OES and concentrations are presented as mg/kg dry weight (DW) of the sample.

Quality assurance

The validity of an analytical method was determined by evaluating the certified reference material (CRM, strawberry leaves, LGC-7162, Community Reference Bureau of the Commission of the European Communities, Brussels, Belgium). The wavelength which produced the most acceptable CRM results with the highest intensity and no interfering elements (Table 1) was selected. Reagent blanks and calibration standards for each element in double-distilled water (within the estimated ranges) that were prepared from 1000 mg/L stock standard solutions (Fluka Analytical, Sigma, Switzerland) were used to produce five-point calibration curves relating to concentration strength. The best linear fit of the curves was chosen. All samples, including calibration standards and blanks, were analysed in 70% nitric acid to eliminate matrix effects and reduce spectral interferences. The CRMs were analysed to verify the correctness and exactness of the calibration curve and were also used to accept or reject the calibration curve. The limit of detection (LOD) and limit of quantification (LOQ) were also determined to validate the method.

The method's accuracy was determined by comparing mean experimental values from three replicates to the certified values for an analyte using the CRMs (Table 1). A two-sample *t*-test (assuming equal variances) confirmed the accuracy as there were no statistically significant differences between the means ($p>0.05$). Method precision, which shows measurement closeness after repeated analysis of an analyte, was estimated by comparing the %RSD for the CRM. This should be within 20% of the true value. The experimental values at the 95% confidence interval were within the appropriate range of that stipulated for the CRM. Therefore, the analytical method was accepted.

Bioaccumulation factor and pollution indicator

The bioaccumulation factor (BAF) defines the ability of plants to accumulate and store heavy metals in the soil. BAF is calculated as the plant's metal concentration ratio to the metal concentration in the corresponding soil (Equation 1).[19]

$$BAF = Metal\frac{plant}{Metal}soil \qquad \text{Equation 1}$$

The geoaccumulation index (I_{geo}) determines the soil's degree of metal pollution in the seven grades ranging from uncontaminated to extremely contaminated (Equation 2).[20]

$$Igeo = log2\left[\frac{Cn}{1.5Bn}\right], \qquad \text{Equation 2}$$

Table 1: The detection limit (DL) of the instrument, limit of quantitation (LOQ), certified and measured values for the certified reference material (CRM – LGC7162 strawberry leaves and soil D081-540, $n = 3$), based on dry mass

Elements	Wavelength (nm)	DL (mg/kg)	LOQ	Plant CRM	Measured	Soil CRM	Measured	Acceptable
As	193.696	0.0100	0.0303	–	–	88.4	120.7±14.42	72–105
Cd	228.802	0.0027	0.0082	–	–	143	154.9±13.6	116–169
Co	228.616	0.007	0.0212	–	–	199	229±52.3	166–233
Cr	267.716	0.0071	0.024	2.15±0.34	2.02±0.04	86.8	88.67±14.81	69.3–104
Cu	324.752	0.0054	0.0164	10	10.35±0.14	268	279±55.1	219–317
Fe	238.204	0.0062	0.0188	818±48	820.79±7.82	12 800	12 170±5388	5380–20 100
Mn	257.61	0.0016	0.0048	171±10	180.19±0.84	425	521±114.5	347–502
Ni	231.604	0.0480	0.1455	2.6± 0.7	2.72±0.06	236	255± 26.8	194–279
Pb	220.353	0.09	0.2727	1.8±0.4	1.78±0.10	97.9	137.52±30.6	80–116
Zn	213.857	0.0018	0.0055	24±5	25.26±0.48	130	143.9±38.1	106–115

where Cn is the elemental concentration, Bn is the total baseline concentration, and 1.5 is used to reduce background value variations to the rock composition variations.

Health risk assessment

The estimated daily intake (EDI) of the toxic metals was measured in mg/kg body weight per day[21]:

$$EDI = [X]x\ IR/Bw, \quad \text{Equation 3}$$

where [X] is the concentration of the toxic metal (mg/kg, DW), IR is the ingestion rate of leafy vegetables per person estimated to be 0.062 kg per day, and Bw is the average human body weight in South Africa, usually 70 kg.[22]

The target hazard quotient (THQ) is the ratio of exposure to a toxic metal with its reference dose (Equation 4), and it estimates the non-carcinogenic risk caused by exposure to the toxic metal.[23,24] For each toxic metal, a THQ<1 indicates low risk, and THQ≥1 shows high risk. RfD_0 is the safe oral reference dose of elements (mg/kg/day) calculated for adults. The RfD_0 values are as follows: As (0.0003), Cd (0.001), and Pb (0.004).

$$THQ = EDI/RfDo \quad \text{Equation 4}$$

Carcinogenic risk (CR) estimates the probability of developing cancer due to exposure to toxic metals over a lifetime (Equation 5).[22,23]

$$CR = EDI\ x\ CPSo, \quad \text{Equation 5}$$

where CPS_0 is the oral slope factor of the toxic metal (mg/kg/body weight/day). The CPS_0 values are as follows: As (1.5), Cd (6.3), and Pb (0.0085).[24] For each toxic heavy metal, a $CR > 1 \times 10^{-4}$ indicates a high probability of carcinogenic risk.[23]

Statistical analysis

A comparative study and grouping of results were performed using a one-way analysis of variance (ANOVA) and Tukey's post-hoc test to determine and evaluate significant differences between means. Statistical analysis was performed using the Statistical Package for the Social Sciences (SPSS) (PASW version 24, IBM Corporation, Cornell, NY, USA).

Results and discussion

Phytochemical investigation

Two phytosterols (stigmasterol and β-sitosterol) and pheophytin *a* were isolated from the leaves and the roots of *C. rotundifolia*. The compounds (Figure 2) were identified using spectroscopic data (NMR, GC-MS/LC-MS, IR, and UV/Vis) and by comparison of the experimental data with values in the literature. Compound **1** (110 mg), a white crystalline powder from the DCM extract of the roots with molecular ion peak at *m/z* 412 [M^+] corresponding to $C_{29}H_{48}O$, was identified as stigmasterol.[25] Stigmasterol is commonly isolated from the genus *Cissus*.[25-28] However, this is the first report on its isolation from *C. rotundifolia*. This compound is known for its antibacterial, antioxidant, anticancer, and cholesterol-lowering properties, which may contribute to the healing properties of *C. rotundifolia*.[29,30] A white powder (35 mg) (compound **2**) with a molecular ion peak at *m/z* 414 corresponding to $C_{29}H_{50}O$ was obtained from the root of *C. rotundifolia*. The ^1H and ^{13}C NMR spectral data confirmed the compound to be β-sitosterol, previously isolated from *C. rotundifolia* and other plants in the genus, such as *C. quadrangularis* and *C. sicyoides*.[25,26] β-sitosterol is known for its antibacterial activity, which may also justify the use of *C. rotundifolia* in wound healing.[31] Compound **3**, isolated as a dark green solid (120 mg) from the DCM extract of the leaves, with molecular ion peak at *m/z* 871 [M^+] corresponding to $C_5H_{74}N_4O_5$, was identified as pheophytin *a*.[32] This is the first isolation of pheophytin *a* from *C. rotundifolia*, but this compound was isolated from *C. assamica* and *C. quadrangularis* in the genus *Cissus*.[33,34] Pheophytin *a* may contribute towards the antioxidant and anti-inflammatory properties of the crude extract of *C. rotundifolia*.[32,34]

Elemental composition

We evaluated the concentrations of 11 metals in *C. rotundifolia* (leaves and stems) and the corresponding soil samples collected from eight different sampling sites in KwaZulu-Natal (Figure 1). Except for sites 4

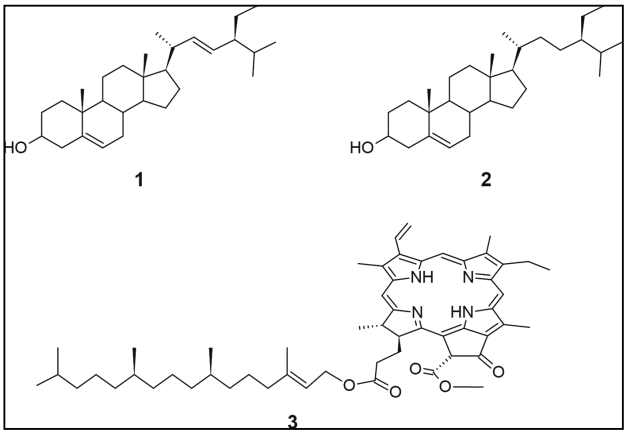


Figure 2: Compounds isolated from *Cissus rotundifolia.*

(Richards Bay) and 5 (Cato Ridge), higher concentrations of Mn were detected in the plant than in the soil (Table 2, BAF>1). This could mean that the plant can take up the required concentrations from the soil to meet its metabolic needs. Alternatively, higher concentrations in aerial plant parts could be due to atmospheric deposition from contaminated sites. *C. rotundifolia* leaves from a mountain in Yemen had an average Mn concentration of 18.9 mg/kg, which is relatively low compared with that in our study (35.1—74.7 mg/kg).[17] However, mountainous regions are considered pristine due to their distance away from anthropogenic sources, which could explain the disparity with our concentrations. A study of the impact of atmospheric deposition on particulate Mn distribution in northwestern Mediterranean surface water[35] showed variations between atmospheric-depositional and settling fluxes based on atmospheric input from high/low productivity seasons. This was also related to more suspended particulate Mn and particulate residence times in standing crops.[35] This study showed that higher concentrations in aerial plant parts could be due to atmospheric deposition, as seen in our study. Another study investigated Mn migration from soil to Scots pine (*Pinus sylvestris* L.) shoots and needles contamination by a steel plant's emissions.[36] Control and contaminated soil concentrations were comparable, but needles at contaminated sites had concentrations of Mn that were 2.5 times higher than those at control sites. This could indicate an increase in metal mobility at contaminated areas with a shift in migration pattern towards Mn accumulation in the aerial parts of the plant, as seen in our study.

For Zn, plant concentrations were higher than soil concentrations at all sites (Table 2). *C. rotundifolia* leaves from a mountain in Yemen[17] had an average Zn concentration of 2.2 mg/kg, and fruits from Shada Mountain in southwest Saudi Arabia[18] had an average Zn concentration of 9.2 mg/kg – these concentrations are relatively low compared to those in our study (19.9—73.1 mg/kg). A study focusing on atmospheric deposition of Zn to forest vegetation in the Tennessee Valley showed

Table 2: Concentration of microelements in *Cissus rotundifolia* (CR) leaves (L), stems (S), and soil from the eight sites (mean (SD), *n*=3), and bioaccumulation factors

Site	Concentration (mg/kg)			Bioaccumulation factor	
	CR Leaves	CR Stems	Soil	[CRL]/[soil]	[CRS]/[soil]
Co					
S1	0.22 (0.18)	0.21 (0.20)	0.1 (0.26)	2.4	2.4
S2	0.03 (0.30)	1.11 (1.67)	0.6 (0.20)	0.1	1.8
S3	0.07 (0.51)	ND	0.6 (0.39)	0.1	ND
S4	ND	0.44 (0.25)	0.7 (0.35)	ND	0.6
S5	ND	0.20 (0.05)	1.1 (0.41)	ND	0.2
S6	0.48 (0.21)	0.41 (0.02)	0.6 (0.65)	0.8	0.7
S7	0.23 (0.02)	0.11 (0.02)	0.3 (0.19)	0.9	0.4
S8	0.01 (0.04)	0.10 (0.04)	0.5 (0.09)	0.0	0.2
Cr					
S1	0.17 (1.05)	1.66 (0.96)	3.0 (0.6)	0.1	0.6
S2	0.82 (0.27)	ND	4.0 (3.4)	0.2	ND
S3	1.01 (0.37)	ND	6.6 (2.94)	0.2	ND
S4	ND	ND	9.7 (0.50)	0.0	ND
S5	ND	ND	9.0 (4.35)	ND	ND
S6	3.42 (0.88)	0.24 (0.36)	4.2 (1.75)	0.8	ND
S7	ND	ND	2.3 (0.75)	ND	ND
S8	ND	ND	6.2 (1.05)	ND	ND
Cu					
S1	10.34 (1.49)	5.9 (0.70)	3.6 (0.33)	2.9	1.7
S2	5.93 (0.08)	4.66 (0.01)	1.8 (0.98)	3.4	2.6
S3	4.29 (0.29)	8.91 (0.07)	3.1 (0.92)	1.4	2.9
S4	4.8 (0.74)	9.2 (1.25)	13.6 (2.45)	0.4	0.7
S5	4.1 (0.92)	8.48 (0.45)	14.3 (0.04)	0.3	0.6
S6	5.52 (0.48)	5.13 (1.09)	1.6 (0.04)	3.4	3.2
S7	5.35 (0.17)	1.96 (0.26)	0.9 (0.31)	5.9	2.2
S8	5.15 (2.1)	3.91 (0.48)	1.8 (0.55)	2.8	2.1

Site	Concentration (mg/kg)			Bioaccumulation factor	
	CR Leaves	CR Stems	Soil	[CRL]/[soil]	[CRS]/[soil]
Fe					
S1	421.5 (86.6)	1565 (431)	1560 (197.9)	0.3	1.0
S2	309.5 (17.7)	119.5 (17.7)	2595 (1619)	0.1	0.0
S3	635.5 (177)	227.5 (13.4)	3320 (876.5)	0.2	0.1
S4	379 (52.4)	118.32 (60.4)	3990 (2022)	0.1	0.0
S5	253.5 (33.2)	124.1 (39.5)	5935 (3288)	0.0	0.0
S6	1032 (53.8)	358.5 (0.71)	2410 (1018)	0.4	0.1
S7	99.8 (10.2)	116.9 (48.2)	1505 (304)	0.1	0.1
S8	178.5 (14.9)	77.35 (13.9)	2855 (261.7)	0.1	0.0
Mn					
S1	74.68 (0.61)	48.97 (2.94)	36 (19.1)	2.1	1.4
S2	51.32 (3.85)	61.43 (0.52)	33 (0.12)	1.5	1.8
S3	49.59 (4.22)	85.95 (8.6)	41 (10.7)	1.2	2.1
S4	35.1 (3.8)	39.55 (10.1)	65 (2.02)	0.5	0.6
S5	48.18 (13.7)	37.97 (2.6)	71 (15.4)	0.7	0.5
S6	65.12 (3.92)	102.4 (6.4)	41 (1.84)	1.6	2.5
S7	44.56 (3.35)	22.24 (3.67)	23 (2.18)	1.9	1.0
S8	36.155 (18.8)	43.11 (4.56)	38 (0.52)	1.0	1.1
Zn					
S1	73.81 (0.26)	35.44 (1.22)	32 (10.9)	2.3	1.1
S2	32.5 (17.8)	42.17 (3.35)	11 (2.56)	2.9	3.8
S3	39.23 (8.15)	37.53 (2.2)	23 (3.7)	1.7	1.6
S4	23.64 (5.75)	38.22 (14.9)	25 (4.1)	1.0	1.5
S5	28.96 (10.7)	52.48 (1.09)	29 (0.93)	1.0	1.8
S6	19.87 (0.19)	31.12 (1.98)	17 (6.5)	1.2	1.8
S7	34.12 (21.2)	23.79 (4.73)	11 (0.39)	3.2	2.2
S8	32.12 (15.6)	23.57 (2.95)	21.1 (1.75)	1.6	1.1

S1, Margate; S2, Ballito; S3, Richards Bay; S4, Shongweni; S5, Cato Ridge; S6, St Lucia; S7, Amanzimtoti; S8, Umgababa

that atmospheric deposition during the growing season contributed to relatively high concentrations due to metal particles deposited on dry leaf surfaces.[37] Another study showed physiological control of trace metal uptake with plant root affinity for Zn declining with increased Zn concentration in the soil solution.[38] High Zn concentrations in our study, similar to Mn, could be due to atmospheric deposition or enhanced uptake mechanisms at the root zone to meet physiological needs.

Total soil Fe ranged from 1505 mg/kg to 5935 mg/kg with a range of 99.8–1032 mg/kg in the leaves and 77.4—1565 mg/kg in the stems (Table 2). Site S5 (Shongweni), followed by S4 (Richards Bay), exhibited the highest soil Fe concentrations. The concentration of Fe in the plants from all sites was lower than the maximum permissible limit for Fe in plants (1000 mg/kg), which confirms the plant's ability to limit uptake, especially at elevated soil concentrations.[39] The Fe concentrations in this study were higher than those in the leaves from the plant obtained from a mountain in Yemen (177.3 mg/kg) and from fruits collected from Shada Mountain in Saudi Arabia (44.7 mg/kg).[17,18]

Higher levels of Cu were detected in the plant than in the soil, except at the environmentally polluted sites (S4 – Shongweni and S5 – Cato Ridge). Cu concentrations in the leaves and stems were mainly below the WHO permissible limit of 10 mg/kg for plants.[40] Cu concentrations obtained in this study (average of 5.1 mg/kg) were comparable to those reported in fruits (3.3 mg/kg), but relatively lower than those reported in leaves (11.6 mg/kg).[17,18] The low soil concentrations at most sites indicate Cu deficiencies, and the high BAFs provide evidence of the plants physiological control according to metabolic requirements, with plant root affinity for Cu rising when low soil Cu provides insufficient supply of the metal.

Generally, low concentrations of Co and Cr were detected in plant and soil samples from all sites with BAFs<1, indicating that the uptake of

these elements was restricted. If detected, Cr in plants ranged from 0.17 mg/kg (S1 – Margate) to 3.42 mg/kg (S6 – St Lucia) and 2.3–9.7 mg/kg in soil. Cr mainly exists in the innocuous trivalent form and noxious hexavalent form. An intake of 2– 5 g of Cr(VI) may be fatal. Total Cr in plant leaves at S6 (St Lucia) is within this range; therefore, consumers from this region should exercise caution as the hexavalent form is mobile and bioavailable under oxidising conditions and may contribute significantly to total concentration.[41] In this study, the average concentrations of the micronutrients in the plant were in decreasing order of Fe>Mn>Se>Zn>Cu>Cr>Ni>Co.

Arsenic concentrations in the soil ranged from 0.87 mg/kg (Cato Ridge) to 2.87 mg/kg (Umgababa), which exceeded the WHO maximum permissible limit of 0.5 mg/kg for As in agricultural soils.[42] The plant did not tend to accumulate As with BAFs<1 at most sites (Table 3). Arsenic in the plant leaves ranged from 0.44 mg/kg to 2.17 mg/kg, which is higher than the WHO/FAO maximum permissible limit of 0.1 mg/kg for plants grown on industrial soils.[43] A study reported the average As concentration in consumed vegetables to be 0.27 mg/kg for an industrial site and 0.05 mg/kg for a non-industrial site.[44] In our study, the average As in plant leaves was 1.18 mg/kg, which is higher than those previously reported. For Cd, the plant exhibited concentrations ranging from 0.57 mg/kg to 0.71 mg/kg in the leaves, which surpasses the maximum permissible limit of 0.2 mg/g for Cd in vegetables.[45] The concentration of Cd in the plant at all sites was comparable to that in soil (BAF almost 1), highlighting the plant's potential as an indicator for Cd pollution (Table 3).

Pb was generally higher in the soil than in the plant, with concentrations in the soil at S5 (Cato Ridge, 8.43 mg/kg) being extremely high relative to those at the other sites (Table 3). However, this did not influence the Pb concentration in the plant, which again indicates the plant's ability to protect itself by excluding toxic metals. The high soil concentrations at Cato Ridge could be due to vehicular emissions from a truck stop and railway line close to the sample collection point. The concentrations of Pb in the leaves from S1 (Margate) and S2 (Ballito) were 1.16 mg/kg and 1.06 mg/kg, respectively. According to the South African Department of Health, these values are above the maximum permissible limit of Pb of 0.3 mg/kg in leafy vegetables.[46] The data obtained for the toxic elements (As, Cd and Pb) suggest that *C. rotundifolia* samples collected from sites near roads and pollution sources are not suitable for human consumption. Although the plant limited uptake of these toxic metals, the concentrations obtained at some sites exceeded the maximum permissible limits imposed by authorities for safety for human consumption. The communities in these areas should exercise caution when consuming this plant to avoid toxic health effects.

Table 3: Concentrations of toxic elements (As, Cd and Pb) in *Cissus rotundifolia* leaves and surrounding soil from each site (mean (SD), n=3), estimated daily intake (EDI), target hazard quotient (THQ), carcinogenic risk (CR) and bioaccumulation factor (BAF)

Site	[Leaves] (mg/kg)	EDI	THQ	CR	[Soil] (mg/kg)	BAF [Leaves]/[Soil]
As						
S1	1.75 (0.39)	0.002	5.152	2.32E-03	1.88 (0.14)	0.9
S2	1.01 (0.66)	0.001	2.977	1.34E-03	2.67 (0.03)	0.4
S3	2.17 (0.07)	0.002	6.407	2.88E-03	1.605 (1.055)	1.4
S4	0.52 (0.82)	0.000	1.541	6.94E-04	1.87 (1.15)	0.3
S5	0.44 (0.125)	0.000	1.303	5.86E-04	0.87 (1.13)	0.5
S6	1.88 (0.925)	0.002	5.536	2.49E-03	2.835 (0.475)	0.7
S8	0.52 (0.635)	0.000	1.527	6.87E-04	2.865 (1.15)	0.2
Cd						
S1	0.64 (0.09)	0.001	0.562	3.54E-03	0.555 (0.065)	1.1
S2	0.63 (0.08)	0.001	0.554	3.49E-03	0.66 (0.04)	0.9
S3	0.68 (0.1)	0.001	0.602	3.79E-03	0.54 (0.03)	1.3
S4	0.68 (0.65)	0.001	0.598	3.77E-03	0.73 (0.05)	0.9
S5	0.64 (0.155)	0.001	0.567	3.57E-03	0.645 (0.035)	1.0
S6	0.63 (0.125)	0.001	0.558	3.52E-03	0.64 (0.63)	1.0
S7	0.65 (0.02)	0.001	0.571	3.60E-03	0.63 (0.085)	1.0
S8	0.75 (0.005)	0.001	0.660	4.16E-03	0.575 (0.06)	1.3
Pb						
S1	1.16 (0.015)	0.001	0.257	8.73E-06	0.975 (0.37)	1.2
S2	1.06 (0.99)	0.001	0.235	7.99E-06	0.34 (0.735)	3.1
S5	0.1 (0.002)	0.000	0.021	7.21E-07	8.415 (2.005)	0.0
S6	0.23 (0.085)	0.000	0.051	1.72E-06	0.96 (0.51)	0.2
S7	0.01 (0.145)	0.000	0.002	5.38E-08	0.61 (0.055)	0.0

S1, Margate; S2, Ballito; S3, Richards Bay; S4, Shongweni; S5, Cato Ridge; S6, St Lucia; S7, Amanzimtoti; S8, Umgababa

Table 4: The estimated contribution of essential elements in *Cissus rotundifolia* (leaves and stems) to dietary intake

DRI (mg/day)			Average concentration (mg/10 g dry mass)		Estimated contribution to RDA (10 g/day) %	
Element	RDA	UL	Stems	Leaves	Stems	Leaves
Ca	1300	2500	366.6	405.5	28.2	31.1
Cu	0.9	8	0.1	0.2	13.3	24.5
Fe	18	45	4.6	7.3	25.5	40.7
Mg	320	350	50.4	91.7	15.8	28.7
Mn	2.3	9	0.4	0.5	18.7	22.7
Se	0.055	0.4	0.4	0.7	669	1251
Zn	11	34	0.8	0.3	6.9	2.7

DRI, daily recommended intake; RDA, recommended dietary allowance; UL, tolerable upper intake level; ND, not detectable

The I_{geo} values obtained for the soil collected from the eight sites indicate that the soil was not contaminated with heavy metals. However, the health risk assessment showed possible non-carcinogenic or carcinogenic risks due to exposure to As and Cd. The THQs for the leaves were in the order of As>Cd>Pb and ranged from 1.30 to 6.41 for As, 0.55 to 0.66 for Cd, and 0.002 to 0.257 for Pb (Table 3). A THQ>1 for arsenic at all sites suggests possible non-carcinogenic risk to human health due to As exposure. The carcinogenic risk for As and Cd through consumption of the leaves was 1.57×10^{-3} and 3.68×10^{-3}, respectively. These values exceed the safe limit for cancer risk (1×10^{-4}) recommended by the USEPA[21], signifying the possibility of developing cancer over a lifetime if the plant is consumed frequently[23]. To the best of our knowledge, this is the first study to report on the metal toxicity of *C. rotundifolia*.

Contribution of C. rotundifolia to the human diet

The elemental concentrations of the leaves and stems were compared to the dietary reference intakes (DRIs) for most individuals to estimate the contribution of the consumption of 10 g of *C. rotundifolia* leaves and stems (based on the dry mass) to the diet (Table 4). For the elements to be considered good contributors to the diet, their concentrations should not exceed the tolerable upper intake levels (ULs).[33] Consumption of 10 g of leaves and stems would contribute 25.5% and 40.7% towards the RDA for Fe, respectively. Fe is an essential metal found in haemoglobin, and its deficiency in humans can lead to ailments such as anaemia, cancer, and heart disease.[47] Local people consuming this plant (especially those diagnosed with anaemia) can supplement their Fe intake, similar to consuming fruits and vegetables such as apricots, olives, and beetroot.[48]

The plant contributes >669% towards the RDA for Se, and the level in the leaves exceeded the UL. However, the concentrations of Se in the stems (0.37 mg/10 g) and leaves (0.69 mg/10 g) were similar to that of Brazil nuts (0.36 mg/10 g)[49], which are known to be a rich source of Se[49]. Se has antioxidant properties and has been reported to lower the risk of certain cancers, such as breast, lung, colon, and prostate cancers.[50] However, too much Se can have detrimental health effects and cause depression.[50] The other essential elements studied contribute adequately to their RDAs and may supplement the diet.

Conclusion

The analysis of *C. rotundifolia* revealed that the plant is rich in phytosterols and essential nutrients. The plant–soil system showed control by the plant to meet physiological needs. Bioaccumulation factors for Cd indicated the plant's potential as an indicator and biological monitor for Cd toxicity. Carcinogenic and non-carcinogenic risk assessments for As, Cd and Pb in the plant signified the possibility of developing cancer over time if consumed from exposed sites. From this study, it can be concluded that *C. rotundifolia* from non-polluted sites or sites away from roads, landfills, and other pollution sources is safe for human consumption. The findings from this study show the synergies when consuming medicinal plants and provide evidence for *C. rotundifolia* as a nutraceutical. Its consumption for therapeutic benefit would not pose metal toxicity risks, and for nutritional value would give the added advantage of ingesting antioxidants.

Acknowledgements

We thank the University of KwaZulu-Natal and the National Research Foundation (NRF) of South Africa for funding this research.

Competing interests

We have no competing interests to declare.

Authors' contributions

B.P.M: Conceptualisation, methodology, data analysis, writing – the initial draft, writing – revisions, and student supervision. C.Z.B.: Methodology, writing – the initial draft, data collection, and data analysis. R.M.: Conceptualisation, methodology, data analysis, student supervision, project management, writing – revisions and funding acquisition.

References

1. Atanasov AG, Zotchev SB, Dirsch VM, Supuran CT. Natural products in drug discovery: Advances and opportunities. Nat Rev Drug Discov. 2021;20(3):200–216. https://doi.org/10.1038/s41573-020-00114-z
2. Hassan A, Rahman S, Deeba F, Mahmud S. Antimicrobial activity of some plant extracts having hepatoprotective effects. J Med Plants Res. 2009;3(1):20–23. https://academicjournals.org/journal/JMPR/article-full-text-pdf/123755814602
3. Obiajunwa E, Adebajo A, Omobuwajo O. Essential and trace element contents of some Nigerian medicinal plants. J Radioanal Nucl Chem. 2002;252(3):473–476. https://doi.org/10.1023/a:1015838300859
4. Zafar M, Khan MA, Ahmad M, Jan G, Sultana S, Ullah K, et al. Elemental analysis of some medicinal plants used in traditional medicine by atomic absorption spectrophotometer (AAS). J Med Plants Res. 2010;4(19):1987–1990. https://doi.org/10.5897/JMPR10.081
5. Ekinci N, Ekinci R, Polat R, Budak G. Analysis of trace elements in medicinal plants with energy dispersive X-ray fluorescence. J Radioanal Nucl Chem. 2004;260(1):127–131. https://doi.org/10.1023/B:JRNC.0000027071.72742.ee
6. Chojnacka K, Chojnacki A, Gorecka H, Górecki H. Bioavailability of heavy metals from polluted soils to plants. Sci Total Environ. 2005;337(1–3):175–182. https://doi.org/10.1016/j.scitotenv.2004.06.009
7. Tuhy Ł, Samoraj M, Michalak I, Chojnacka K. The application of biosorption for production of micronutrient fertilizers based on waste biomass. Appl Biochem Biotechnol. 2014;174(4):1376–1392. https://doi.org/10.1007/s12010-014-1074-0
8. Rajkumar M, Prasad MNV, Swaminathan S, Freitas H. Climate change driven plant–metal–microbe interactions. Environ Int. 2013;53:74–86. https://doi.org/10.1016/j.envint.2012.12.009

9. Alqahtani J, Formisano C, Chianese G, Luciano P, Stornaiuolo M, Perveen S, et al. Glycosylated phenols and an unprecedented diacid from the Saudi plant *Cissus rotundifolia*. J Nat Prod. 2020;83(11):3298–3304. https://doi.org/10.1021/acs.jnatprod.0c00597

10. Onyechi UA, Ibeanu VN. Effects of diets containing *Cissus rotundifolia* flour on lipid profile of rats and postprandial glucose levels of normoglycemic human adults. Afr J Biotechnol. 2016;15(14):557–564. https://doi.org/10.5897/AJB2015.14677

11. Al-Fatimi M, Wurster M, Schröder G, Lindequist U. Antioxidant, antimicrobial and cytotoxic activities of selected medicinal plants from Yemen. J Ethnopharmacol. 2007;111(3):657–666. https://doi.org/10.1016/j.jep.2007.01.018

12. Alshawsh MA, Mothana RA, Al-Shamahy HA, Alsllami SF, Lindequist U. Assessment of antimalarial activity against *plasmodium falciparum* and phytochemical screening of some Yemeni medicinal plants. Evid Based Complementary Altern Med. 2009;6(4):453–456. https://doi.org/10.1093/ecam/nem148

13. Mziray AM, Maina CI, Kaingu CK. Anti-fertility properties of *Cissus rotundifolia* (Forssk.) Vahl. Extract using female Wistar rats. Discovery Phytomedicine. 2020;7(2):58–64. https://doi.org/10.15562/phytomedicine.2020.120

14. Said A, Aboutabl EA, Melek FR, Abdel Jaleel Raheem Abdel Jaleel G, Raslan M. Phytoconstituents profiling of *Cissus rotundifolia* (Forssk.) Vahl. by HPLC-MS/MS, and evaluation of its free radical scavenging activity (DPPH) and cytotoxicity. Trends Phytochem Res 2018;2(2):65–74.

15. AL-Bukhaiti WQ, Noman A, Mahdi AA, Ali AH, Abed SM, Rashed MM, et al. Profiling of phenolic compounds and antioxidant activities of *Cissus rotundifolia* (Forssk.) as influenced by ultrasonic-assisted extraction conditions. J Food Meas Charact. 2019;13(1):634–647. https://doi.org/10.1007/s11694-018-9976-0

16. Korish M. Nutritional evaluation of wild plant *Cissus rotundifolia*. Ital J Food Sci. 2016;28(1):43–49. https://doi.org/10.14674/1120-1770/ijfs.v456

17. Al-Bukhaiti WQ, Noman A, Mahdi AA, Abed SM, Ali AH, Mohamed JK, et al. Proximate composition, nutritional evaluation and functional properties of a promising food: Arabian wax *Cissus* (*Cissus rotundifolia* Forssk) leaves. J Food Sci Technol. 2019;56(11):4844–4854. https://doi.org/10.1007/s13197-019-03947-8

18. Hegazy AK, Mohamed AA, Ali SI, Alghamdi NM, Abdel-Rahman AM, Al-Sobeai S. Chemical ingredients and antioxidant activities of underutilized wild fruits. Heliyon. 2019;5(6), e01874. https://doi.org/10.1016/j.heliyon.2019.e01874

19. Oti WO. Bioaccumulation factors and pollution indices of heavy metals in selected fruits and vegetables from a derelict mine and their associated health implications. Int J Environ Sustain. 2015;4(1).

20. Muller G. Index of geoaccumulation in sediments of the Rhine River. Geojournal. 1969;2:108–118.

21. US Environmental Protection Agency. USEPA Regional Screening Levels (RSLs) – Generic tables [webpage on the Internet]. No date [updated May 2023; cited 2023 Jun 05. Available from: https://www.epa.gov/risk/regional-screening-levels-rsls-generic-tables

22. Bo S, Mei L, Tongbin C, Zheng Y, Yunfeng X, Xiaoyan L, et al. Assessing the health risk of heavy metals in vegetables to the general population in Beijing, China. J Environ Sci. 2009;21(12):1702–1709. https://doi.org/10.1016/S1001-0742(08)62476-6

23. Javed M, Usmani N. Accumulation of heavy metals and human health risk assessment via the consumption of freshwater fish Mastacembelus armatus inhabiting, thermal power plant effluent loaded canal. SpringerPlus. 2016;5(1), Art. #776. https://doi.org/10.1186/s40064-016-2471-3

24. Kortei NK, Koryo-Dabrah A, Akonor PT, Manaphraim NYB, Ayim-Akonor M, Boadi NO, et al. Potential health risk assessment of toxic metals contamination in clay eaten as pica (geophagia) among pregnant women of Ho in the Volta Region of Ghana. BMC Pregnancy Childbirth. 2020;20(1), Art. #160. https://doi.org/10.1186/s12884-020-02857-4

25. Chaturvedula VSP, Prakash I. Isolation of stigmasterol and β-sitosterol from the dichloromethane extract of *Rubus suavissimus*. Int Curr Pharm J. 2012;1(9):239–242. https://doi.org/10.3329/icpj.v1i9.11613

26. Sharma N, Nathawat R, Gour K, Patni V. Establishment of callus tissue and effect of growth regulators on enhanced sterol production in *Cissus quadrangularis* L. Int J Pharmacol. 2011;7(5):653–658. https://doi.org/10.3923/ijp.2011.653.658

27. Saifah E, Vaisiriroj V, Kelley CJ, Higuchi Y. Constituents of the roots of *Cissus rheifolia*. J Nat Prod. 1987;50(2):328. https://doi.org/10.1021/np50050a057

28. Atiku I, Musa A, Sule M. Isolation of stigmasterol from methanolic extract of *Cissus cornifolia* Baker (Planch). Niger J Pharm Sci. 2013;12(1).

29. Pierre LL, Moses MN. Isolation and characterisation of stigmasterol and β-sitosterol from *Odontonema strictum* (acanthaceae). J Innov Pharm Biol Sci. 2015;2(1):88–95.

30. Kaur N, Chaudhary J, Jain A, Kishore L. Stigmasterol: A comprehensive review. Int J Pharm Sci Res. 2011;2(9):2259.

31. Beltrame FL, Pessini GL, Doro DL, Dias Filho BP, Bazotte RB, Cortez DAG. Evaluation of the antidiabetic and antibacterial activity of *Cissus sicyoides*. Braz Arch Biol Technol. 2002;45:21–25. https://doi.org/10.1590/S1516-89132002000100004

32. Ogunlaja OO, Moodley R, Baijnath H, Jonnalagadda SB. Chemical constituents and in vitro antioxidant activity of crude extracts and compounds from leaves and stem bark of *Ficus Burtt-Davyi*. Acta Pol Pharm. 2016;73(6):1593–1600.

33. Chan Y-Y, Wang C-Y, Hwang T-L, Juang S-H, Hung H-Y, Kuo P-C, et al. The constituents of the stems of *Cissus assamica* and their bioactivities. Molecules. 2018;23(11):2799. https://doi.org/10.3390/molecules23112799

34. Pandey S, Parmar S, Shukla MS, Sharma VS, Dwivedi A, Pandey A, et al. Phytochemical and pharmacological investigation of *Cissus quadrangularis* L. Herb Med J. 2022;7(2).

35. Davies JE, Buat-Ménard P. Impact of atmospheric deposition on particulate manganese and aluminium distribution in northwestern Mediterranean surface water. Glob Planet Change. 1990;3(1):35–45. https://doi.org/10.1016/0921-8181(90)90054-G

36. Zaitsev GA, Dubrovina OA, Shainurov RI. Iron and manganese migration in "soil–plant" system in Scots pine stands in conditions of contamination by the steel plant's emissions. Sci Rep. 2020;10(1), Art. #11025. https://doi.org/10.1038/s41598-020-68114-y

37. Lindberg SE, Harriss RC, Turner RR. Atmospheric deposition of metals to forest vegetation. Science. 1982;215(4540):1609–1611. https://doi.org/10.1126/science.215.4540.1609

38. Moodley R, Koorbanally N, Jonnalagadda SB. Elemental composition and fatty acid profile of the edible fruits of Amatungula (*Carissa macrocarpa*) and impact of soil quality on chemical characteristics. Anal Chim Acta. 2012;730:33–41. https://doi.org/10.1016/j.aca.2011.11.066

39. Hiroki M. Effects of heavy metal contamination on soil microbial population. Soil Sci Plant Nutr. 1992;38(1):141–147. https://doi.org/10.1080/00380768.1992.10416961

40. Ogunlaja OO, Moodley R, Baijnath H, Jonnalagadda SB. Elemental distribution and health risk assessment of the edible fruits of two ficus species, *Ficus sycomorus* L. and *Ficus burtt-davyi* Hutch. Biol Trace Elem Res. 2020;198(1):303–314. https://doi.org/10.1007/s12011-020-02048-4

41. Kapoor RT, Bani Mfarrej MF, Alam P, Rinklebe J, Ahmad P. Accumulation of chromium in plants and its repercussion in animals and humans. Environ Pollut. 2022;301, Art. #119044. https://doi.org/10.1016/j.envpol.2022.119044

42. Tasrina R, Rowshon A, Mustafizur A, Rafiqul I, Ali M. Heavy metals contamination in vegetables and its growing soil. J Environ Anal Chem. 2015;2(142):2–6. http:doi.org/10.4172/2380-2391.1000142

43. Secretariat, Codex Alimentarius Commission, Programme JFWFS. Report of the fifth session of the Codex Committee on Contaminants in foods. Rome: Codex Alimentarius Commission; 2011.

44. Haque MM, Niloy NM, Khirul MA, Alam MF, Tareq SM. Appraisal of probabilistic human health risks of heavy metals in vegetables from industrial, non-industrial and arsenic contaminated areas of Bangladesh. Heliyon. 2021;7(2), e06309. https://doi.org/10.1016/j.heliyon.2021.e06309

45. Xiao R, Wang S, Li R, Wang JJ, Zhang Z. Soil heavy metal contamination and health risks associated with artisanal gold mining in Tongguan, Shaanxi, China. Ecotoxicol Environ Saf. 2017;141:17–24. https://doi.org/10.1016/j.ecoenv.2017.03.002

46. Gounden T, Moodley R, Jonnalagadda SB. Elemental analysis and nutritional value of edible *Trifolium* (clover) species. J Environ Sci Health B. 2018;53(8):487–492. https://doi.org/10.1080/03601234.2018.1462923

47. Camaschella C. Iron deficiency. Blood. 2019;133(1):30–39. https://doi.org/10.1182/blood-2018-05-815944

48. Gounden T. Elemental distribution in selected edible *Trifolium* species (clover) and the impact of soil quality on the chemical characteristics of *Trifolium dubium* [MSc thesis]. Durban: University of KwaZulu-Natal; 2017. http://hdl.handle.net/10413/15691

49. Moodley R, Kindness A, Jonnalagadda SB. Elemental composition and chemical characteristics of five edible nuts (almond, Brazil, pecan, macadamia and walnut) consumed in southern Africa. J Environ Sci Health B. 2007;42(5):585–591. https://doi.org/10.1080/03601230701391591

50. Tinggi U. Selenium: Its role as antioxidant in human health. Environ Health Prev Med. 2008;13(2):102–108. https://doi.org/10.1007/s12199-007-0019-4

AUTHORS:
Rima Beesoo[1,2,3,4]
Ranjeet Bhagooli[1,2,5,6,7,8]
Theeshan Bahorun[3]
Vidushi S. Neergheen[3,4]

AFFILIATIONS:
[1]Department of Biosciences and Ocean Studies, Faculty of Science, University of Mauritius, Réduit, Mauritius
[2]Pole of Research Excellence, Sustainable Marine Biodiversity Research Group, University of Mauritius, Réduit, Mauritius
[3]Biopharmaceutical Unit, Centre for Biomedical and Biomaterials Research, University of Mauritius, Réduit, Mauritius
[4]Department of Health Sciences, Faculty of Science, University of Mauritius, Réduit, Mauritius
[5]The Biodiversity and Environment Institute, Réduit, Mauritius
[6]The Society of Biology (Mauritius), Réduit, Mauritius
[7]Institute of Oceanography and Environment (INOS), University of Malaysia, Terengganu, Malaysia
[8]Department of Marine Science, Faculty of Fisheries and Marine Science, Diponegoro University, Jalan Prof. Soedarto SH, Kampus Tembalang, Semarang, Indonesia

CORRESPONDENCE TO:
Rima Beesoo

EMAIL:
r.beesoo@gmail.com

DATES:
Received: 08 Apr. 2022
Revised: 19 Jan. 2023
Accepted: 07 Feb. 2023
Published: 08 Aug. 2023

HOW TO CITE:
Beesoo R, Bhagooli R, Bahorun T, Neergheen VS. In vitro cytotoxic and apoptotic activity of the Mauritian marine sponge *Neopetrosia exigua*. S Afr J Sci. 2023;119(7/8), Art. #13745. https://doi.org/10.17159/sajs.2023/13745

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

DATA AVAILABILITY:
☐ Open data set
☐ All data included
☒ On request from author(s)
☐ Not available
☐ Not applicable

EDITORS:
Priscilla Baker
Amanda-Lee Manicum

KEYWORDS:
apoptosis, cytotoxic, marine sponge, *Neopetrosia exigua*, Mauritius

FUNDING:
Mauritius Research Council, Global Young Academy

In vitro cytotoxic and apoptotic activity of the Mauritian marine sponge *Neopetrosia exigua*

Marine sponges belonging to the genus *Neopetrosia* represent a quasi-inexhaustible source of novel cytotoxic compounds. Yet studies delineating their molecular mechanisms of action in cancer cells remain scarce. We investigated the cytotoxic and apoptosis inducing potential of the Mauritian marine sponge *Neopetrosia exigua* derived crude extract, hexane and ethyl acetate fraction. Their cytotoxic activity was screened against four cancer cell lines and two non-malignant cell lines via the Alamar Blue metabolic assay. The level of intracellular reactive oxygen species (ROS) production, endogenous antioxidant enzyme activity (catalase and superoxide dismutase) and mitochondrial membrane potential were determined. The ability of the active extract to induce apoptosis in cancer cells and modulate the expression levels of apoptotic markers (caspases and polyADP-ribose polymerase (*PARP*)) was further evaluated via western blot. The ethyl acetate fraction (NEEAF) displayed the highest inhibitory effect with an IC_{50} of 6.87 µg/mL against the liver hepatocellular carcinoma cell line (HepG2). Mechanistically, NEEAF induced morphological hallmarks characteristic of apoptosis, increased ROS production, decreased catalase and superoxide dismutase activity and mitochondrial membrane depolarisation in a concentration-dependent manner compared to the control ($p<0.05$). In addition, NEEAF induced the activation of caspase-9, -7, -3 and cleavage of PARP. Overall, this study provides biochemical evidence for oxidative stress-mediated cytotoxicity and apoptosis in HepG2 cells by NEEAF. Further in-depth investigations are needed to isolate the active constituents, which may potentially lead to the development of novel anticancer therapeutics.

Significance:

- Marine sponges represent an untapped goldmine of structurally unique compounds with interesting anticancer properties.
- This important initial investigative work will set the stage for more in-depth mechanistic studies and chemical characterisation of potentially novel bioactive compounds from the genus *Neopetrosia*.
- This work will also help to strengthen frameworks oriented towards the conservation of *Neopetrosia* species in the Western Indian Ocean region.

Introduction

Marine sponges represent promising bio-factories of structurally unique pharmacological compounds, many of which have been used as templates for the development of novel anticancer drugs.[1] An important feature of some sponge-derived compounds is their ability to induce apoptosis in cancer cells as their modus operandi.[2] Apoptosis is a physiological defence mechanism that involves the removal of unwanted, old, or injured cells, including cancer cells.[3] Over the last two decades, several crude extracts and marine natural products exhibiting anticancer activity through apoptosis have been isolated from marine sponges.[4-7] The latest marine-derived drug is Halaven® (eribulin mesylate), an analogue of the sponge metabolite halichondrin B derived from the sponge *Halichondria* sp. This drug was approved in 2010 for the treatment of advanced metastatic cancer. Growing evidence indicates that this drug targets signalling intermediates in apoptosis-inducing pathways, which appears to be associated with its effectiveness in modulating the process of carcinogenesis. In addition, more apoptosis-inducing compounds derived from sponge extracts such as hemiasterlins and spongistatins are currently at different stages of clinical trials.[2] Hence, the discovery of active extracts and/or novel compounds that target different cellular signaling cascades in apoptosis is paramount, especially if the process is selective to cancer cells.

Marine sponges belonging to the genus *Neopetrosia* have recently received increasing attention in natural product chemistry due to their novel bioactive metabolites with remarkable cytotoxic activities, namely neopetrocyclamines[8], neopetrosiquinones[9], araguspongines[10], and renieramycins[11]. However, to date, reports on their underlying molecular mechanisms of action remain unexplored. The genus *Neopetrosia* is still yielding bioactive extracts/compounds with promising bioactivities[12,13], and this was the impetus for the investigation of a member of this genus. Ecological and seasonal changes affect several abiotic factors such as salinity, pH, and temperature, as well as biotic factors including epifaunal diversity, all of which are actively involved in the biosynthesis of this sponge's natural products.[14] Thus, there is a high need to investigate the pharmacological activities of *Neopetrosia* species from different parts of the ocean, especially as local conditions could be an important factor in their bioactivity.

The Republic of Mauritius is a maritime country with a substantial potential for utilising marine organisms that are not yet fully utilised as sources of bioactive substances. In recognition of its unique marine environment with a vast pool of untapped marine biological resources, the vision of the Mauritian government is to enhance the island's great potential as an ocean state. In particular, the exclusive economic zone of Mauritius harbours a diverse assemblage of *Neopetrosia* species and thus offers a fruitful opportunity for the discovery of new bioactive agents.

However, their population density has significantly declined over the last few years (Bhagooli R 2019, oral communication, January 14personal communication, 2019). As part of our screening programme to search for bioactive sponge extracts that can spawn avenues in the discovery of new anticancer leads, herein we report on the in vitro cytotoxic and apoptototic activity of the crude extract and fractions (hexane, ethyl acetate and aqueous fractions) from the Mauritian sponge *Neopetrosia exigua*. This preliminary screening will subsequently set the stage for more in-depth mechanistic studies and chemical characterisation of active compounds. Furthermore, it will also strengthen frameworks oriented towards the conservation of *Neopetrosia* species in Mauritian waters. This goal to promote marine bio-discovery research and sustainable use of the ocean resources in Mauritius through aquaculture is in line with Sustainable Development Goal 14.[15]

Materials and methods

Animal material

The marine sponge *Neopetrosia exigua* (Figure 1) was collected whilst snorkelling off the Amber Island, Republic of Mauritius. The sponge was transferred to the lab under seawater, cleaned of debris and frozen at –80 °C until use. The sponge species was identified at the University of Mauritius and the taxonomy was confirmed using the World Porifera Database.[16]

Extract preparation

The sponge tissue was freeze dried, ground into powder (250 g) and extracted exhaustively with dichloromethane and methanol (1:1) for 48 h. The filtrate was flash evaporated under vacuum (LABORATA 4003,

Figure 1: Marine sponge *Neopetrosia exigua* collected from Mauritian waters.

Heidolph, Germany) to yield the crude extract (NECE), which was further partitioned to obtain the *n*-hexane (NEHF) and ethyl acetate (NEEAF) fractions.

Cell culture

The cancer and non-malignant cell lines were purchased from the American Type Culture Collection (USA). The cancer cell lines HepG2 (hepatocellular carcinoma), HeLa (cervical adenocarcinoma), and HCT116 (colorectal carcinoma), and the non-malignant cell lines RPE-1 (retinal epithelium) and MRC-5 (lung fibroblast) were grown in Dulbecco's Modified Eagle Medium (DMEM) while OE33 (oesophageal adenocarcinoma) was cultured in RPMI-1640. The media were supplemented with 10% foetal bovine serum, 2 mM l-glutamine and 1% Penstrep. Cells were maintained at 37 °C in an atmosphere of 5% CO_2 and 95% humidity. All the reagents used for cell culture were purchased from Gibco Life Technologies, UK.

Cytotoxic assay

The cells were seeded at a cell density of 1×10^4 cells/well in a 96-well plate. After 24 h, cells were treated with the marine sponge crude extract and fractions at different concentrations (0.78, 1.56, 3.13, 6.25, 12.5, 25, 50 μg/mL). The negative control received the vehicle dimethylsulfoxide (DMSO, 0.05%). Cytotoxicity was assayed at 24 h after extract treatment. A volume of 10 μL of alamarBlue dye (Thermo Fisher Scientific, USA) was added to each well, following which the plates were incubated at 37 °C for 4 h. The absorbance was measured at 570 nm and 600 nm in a multiplate reader (Biotek Synergy HT, USA). Etoposide (25–0.156 μg/mL) was used as a positive control. The cytotoxicity was expressed as IC_{50}, and presented as mean±SD from three independent experiments. The selectivity index (SI) was obtained by dividing the IC_{50} value for the non-malignant cell lines by the value of the IC_{50} for cancer cell lines. The SI value indicates the specificity of the sponge extracts for cancer cells. An SI value greater than 2 indicates that an extract/compound is more toxic to cancer cells than to normal cells.[17]

Morphological determination of apoptosis by Hoechst 33342 staining

The Hoechst 33342 dye (2′-[4-ethoxyphenyl]-5-[4-methyl-1-piperazinyl]-2,5′-bi-1H-benzimidazole trihydrochloride trihydrate) (Sigma Chemicals, UK) was used to analyse morphological signs of apoptosis in HepG2 cells treated with the sponge extract. The cells were seeded at 3×10^4/well in a 24-well plate and exposed to the sponge extract at different concentrations for 24 h. The cells were washed with 1X PBS (phosphate-buffered saline) thrice and fixed with 4% paraformaldehyde for 30 min before staining with Hoechst (1 μg/mL) for 15 min in the dark. The cellular apoptotic features were analysed under an inverted fluorescence microscope (EVOS fluorescence microscope, Life Technologies).

Measurement of intracellular reactive oxygen species production

The potential influence of reactive oxygen species (ROS) on the apoptosis induction in the treated cells was further examined. HepG2 cells were seeded at 1×10^4 cells/well in a dark 96-well plate and subsequently treated with the sponge extract (0.78–50 μg/mL) for 24 h. After incubation, the treated and untreated cells (control) were washed with 1X PBS and incubated with 100 μL of 25 μM dichloro-dihydro-fluorescein diacetate (DCFH-DA) (Sigma Chemicals, UK) for 45 min at 37 °C in the dark. Fluorescence was measured at excitation and emission wavelengths of 485 nm and 520 nm, respectively. Results were expressed as a percentage of the control.

Measurement of endogenous antioxidant enzyme activities

HepG2 cells were seeded at a density of 2×10^5 cells/well in a six-well plate and then incubated with the sponge extract (0.78–50 μg/mL). After 24 h, the cells were washed with 1X PBS, mixed with Complete™ lysis–M buffer reagent (Roche Diagnostics GmbH, Mannheim, Germany) and centrifuged at 10 000 g for 10 min at 4 °C. The protein content of the lysate supernatant was obtained by the Bradford test. The effect of the extract on superoxide dismutase and catalase antioxidant enzyme activity was assessed using commercial kits (BioVision Inc. Mountain View, CA, USA).[18]

Measurement of mitochondrial membrane potential

Changes in the mitochondrial membrane potential ($\Delta\Psi$m) were detected using the JC-1 dye (JC-1-5, 5′, 6, 6′-tetrachloro-1, 1′, 3, 3′ tetra ethylbenzimidazolcarbocyanine iodide) (Thermo Fisher Scientific, USA). HepG2 cells were seeded at 1×10^4 cells/well in a 96-well dark plate and treated with the sponge extract (0.78–50 μg/mL) for 24 h. After incubation, the cells were stained with 2 μM JC-1 at 37 °C for 45 min. Untreated cells and etoposide were used as negative and positive controls, respectively. Fluorescence intensity was monitored at 485 nm (excitation)/528 nm (emission) and 540 nm (excitation)/590 nm (emission). Changes in the ratio between the measurements compared to the control are indicative of changes in mitochondrial membrane potential.

Western blot

HepG2 cells were seeded into a six-well plate at 2×10^5 cells/well and treated with the sponge extract at 12.5 and 50 μg/mL for 24 h. The cells were lysed in Complete™ lysis–M buffer reagent (Roche Diagnostics GmbH, Mannheim, Germany), centrifuged at 10 000 × g for 10 min at 4 °C and the protein concentration of the lysates was determined using the Bradford test. The lysates were heated for 5 min at 100 °C. A total of 5 μg of protein extract was subjected to 5% stacking and 12% resolving sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE), then transferred onto a polyvinylidenedifluoride (PVDF) membrane, blocked with 100% methanol for 30 s, incubated with the primary antibodies (1:1000) for 24 h at 4 °C, and finally incubated with HRP markers-conjugated secondary goat-anti-rabbit antibodies (1:10000 dilution) in the dark at room temperature for 2 h. The blots were analysed for band densities using Image Lab software (Bio-Rad Laboratories, Inc). The relative protein expression was normalised to α-tubulin. The primary antibodies were rabbit polyclonal antibodies of caspase-9, caspase-7, caspase-3, PARP and α-tubulin (Cell Signaling Technology, USA). The relative protein expression was normalised to α-tubulin. The results were expressed as mean±SD.

Data analysis

Data were expressed as mean±SD (n=3) from three independent assays. Statistical analysis was performed using Prism, version 5.01, from GraphPad Software (USA). The results were analysed using a one-way analysis of variance followed by a least significant difference test; $p<0.05$ was considered to be statistically significant.

Results and discussion

The crude extract and fractions of the marine sponge *N. exigua* displayed dose-dependent cytotoxic effects in all of the tested cancer cell lines (Figure 2). The ethyl acetate (NEEAF) and hexane (NEHF) fractions recorded the highest growth inhibitory activity against HepG2 cells with IC_{50} values of 6.87±0.78 ug/mL and 8.17±0.99 ug/mL, respectively (Table 1). Interestingly, NEEAF was 4.54- and 4.92-fold less toxic with respect to the immortalised normal RPE-1 and MRC-5 cell lines, respectively. This indicates that NEEAF has the potential to be further developed into a less toxic therapeutic candidate against liver cancer. The cytotoxic effects elicited by NEEAF may be ascribed to the previously characterised bioactive constituents isolated from *N. exigua,* such as araguspongines[10], 5α, 8α-epidioxy sterols[10], exiguamine A[19], demethylxestopongin B[20] and papuamine[8].

We further investigated the mechanistic pathways underlying the cytotoxic activity of NEEAF in HepG2 cells. Morphological changes characteristic of apoptosis, such as membrane blebbing and formation of apoptotic bodies (only fluorescent microscopic images corresponding to 12.5 and 50 μg/mL are shown in Figure 3), were observed in Hoescht 33342 stained HepG2 cells treated with NEEAF while the control cells appeared

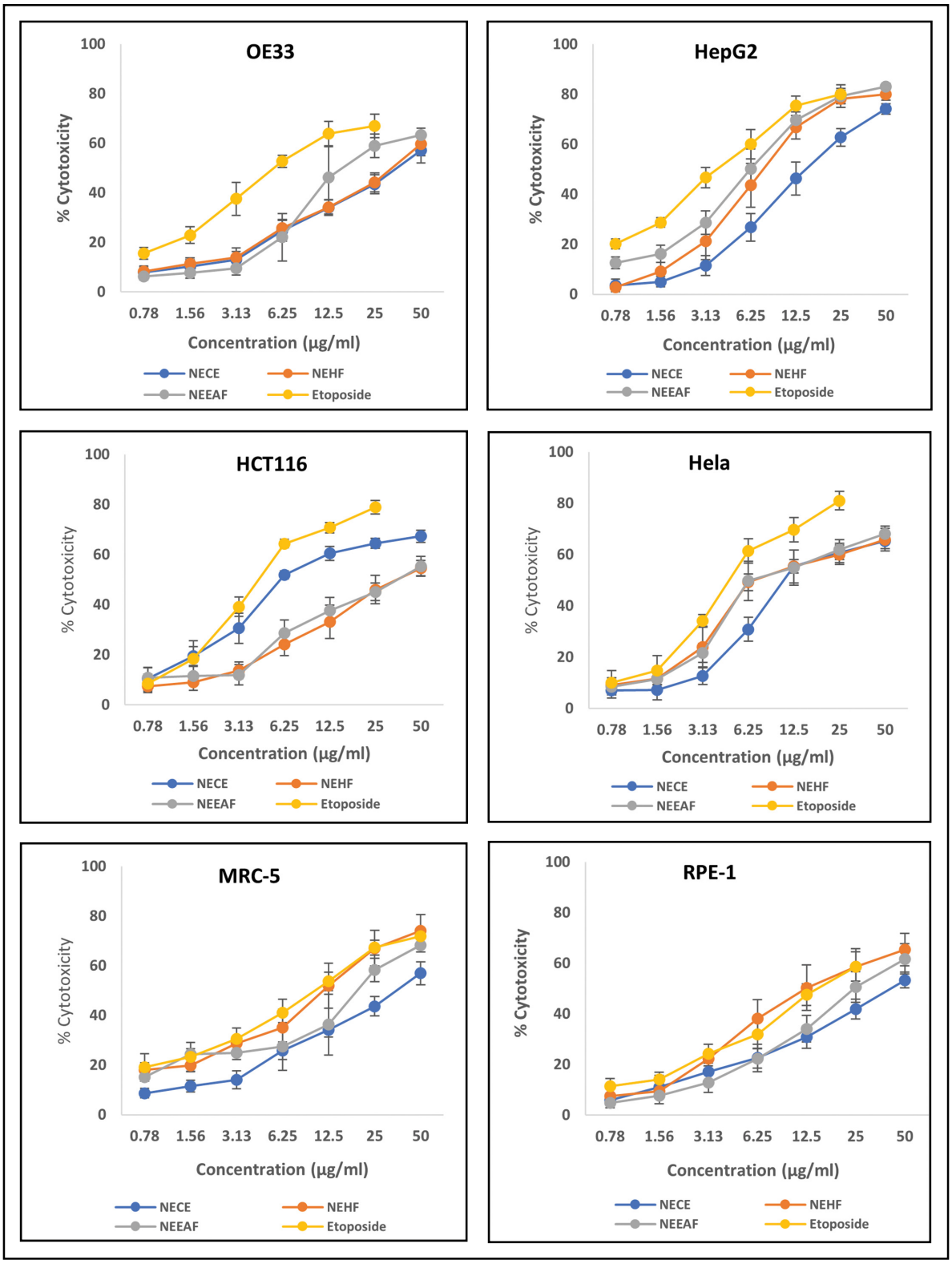


Figure 2: Dose-dependent cytotoxic activity of NECE, NEHF and NEEAF. The results are expressed as the percentage of the negative control (untreated cells). Cytotoxic activity of the extracts was compared to the reference drug, etoposide. Data are mean±SD (n=3) in three independent experiments.

normal with round and homogeneous nuclei. The implications of ROS in various physiological and pathological processes, including apoptosis, have been reported extensively.[21-23] There is compelling evidence indicating that when ROS generation overcomes cellular antioxidant defence mechanisms, it triggers apoptosis in cancer cells.[24-26] In this view, the level of ROS production was measured to probe its involvement in NEEAF-induced apoptosis. In general, a concentration-dependent ROS production from 122.00±7.41% to 298.26±58.30% was observed in HepG2 cells relative to control when treated with 0.78–50 µg/mL NEEAF (p<0.05) (Figure 4a).

Table 1: IC_{50} value of *Neopetrosia exigua* derived crude extract and fractions against selected human cancer and normal cell lines and their selectivity index (SI)

	Cytotoxic activity (IC_{50} μg/mL)					
	Cancer cell lines				Normal cell lines	
	OE33	HepG2	HeLa	HCT116	MRC-5	RPE-1
NECE	34.47±5.19^{a***} (1.32)/**1.18**	13.96±1.73^{a**} (3.27)/**2.92**	16.64±5.25^{a**} (2.74)/**2.45**	12.53±1.43^{b**} (2.28)/**2.96**	45.67±10.45^{a***}	40.79±3.97^{a***}
NEHF	31.67±5.33^{a***} (0.36)/**0.38**	8.17±0.99^{b*} (1.39)/**1.48**	12.69±2.62^{b**} (0.89)/**0.96**	35.74±6.70^{a***} (0.32)/**0.34**	11.33±3.12^{b}	12.13±2.09^{c}
NEEAF	18.43±3.71^{b**} (1.84)/**1.69**	6.87±0.78^{c*} (4.92)/**4.54**	12.29±3.31^{b**} (2.76)/**2.54**	34.61±8.56^{a***} (0.98)/**0.81**	33.83±7.41^{b***}	31.22±4.61^{b***}
Etoposide	6.74±0.66 (1.82)/**2.40**	4.49±0.72 (2.73)/**3.60**	5.47±0.66 (2.28)/**2.96**	4.93±0.35 (2.48)/**3.49**	12.24±2.77	16.17±3.93

NECE, crude extract; NEHF, hexane fraction; NEEAF, ethyl acetate fraction.

Data represent mean IC_{50} values ± SD of triplicate in three independent experiments. One-way ANOVA at 5% significance level; different superscripts on the mean values column wise represent significant differences among the sponge crude extract and fractions ($p<0.05$). Asterisks (*) represent significant differences between the extracts and etoposide (positive control), $^{*}p\leq0.05$, $^{**}p\leq0.01$, $^{***}p\leq0.001$. In bracket (): selectivity index (SI) determined as a ratio of the IC_{50} of the immortalised MRC-5 cell line to the IC_{50} of cancer cell lines. In bold black: selectivity index determined as a ratio of the IC_{50} of the immortalised RPE-1 cell line to the IC_{50} of cancer cell lines.

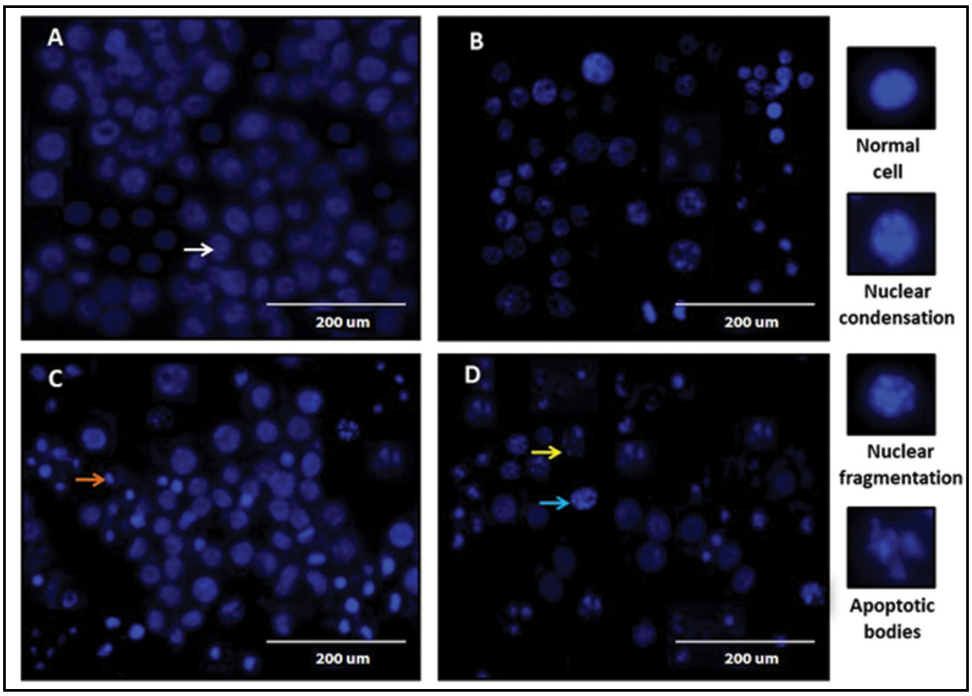


Figure 3: Representative fluorescence images of HepG2 cells treated with NEEAF. HepG2 cells were treated with 0.05% DMSO (A), 12.5 μg/mL etoposide (B), 12.5 μg/mL NEEAF (C) and 50 μg/mL NEEAF (D) for 24 h. White arrow: normal nuclei; orange arrow: nuclear condensation; blue arrow: nuclear fragmentation; yellow arrow: apoptotic bodies.

ROS-mediated cytotoxicity may also be achieved by using agents that can inhibit the cellular antioxidant defence mechanisms; the use of an extract or compound that promotes a rise in intracellular ROS generation, in conjunction with antioxidant system inhibitors, might be a promising strategy towards developing more successful anticancer therapeutics.[27] Along this line, we determined the activities of antioxidant enzymes, primarily SOD and CAT, biomarkers of oxidative stress.[28] Upon treatment with 0.78–50 μg/mL NEEAF, SOD activity decreased from 30.37±4.58 to 9.76±3.95 U/mg of protein, which was about three-fold lower than that of the untreated cells (Figure 4b). Unlike SOD, NEEAF induced significantly higher CAT activity (1.20 ± 0.16 U/mg of protein), particularly at 1.56 μg/mL compared to control (0.85 ± 0.15 U/mg of protein) ($p<0.05$). However, treated cells lost the ability to maintain ROS/antioxidant balance, and a decrease in antioxidant activity was observed from 3.13 to 50 μg/mL of NEEAF (Figure 4c). Therefore, oxidative stress via an increase in ROS production might be a plausible factor leading to NEEAF-induced apoptosis in HepG2 cells. This was further supported by the collapse of the mitochondrial membrane potential (ΔΨm) which is considered to be another key event in the ROS-mediated apoptotic pathway.[22] Initiation of the apoptotic pathway is often stimulated by an increase in the permeability of the mitochondrial membrane, and release of cytochrome *c* which leads to the activation of the caspase cascades.[3] NEEAF depolarised the mitochondrial membranee potential as shown by a significant dose-dependent decline in the JC-1 aggregates/JC-1

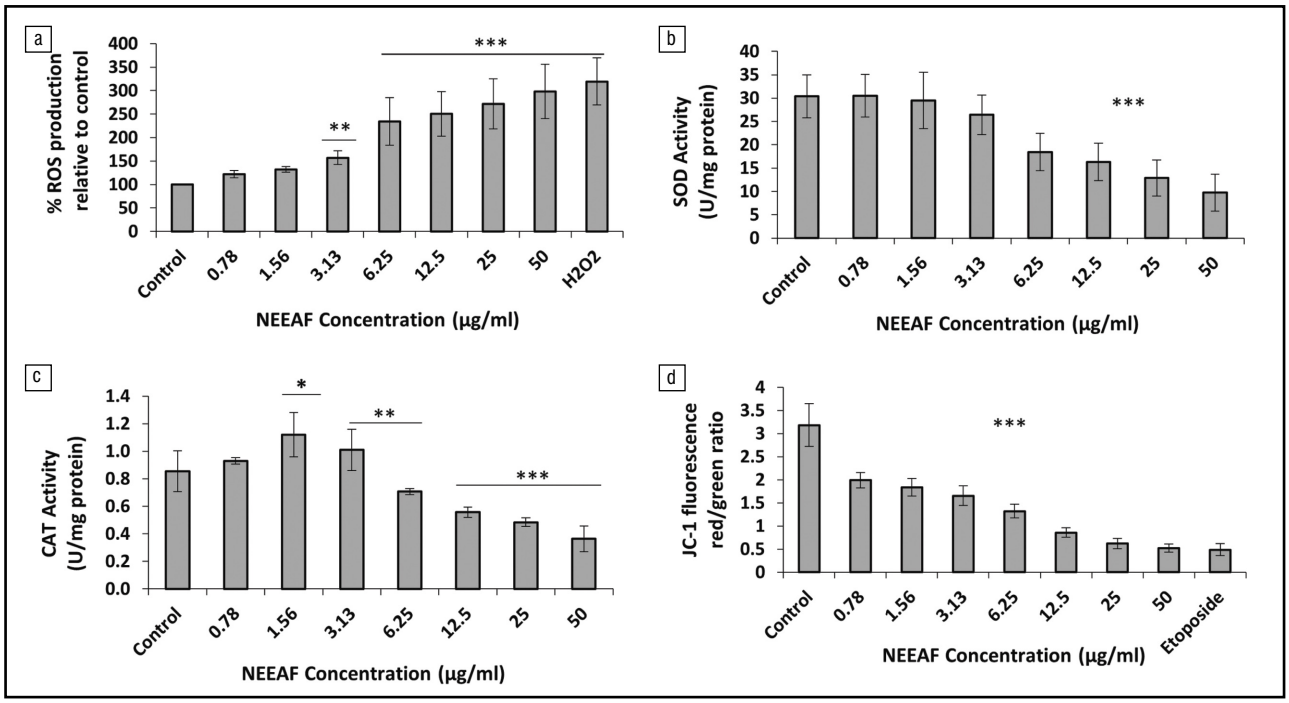


Figure 4: Induction of intracellular ROS production, endogenous antioxidant activities and change in mitochondrial membrane potential by NEEAF. (a) NEEAF induced a dose-dependent increase in intracellular ROS production. HepG2 cells were treated with NEEAF (0.78–50 µg/mL) for 24 h and fluorescence was detected using 25 µM DCFH-DA dye. (b and c) Superoxide dismutase (SOD) and catalase (CAT) antioxidant enzyme activities decreased in HepG2 cells after treatment with NEEAF (0.78–50 µg/mL) for 24 h. (d) NEEAF induced a decline in mitochondrial membrane potential (indicated by fluorescence intensity of JC-1) of HepG2 cells after incubation with NEEAF (0.78–50 µg/mL) measured for 24 h. Data are presented as mean±SD (n=3) in three independent experiments and normalised to cell number/total protein content. *p<0.05, **p<0.01, ***p<0.001 vs. control

monomers ratio from 0.78 ug/mL to 50 ug/mL compared to the control (p<0.05) (Figure 4d). This result is in accordance with the increased ROS production triggered by NEEAF and suggests its potent induction of the mitochondrial pathway leading to apoptosis.

The caspase family proteases play a vital role in the mechanism of apoptosis.[29,30] In particular, the executor caspase-3 is a key regulator of apoptosis, which is activated by the initiator caspase-9, during the mitochondrial pathway of apoptosis. Caspase-3 and -7 also participate in the proteolytic cleavage of PARP, which usually leads to cellular disassembly and serves as a key marker of apoptosis.[31] As shown in Figure 5a and 5b, treatment with NEEAF at 12.5 and 50 ug/mL downregulated pro-caspase-9, pro-caspase-7, pro-caspase-3 and PARP protein levels in HepG2 cells in a dose-dependent manner (p<0.05). Similarly, it also up-regulated the expression level of cleaved caspase-9 (1.8–2.8 folds), 7 (9.9–18.5 folds) and 3 (8.2–9.8 folds) as well as PARP (15.7–19.3 folds) at 25 µg/mL and 50 µg/mL relative to their respective control (p<0.05). While several compounds with interesting apoptotic activity have been previously

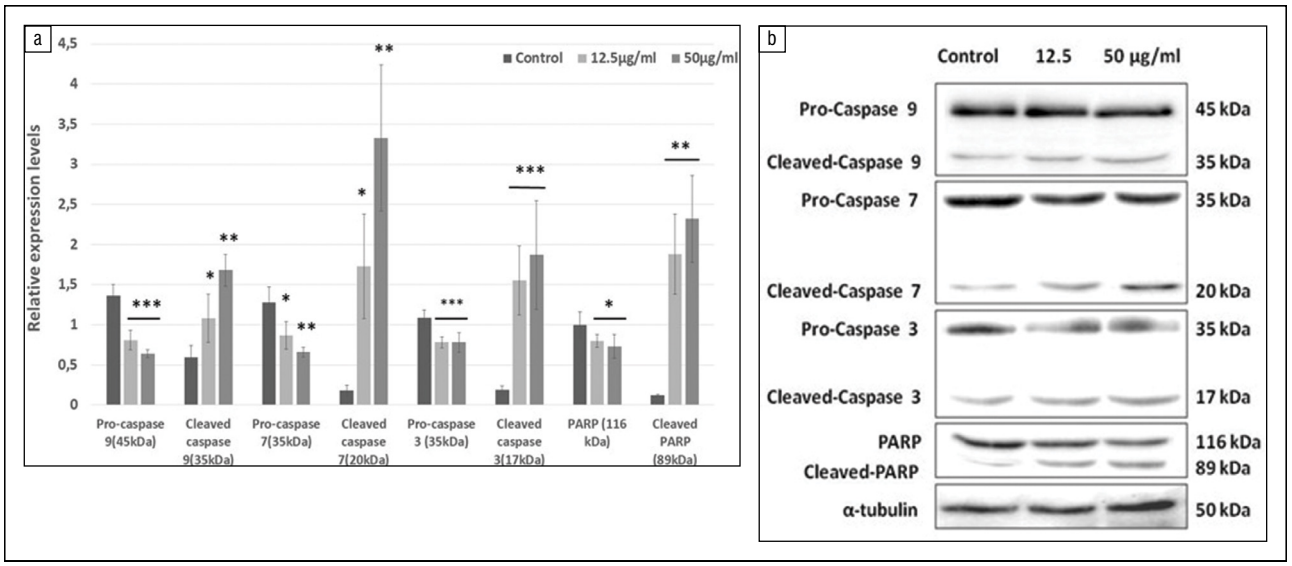


Figure 5: Effect of NEEAF on the expression levels of pro-apoptotic proteins as revealed by western blot analysis. (a) Densitometric-intensity data showing fold changes of caspase-9, caspase-7 and caspase-3 and PARP induced by NEEAF (12.5 μg/mL and 50 μg/mL) in HepG2 cells relative to their respective controls. (b) Representative blots showing fold changes of caspase-9, caspase-7 and caspase-3 and PARP induced by NEEAF (12.5 μg/mL and 50 μg/mL) in HepG2 cells relative to their respective controls. Data are presented as mean±SD (n=3) in three independent experiments. *p<0.05, **p<0.01, ***p<0.001 vs. control.

identified from marine sponges, so far, studies focusing on the mechanisms of cytotoxic agents derived from *Neopetrosia* species are limited. In 2001, Fujiwara and team reported the isolation of the polyketide halenaquinone from the Okinawan sea sponge *Neopetrosia* sp. which subsequently induced apoptosis in PC-12 cells by inhibiting the activity of phosphatidylinositol 3-kinase.[32] Renieramycin J, a new tetrahydroisoquinoline alkaloid, isolated from the Japanese *Neopetrosia* sp., induced changes in the morphology of rat 3Y1 cells which were characteristic of RNA/protein synthesis inhibitors.[11] Its analogue renieramycin M, isolated from *Xestospongia* sp., was later shown to induce apoptotic activity in non-small cell lung cancer cells via the p53-dependent pathway.[33] In light of these findings, we can thus postulate that NEEAF may yield compounds with interesting apoptotic activity.

Conclusion

Overall, this study provides a new understanding into the mode of action underlying the cytotoxicity of the constituents present in the Mauritian sponge *N. exigua* extract. Its promising inhibitory effect against HepG2 cells was mainly due to apoptosis induction via the ROS-mediated mitochondrial pathway. Nevertheless, further investigations are necessary to explore the mechanistic pathways by which NEEAF modulate the expression of other pro/anti-apoptotic proteins in order to probe its bioefficacy as anticancer therapeutics. This will provide a broad insight into its various molecular targets in the process of carcinogenesis. These data also warrant further in-depth investigations into the chemical characterisation of bioactive compounds in NEEAF to unravel their potential pharmacological applications.

Acknowledgements

This work was supported by the Mauritius Research Council under the National Research and Innovation Chair Programme to T.B. and the North-South Interdisciplinary Grant of the Global Young Academy awarded to V.S.N. and Alexander Kagansky. We thank the Albion Fisheries Research Centre, Ministry of Blue Economy, Marine Resources, Fisheries and Shipping for granting permission to collect marine sponges in Mauritian waters. We dedicate this paper to the memory of Dr Alexander Kagansky from the Synthetic Epigenetics Laboratory, MRC Human Genetics Unit, Institute of Genetics and Molecular Medicine, University of Edinburgh, Scotland, who contributed to the design of the cytotoxic study of the marine *Neopetrosia exigua* derived extracts.

Competing interests

We have no competing interests to declare.

Authors' contributions

All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by R.B. The first draft of the manuscript was written by R.B. and all authors commented on proceeding versions of the manuscript. All authors read and approved the final manuscript.

References

1. Ghareeb MA, Tammam MA, El-Demerdash A, Atnasov, AG. Insights about clinically approved and preclinically investigated marine natural products. Curr Res Biotechnol. 2020;2:88–102. https://doi.org/10.1016/j.crbiot.2020.09.001
2. Beesoo R, Neergheen-Bhujun VS, Bhagooli R, Bahorun T. Apoptosis inducing lead compounds isolated from marine organisms of potential relevance in cancer treatment. Mutat Res Fund Mol Mech Mut. 2014;768:84–97. https://doi.org/10.1016/j.mrfmmm.2014.03.005
3. Kim C, Kim B. Anti-cancer natural products and their bioactive compounds inducing ER stress-mediated apoptosis: A review. Nutrients. 2018;10:1021. https://doi.org/10.3390/nu10081021
4. Bae W, Lim HK, Kim KM, Cho H, Lee SY, Jeong CS, et al. Apoptosis-inducing activity of marine sponge *Haliclona* sp. extracts collected from Kosrae in nonsmall cell lung cancer A549 cells. Evid Based Complement Alternat Med. 2015;2015, Art. #717959. https://doi.org/10.1155/2015/717959
5. Surti M, Patel M, Redhwan A, Al-Keridis LA, Adnan M, Alshammari N, et al. Ilimaquinone (marine sponge metabolite) induces apoptosis in HCT-116 human colorectal carcinoma cells via mitochondrial-mediated apoptosis pathway. Mar Drugs. 2022;20:582. https://doi.org/10.3390/md20090582
6. Ciftci HI, Can M, Ellakwa DE, Suner SC, Ibrahim MA, Oral A, et al. Anticancer activity of Turkish marine extracts: A purple sponge extract induces apoptosis with multitarget kinase inhibition activity. Invest New Drugs. 2020;38:1326–1333. https://doi.org/10.1007/s10637-020-00911-8
7. Morais SR, Chitra K, Jeyabalan S, Wong LS, Sekar M, Chidambaram K, et al. Anticancer potential of *Spirastrella pachyspira* (marine sponge) against SK-BR-3 human breast cancer cell line and in silico analysis of its bioactive molecule sphingosine. Front Mar Sci. 2022;9:950880. https://doi.org/10.3389/fmars.2022.950880
8. Liang Z, Sulzmaier F, Yoshida W, Kelly M, Ramos J, Williams P. Neopetrocyclamines A and B, polycyclic diamine alkaloids from the sponge *Neopetrosia* cf *exigua*. J Nat Prod. 2015;78:543–547. https://doi.org/10.1021/np500759r
9. Winder PL, Baker HL, Linley P, Guzmán EA, Pomponi SA, Diaz MC, et al. Neopetrosiquinones A and B, sesquiterpene benzoquinones isolated from the deep-water sponge *Neopetrosia* cf. *proxima*. Bioorg Med Chem. 2011;19:6599–6603. https://doi.org/10.1016/j.bmc.2011.09.026
10. Liu H, Mishima Y, Fujiwara T, Nagai H, Kitazawa A, Mine Y, et al. Isolation of araguspongine M, a new stereoisomer of an araguspongine/xestospongin alkaloid, and dopamine from the marine sponge *Neopetrosia exigua* collected in Palau. Mar Drugs. 2004;2:154–163. https://doi.org/10.3390/md204154
11. Oku N, Matsunaga S, Van Soest RWM, Fusetani N. Renieramycin J a highly cytotoxic tetrahydroisoquinoline alkaloid from a marine sponge *Neopetrosia* sp. J Nat Prod. 2003;66:1136–1139. https://doi.org/10.1021/np030092g
12. Beesoo R, Bhagooli R, Neergheen-Bhujun VS, Li WW, Kagansky A, Bahorun T. Antibacterial and antibiotic potentiating activities of tropical marine sponge extracts. Comp Biochem Physiol C Toxicol Pharmacol. 2017;196:81–90. https://doi.org/10.1016/j.cbpc.2017.04.001
13. Chen B, Huan XJ, Miao ZH, De Voogd NJ, Gu YC, Wang CY, et al. Uncommon bis- quinolizidine alkaloids from the Hainan sponge *Neopetrosia chaliniformis*. Chin J Chem. 2021;39:1838–1842. https://doi.org/10.1002/cjoc.202100091
14. Aktas N, Genç Y, Gozcelioglu B, Konuklugil B, Harput U. Radical scavenging effect of different marine sponges from Mediterranean coasts. Rec Nat Prod. 2013;7:96–104.
15. United Nations. Goal 14: Conserve and sustainably use the oceans, seas and marine resources for sustainable development [webpage on the Internet]. No date [cited 2022 Dec 25]. Available from: https://sdgs.un.org/goals/goal14
16. Van Soest RWM, Boury-Esnault N, Hooper JNA, Rützler K, De Voogd NJ, Alvarez B, et al. World Porifera Database [database on the Internet]. No date [cited 2021 Jan 05]. Available from: http://www.marinespecies.org/porifera
17. Tantengco GOA, Limbo AC, Marco Nemesio Montaño E, Jacinto DS. Cytotoxic activity of crude extract and fractions from *Sargassum siliquosum* (JG Agardh) and other seaweeds against selected human cancer cell lines. Int J Bios. 2015;7:207–215.
18. BioVision. Catalase activity colorimetric/fluorometric assay kit. Milpitas, CA: BioVision; 2017. Available from: https://www.biovision.com/documentation/datasheets/K773.pdf
19. Brastianos H, Vottero E, Patrick B, Van Soest RWM, Matainaho T, Mauk A, et al. Exiguamine A, an indoleamine-2, 3-dioxygenase (IDO) inhibitor isolated from the marine sponge *Neopetrosia exigua*. J Am Chem Soc. 2006;128:16046–16047. https://doi.org/10.1021/ja067211+
20. Li Y, Qin S, Guo Y, Gu Y, VanSoest RWM. 9′-Epi-3β,3′β-dimethyl xestospongin C,a new macrocyclic diamine alkaloid from the Hainan sponge *Neopetrosia exigua*. Planta Medica. 2010;77:179–181. https://doi.org/10.1055/s-0030-1250164
21. Circu ML, Aw TY. Reactive oxygen species, cellular redox systems, and apoptosis. Free Radic Biol Med. 2010;48:749–762. https://doi.org/10.1016/j.freeradbiomed.2009.12.022
22. Redza-Dutordoir M, Averill-Bates DA. Activation of apoptosis signalling pathways by reactive oxygen species. Biochim Biophys Acta. 2016;1863:2977–2992. https://doi.org/10.1016/j.bbamcr.2016.09.012
23. Bauer D, Werth F, Nguyen HA, Kiecker F, Eberle J. Critical role of reactive oxygen species (ROS) for synergistic enhancement of apoptosis by vemurafenib and the potassium channel inhibitor TRAM-34 in melanoma cells. Cell Death. 2017; 8:e2594. https://doi.org/10.1038/cddis.2017.6
24. Liao YJ, Bai HY, Li ZH, Zou J, Chen JW, Zheng F, et al. Longikaurin A, a natural ent-kaurane, induces G2/M phase arrest via downregulation of Skp2 and apoptosis induction through ROS/JNK/c-Jun pathway in hepatocellular carcinoma cells. Cell Death Dis. 2014;5:e1137. https://doi.org/10.1038/cddis.2014.66

25. Xie P, Fuji I, Zhao J, Shinohara M, Matsukura M. A novel polysaccharide derived from algae extract induces apoptosis and cell cycle arrest in human gastric carcinoma MKN45 cells via ROS/JNK signaling pathway. Int J Oncol. 2016;49:1561–1568. https://doi.org/10.3892/ijo.2016.3658

26. Marvibaigi M, Amini N, Supriyanto E, Abdul Majid FA, Kumar Jaganathan S, Jamil S, et al. Antioxidant activity and ROS-dependent apoptotic effect of *Scurrula ferruginea* (Jack) danser methanol extract in human breast cancer cell MDA-MB-231. PLoS ONE. 2016;11:e0158942. https://doi.org/10.1371/journal.pone.0158942

27. Liou GY, Storz P. Reactive oxygen species in cancer. Free Radic Res. 2010;44:479–496. https://doi.org/10.3109/10715761003667554

28. Hassan K, Elobeid MA, Virk P, Omer SA, El Amin M, Daghestani MA, et al. Bisphenol A induces hepatotoxicity through oxidative stress in rat model. Oxid Med Cell Longev. 2012;2012, Art. #194829. https://doi.org/10.1155/2012/194829

29. Kuo YJ, Yang JS, Lu CC, Chiang SY, Lin JG, Chung JG. Ethanol extract of *Hedyotis diffusa* willd upregulates G0/G1 phase arrest and induces apoptosis in human leukemia cells by modulating caspase cascade signaling and altering associated genes expression was assayed by cDNA microarray. Environ Toxicol. 2015;30:1162–1177. https://doi.org/10.1002/tox.21989

30. Panicker NG, Balhamar SOMS, Akhlaq S, Qureshi M, Rizvi TS, Al-Harassi A, et al. Identification and characterization of the caspase-mediated apoptotic activity of *Teucrium mascatense* and an isolated compound in human cancer cells. Molecules. 2019;24:977. https://doi.org/10.3390/molecules24050977

31. Vo PHT, Nguyen TDT, Tran HT, Tran HT, Nguyen YN, Doan MT, et al. Cytotoxic components from the leaves of *Erythrophleum fordii* induce human acute leukemia cell apoptosis through caspase 3 activation and PARP cleavage. Bioorg Med Chem Lett. 2021;31:127673. https://doi.org/10.1016/j.bmcl.2020.127673

32. Fujiwara H, Matsunaga K, Saito M, Hagiya S, Furukawa K, Nakamura H, et al. Halenaquinone, a novel phosphatidylinositol 3-kinase inhibitor from a marine sponge, induces apoptosis in PC12 cells. Eur J Pharmacol. 2021;413:37–45. https://doi.org/10.1016/S0014-2999(00)00944-4

33. Halim H, Chunhacha P, Suwanborirux K, Chanvorachote P. Anticancer and antimetastatic activities of renieramycin M, a marine tetrahydroisoquinoline alkaloid, in human non-small cell lung cancer cells. Anticancer Res. 2011;31:193–201. https://pubmed.ncbi.nlm.nih.gov/21273598/

AUTHORS:
Hafizu Muhammed[1]
Joshua B. Balogun[1]
Musa M. Dogara[1]
Babatunde Adewale[2]
Abdulganiyu A. Ibrahim[3]
Chinedu B. Okolugbo[4]
Graham Jackson[5]

AFFILIATIONS:
[1]Department of Biological Sciences, Federal University, Dutse, Nigeria
[2]Public Health and Epidemiology Department, Nigeria Institute for Medical Research, Lagos, Nigeria
[3]Department of Nutrition and Dietetics, Jigawa State Polytechnic, Dutse, Nigeria
[4]Department of Animal and Environmental Biology, Delta State University, Abraka, Nigeria
[5]Department of Chemistry, University of Cape Town, Cape Town, South Africa

CORRESPONDENCE TO:
Graham Jackson

EMAIL:
graham.jackson@uct.ac.za

DATES:
Received: 26 Apr. 2022
Revised: 13 Feb. 2023
Accepted: 07 Mar. 2023
Published: 08 Aug. 2023

HOW TO CITE:
Muhammed H, Balogun JB, Dogara MM, Adewale B, Ibrahim AA, Okolugbo CB, et al. Co-infection of urogenital schistosomiasis and malaria and its association with anaemia and malnutrition amongst schoolchildren in Dutse, Nigeria. S Afr J Sci. 2023;119(7/8), Art. #13846. https://doi.org/10.17159/sajs.2023/13846

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

DATA AVAILABILITY:
☐ Open data set
☒ All data included
☐ On request from author(s)
☐ Not available
☐ Not applicable

EDITOR:
Pascal Bessong

KEYWORDS:
urogenital schistosomiasis, malaria, anaemia, malnutrition, co-infection

FUNDING:
None

Co-infection of urogenital schistosomiasis and malaria and its association with anaemia and malnutrition amongst schoolchildren in Dutse, Nigeria

Schistosomiasis is a neglected tropical disease. Sub-Saharan Africa accounts for 93% of the world's 207 million schistosomiasis cases. Urogenital schistosomiasis and malaria are both public health problems in Nigeria, where they are endemic. We determined the co-prevalence of urogenital schistosomiasis and malaria in schoolchildren and assessed its implication on anaemia and malnutrition. This cross-sectional study was conducted amongst primary schoolchildren in the Warwade, Saya Saya and Jigawar Daha villages of Nigeria. Urine samples were collected to detect *Schistosoma haematobium* eggs, and finger prick blood was used for haemoglobin concentration and malaria diagnosis. Nutritional status was assessed using anthropometric measurements and a pre-tested questionnaire. The overall prevalence and density of *S. haematobium* were 27.7% and 9 eggs/10 mL, respectively, with significant differences between villages and sexes. The prevalence of malaria and infection density was 10.4% and 330 mps/μL, respectively. Co-infection prevalence was 3.3%. Anaemia prevalence was 66%, with significant variation across villages and between sexes. Prevalence of stunting, underweight, and wasting was 41.7%, 46%, and 29.7%, respectively. Mean haemoglobin concentrations in *Plasmodium* and children co-infected with urogenital schistosomiasis were significantly lower than those who were negative for the infection. No significant association was observed between malnutrition and single or co-infection of urogenital schistosomiasis and malaria. After adjusting for variables associated with anaemia, village of residence remained a significant predictor of anaemia. Water contact activities, such as fishing, swimming, and irrigation, emerged as independent risk factors of *S. haematobium* infection.

Significance:

Urogenital schistosomiasis and malaria infections are prevalent in communities around Warwade dam in Dutse, Nigeria, and cause anaemia. Continuous monitoring, proper treatment and regular intervention is desirable in the communities.

Introduction

The two most common tropical parasitic diseases in sub-Saharan Africa are schistosomiasis and malaria, with both being a serious public health problem.[1] The co-infection of malaria and helminth parasites in Africa is as a result of their high proportions, overlapping distribution, transmission methods and other risk factors.[2] There are three prevalent forms of schistosomiasis in Nigeria: *Schistosoma haematobium, S. mansoni*, and *S. intercalatum*.[3] Approximately two-thirds of the schistosomiasis cases are due to infection caused by *S. haematobium*, which represents an important cause of severe urinary tract disease.[4] Previous studies have confirmed that urogenital schistosomiasis and malaria infections are endemic in Nigeria.[5] Co-infections of these parasites lead to reduced learning, reduced school achievements and poor development in children.[6] Epidemiological studies have highlighted that individuals co-infected with more than one parasite species are at risk of increased morbidity, increased severity of the infecting parasite species and increased susceptibility to other infections.[7] In sub-Saharan Africa, anaemia and malnutrition are the most frequent conditions encountered in field surveys and parasitic infections are among the major causes of these conditions.[8] Mechanisms through which parasitic infections cause anaemia and malnutrition include damage to intestinal mucosa which results in impaired digestion and absorption of nutrients, intestinal blood loss, interference with processes leading to production of blood cells in the bone marrow, and impaired immune development.[9] The control of schistosomiasis and malaria infections has become a concern for many governments and has the support of donors (including international organisations) following the London Declaration of 2012.[10] Identifying areas where these infections occur, and studying their co-infection, risk factors and implication on anaemia and nutritional status will increase the efficiency and proper implementation of control or elimination programmes.[11] Based on previous studies, urinary schistosomiasis is prevalent in Jigawa State[12], and work has been done on its co-infection with malaria and its effect on anaemia and nutritional status. In this study, we sought to determine whether there is a significant co-infection of urogenital schistosomiasis and malaria and its risk of anaemia and malnutrition amongst schoolchildren in the areas around Dutse in Jigawa State, Nigeria.

Materials and methods

Study area

The study was conducted from June to August 2021 in Dutse, northwestern Nigeria. The study areas were Warwade, Saya Saya and Jigawar Daha rural localities about 15–19 km south of Dutse. Warwade lies between latitude 11˚44′3″ N and longitude 9˚13′38″ E; Saya Saya lies between latitude 11˚44′50″ N and longitude 9˚12′2″ E; and Jigawar Daha lies between latitude 11˚45′28″ N and longitude 9˚11′39″ E (Figure 1). The communities were purposely selected because of their closeness to the major dam in the area (Warwade Dam). The Dam is used for irrigation, fishing, recreation and other domestic purposes. The relief of the area is flat with little undulation. The average annual temperature in Dutse is 26.8 °C. The warmest month, on average, is April, which has an average temperature of 31 °C. The coolest month on average is January, with an average temperature of 21.7 °C. The annual average precipitation in Dutse is 729 mm.[13]

Study population, sample size estimation and sampling technique

The study population consisted of primary schoolchildren of the selected communities. The children were selected from class 1–6 and within the age range of 5–15 years, using stratified random sampling. The sample size (n) was estimated using the formula[14]:

$$n = Z^2 \times Pq \div d^2$$

where n is the sample size required, d is the acceptable margin of error (5%), Z is the standard normal deviate of 1.96 at the 95% confidence level, P is the prevalence of urogenital schistosomiasis reported in Dutse[15] (10.0%) or prevalence of malaria (43.7%)[16]. The proportion of negative urogenital schistosomiasis or malaria (q) is given by (q = 1-p). A minimum sample size of 258 was obtained from the average (138+378/2) of the calculated sample size of urogenital schistosomiasis (138) and malaria (378). However, due to the incidence of non-compliance and loss of samples, 300 samples were analysed in the study.

Sample collection and analysis

In the field, the middle finger of each enrolled pupil was cleaned with an alcohol swab and pricked with a sterile lancet to obtain a thick blood film on a labelled clean glass slide for the determination of malaria parasites. Haemoglobin levels were determined from another drop of blood on a Hb(Haemoglobin) test strip and Bioaid haemoglobin meter (Hangzhou Bosure Biotech Co. Ltd, China). The Hb value displayed was recorded to the nearest 0.1 g/dL. According to the World Health Organization (WHO), children of 5 to 11 years of age have a normal haemoglobin concentration with values of ≥11.5 g/dL and mild, moderate or severe anaemia with haemoglobin concentrations of 11.4–11.0 g/dL, 10.9–8.0 g/dL and ≤8.0 g/dL, respectively. Children of 12 to 14 years of age have a normal haemoglobin concentration with values of ≥12.0 g/dL and mild, moderate or severe anaemia with values of 11.9–10.9 g/dL, 10.9–8.0 g/dL and ≤8.0 g/dL, respectively. People of 15 years and older have a normal haemoglobin concentration with values of ≥13.0 g/dL and mild, moderate or severe anaemia with values of 12.9–11.0 g/dL, 10.9–8.0 g/dL and ≤8.0 g/dL, respectively.[17] Each of the enrolled children was provided with a clean, labelled, screw-capped plastic container and instructed to urinate into the container between 10:00 and 14:00[18] and to close the cap tightly.

Urine analysis

Urine samples with gross haematuria were observed visually while micro-haematuria was detected using Medi-Test Combi 9 urinalysis test strips (Macherey-Nagel, Germany). Fltration was used to determine the presence or absence of *S. haematobium* eggs in the urine samples. About 10 mL of urine was transferred into a beaker after gentle shaking of the urine sample container. With the aid of a pipette dropper, 2 to 3 drops of eosin were added to the beaker and mixed. The mixture was drawn up using a 10 mL syringe. A labelled Whatman® Nucleopore filter was inserted into a Millipore and fitted tightly. Keeping the syringe and the Millipore in a vertical direction, the plunger of the syringe containing the urine was depressed. Thereafter, 20 mL of Lugol's iodine was also flushed through the filter holding the filter paper. The filter unit was then unscrewed and the filter allowed to air dry. The whole filter was viewed under a microscope at low objective power (objective x4). Infection was recorded as the number of eggs per 10 mL of urine.[18,19] The intensity of infection was categorised as either heavy (>50 eggs/10 mL of urine) or light (<50 eggs/10 mL of urine).[7,18]

Malaria parasite diagnosis

Thick blood films were dehaemoglobinised in water and stained with 10% Giemsa solution for 10 min, rinsed in water and air dried. The

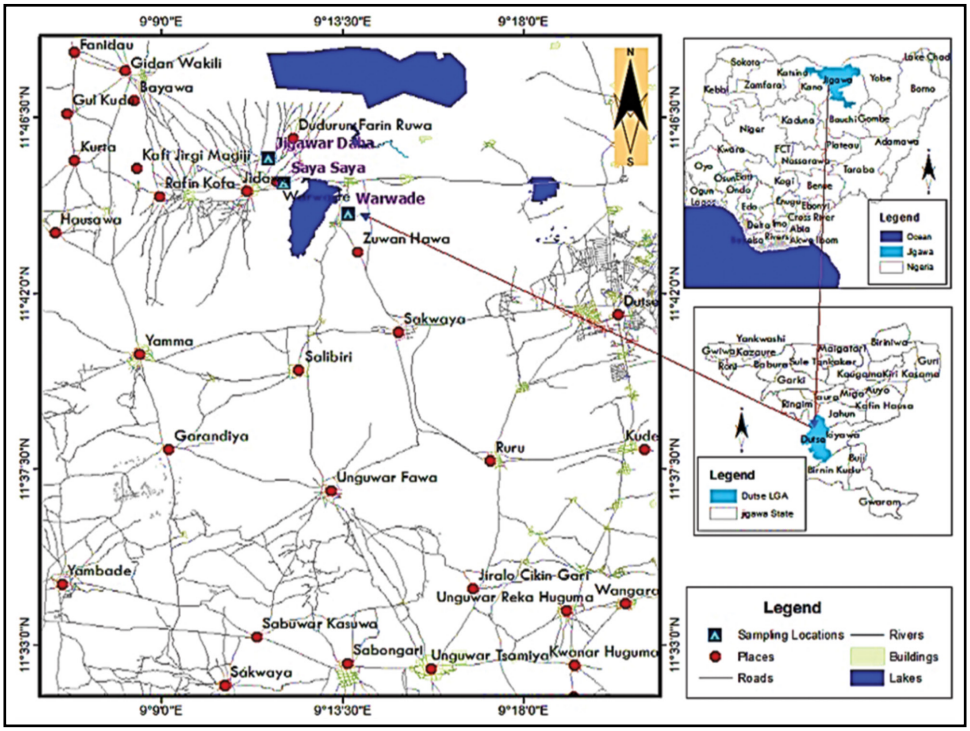


Figure 1: Map of Dutse local government area (LGA) in Jigawa State showing the sampling location.

stained films were viewed using 100 immersion oil objective. Slides were termed positive when asexual forms (trophozoites) and/or gametocytes of *Plasmodium* spp. were observed. The malaria parasites per microlitre of blood was determined by multiplying the average number of parasites per high power field (100x objective and 10x eye piece) by a factor of 500. Between 30 and 50 fields were examined.[20,18] Parasitaemia was classified as low (≤ 500 parasites/μL of blood), moderate (501–5000 parasites/μL of blood) or high (>5000 parasites/μL of blood).[7]

Questionnaire administration and anthropometric measurements

Using a simple pre-tested questionnaire, we interviewed the enrolled children, with the help of the schoolteachers, to obtain their sociodemographic data and associated risk factors for urogenital schistosomiasis and malaria. Age, height, weight and mid-upper arm circumference (MUAC) were recorded to determine participants' anthropometric indices. Height and weight were measured to the nearest 0.1 cm and 0.5 kg using a stadiometer and scale, respectively. MUAC was measured to the nearest 0.1 cm using a graduated tape. These parameters were used to calculate nutritional indices for stunting (height-for-age), wasting (body mass index (BMI)-for-age) and underweight (weight for age). The indices were computed as z-scores based on the WHO growth reference curves.[21] Height-for-age z-score (HAZ), BMI-for-age z-score (BAZ) and weight-for-age z-score (WAZ) were calculated. Children whose HAZ, BAZ and/or WAZ were more than two standard deviations below normal were considered stunted, wasted and/or underweight, respectively.[22]

Data analysis

All collected data were entered into Microsoft Excel (MS Excel 2016) and were checked manually for their completeness. The data were further analysed using the IBM-statistical package for the social sciences (SPSS) version 25 (IBM—SPSS, Inc, Chicago, IL, USA). Descriptive measures such as the mean, standard deviation (SD), median (interquartile), frequencies, and proportions were used to summarise the data. Differences in proportions between populations were obtained using the chi-square (χ^2) test and one-way analysis of variance (ANOVA). Median (interquartile) parasite density of *S. haematobium* and *Plasmodium falciparum* by village, sex and age group were compared using the Kruskal–Wallis test, Fisher's exact test and Mann–Whitney test. Mean (SD) Hb levels were compared using a one-way ANOVA and independent t-test where appropriate. Multivariate logistic regression analysis was used to obtain the predictors of anaemia, while both bivariate and multivariate regression analyses were used in determining risk factors associated with the transmission of *Plasmodium* sp. and *S. haematobium.*

Ethical approval

The protocol of the study was reviewed and approved by the Health Research Committee of the Ministry of Health, Jigawa State (JHREC/2021/038). The State Universal Basic Education Board (SUBEB) of Jigawa State gave administrative approval to conduct the research (reference number SUBEB/ADM/23/vol.1). Participation was voluntary, and a participant could decide to halt their participation in the study at any time without any penalty. The study complied with the institutional guidelines, rules and regulations of the Nigerian National Code for Health Research Ethics.

Results

A total of 300 schoolchildren from three communities – Jigawar Daha (115; 38.3%), Warwade (114; 38%) and Saya Saya (71; 23.7%) – were examined during the study. The demographic characteristics of the schoolchildren are presented in Table 1. More than half (57%) of the participants were female. Jigawar Daha had the highest proportion of female children, while male children were significantly higher in Warwade Primary School (p=0.006). The mean age of all participants was 9.8±2.6 years and the predominant age group was 8 to 11 years (57.3%). Principal water contact activities are depicted in Figure 2. The overall sociodemographics of the study participants are shown in Table 1.

Table 1: Sociodemographic characteristics of the study participants (*n* (%))

Variable	All participants	Warwade	Saya Saya	Jigawar Daha	*p*-value
Sex					0.006*
Female	171 (57)	52 (45.6)	43 (60.6)	76 (66.1)	
Male	129 (43)	62 (54.4)	28 (39.4)	39 (33.9)	
Age (years)					< 0.001*α
mean ± SD	9.8 ± 2.6	9.7 ± 2.0	8.2 ± 2.5	11.0 ± 2.4	
5–7	53 (17.7)	18 (15.8)	29 (40.8)	6 (5.2)	
8–11	172 (57.3)	76 (66.7)	34 (47.9)	62 (53.9)	
12–15	75 (25)	20 (17.5)	8 (11.3)	47 (40.9)	
Educational level					0.28
Primary 1	47 (15.7)	16 (14)	17 (23.9)	14 (12.2)	
Primary 2	50 (16.7)	17 (14.9)	13 (18.3)	20 (17.4)	
Primary 3	47 (15.7)	20 (17.5)	10 (14.1)	17 (14.8)	
Primary 4	52 (17.3)	20 (17.5)	8 (11.3)	24 (20.9)	
Primary 5	45 (15)	21 (18.4)	12 (16.9)	12 (10.4)	
Primary 6	59 (19.7)	20 (17.5)	11 (15.5)	28 (24.3)	
Parent occupation					0.15
Farmer	224 (74.7)	79 (69.3)	50 (70.4)	95 (82.6)	
Trader	47 (15 7)	20 (17.5)	14 (19.7)	13 (11 3)	
Artisan	20 (6.7)	12 (10.5)	5 (7.0)	3 (2.6)	
Civil servant	9 (3.0)	3 (2.6)	2 (2.8)	4 (3.5)	

**Significant at p<0.05; chi-square test and one-way ANOVA*α*

Prevalence, parasite density, single and co-infection of urogenital schistosomiasis and malaria

The prevalence of single infection with *S. haematobium* or *Plasmodium* spp. and co-infection were 27.7% (95% CI = 22.6–32.8%), 10.3% (95% CI = 6.7–13.7%) and 3.3% (95% CI = 1.3–5.3%), respectively (Table 2). The prevalence of single infection with *S. haematobium* in schoolchildren varied significantly from village to village (p<0.001) and between sexes (p=0.001), but not among age groups (p=0.45). The prevalence of single infection with *S. haematobium* was higher among male schoolchildren (37.2%). Also, the prevalence at Warwade (50.9%) was about three times that at Saya saya (16.9%), and 4.5 times that at Jigawar Daha (11.3%). Regarding single infection with *Plasmodium* spp. and co-infection with *Plasmodium* spp. and *S. haematobium*, there was no significant variation with sex or age, or across villages (Table 2). Out of 300 schoolchildren, 75 (25%) had light infection of *S. haematobium*, while only 8 (2.7%) had heavy infections. The median (interquartile) of *S. haematobium* parasite density was 9 (4 to 21), with significant difference among the villages (p<0.001) but not between sex (p=0.27) or among age groups (p=0.49). *Plasmodium* spp. parasitaemia was low (≤ 500 parasites/μL of blood) in 26 (8.7%) children and moderate (501–5000 parasites/μL of blood) in 5 (1.7%) children. Parasite density ranged from 50 to 3250 parasites per microlitre of blood, with a median (interquartile) of 330 (175 to 465).There was no significant variation in infection intensity across villages (p=0.68), between sexes (p=0.21) or among age groups (p=0.93).

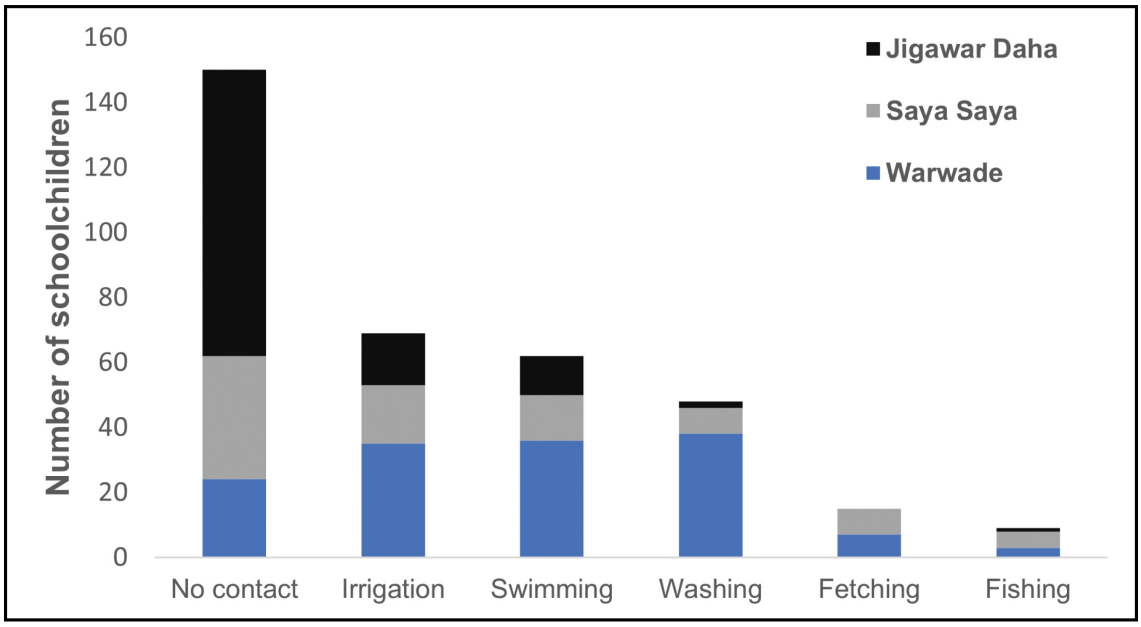


Figure 2: Water contact activities in schoolchildren stratified by villages.

Table 2: Prevalence, parasite density, single and co-infection of urogenital schistosomiasis and malaria in schoolchildren by village, sex and age

Variable	N	Single *S. haematobium* infection			Single *Plasmodium* spp. infection			Co-infection prevalence
		Prevalence *n* (%)	Parasite density (eggs/10 mL urine)		Prevalence *n* (%)	Parasite density (eggs/10 mL urine)		
			Range	Median (IQR)		Range	Median (IQR)	*n* (%)
Village	300	83 (27.7)	1 to 68	9 (4–21)	31 (10.4)	50–3250	330 (175–465)	10 (3.3)
Warwade	114	58 (50.9)	2 to 68	14.5 (6–26.5)	7 (6.1)	90–475	335 (190–465)	4 (3.5)
Saya Saya	71	12 (16.9)	2 to 27	6.5 (4.5–14.8)	7 (9.9)	135–430	325 (170–335)	2 (2.8)
Jigawar Daha	115	13 (11.3)	1 to 7	2 (1.5–4.0)	17 (14.8)	20–3250	395 (155–535)	4 (3.5)
p-value		< 0.001*		< 0.001*[c]	0.1		0.68[c]	0.96
Sex								
Female	171	35 (20.5)	1 to 64	7 (3–17)	19 (11.1)	50–3250	335 (220–5325)	6 (3.5)
Male	129	48 (37.2)	1 to 68	9 (4.3–25.8)	12 (9.3)	80–475	210 (145–430)	4 (3.1)
p-value		0.001*		0.27[e]	0.61		0.21[e]	0.85
Age group								
5–7	53	11 (20.8)	1 to 28	17 (8–23)	9 (17.0)	130–535	335 (182.5–470)	1 (1.9)
8–11	172	51 (29.7)	1 to 68	9 (4–21)	17 (9.9)	50–3250	330 (152.5–452.5)	9 (5.2)
12–15	75	21 (28.0)	1 to 64	6 (3 to 26)	5 (6.7)	135 to 560	230 (155–477.5)	0
p-value		0.45		0.49[c]	0.16		0.93[c]	0.07[d]

*p-value was obtained using chi-square test, Kruskal–Wallis test[c], Fisher's Exact test[d] and Mann–Whitney test[e]; *significant at p<0.05*

Prevalence of anaemia and malnutrition

The overall prevalence of anaemia was 66% (95% CI = 60.6–71.4%). For severe, moderate and mild anaemia the prevalence was 5.7% (95% CI = 3.1–8.3%), 40.3% (95% CI = 34.7–45.9%) and 20% (95% CI = 15.5–24.5%), respectively (Figure 3). The prevalence of anaemia differed significantly among villages ($p<0.001$), between sexes ($p=0.03$) and among age groups ($p=0.003$). The anaemia prevalence was highest in Warwade, while Saya Saya had the lowest prevalence (Table 3). A higher prevalence of anaemia was observed in male and older children (12–15 years).

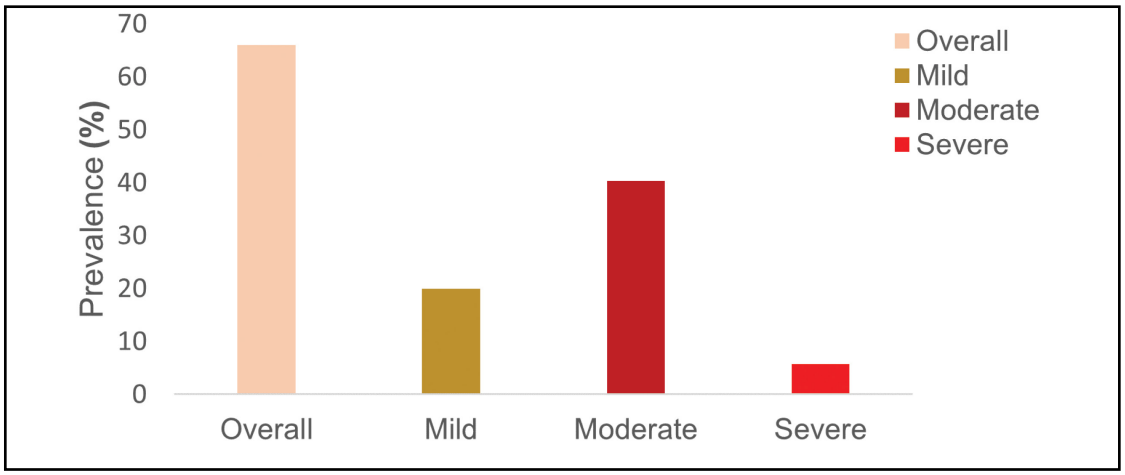


Figure 3: Prevalence of anaemia amongst schoolchildren in the studied areas.

Table 3: Prevalence of anaemia and malnutrition in schoolchildren by village, sex and age

Variable	N	Anaemia			Prevalence of malnutrition n (%)		
		Prevalence	Hb concentration, g/dL		Stunting	Underweight	Wasting
		n (%)	Range	Mean (SD)	*n* (%)	*n* (%)	*n* (%)
Village	300	198 (66)	3.6–19.9	10.9 (1.8)	125 (41.7)	138 (46.0)	89 (29.7)
Warwade	114	85 (74.6)	7.3–14.6	10.9 (1.2)	53 (46.5)	63 (55.3)	41 (36.0)
Saya Saya	71	33 (46.5)	3.6–14.7	11.5 (1.9)	33 (46.5)	26 (36.6)	16 (22.5)
Jigawar Daha	115	80 (69.6)	3.6–19.9	10.6 (2.2)	39 (33.9)	49 (416)	32 (27.8)
p-value		<0.001*		0.01*[a]	0.10	0.03*	0.13
Sex							
Female	171	104 (60.8)	3.6–19.9	10.9 (2.0)	65 (38.0)	79 (46.2)	52 (30.4)
Male	129	94 (72.9)	3.6–14.7	10.9 (1.5)	60 (46.5)	59 (45.7)	37 (28.7)
p-value		0.03*		0.95[b]	0.14	0.94	0.75
Age group							
5–7	53	29 (54.7)	7.3–14.7	11.4 (1.2)	13 (24.5)	6 (11.3)	8 (15.1)
8–11	172	108 (62.8)	3.8–14.4	10.9 (1.8)	70 (40.7)	76 (44.2)	44 (25.6)
12–15	75	61 (81.3)	3.6–19.9	10.7 (2.3)	42 (56.0)	56 (74.7)	37 (49.3)
p-value		0.003*		0.06[a]	0.002*	<0.001*	<0.001*

*p-value obtained using chi-square test, one-way ANOVA[a] and independent t-test[b]; *significant at p<0.05*

Hb, haemoglobin; SD, standard deviation

Urinary schistosomiasis and malaria co-infection and association with anaemia and malnutrition

The proportion of anaemic schoolchildren with *S. haematobium* or *Plasmodium* spp. or co-infection with both parasites was not significantly higher than in those without these infections (Table 4). No significant variation in haemoglobin levels was observed between children with and without *S. haematobium* infection (10.9 vs 10.9 g/dL; p=0.91); however, mean haemoglobin levels in children with malaria infection were significantly lower than in uninfected children (9.7 vs 11.0 g/dL; p=0.004). Also, as shown in Figure 4, haemoglobin concentrations decreased significantly with increasing malaria parasite density (p<0.001). Regarding co-infection, mean haemoglobin levels were significantly lower in children infected with both parasites (9.6 vs 10.9 g/dL; p=0.03). There was no significant association between single infection with schistosomiasis or malaria or with co-infection and nutritional status. The results in Table 4 indicate that cases of malnutrition did not significantly differ between infected and non-infected children in the study areas.

Predictors of anaemia

Variables that showed significant associations (p<0.05) with anaemia were considered for multivariate logistic regression analysis. The results in Table 5 indicate that only village of residence remained a significant predictor of anaemia among schoolchildren. The odds of anaemia among schoolchildren was 3.4 times higher at Warwade and 2.6 times higher at Jigawar Daha compared to Saya Saya. Moreover, male children were 1.7 times more likely to be anaemic than female children, and children aged 12–15 years and 8–11 years were, respectively, 1.6 and 1.4

Table 4: Urinary schistosomiasis, malaria and co-infection association with anaemia and malnutrition

Variable	N	Anaemia			Prevalence of malnutrition *n* (%)		
		Prevalence	Hb concentration, g/dL		Stunting	Underweight	Wasting
		n (%)	Range	Mean (SD)	*n* (%)	*n* (%)	*n* (%)
S. haematobium							
Negative	217	138 (63.6)	3.6–19.9	10.9 (1.9)	86 (39.6)	95 (43.8)	63 (29.0)
Positive	83	60 (72.3)	5.5–14.7	10.9 (1.5)	39 (47.0)	43 (51.8)	26 (31.3)
p-value		0.16		0.91[b]	0.25	0.21	0.69
Plasmodium spp.							
Negative	269	175 (65.1)	3.6–19.9	11.0 (1.7)	112 (41 .6)	127 (47.2)	81 (30.1)
Positive	31	23 (74.2)	3.8–13.0	9.7 (24)	13 (41.9)	11 (35.5)	8 (25.8)
p-value		0.31		0.004*[b]	0.97	0.22	0.62
Co-infection							
Negative	290	191 (65.9)	3.6–19	10.9 (1.8)	121 (41.7)	135 (46.6)	86 (29.7)
Positive	10	7 (70.0)	5.5–12.3	9.6 (2.5)	4 (40.0)	3 (30.0)	3 (30.0)
p-value		0.79		0.03*[b]	0.91	0.30	0.98

*p-value obtained using chi-square test and independent t-test[b]; *significant at p<0.05*

Hb, haemoglobin

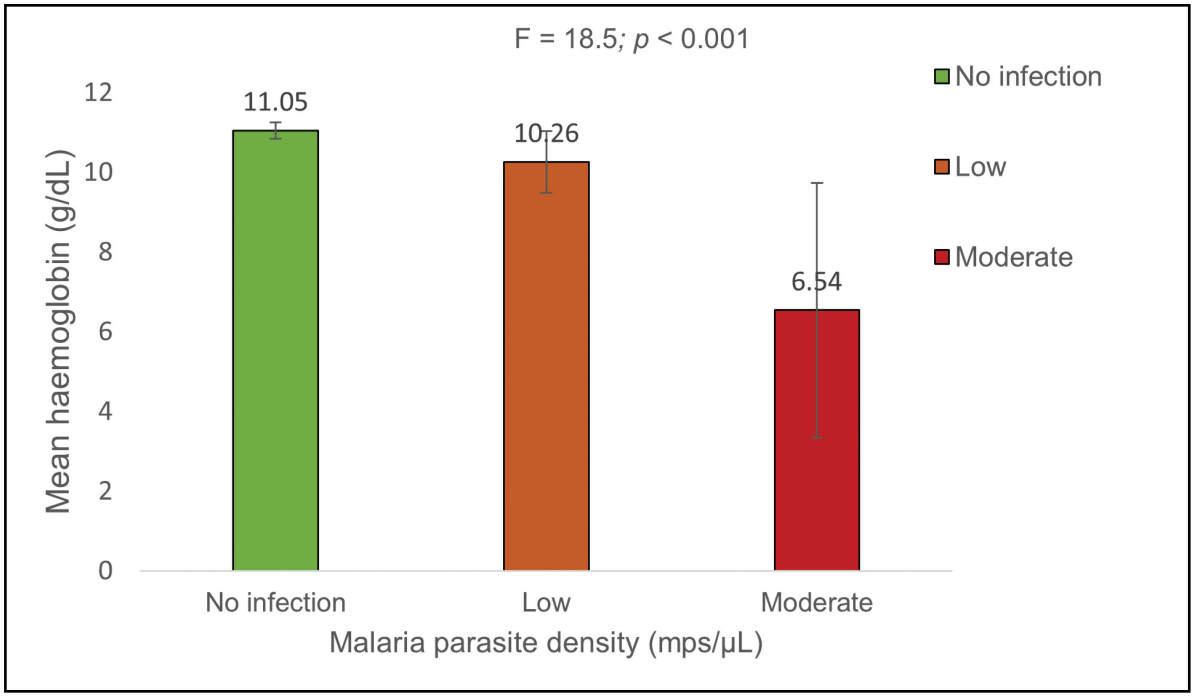


Figure 4: Association between haemoglobin levels and malaria parasite density.

times more likely to be anaemic than the 5–7-year cohort. Stunted and underweight children had almost 2 times the risk of being anaemic when compared to non-stunted children in the studied population (Table 5).

Risk factors associated with the transmission of urogenital schistosomiasis and malaria

At bivariate level, water contact activities (except for fetching) showed significant association (irrigation: p=0.01; swimming: p<0.001; washing: p=0.002, fishing; p=0.01) with the transmission of *S. haematobium* infection (Table 6). After adjusting for village of residence, sex and age, water contact activities that emerged as independent significant risk factors of *S. haematobium* infection included fishing, swimming and irrigation. Engaging in fishing activities was associated with an 8.3-fold increased likelihood of *S. haematobium* infection relative to those who did not fish. Also, the odds of *S. haematobium* infection among schoolchildren who engaged in swimming was 4.5 times greater than in those children who did not swim. In addition, schoolchildren engaged in irrigation activities were 2.3 times more likely to be infected with *S. haematobium* compared to those who did not (Table 6). Regarding *Plasmodium* spp. infection, the use of insecticides and insecticide treated nets decreased the odds of

Table 5: Predictors of anaemia in schoolchildren in the studied communities

Predictor	Bivariate analysis			Multivariate analysis		
	cOR	95% CI	*p*-value	aOR	95% CI	*p*-value
Area						
Saya Saya (ref.)	1.0			1.0		
Jigawar Daha	2.6	1.4 – 4.9	0.002*	2.3	1.2 – 4.7	0.02*
Warwade	3.4	1.8 – 6.3	<0.001	3.1	1.6 – 6.2	0.001*
Sex						
Female (ref.)	1.0			1		
Male	1.7	1.1 – 2.8	0.03*	1.6	0.95 – 2.8	0.08
Age group (years)						
5–7 (ref.)	1.0			1.0		
8–11	1.4	0.75 – 2.6	0.29	1.1	0.52 – 2.2	0.83
12–15	3.6	1.6 – 8.0	0.002*	2.5	0.94 – 6.8	0.07
S. haematobium						
Negative (ref.)	1.0			–		
Positive	1.5	0.86 – 2.6	0.16	–	–	–
Plasmodium sp.						
Negative (ref.)	1.0			–		
Positive	1.5	0.67 – 3.6	0.31	–	–	–
Co-infection						
Negative (ref.)	1.0			–		
Positive	1.2	0.31 – 4.8		–	–	–
Haematuria						
Negative (ref.)	1.0			–		
Positive	1.5	0.52 – 4.2	0.47	–	–	–
Stunted						
No (ref.)	1.0			1.0		
Yes	1.8	1.1 – 3.0	0.02*	1.7	0.92 – 3.0	0.09
Underweight						
No (ref.)	1.0			1.0		
Yes	1.7	1.1 – 2.8	0.03*	1.01	0.55 – 1.9	0.98
Wasted						
No (ref.)	1.0			–		
Yes	0.83	0.49 – 1.4	0.47	–	–	–

*cOR, crude odds ratio; aOR, adjusted odds ratio; ref., reference; CI, confidence interval; *significant at p<0.05*

Table 6: Risk factors associated with urogenital schistosomiasis infection in schoolchildren

Risk factors	Bivariate analysis			Multivariate analysis		
	cOR	95% CI	*p*-value	aOR	95% CI	*p*-value
Village						
Saya Saya (ref.)	1.0			1.0		
Jigawar Daha	0.63	0.27 – 1.5	0.28	1.04	0.40 – 2	0.94
Warwade	5.1	2.5 – 10	<0.001*	5.5	2.4 – 12.5	< 0.001*
Sex						
Female (ref.)	1.0			1.0		
Male	2.3	1.4 – 3.9	0.001*	1.1	0.60 – 2.1	0.72
Age group (years)						
5–7 (ref.)	1.0			–		
8–11	1.6	0.77 – 3.4	0.21	–	–	–
12–15	1.5	0.65 – 3.4	0.35	–	–	–
Irrigation						
No (ref.)	1.0			1.0		
Yes	2.2	1.2 – 3.9	0.01*	2.3	1.2 – 4.6	0.01*
Swimming						
No (ref.)	1.0			1.0		
Yes	5.6	3.1 – 10.2	< 0.001*	4.5	2.3 – 9.1	< 0.001*
Washing						
No (ref.)	1.0			1.0		
Yes	2.6	1.4 – 5.0	0.002*	1.4	0.68 – 3.0	0.35
Fetching						
No (ref.)	1.0			–		
Yes	1.3	0.44 – 4.0	0.62	–	–	–
Fishing						
No (ref.)	1.0			1.0		
Yes	5.6	1.4 – 22.8	0.01*	8.3	1.6 – 43.0	0.01*

*cOR, crude odds ratio; aOR, adjusted odds ratio; ref., reference; CI, confidence interval; *significant at p<0.05*

Plasmodium spp. infection among schoolchildren (Table 7). However, these risk factors (use of insecticides and treated nets) did not attain statistical significance in the studied population, in both bivariate and multivariate analyses (Table 7).

Discussion

Schistosomiasis and malaria have an adverse effect on cognitive development, leading to diminished educational performance and absenteeism.,[2324] Both urogenital schistosomiasis and malaria lead to anaemia and growth retardation in children.[25] From the findings in this study, urogenital schistosomiasis is highly prevalent (≥50%) in Warwade, whereas it is moderately prevalent (≥10%) in Saya Saya and Jigawar Daha. This is because Warwade is the closest community to the Warwade Dam, followed by Saya Saya and Jigawar Daha and the closer the children are to the water body the higher the probability of being infected with schistosomiasis. This shows that there is a high transmission of urogenital schistosomiasis around the Warwade Dam in Dutse. The high prevalence and infection intensity in schoolchildren in the area is an indicator that the rest of the population are at high risk of infection. The 50.9% prevalence of urogenital schistosomiasis in Warwade is greater than the prevalence reported by Alhaji et al.[26] who found a 12.3% prevalence in the Warwade community using a sedimentation technique. This difference is because we sampled primary schoolchildren and adolescents and employed a filtration technique. The overall prevalence of *S. haematobium* (27.7%) in the studied villages is also greater than the prevalence found in a recent study by Dogara et al.[15] who reported a 10% prevalence of *S. haematobium* in the Dutse metropolis using a sedimentation method. This difference is due to the prximity of our studied villages to Warwade Dam and because we employed a filtration technique which is more sensitive than a sedimentation method.[19]

Moreover, the 27.7% prevalence of *S. haematobium* in the studied communities is higher than the prevalence reported by David et al.[27] in Gwaram, Jigawa State. However, the overall prevalence of *S. haematobium* in this study is lower than the overall prevalence found by Awosolu et al.[28] in southwestern Nigeria. Amaechi et al.[29] recorded an

Table 7: Risk factors associated with malaria infection in schoolchildren

Risk factor	Bivariate analysis			Multivariate analysis		
	cOR	95% CI	*p*-value	aOR	95% CI	*p*-value
Village						
Saya Saya (ref.)	1.0			1.0		
Jigawar Daha	1.6	0.62 – 4.0	0.33	3.3	1.1 – 10.0	0.04*
Warwade	0.60	0.20 – 1.8	0.36	0.92	0.28 – 3.0	0.88
Sex						
Female (ref.)	1.0				1.0	
Male	0.82	0.38 – 1.8	0.61	0.80	0.35 – 1.8	0.60
Age group (years)						
5–7 (ref.)	1.0			1.0		
8–11	0.54	0.22 – 1.3	0.16	0.34	0.12 – 0.94	0.04*
12–15	0.35	0.11 – 1.1	0.08	0.15	0.04 – 0.59	0.01
Use of ITN						
No (ref.)	1.0			1.0		
Yes	0.64	0.30 – 1.4	0.21	0.85	0.38 – 1.9	0.69
Use of insecticide						
No (ref.)	1.0			1.0		
Yes	0.65	0.31 – 1.4	0.26	0.54	0.24 – 1.2	0.13

*cOR, crude odds ratio; aOR, adjusted odds ratio; ref., reference; CI, confidence interval; ITN, insecticide-treated net; *significant at p<0.05*

overall prevalence of 50.8% in southeastern Nigeria. The prevalence and infection intensity of *S. haematobium* varied by village. This is probably due to the level of exposure and the distance of the communities to the main water body (Warwade Dam) in the area. However, the prevalence and intensity also varied by sex; being more prevalent in male than female children; this is due to water contact activity/behaviour and susceptibility to infection in relation to sex. This is consistent with the work in Ondo State[30] in southeastern Nigeria[31] and in The Gambia.[32] However this is contrary to the findings of Otuneme et al.[33] and Hassan et al.[34] who recorded higher prevalence in female participants in southwestern Nigeria. Although there is no significant difference between prevalence and infection intensity among age groups, children of 8–11 years and 11–15 years have relatively higher prevalence and intensity when compared to the 5–7-year cohort. This is comparable to the research in Côte d'Ivoire where higher prevalence was found in 8–15-year-olds.[35] Uweh et al.[36] also reported higher prevalence in the 11–15-year age group in Benue State and another higher prevalence was found in a fishing community in Kebbi State within the age group of 8–10 years.[37]

The total prevalence of malaria of 10.3% reported in this study was lower than the prevalence (51%) reported by Dogara and Ocheje[38] in Dutse General Hospital. However, the relatively low prevalence of malaria in the studied villages may be attributed to the season in which the research was conducted; high prevalence of malaria in northern Nigeria usually occurs in the rainy season (August–November). Also, our low prevalence is because ours was not a hospital-based study; many of the infections are asymptomatic with moderate parasitaemia. A very low co-infection (3.3%) was detected between *S. haematobium* and *Plasmodium* spp., showing no significant association between urogenital schistosomiasis and malaria infection in our study. Co-infection prevalence in this research is lower when compared to other studies in Nigeria. Morenikeji et al.[39] reported 57.1% co-infection prevalence in a rural community in southwestern Nigeria; Oladele et al.[40] found 16% co-infection prevalence in Ogun State. However, co-infection is lower than that reported by Nyarko et al.[41] in Ghanian schoolchildren and Deribew et al.[42] in Ethiopian schoolchildren (0.9% and 2.84%, respectively). Notwithstanding this, in a recent study by Sumbele et al.[7] in Cameroon, a relatively higher 8.3% co-infection prevalence of *S. haematobium* and *P. falciparum* was reported. Moreover, *S. haematonium* and *Plasmodium* may modulate the effect of each other within their host. According to Lyke et al.[43] *S. haematobium* exerts a persistent stimulatory effect on the host immune system, protecting children against uncomplicated *P. falciparum.* Conversely, schistosomiasis can have a negative effect on host response to malaria, including increased susceptibility to *Plasmodium* infection and increased severity of disease, especially among children.[44] Oladele et al.[40] found most co-infected school-aged children had malnutrition, impaired cognitive development, splenomegaly and fatigue, resulting in poor school performance and reduced overall physical work capacity.

Our study revealed that an overall 66% of the schoolchildren were anaemic. Warwade, Saya Saya and Jigawar Daha had an anaemia prevalence of 74.6%, 46.5% and 69.6%, respectively. This showed that the schoolchildren in the villages were severely anaemic – an indication of a huge public health problem.[17] The higher prevalence of anaemia in schoolchildren is comparable to the findings o Nyarko et al.[41] in Ghana, Deribew et al.[42] in Ethiopia and Sumbele et al.[7] in Cameroon who reported an anaemia prevalence of 59.9%, 81.8% and 74.4%, respectively. However, in northwestern Nigeria, a lower anaemia prevalence of 11.7% was reported by Oladele et al.[40] Although there are many causes of anaemia, the high prevalence of anaemia in Warwade, Saya Saya and Jigawar Daha is probably due to the high prevalence and infection intensity of *S. haematobium* and *Plasmodium.* Therefore, this study provides further evidence that parasitic infections are associated with anaemia. Results of this research revealed a mean haemoglobin concentration of 10.9±1.8 g/dL. This is because the prevalence of moderate anaemia is greater than that of severe and mild anaemia. Although in this study there was no significant difference in haemoglobin levels in children with and without *S. haematobium* infection, the mean

haemoglobin level in children with malaria was significantly lower than that in uninfected children, and haemoglobin concentration decreased significantly with increasing malaria parasite density. This result is consistent with the findings of Konaté et al.[45] in Mali, Starck et al.[46] in Burkina Faso and Ehiem et al.[47] in Ghanaian schoolchildren. Moreover, mean haemoglobin levels were significantly lower in children co-infected with *S. haematobium* and *Plasmodium.* This is due to the combined loss of red blood cells (erythrocytes) as a result of *S. haematobium* and *Plasmodium* infections. This result is also comparable to the findings of Kinung'hi et al.[48] in Tanzania, Sumbele et al.[7] in Cameroon and Edosomwan et al.[49] in Nigeria. We also found a higher prevalence of anaemia in male children than in female children. This is probably due to the higher prevalence and intensity of *S. haematobium* in the male participants in the study. Nevertheless, Warwade (village of residence) was the only significant predictor of anaemia. This is likely due to the severe and prolonged burden of *S. haematobium* infection (>50% prevalence) and malnutrition in Warwade. Being stunted and underweight increased the risk of anaemia in the schoolchildren.

Accordance to the WHO classification of the severity of malnutrition, malnutrition (stunting, underweight and wasting) is highly prevalent in the studied population.[50] Millions of children in the world suffer from malnutrition; although the causes of malnutrition are multifactorial, studies have indicated that malaria and urogenital schistosomiasis increase the risk of malnutrition[51]. Thus *S. haematobium* and *Plasmodium* infections are detrimental to growth and development of children, which could lead to attention deficit, school absenteeism and reduced cognitive ability.[50] Other studies in Nigeria have found lower prevalence of stunting, underweight and wasting: Ayogu et al.[52] in southeastern Nigeria, Umeokonkwo et al.[53] in Abakaliki metropolis and Ajakaye and Ibukunoluwa[54] in Edo State. The prevalence of malnutrition (stunting, underweight and wasting) detected in this study is also higher when compared to the findings in Angola[55], in northwestern Tanzania[48] and in Cameroon[7].

Underweight varied significantly from village to village, while stunting, underweight and wasting varied significantly by age group. This finding is probably due to nutritional and environmental stress, as older children (12–15 years) are more hyperactive than those much younger (5–7 years). We did not discern a significant association between single or multiple parasite infections (*S. haematonium* and *Plasmodium*) and malnutrition. Although an association between single and multiple parasite infections and malnutrition have been reported in Kenya and Angola[56,55], some studies in Nigeria and Tanzania reported no association[57,49]. The lack of association between *S. haematobium, Plasmodium* sp. or co-infection of the parasites and malnutrition is likely due to other factors that are associated with undernutrition, maybe socio-economic factors and other infections.

Water contact activities that aided the transmission of urogenital schistosomiasis in this study were fishing, swimming and irrigation. This is line with the research of Singh et al.[58] in Sokoto, Mafiana et al.[59] in Ogun State and N'Guessan et al.[60] in Mauritania. In this study, children who were engaged in fishing had a 8.3-fold greater likelihood of urogenital schistosomiasis infection relative to those who did not fish. Also, the odds of *S. haematobium* infection among schoolchildren who engaged in swimming was 4.5 times greater than those who did not swim. In addition, schoolchildren engaged in irrigation activities were 2.3 times more likely to be infected with *S. haematobium* compared to those who did not engage in irrigation. Again, children in Warwade are 5.5 times more likely to have *S. haematobium* infection when compared to children in Saya Saya and Jigawar Daha. The use of insecticides and insecticide-treated nets decreased the odds of *Plasmodium* sp. infection among schoolchildren but not statistically significantly in the studied population, in both bivariate and multivariate levels of analyses.

Conclusions

Schistosomiasis is a neglected tropical disease; the largest burden of which is found in sub-Saharan Africa, which accounts for ~93% of the world's ~207 million schistosomiasis cases. In Nigeria, both schistosomiasis and malaria are diseases of public health concern which affect mainly schoolchildren. *S. haematobium* and *Plasmodium* spp infections were prevalent in the schools in our study. Although a low co-infection was observed, a high prevalence of anaemia and malnutrition was found in the studied areas. Despite the fact that the parasite infections (*S.haematobium, Plasmodium* sp. and co-infection) were not significantly correlated with malnutrition (stunting, underweight or wasting), there was, however, a significant difference in *Plasmodium* spp. single infection as well as co-infection with anaemia. Village of residence (Warwade and Saya Saya) was an independent significant predictor of anaemia, while being in the 12–15-year age group, being male, stunted and underweight increased the odds of anaemia. Swimming, irrigation and fishing were independent significant risk factors of urogenital schistosomiasis infection, while the use of insecticide and insecticide-treated nets decreased the odds of malaria infection. In view of the findings of this study, there is need for large-scale interventions in the communities within and around the Warwade Dam area via mass drug administration.

Acknowledgement

We appreciate the involvement of the villages and school heads of Warwade, Saya Saya and Jigawar Daha in ensuring a problem-free sample collection.

Competing interests

We have no competing interests to declare.

Authors' contributions

J.B.B.: Conceptualisation, writing – initial draft, data collection, parasitological analysis, writing – final draft. B.A.: Conceptualisation, writing – initial draft. H.M.: Writing – initial draft, data collection, parasitological analysis, nutritional analysis, writing – final draft. M.M.D.: Writing – initial draft, parasitological analysis. C.B.O.: Data collection, parasitological analysis. A.A.I.: Nutritional analysis. G.J.: Writing – final draft.

References

1. Sangweme DT, Midzi N, Zinyowera-Mutapuri S, Mduluza T, Diener-West M, Kumar N. Impact of schistosome infection on *Plasmodium falciparum* malariometric indices and immune correlates in school age children in Burma Valley, Zimbabwe. PLoS Negl Trop Dis. 2010;4(11):e882. https://doi.org/10.1371/journal.pntd.0000882

2. Njua-Yafi C, Achidi EA, Anchang-Kimbi JK, Apinjoh TO, Mugri RN, Chi HF, et al. Malaria, helminths, co-infection and anaemia in a cohort of children from Mutengene, south western Cameroon. Malar J. 2016;15(1), Art. #69. https://doi.org/10.1186/s12936-016-1111-2

3. Ekpo UF, Hürlimann E, Schur N, Oluwole AS, Abe EM, Mafe MA, et al. Mapping and prediction of schistosomiasis in Nigeria using compiled survey data and Bayesian geospatial modelling. Geospat Health. 2013;7(2):355–366. https://doi.org/10.4081/gh.2013.92

4. Hotez PJ, Kamath A. Neglected tropical diseases in sub-Saharan Africa: Review of their prevalence, distribution, and disease burden. PLoS Negl Trop Dis. 2009;3(8):e412. https://doi.org/10.1371/journal.pntd.0000412

5. Afolabi MO, Ale BM, Dabira ED, Agbla SC, Bustinduy AL, Ndiaye JL, et al. Malaria and helminth co-infections in children living in endemic countries: A systematic review with meta-analysis. PLoS Negl Trop Dis. 2021;15(2):e0009138. https://doi.org/10.1371/journal.pntd.0009138

6. Brooker SJ, Pullan RL, Gitonga CW, Ashton RA, Kolaczinski JH, Kabatereine NB, et al. Plasmodium–helminth coinfection and its sources of heterogeneity across east Africa. J Infect Dis. 2012;205(5):841–852. https://doi.org/10.1093/infdis/jir844

7. Sumbele IU, Otia OV, Francis L, Bopda OS, Ebai CB, Ning TR, et al. Confounding influences of malnutrition and *Plasmodium falciparum* and *Schistosoma haematobium* infections on haematological parameters in school children in Muyuka, Cameroon. BMC Infect Dis. 2021;21(1), Art. #477. https://doi.org/10.1186/s12879-021-06201-9

8. Crompton DW, Nesheim MC. Nutritional impact of intestinal helminthiasis during the human life cycle. Annu Rev Nutr. 2002;22(1):35–59. https://doi.org/10.1146/annurev.nutr.22.120501.134539

9. World Health Organization (WHO). Helminth control in school-age children: A guide for managers of control programmes. Geneva: World Health Organization; 2011. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3923741

10. World Health Organization. London Declaration on NTDs [document on the Internet]. c2012 [cited 2021 Nov 13]. Available from: https://www.who.int/neglected_diseases/London_Declaration_NTDs.pdf

11. Linehan M, Hanson C, Weaver A, Baker M, Kabore A, Zoerhoff KL, et al. Integrated implementation of programs targeting neglected tropical diseases through preventive chemotherapy: Proving the feasibility at national scale. Am J Trop Med Hyg. 2011;84(1):5. https://doi.org/10.4269/ajtmh.2012.11-1589

12. Balogun JB, Babatunde A, Balogun SU, Lawan A, Haladu SI, Dogara MM, et al. Prevalence and associated risk factors of urinary schistosomiasis among primary school pupils in the Jidawa and Zobiya communities of Jigawa State, Nigeria. Ann Glob Health. 2022;88(1):71. https://doi.org/10.5334/aogh.3704

13. Tasiu Y. Evaluation of the status of water resources and infrastructure for community development in Warwade, Dutse, Jigawastate, Nigeria. Journal. fudutsinma.edu.ng. c2018 [cited 2021 Nov 03]. Available from: http://journal.fudutsinma.edu.ng/index.php/fjs/article/view/436/0

14. Thrustfield M. Veterinary epidemiology. 3rd ed. Oxford: Blackwell Science Ltd; 2007. p. 233–250.

15. Dogara MM, Ahmad S, Balogun BJ, Dawaki SS, Mustapha MB, Abdurrahman AU, et al. Schistosomiasis and associated risk factors among school-aged children in northern Nigeria. Int J Transl Med Res Public Health. 2020;4(2):103–111. https://doi.org/10.21106/ijtmrph.146

16. Sa'idu HI, Shiaka GP, Balogun JB. Prevalence and pattern of severe malaria among children in two general hospitals, Jigawa State-Nigeria. Fudma J Sci. 2021;5(2):511–518. Available from: https://fjs.fudutsinma.edu.ng/index.php/fjs/article/download/664/514/

17. Murphy JF. Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity. Vitamin and mineral nutrition information system. Geneva: World Health Organization; 2011. Available from: https://apps.who.int/iris/bitstream/handle/10665/85839/WHO_NMH_NHD_MNM_11.1_eng.pdf

18. Cheesbrough M. District laboratory practice in tropical countries: Part 2. Cambridge: Cambridge University Press; 2005. p. 236–295.

19. World Health Organization. Bench aids for the diagnosis of intestinal parasites. Geneva: World Health Organization; 2019. Plate 4–11.

20. Greenwood BM, Armstrong JR. Comparison of two simple methods for determining malaria parasite density. Trans R Soc. 1991;85(2):186–188. https://doi.org/10.1016/0035-9203(91)90015-q

21. World Health Organization. WHO child growth standards: Length/height-for-age, weight-for-age, weight-for-length, weight-for-height and body mass index-for-age: Methods and development. Geneva: World Health Organization; 2006.

22. Blössner M, Siyam A, Borghi E, Onyango A, De Onis M. WHO AnthroPlus for personal computers manual: Software for assessing growth of the world's children and adolescents. Geneva: World Health Organization; 2009.

23. Ezeamama AE, Bustinduy AL, Nkwata AK, Martinez L, Pabalan N, Boivin MJ, et al. Cognitive deficits and educational loss in children with schistosome infection — a systematic review and meta-analysis. PLoS Negl Trop Dis. 2018;12(1):e0005524. https://doi.org/10.1371/journal.pntd.0005524

24. Miller K, Lori J, Liu X, Boivin M, Giordani B. The cognitive burden of severe malaria in the Ugandan classroom and the effects of a computerized intervention. Appl Nurs Res. 2022;63:151551. https://doi.org/10.1016/j.apnr.2021.151551

25. Stecher CW, Sacko M, Madsen H, Wilson S, Wejse C, Keita AD, et al. Anemia and growth retardation associated with *Schistosoma haematobium* infection in Mali: A possible subtle impact of a neglected tropical disease. Trans R Soc. 2017;111(4):144–153. https://doi.org/10.1093/trstmh/trx037

26. Alhaji GK, Dogara MM, Balogun JB, Abubakaar MM, Sufi MA, Dawaki SS, et al. Prevalence of schistosomiasis in Warwade community, Jigawa State, Nigeria. DUJOPAS. 2002;7(3b):10–23. https://doi.org/10.4314/dujopas.v7i3b.2

27. David J, Panda SM, Samaila AB. Epidemiological study of schistosomiasis in Basirka, Gwaram local government area, Jigawa State, Nigeria. J Pure Appl Sci. 2021;21(1):144–157. https://doi.org/10.5455/sf.98312

28. Awosolu OB, Shariman YZ, Haziqah MT F, Olusi TA. Will Nigerians win the war against urinary schistosomiasis? Prevalence, intensity, risk factors and knowledge assessment among some rural communities in Southwestern Nigeria. Pathogens. 2020;9(2):128. https://doi.org/10.3390/pathogens9020128

29. Amaechi EC. Urinary schistosomiasis among school age children in some rural communities of Abia State, South Eastern Nigeria. Anim Res Int. 2014;11(2):1953–1957. https://www.unn.edu.ng/wp-content/uploads/2016/06/5_-Amaechi.pdf

30. Onifade OE, Oniya MO. Prevalence of urinary schistosomiasis and efficacy of praziquantel: A case study of school pupils in Oke-Igbo, Ondo state, Nigeria. J Epidemiol. 2018;95:13. https://doi.org/10.1016/j.parepi.2016.03.006

31. Afiukwa FN, Nwele DE, Uguru OE, Ibiam GA, Onwe CS, Ikpo AU, et al. Transmission dynamics of urogenital schistosomiasis in the rural community of Ebonyi State, South Eastern Nigeria. Parasitol Res. 2019;2019, Art. #7596069. https://doi.org/10.1155/2019/7596069

32. Joof E, Sanyang AM, Camara Y, Sey AP, Baldeh I, Jah SL, et al. Prevalence and risk factors of schistosomiasis among primary school children in four selected regions of The Gambia. PLoS Negl Trop. Dis. 2021;15(5):e0009380. https://doi.org/10.1371/journal.pntd.0009380

33. Otuneme OG, Obebe OO, Sajobi TT, Akinleye WA, Faloye TG. Prevalence of Schistosomiasis in a neglected community, South western Nigeria at two points in time, spaced three years apart. Afr Health Sci. 2019;19(1):1338–1345. https://doi.org/10.4314/ahs.v19i1.5

34. Hassan AO, Amoo AO, Deji-Agboola AM, Akinwale OP, Gyang PV, Adeleke MA. Current status of urinary schistosomiasis in communities around the Erinle and Eko-Ende Dams and the implications for schistosomiasis control in Nigeria. S Afr J Infect Dis. 2014;29(4):137–140. https://doi.org/10.1080/23120053.2014.11441588

35. Yapi YG, Briët OJ, Diabate S, Vounatsou P, Akodo E, Tanner M, et al. Rice irrigation and schistosomiasis in savannah and forest areas of Côte d'Ivoire. Acta Trop. 2005;93(2):201–211. https://doi.org/10.1016/j.actatropica.2004.11.005

36. Uweh PO, Omudu EA, Onah IE. Current status of schistosomiasis amongst school children in Igedeland, Benue State, Nigeria. Niger Ann Pure Appl Sci. 2015;6:16–21. https://doi.org/10.46912/napas.3

37. Rikoto JA, Danladi YK. Urinary schistosomiasis among school age children of Sarkawa fishing community in Yauri, Kebbi State. Equ J Sci Technol. 2013;1:1–5. Available from: http://equijost.com/?mno=302643615

38. Dogara MM, Ocheje AJ. Prevalence of malaria and risk factors among patients attending Dutse General Hospital, Jigawa State, Nigeria. Int J Pub Environ Health. 2016;11:270–277. Available from: https://journalissues.org/irjpeh/wp-content/uploads/sites/8/2016/11/Ocheje-and-dogara.pdf

39. Morenikeji OA, Eleng IE, Atanda OS, Oyeyemi OT. Renal related disorders in concomitant *Schistosoma haematobium–Plasmodium falciparum* infection among children in a rural community of Nigeria. J Infect Public Health. 2016;9(2):136–142. https://doi.org/10.1016/j.jiph.2015.06.013

40. Oladele VS, Awobode HO, Anumudu CI. Subtle morbidities associated with malaria co-infection with schistosomiaisis among children in South-West Nigeria. Afr J Med Med Sci. 2014;43Suppl:125–135.

41. Nyarko R, Torpey K, Ankomah A. *Schistosoma haematobium*, *Plasmodium falciparum* infection and anaemia in children in Accra, Ghana. Trop Dis Travel Med Vaccines. 2018;4(1):1–6. https://doi.org/10.1186/s40794-018-0063-7

42. Deribew K, Tekeste Z, Petros B. Urinary schistosomiasis and malaria associated anemia in Ethiopia. Asian Pac J Trop Biomed. 2013;3(4):307–310. https://doi.org/10.1016/s2221-1691(13)60068-4

43. Lyke KE, Dabo A, Arama C, Diarra I, Plowe CV, Doumbo OK, et al. Long-term maintenance of CD4 T-cell memory responses to malaria antigens in Malian children coinfected with *Schistosoma haematobium*. Front Immunol. 2018;8:1995. https://doi.org/10.3389/fimmu.2017.01995

44. Inyang-Etoh PC, Agorye AH, Okpokam DC, Opara-Osuoha U. Prevalence of malaria and intestinal parasites coinfections and their effects on haemoglobin levels among school aged children in Bebuatsuan clan, Obudu, cross River State, Nigeria. J Med Allied Sci. 2017;7(2):92–98. https://doi.org/10.5455/jmas.259017

45. Konaté D, Diawara SI, Touré M, Diakité SA, Guindo A, Traoré K, et al. Effect of routine seasonal malaria chemoprevention on malaria trends in children under 5 years in Dangassa. Mali Malar J. 2020;19(1):1–6. https://doi.org/10.1186/s12936-020-03202-y

46. Starck T, Bulstra CA, Tinto H, Rouamba T, Sie A, Jaenisch T, et al. The effect of malaria on haemoglobin concentrations: A nationally representative household fixed-effects study of 17,599 children under 5 years of age in Burkina Faso. Malar J. 2021;20(1), Art. #416. https://doi.org/10.1186/s12936-021-03948-z

47. Ehiem RC, Nanse FA, Adu-Frimpong M, Mills-Robertson FC. Parasitaemia estimation and prediction of hepatocellular dysfunction among Ghanaian children with acute malaria using haemoglobin levels. Heliyon. 2021;7(7):e07445. https://doi.org/10.1016/j.heliyon.2021.e07445

48. Kinung'hi SM, Mazigo HD, Dunne DW, Kepha S, Kaatano G, Kishamawe C, et al. Coinfection of intestinal schistosomiasis and malaria and association with haemoglobin levels and nutritional status in school children in Mara region, Northwestern Tanzania: A cross-sectional exploratory study. BMC Res Notes. 2017;10(1), Art. #583. https://doi.org/10.1186/s13104-017-2904-2

49. Edosomwan EU, Evbuomwan IO, Agbalalah C, Dahunsi SO, Abhulimhen-Iyoha BI. Malaria coinfection with neglected tropical diseases (NTDs) in children at internally displaced persons (IDP) camp in Benin City, Nigeria. Heliyon. 2020;6(8):e04604. https://doi.org/10.1016/j.heliyon.2020.e04604

50. De Onis M, Blossner M, World Health Organization. WHO global database on child growth and malnutrition. Geneva: World Health Organization; 1997. Available from: https://apps.who.int/iris/bitstream/handle/10665/63750/WHO_NUT_97.4.pdf%3bjsessionid%3dE8549D8861C7DC1EA9079BF6B8F98AD5%3fsequence%3d1

51. Munisi DZ, Buza J, Mpolya EA, Kinung'hi SM. *Schistosoma mansoni* infections, undernutrition and anaemia among primary schoolchildren in two onshore villages in Rorya District, North-Western Tanzania. PLoS One. 2016;11(12):e0167122. https://doi.org/10.1371/journal.pone.0167122

52. Ayogu RN, Afiaenyi IC, Madukwe EU, Udenta EA. Prevalence and predictors of under-nutrition among school children in a rural South-eastern Nigerian community: A cross sectional study. BMC Public Health. 2018;18(1):1–9. https://doi.org/10.1186/s12889-018-5479-5

53. Umeokonkwo AA, Ibekwe MU, Umeokonkwo CD, Okike CO, Ezeanosike OB, Ibe BC. Nutritional status of school age children in Abakaliki metropolis, Ebonyi State, Nigeria. BMC Pediatr. 2020;20(1):1–9. https://doi.org/10.1186/s12887-020-1994-5

54. Ajakaye OG, Ibukunoluwa MR. Prevalence and risk of malaria, anemia and malnutrition among children in IDPs camp in Edo State, Nigeria. Parasite Epidemiol Control. 2020;8:e00127. https://doi.org/10.1016/j.parepi.2019.e00127

55. Sousa-Figueiredo JC, Gamboa D, Pedro JM, Fançony C, Langa AJ, Soares Magalhães RJ, et al. Epidemiology of malaria, schistosomiasis, geohelminths, anemia and malnutrition in the context of a demographic surveillance system in northern Angola. PLoS One. 2012;7(4):e33189. https://doi.org/10.1371/journal.pone.0033189

56. Bustinduy AL, Parraga IM, Thomas CL, Mungai PL, Mutuku F, Muchiri EM, et al. Impact of polyparasitic infections on anemia and undernutrition among Kenyan children living in a *Schistosoma haematobium*-endemic area. Am J Trop Med Hyg. 2013;88(3):433. https://doi.org/10.4269/ajtmh.12-0552

57. Houmsou RS, Amuta EU, Sar TT. Profile of an epidemiological study of urinary schistosomiasis in two local government areas of Benue state, Nigeria. Int J Med Biomed Res. 2012;1(1):39–48. https://www.ajol.info/index.php/ijmbr/article/view/91826

58. Singh K, Muddasiru D, Singh J. Current status of schistosomiasis in Sokoto, Nigeria. Parasite Epidemiol Control. 2016;1(3):239–244. https://doi.org/10.1016/j.parepi.2016.08.003

59. Mafiana CF, Ekpo UF, Ojo DA. Urinary schistosomiasis in preschool children in settlements around Oyan Reservoir in Ogun State, Nigeria: Implications for control. Trop Med Int Health. 2003;8(1):78–82. https://doi.org/10.1046/j.1365-3156.2003.00988.x

60. Gbalégba NG, Silué KD, Ba O, Ba H, Tian-Bi NT, Yapi GY, et al. Prevalence and seasonal transmission of *Schistosoma haematobium* infection among school-aged children in Kaedi town, southern Mauritania. Parasites Vectors. 2017;10(1):1–2. https://doi.org/10.1186/s13071-017-2284-4

AUTHORS:
Khatija Ahmed[1,2]
Mookho Malahleha[1,3]
Thulisile E. Mbatsane[1]
Dineo Thindisa[1]
Veronique C. Bailey[1]
Ishen Seocharan[4]
Athmanundh Dilraj[1]

AFFILIATIONS:
[1]Setshaba Research Centre, Pretoria, South Africa
[2]Department of Medical Microbiology, University of Pretoria, Pretoria, South Africa
[3]Synergy Biomed Research Institute, East London, South Africa
[4]South African Medical Research Council, Durban, South Africa

CORRESPONDENCE TO:
Athmanundh Dilraj

EMAIL:
ADilraj@setshaba.org.za

DATES:
Received: 21 Apr. 2022
Revised: 09 Oct. 2022
Accepted: 15 Mar. 2023
Published: 08 Aug. 2023

HOW TO CITE:
Ahmed K, Malahleha M, Mbatsane TE, Thindisa D, Bailey VC, Seocharan I, et al. Addressing missed visits to improve retention of young South African women in clinical trials. S Afr J Sci. 2023;119(7/8), Art. #13809. https://doi.org/10.17159/sajs.2023/13809

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

DATA AVAILABILITY:
☐ Open data set
☐ All data included
☒ On request from author(s)
☐ Not available
☐ Not applicable

EDITOR:
Pascal Bessong

KEYWORDS:
clinical trials, retention, protocol adherence, missed visits

FUNDING:
Bill and Melinda Gates Foundation (OPP1032115), US Agency for International Development (AID-OAA-A-15–00045), Swedish International Development Cooperation Agency (2017/762965–0), South Africa Medical Research Council, United Nations Population Fund

Addressing missed visits to improve retention of young South African women in clinical trials

In clinical trials, a vital protocol requirement for participants is adherence to scheduled visits. A substantial number of missed visits and the resultant missing data could affect generalisability of the findings and undermine the scientific conclusions. We aimed to investigate the extent of and reasons for missed visits in the Evidence for Contraceptive Options and HIV Outcomes (ECHO) trial in order to optimise recruitment and retention practices. Despite being a multi-country study, we investigated missed visits only at Setshaba Research Centre in Soshanguve, Tshwane, South Africa. Of 810 participants enrolled at Setshaba Research Centre, 94 (11.6%) participants missed visits and 231 missed visits were recorded. Of the 94 participants who missed visits, 53 (56.4%) missed at least two visits; 37 (39.4%) missed three or more visits, and of these, 32 (86.5%) missed at least two visits for the same reason. Overall, the main reasons for missed visits were: participant had to work (60; 26.0%), unable to contact participant (60; 26.0%), participant relocated (32; 13.9%), and participant travelled out of area (23; 10%). The large proportion of participants who missed two or more visits indicates that participants who miss a single visit are likely to miss even more, often for the same reason. Site staff need to be vigilant to detect any trends in missed visits early and innovative in developing personalised strategies to minimise missed visits and retain participants until completion of their scheduled visits.

Significance:

- Despite trial site staff developing strategies to minimise missed visits, they will not be able to anticipate all scenarios.
- Participants' work commitments, loss of contact with participants, and participants' travel/relocation to distant areas were the main reasons for missing visits, and site staff need to consider the potential for these to arise during the course of the study when assessing potential participants at enrolment and at each follow-up visit.
- Case report forms designed for multi-country studies should be adapted to reflect the most likely reasons for missed visits for the local situation, so that trends in missed visits can be identified and addressed early.

Introduction

High participant retention is a critical element of clinical trials which can only be achieved by ensuring that participants comply with protocol requirements. A vital protocol requirement for participants is adherence to their scheduled visits. Retention of enrolled subjects is essential for both scientific and economic reasons. Missed visits can not only compromise the safety of participants but also impact on the study data and trial outcomes. Substantial instances of missed visits and the resultant missing data are serious problems and can affect generalisability of the results, significantly bias the results, reduce study power and undermine the scientific trustworthiness of causal conclusions from clinical trials.[1,2] Keeping participants in a trial ultimately helps keep a study on track, saving time, money and resources in the process.[3]

Missed visits are frequently unintentional and could be due to factors outside a participant's control. Women, in particular, perform multiple social roles such as being members of kinship networks, wives, girlfriends, caregivers, income earners, and scholars. Fulfilling these roles often compromises their ability to attend clinic visits.[4] Other reasons provided by participants for missing visits were being out of the study area, being "busy" or unavailable because of work, adverse events, not wanting to undergo HIV testing at each quarterly visit, long waiting periods at the clinics, negative rumours in the community regarding the study, disapproval of the study by partners and/or parents, relocation for employment, financial constraints, forgetting visits, incarceration, and unstable housing.[3,5-8] However, most of the studies in which these reasons were given were conducted over a decade ago and were outside of South Africa. Reasons for missing visits are not universal and differ from place to place and over time. Therefore, there is a need to conduct such a study in South Africa.

An opportunity to investigate reasons for missed visits at a South African research site arose in the Evidence for Contraceptive Options and HIV Outcomes (ECHO) study – a randomised clinical trial which compared HIV incidence and contraceptive benefits in women using depot medroxyprogesterone acetate, levonorgestrel implant, and copper intrauterine devices.[9] The study site, Setshaba Research Centre, was one of 12 sites across eastern and southern Africa that participated in this study. Setshaba Research Centre is situated in Soshanguve, which is about 100 km from Johannesburg in the district of Tshwane and province of Gauteng, South Africa. The catchment area is made up of peri-urban and informal settlements and is densely populated with a total population of close to a million people. The centre is located in an area of high disease burden, especially HIV/TB, with the incidence of HIV in studies among women conducted at Setshaba Research Centre ranging from 3–6 per 100 person

years. The community has a high proportion of residents who have not completed secondary schooling, a high unemployment rate, and is poorly resourced, which influence the social, economic and structural factors that contribute to the high disease burden.[10,11] It is against this backdrop of demographic and socio-economic factors that we aimed to investigate the extent of missed visits and the reasons thereof in the ECHO trial, in order to revise and optimise recruitment and retention practices for future studies.

Methods

This study involved a secondary analysis of data from the ECHO study. The 12 sites in the ECHO study were from across South Africa, Kenya, Eswatini, and Zambia. A total of 7830 HIV-uninfected women between 16 and 35 years of age were enrolled in the study. Despite this being a multi-country study, the aspect on missed visits was investigated only at Setshaba Research Centre, Soshanguve, in Tshwane, South Africa. Setshaba Research Centre's site-specific data for screening, enrolment and missed visits were obtained via the ECHO Research Manager, International Clinical Research Center, Department of Global Health, University of Washington, USA. Setshaba Research Centre enrolled 810 participants in 2016–2017 and they were followed up at 1, 3, 6, 9, 12, 15 and 18 months.

During the ECHO trial, Setshaba Research Centre implemented standard retention practices developed from experience in previous trials to minimise missed visits and maximise retention efforts. Relevant processes were outlined to participants prior to enrolment and reinforced during the course of the study. These included, but were not limited to, use of appointment cards, reminder calls, visit trackers and short message services (SMSs) sent to the participants, after-hour and weekend clinic visits based on participants' needs, home visits, clinic waiting area adherence discussions and collection of locator information during screening, verified and updated, if required, at each subsequent visit. Retention efforts were supplemented with an innovative strategy of quarterly informal social participant events, called 'chilling sessions', for participants to engage with staff and other participants in a relaxed setting, which provided a platform for education and information sharing as well as for addressing participant concerns, myths and fears.

When a participant missed a scheduled visit according to the visit window, the Missed Visit case report form (CRF) was used to document the relevant details. This form was sent to the data centre only once the visit window had closed and it was confirmed that the participant had missed the visit. The Missed Visit study-specific CRF contained the following information: (1) reason for missed visit (specified as follows: Unable to contact participant; Unable to schedule appointment within allowable window; Participant refused visit; Participant incarcerated; Participant admitted to a healthcare facility; Participant travelled out of area; Participant forgot; Participant did not have money; and Other (Specify)); (2) steps taken to address the missed visit (corrective action plan); and (3) additional comments.

For this sub-study, the data were analysed using descriptive statistics. The information provided by the data centre was analysed using Microsoft Excel for Office 365 and Epi Info 7. Data from the Missed Visit study-specific CRF for all those who missed visits were analysed to determine the proportion of participants who missed visits and the frequency and reasons for missing visits. Data from the screening and enrolment visit CRFs were used to report on the demographics of participants. The results are reported as frequencies and percentages.

With respect to the specified reasons 'Unable to contact participant' and 'Unable to schedule appointments within allowable window', several of these were changed in the analyses to the actual reason why participants missed their visits where staff eventually managed to contact the participant or their alternative contact person, or by conducting home visits as part of their corrective action. 'Phone only rings' as stated under 'Other (Specify)' was recoded and combined in the analyses with the specified reason 'Unable to contact participant'. If an entry recorded under 'Other (Specify)' suggested that the participant could not come to the site because of issues related to work, this was recoded as 'Work commitments'. Staff would attempt to verify the inability of the participant to come to the site because of work commitments by following up with the alternative contact person, if this information had been provided by the participant, or by conducting home visits.

Sociodemographic comparisons were made between those who missed visits and those who did not; a significance level of 0.05 was used. The ECHO study was approved by the South African Health Products Regulatory Authority and ethical clearance was provided by the University of the Witwatersrand Human Research Ethics Committee (FHI 360 Study number 523201; Ethics reference number: 141112). Informed consent was obtained from participants at screening and enrolment.

Results

Of the 810 participants enrolled at Setshaba Research Centre, missed visits were reported for 94 (11.6%) participants over the course of the study. The mean age (±standard deviation) of the 94 participants was 23.5 (±3.7) years; 40 (42.6%) did not complete secondary school, 42 (44.6%) completed secondary school and 12 (12.8%) attended a post-secondary institution; 92 (97.9%) were never married and only 15 (16%) were living with a partner; however, 69 (73.4%) had partner support.

Throughout study participation, 231 missed visits were recorded for the 94 participants. Thus, some participants had multiple missed visits. Of the 94 participants who missed visits, 53 (56.4%) missed at least two visits and 37 (39.4%) missed three or more visits. Of those who missed three or more visits, 32 (86.5%) missed at least two visits for the same reason. The majority (74.9%) of missed visits occurred in the second half of the study. The reasons for the missed visits are presented in Table 1.

Of the 231 missed visits, 93 (40.3%) were for specified reasons and 138 (59.7%) were for 'Other' reasons. The main reasons for missing visits were work commitments (60/231; 26.0%) and being unable to contact participants (60/231; 26%), which together accounted for just over half of all missed visits (Table 1). Additional frequently stated reasons were: 'Relocated' (32/231; 13.9%), 'Travelled out of area' (23/231; 10%), 'Unable to honour visit as promised' (20/231; 8.7%) and 'School commitments' (16/231; 6.9%).

There were no missed visits recorded for some of the specified reasons listed on the CRF, namely: Participant incarcerated, Participant admitted to a healthcare facility, Participant forgot, and Participant did not have money.

Table 1: Reasons for missed visits in the Evidence for Contraceptive Options and HIV Outcomes (ECHO) trial at the Setshaba Research Centre trial site, Soshanguve, South Africa ($N = 231$)

Reason for missed visit	*n*	%
Unable to attend because of work	60	26.0
Unable to contact the participant	60	26.0
Participant relocated	32	13.9
Travelled out of the area	23	10.0
Unable to honour visit as promised	20	8.7
School commitments	16	6.9
Participants wanting to withdraw from study participation	5	2.1
Other commitments	4	1.7
Participants wanting to have a baby	3	1.3
Participants experiencing side effects	3	1.3
Unable to schedule appointments within the allowable window	2	0.9
Other	3	1.3

The majority (63/94; 67.0%) of participants who missed visits were younger than 25 years old. There were no significant differences between those who missed visits and those who did not with respect to receiving financial support from partners, earning an income, education (completed secondary education or more versus primary education or less) or marital status (all $p>0.05$).

Discussion

The large proportion of participants who missed two or more visits indicates that participants who miss a visit are likely to miss even more visits. Furthermore, most participants who missed multiple visits missed more than once for the same reason. It is therefore important to identify those participants and their missed visit trends early in the trial so that personalised strategies to minimise missed visits can be devised and implemented. Corrective action to get participants to adhere to their visit schedule was taken as deemed appropriate for the reason of the missed visit. Some of the actions taken were: (1) study staff performed home visits; (2) transport to the research site was arranged; (3) some participants who had relocated but were still interested in continuing with the study were offered extra reimbursement to enable them to fulfil their visit; and (4) letters to confirm study visit attendance were also issued, if requested by the participant.

Despite strategies in place to minimise missed visits at the time of this study, site staff nevertheless faced some challenges. The main reason given for over a quarter of missed visits was that participants were unable to take time off from work, including over weekends. In the VOICE Study conducted in Johannesburg, South Africa, a random sub-sample of 102 female trial participants, 18 to 40 years of age, were interviewed to explore factors that shaped trial participation and adherence to study protocols. It was found that clinic visits interfered with household responsibilities as well as work or school, and trial participants encountered difficulties when trying to balance their commitment to these different activities.[4] Therefore, study staff need to be thorough at screening in establishing whether work commitments may make it difficult for participants to adhere to their visit schedule. Even then participants can become employed only after screening or enrolment or might change jobs, which can subsequently impact on their ability to adhere to the visit schedule. Moreover, we found that some of the participants who relocated did so because their place of employment was far from the trial site and this could not be anticipated at enrolment. Soshanguve is largely a residential area with little opportunity for employment. Thus, many local residents seek employment outside the catchment area in the cities of Johannesburg and Tshwane or in neighbouring provinces. It is therefore not surprising that a large proportion of missed visits were due to participants relocating or travelling out of the area, which could have been to seek employment or for personal reasons. Therefore, participants need to be informed at the outset and reminded during the course of the study to notify study staff of any change in their employment status so that staff can proactively assess the impact thereof on adhering to study visits and plan accordingly. In addition, if an enrolled participant relocates, it is important that the site staff are notified and arrangements can be made to accommodate their visits with extra reimbursement for travel costs and/or weekend visits.

The development and implementation of retention strategies is an ongoing and dynamic process. Yet despite site staff implementing retention measures, certain circumstances beyond their immediate control can derail plans. This was clearly illustrated by the COVID-19 pandemic which highlighted the potential for missed visits due to lockdown measures such as restricted inter-provincial travel, or reluctance by participants to use public transport to come to the research site because of the fear of exposure to the coronavirus from other symptomatic or asymptomatic passengers, particularly during each wave of the pandemic. Thus, site staff had to develop innovative means in a short time to minimise missed visits and maintain high retention to keep trials on track. Clinical trial implementation has been rigid in requiring lengthy in-person visits, which have both cost and inconvenience implications for participants and the research team alike. As a result of the COVID-19 pandemic, sponsors and researchers were forced to become more adaptable and to consider collecting data through other mechanisms, such as telephonically and online, e.g. if no specimens need to be collected, technology can be employed for virtual visits.

Inability to make contact with participants was the other major reason for missed visits. A study team member would generally attempt to contact participants to remind them of their upcoming visit; however, participants would sometimes fail to respond to the call, leading them to not honour their visit. A subsequent study conducted at this site suggested that this could partly be due to participants experiencing physical, financial and technological challenges with their cell phones, such as the battery running flat frequently, their cell phones being unreliable, or lack of data due to high data costs, poor network signal at home and use of applications.[12] Additional strategies implemented by site staff included sending SMSs alerting them of their visit and calling after hours. It was also perceived that some participants may not have answered their phones for fear of the site's number being recognised by an unsupportive partner/family member or employer. Therefore, site staff would sometimes call participants from a number other than the site contact number stated in the consent form, such as a staff member's personal phone or another site office number.

This study also highlights the need for greater reflection on the design of CRFs. It is noteworthy that an inability to honour the visit due to work was recorded under 'Other reasons', as this circumstance accounted for over a quarter of all missed visits. Work commitments should have been anticipated as a potential reason for being unable to honour a scheduled visit. Therefore, this should be listed as a specified reason for missing a visit, particularly if the study population includes employed people. It is more difficult to keep track of reasons for missed visits if these are recorded under 'Other reasons'. Nonetheless, in the absence of regular reports on missed visits from the data management centre, it remains the responsibility of the site research team to discuss each missed visit when this becomes known and to take corrective action. CRFs also need to be modified for the local situation by removing those specified reasons for which no entries were made, as these could be recorded under 'Other reasons' if they do occur in future. This highlights the need for experienced site research staff to be consulted during protocol and CRF design in multi-site trials.

Interestingly, a higher proportion of younger participants missed visits. This could possibly be due to the schoolchildren missing visits and skewing the data towards younger age. It is understandable that participants who are attending school would find it difficult to honour their visit at certain times, particularly during examinations, as this would be their priority. From experience with previous studies, site staff were aware of the difficulty that scholars faced to honour their visit during examinations. Bearing this in mind, the recruitment team was informed that school-going potential participants should not be included in the study. In addition, while this was not a protocol-exclusion criterion, study staff responsible for screening and enrolment tried to exclude participants who were attending school by asking them if they were attending high school before the informed consent form was signed. If participants did not disclose this information, they were then enrolled in the study. Therefore, despite school attendance not being an exclusion criterion, young participants need to be scrutinised further to confirm that they are not attending school to avoid the potential problems of missed visits due to schooling priorities.

The finding that three-quarters of the missed visits occurred in the second half of the trial could indicate participation fatigue, as participants were required to be in the study for 18 months. Other factors, such as side effects and personal commitments, played a minor role in missed visits. The method of contraception did not impact upon retention; at 18 months of follow-up, retention was similar in the three arms of the trial (91.3%, 94.3% and 94.7%).[9]

Our research site conducts other clinical trials with female participants of similar demographic and socio-economic status. Furthermore, other research centres in South Africa conduct clinical trials in peri-urban areas that are similar to Soshanguve. Thus, the findings of this study are relevant to other clinical trials and to research centres in South Africa

as a whole, as well as to clinical trials in other low-to-middle-income countries.

Limitations of the study

The information on missed visits, more especially that due to unspecified 'Other reasons', was not recorded systematically. Each case required careful analysis to determine the reason for the missed visit. Depending on the completeness of information and the interpretation of the study staff, there could have been misclassification of the reason. For example, many participants could not be contacted, and these missed visits were recorded under the specified reason of 'Unable to contact participant'; however, for a substantial number of participants, 'phone only rings' was the reason recorded under 'Other (Specify)'. Furthermore, for some participants for whom initial reasons for missed visits were recorded as 'Unable to contact participant' or 'Phone only rings' or 'Unable to honour visit as promised', the reasons were changed to 'Participant relocated' or 'Unable to attend because of work' or 'School commitments' when staff eventually managed to contact the participant or their alternative contact person, or conducted home visits as part of their corrective action. Therefore, it is likely that for more participants who could not be contacted, the actual reasons for missed visits could be relocation, work or school commitments. The need for recoding of some reasons for missed visits captured on CRFs in the 'Other' field/option could have been detected during earlier site quality-control activities and staff could have been trained/retrained appropriately on CRF completion to avoid having to address this issue during analysis.

Another limitation of the study is that information on the corrective actions taken based on the information gained about non-attendance was not captured systematically, making it difficult to quantify the extent of the improvement in retention. However, the retention rate at the end of the study was 92%, which was above the standard of 90%.

Conclusions

Retention of participants in clinical trials is largely dependent upon recruitment of participants who will follow protocol requirements, with adherence to scheduled visits being one such requirement. Work commitments, inability to make contact with participants, and travel or relocation to areas distant from the research site were the main reasons for missing visits. Furthermore, we found that participants who missed a visit were likely to miss even more visits, often for the same reason. Therefore, site staff should be vigilant to detect any trends in missed visits early and develop personalised strategies to minimise missed visits and retain participants until completion of their scheduled visits. Despite trial site staff developing strategies to minimise missed visits, they will not be able to anticipate all scenarios and may still face some challenges. This has been highlighted most recently by the COVID-19 pandemic – an unprecedented event that caused chaos and immense hardship at local, national and global levels, simultaneously placing clinical trials and participants' safety at risk. Thus, sponsors and researchers must move away from rigid in-person study visits and be more flexible in collecting data, such as through virtual or telephonic means, to ensure that study visits are not compromised. In addition, research staff should investigate the different types of technological tools used by participants and the extent of their use to determine which alternative options of data collection would best suit the participants in their catchment area.

Longer duration studies require more innovative approaches and need more individualised retention strategies. Furthermore, case report forms designed for multi-country studies should be adapted to reflect the most likely reasons for missed visits for the local situation, and this may help in identifying trends in missed visits early.

Acknowledgements

This work and the Evidence for Contraceptive Options and HIV Outcomes (ECHO) Study were made possible by the combined generous support of the Bill & Melinda Gates Foundation (grant OPP1032115), the American people through the US Agency for International Development (grant AID-OAA-A-15–00045), the Swedish International Development Cooperation Agency (grant 2017/762965–0), the South African Medical Research Council, and the United Nations Population Fund. Contraceptive supplies were donated by the Government of South Africa and US Agency for International Development. The contents of this paper are solely the responsibility of the authors and do not necessarily reflect the views, decisions, or policies of the institutions with which they are affiliated, the ECHO trial funders, or the supporting governments.

Data availability

Access to the data from this ancillary study of the ECHO Study may be requested through submission of a research concept to the principal author: KAhmed@setshaba.org.za. The concept must include the research question, data requested, analytic methods, and steps taken to ensure ethical use of the data. Access will be granted if the concept is evaluated and found to have scientific merit and if sufficient data protections are in place. As of the time of publication, data access applications are in process with the governing institutional review boards of the ECHO Study to make de-identified data from the primary ECHO data set publicly available.

Competing interests

We have no competing interests to declare.

Authors' contributions

K.A.: Conceptualisation, study design, oversight and leadership, writing – revisions. M.M.: Conceptualisation, study design, project leadership, writing – revisions. T.E.M., D.T.: Data collection, writing – revisions. V.C.B.: Management and coordination of research activities, writing – revisions. I.S.: Data analysis, writing – revisions. A.D.: Conceptualisation, study design, data analysis, writing – initial draft, coordinated revisions and finalised the manuscript. All authors reviewed and approved the final manuscript.

References

1. Fewtrell MS, Kennedy K, Singhal A, Martin RM, Ness A, Hadders-Algra M, et al. How much loss to follow-up is acceptable in long-term randomised trials and prospective studies? Arch Dis Child. 2008;93:458–461. https://doi.org/10.1136/adc.2007.127316
2. Singhal R, Rana R. Intricacy of missing data in clinical trials: Deterrence and management. Int J Appl Basic Med Res. 2014; 4:S2–S5.
3. Advarra. Retention in clinical trials: Keeping patients on protocols [webpage on the Internet]. c2021 [cited 2022 Apr 21]. Available from: https://www.advarra.com/resource-library/retention-in-clinical-trials-keeping-patients-on-protocols/
4. Magazi B, Stadler J, Delany-Moretlwe S, Montgomery E, Mathebula F, Hartmann M, et al. Influences on visit retention in clinical trials: Insights from qualitative research during the VOICE trial in Johannesburg, South Africa. BMC Women's Health 2014;14:88. https://doi.org/10.1186/1472-6874-14-88
5. Gappoo S, Montgomery ET, Gerdts C, Naidoo S, Chidanyika A, Nkala B, et al. Novel strategies implemented to ensure high participant retention rates in a community based HIV prevention effectiveness trial in South Africa and Zimbabwe. Contemp Clin Trials. 2009;30:411–418. https://doi.org/10.1016/j.cct.2009.05.002
6. Liew SM, Tong SF, Lee VKM, Ng CJ, Leong KC, Teng CL. Text messaging reminders to reduce non-attendance in chronic disease follow-up: A clinical trial. Br J Gen Pract. 2009;59;916–920. https://doi.org/10.3399/bjgp09X472250
7. Raya PM, García RR, De Sánchez AT. Patients' retention strategies in clinical trials. Endocrinol Nutr. 2015;62:475–477. https://doi.org/10.1016/j.endonu.2015.08.001
8. Haley DF, Lucas J, Golin CE, Wang J, Hughes JP, Emel L, et al. On behalf of the HPTN 064 Study Team. Retention strategies and factors associated with missed visits among low income women at increased risk of HIV acquisition in the US (HPTN 064). AIDS Patient Care STDs. 2014;28:206–217. https://doi.org/10.1089/apc.2013.0366

9. Evidence for Contraceptive Options and HIV Outcomes (ECHO) Trial Consortium. HIV incidence among women using intramuscular depot medroxyprogesterone acetate, a copper intrauterine device, or a levonorgestrel implant for contraception: A randomised, multicentre, open-label trial. Lancet. 2019;394(10195):303–313. https://doi.org/10.1016/S0140-6736(19)31288-7

10. Massyn N, Tanna G, Day C, Ndlovu N. District health barometer: District health profiles 2017/18. Durban: Health Systems Trust; 2018. Available from: https://www.hst.org.za/publications/District%20Health%20Barometers/District%20Health%20Barometer-District%20Health%20Profiles%2020172018.pdf

11. Statistics South Africa (Stats SA). City of Tshwane: Key statistics [webpage on the Internet]. c2022 [cited 2023 Aug 04]. Available from: https://www.statssa.gov.za/?page_id=1021&id=city-of-tshwane-municipality

12. Mapetla K, Malahleha M, Van Niekerk N, Thindisa D, Mpete L, Ahmed K, et al. Establishing communication challenges and preferences among clinical trial participants in an under-resourced setting to improve adherence to study visits and participant retention. Clin Trials. 2022;19:81–85. https://doi.org/10.1177/17407745211062077

AUTHORS:
Jonathan A. Holmes[1]
Jennifer M. Fitchett[2]

AFFILIATIONS:
[1]Environmental Change Research Centre, Department of Geography, University College London, London, UK
[2]School of Geography, Archaeology and Environmental Studies, University of the Witwatersrand, Johannesburg, South Africa

CORRESPONDENCE TO:
Jonathan Holmes

EMAIL:
j.holmes@ucl.ac.uk

DATES:
Received: 02 Oct. 2022
Revised: 16 Feb. 2023
Accepted: 14 Mar. 2023
Published: 08 Aug. 2023

HOW TO CITE:
Holmes JA, Fitchett JM. Construction and testing of a low-cost device for the collection of rainfall samples destined for stable isotope analysis. S Afr J Sci. 2023;119(7/8), Art. #14914. https://doi.org/10.17159/sajs.2023/14914

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

DATA AVAILABILITY:
☐ Open data set
☒ All data included
☐ On request from author(s)
☐ Not available
☐ Not applicable

EDITOR:
Teresa Coutinho

KEYWORDS:
precipitation, oxygen isotopes, hydrogen isotopes, rainfall collector, southern Africa

FUNDING:
WITS-UCL Strategic Partnership seed fund

Construction and testing of a low-cost device for the collection of rainfall samples destined for stable isotope analysis

Oxygen- and hydrogen-isotope ratios in rainfall provide important hydroclimatic information, yet despite a global network of rainfall isotope measurements, significant geographical gaps exist in data coverage, with only three long-term stations spanning the southern African region. Project-based, *ad hoc* collections of rainfall for isotope analysis can improve this coverage. However, all rainfall samples that are destined for stable isotope analysis must be collected in such a way to avoid evaporation and resultant isotope fractionation. While such rainwater collectors are available commercially, both the product and shipping are prohibitively costly. We describe the construction of a simple rainfall collector using a design from the literature and materials that are readily available in South African hardware stores. Our rainwater collector can be constructed for the much lower cost of just under ZAR820 in comparison with the cost of ZAR9300 inclusive of shipping from commercial outlets (2022 prices). Our design modifications have the added advantage of portability, with the rainwater collector housed in a bucket with a handle. The device was tested by comparing its performance, in terms of evaporative water loss and isotopic fractionation, with that of an open bottle, using tap water in both cases. Testing confirmed that the collector prevented evaporation over a one-week period, indicating that it is suitable for weekly or more frequent sampling of rainfall. Although the design described was based on materials procured in South Africa, it could easily be adapted for construction elsewhere.

Significance:

- Hydrogen and oxygen isotope composition of rainfall provides valuable climatic information.
- Rainwater collectors for stable isotope samples must prevent evaporation, as evaporation will alter the isotopic signature.
- We describe the construction and testing of a bespoke, low-cost and portable device that can be used to collect rainfall samples destined for oxygen- and hydrogen-isotope analysis without significant evaporation.

Introduction

Oxygen- and hydrogen-isotope ratios ($^{18}O/^{16}O$ and $^{2}H/^{1}H$, respectively) of precipitation are valuable tracers of hydroclimatic processes used extensively to investigate meteorology and climate on short (e.g. hourly[1]) to long (inter-decadal[2]) timescales and on spatial scales ranging from local[3] to global[4].

The Global Network of Isotopes in Precipitation (GNIP) is an extensive database of precipitation isotope data hosted by the International Atomic Energy Agency (IAEA). It currently contains precipitation isotope data for more than 1000 sites across around 125 countries. Most of the measurements relate to integrated monthly precipitation samples[5], although data from more frequent sampling (e.g. daily or event based) are available for some localities both in GNIP, and in other published studies[6,7]. In southern Africa, there are GNIP stations in Pretoria, Cape Town and Windhoek. Precipitation stable-isotope data have been used to investigate moisture sources in a number of studies, including on the Tibetan Plateau[8], in subtropical China[9], the Indian Himalaya[10], North and South America[11,12], Europe[13] and Africa[14].

The climatic heterogeneity and different moisture sources across southern Africa make this an excellent region in which to use water isotope data to track moisture source. The region is characterised by three distinct rainfall zones – a winter-rainfall zone driven by the southern Westerlies, which is constrained to the southwestern tip of the subcontinent, a year-round-rainfall zone that covers much of the southern and western coastlines of South Africa, and the remainder of the region characterised by convective summer rainfall.[15] Total rainfall amount decreases from east to west across the subcontinent, influenced by the warm Agulhas Current off the east coast, and the cold Benguela off the west coast, and modulated by local topography.[16] Synoptic shifts have been detected in the location of the rainfall zones over recent decades, driven by changes in moisture sources.[17] These shifts, particularly in the winter-rainfall zone, have been implicated in the severe 'Day Zero' drought in Cape Town, which spanned 2015–2018.[18] These impacts highlight the importance of monitoring moisture sources for the region, particularly under climate change.

Previous work has demonstrated the value of precipitation isotope data for investigating moisture source. Lekete and Abiye[19] tracked moisture sources from a daily precipitation isotope record from Johannesburg, showing that the isotopic composition of discrete rainfall events was determined by rainfall source and trajectory together with modifications during transport, but correlation with air temperature and rainfall amount was weak. In addition, there are multi-decadal monthly GNIP data from two sites (Cape Town and Pretoria) in mainland South Africa,

and from Windhoek in Namibia.[20] Additional daily or event-based data sets of varying timespan have been published for several other sites from South Africa[21-23] and Namibia[24-25]. Harris et al.[21] also found a weak correlation with air temperature and rainfall amount for Cape Town precipitation, with distinct isotopic signatures for storm, hail and snowfall versus other types of precipitation. Braun et al.[22] noted that the isotopic composition of precipitation in the year-round precipitation zone of the South African south coast reflected complex interactions between temperate and tropical to subtropical air masses. Durowoju et al.[23] found strong seasonal variations in the isotope values of precipitation from the Limpopo Province of South Africa, and noted that recycled moisture from surface water and evapotranspiration had a significant influence on its composition. Kaseke et al.[24,25] found that the isotopic composition of precipitation across Namibia reflected moisture source, but also showed significant local modifications. Despite these studies, the geographical coverage of precipitation data across southern Africa remains sparse, particularly by international standards: for example, there are relatively dense networks in Europe, other parts of Africa, South America and parts of China.[20] Additional data, from either monthly or more frequent sampling at strategic locations, could be invaluable in improving the network of data for southern Africa, to address critical meteorological, climatological and water-resource questions.

The collection of additional data requires an expansion of the network of stations at which precipitation is collected for stable isotope analysis, whether temporary or permanent. Devices for the collection of precipitation destined for stable isotope analysis must be designed to minimise post-collection evaporation of the water, because evaporation leads to enrichment of the heavy isotopes (^{18}O and ^{2}H) because of fractionation. This fractionation in turn leads to the modification of the original precipitation isotope ratios, rendering the data of limited value, particularly for the identification of moisture sources. Event-based precipitation samples, which are generally transferred to sealed containers shortly after the precipitation event, are less susceptible to such modification, although evaporative enrichment may still occur over a few hours, especially in the warm, windy and low humidity conditions prevalent in much of southern Africa. For longer sampling intervals of weeks to months, prevention of evaporation is critical. Different collector designs have been employed to prevent or minimise evaporation.[26] However, these collectors can be expensive to purchase commercially, especially in studies for which collectors need to be installed in multiple locations, and may be impractical in situations where portability is required. Here, we describe the construction of a simple, relatively low-cost rainfall sampler for use in South Africa that uses tube dip-in with pressure equilibration to minimise evaporation.[26] We present the results of the testing of this rainwater collector, confirming its efficacy in preventing evaporation, and thus retaining the stable isotope ratios of the source waters.

Rainwater collector design and construction

We adapted a design described by Gröning et al.[27] by using materials procured locally from builders' merchants and other hardware stores in South Africa, to produce a collector inexpensively (Table 1). Tube dip-in with pressure equilibration[26] involves a collecting funnel to which is attached a small-diameter tube that extends to the base of a collection bottle. When even a small amount of rainwater is collected in the collecting bottle, the outlet of the collection tube is submerged, minimising any evaporative loss of water back out through the collecting funnel. An outlet from the collecting bottle allows pressure equilibration with the outside atmosphere; to prevent water vapour loss through this outlet, a long (10 m) small-diameter (5 mm internal diameter) tube is wound around an inner housing. A version of the above[27] is now also available commercially (www.rainsampler.com) and used by a number of IAEA GNIP stations, but is relatively expensive (over ZAR3500 at the time of writing in September 2022) and the shipping costs to African countries are quite prohibitive (for example, around ZAR5800 to South Africa).

Our collector consists of a funnel and collecting bottle, the dip-in tube extending from the funnel to the bottom of the collecting bottle, an inner housing that surrounds the collecting bottle, and then an outer housing in which the entire construction sits except for the collecting funnel and support (Figure 1A–D). The funnel used here was a Formosa ISO9001 Large Funnel Number 8095 with an internal diameter of 140 mm. A length of 4 mm internal diameter PVC aquarium tubing fixed into the narrow funnel tube with epoxy putty was used for the dip-in or inlet tube. To ensure that the inlet tube reached the bottom of the collecting vessel, we weighted it using three steel washers held in place with epoxy resin. The equilibration tube extends from the outlet in the lid of the collecting bottle and is wound around the inner housing to accommodate the entire 10 m length; it was held in place with duct tape. The inner housing was made from underground drainage pipes used in the building trade, and the outer housing was a 20 L household bin with a lid and handle. The provision of a handle makes the entire construction easily portable, but as the bucket is relatively lightweight, we would recommend placing a heavy weight such as a brick in the base of the bin to hold it in place,

Table 1: Components used in the rain collector construction and their cost

Component	Number required	Cost per item[a]	Total cost
Addis 20 L bucket and lid	1	159.90	159.90
Plastic 140 mm diameter funnel	1	8.90	8.90
Marley underground pipe PVC twin-wall push-fit double socket, 110 mm diameter	2	57.00	114.00
Marley underground pipe PVC plain stopend, 110 mm diameter	2	43.00	86.00
PVC tubing 6 mm (1 m length)	10	15.00	150.00
PVC tubing 5 mm (1 m length)	0.3	15.00	4.50
Epoxy putty, pack	1	55.00	55.00
Epoxy resin[b]	1	150.00	150.00
PVC-pipe adhesive[b]	1	54.00	54.00
Metal washers	3	1.50	4.50
Duct tape, roll[b]	1	29.00	29.00
Total cost per collector			ZAR815.80

[a] *Prices are in South African rand (ZAR), as of June 2022*

[b] *One item is sufficient for several collectors*

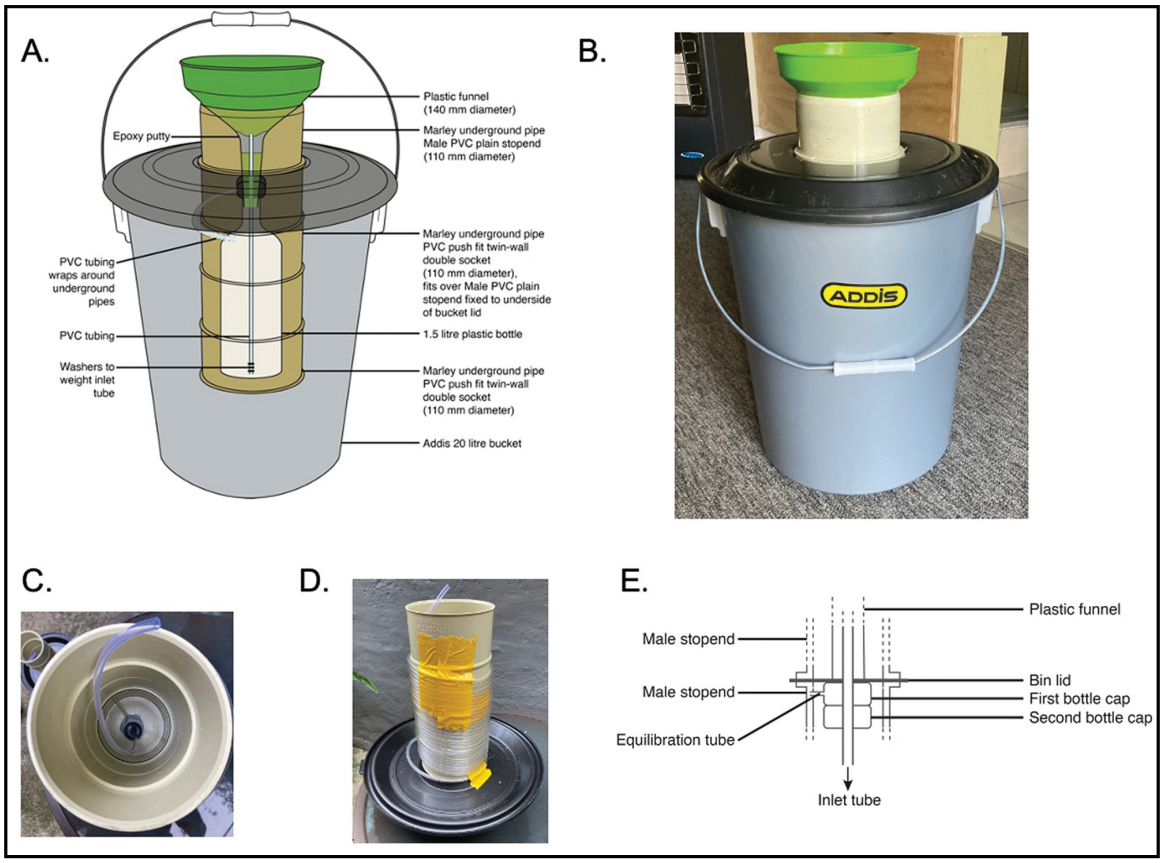


Figure 1: Rain collector design. (A) Diagram of the completed collector, showing component parts. The equilibration tubing is only partially shown, for sake of clarity, but in reality extends around the inner tube, as shown in (D). (B) External view of collector. (C) View from the bottom of the inner tubing showing inlet tube: collecting bottle removed. (D) View of inner tubing with equilibration tubing, fixed to lid, shown inverted. (E) Detail of the bottle-lid construction.

especially in windy environments. The collection bottle was a recycled 1.5 litre sparkling water bottle, which was thoroughly cleaned and dried before use in the rainwater collector. Two lids were used – one to house the collection bottle and a second to accommodate the equilibration tube. These were affixed to the outer housing lid and to each other using epoxy resin (Figure 1E). We used epoxy putty to secure the collection tube into the collecting funnel. The choice of components was dictated both by the required design and by availability in builders' merchants and other hardware stores in Johannesburg, South Africa.

Testing the rainwater collector

We tested the collector prior to installation to ensure that it minimised the effects of evaporation. A 300 mL aliquot of Johannesburg tap water was placed in the collector and 300 mL tap water was placed in an open bottle, and both were left in a well-ventilated indoor space in Johannesburg, with temperatures ranging from nighttime lows of ~6 °C to daytime highs of ~18 °C, for one week spanning 12–18 June 2022. Because rain fell during that week, it was not possible to perform the test outdoors, as both the open bottle and the rainwater collector samples would have been contaminated by the isotopic signature of that rainfall. On the evening of 18 June 2022, water was removed from the collector and the open bottle and transferred to 10 mL polyethylene sample bottles, which were completely filled and sealed with electrical tape to prevent evaporation. Duplicates were collected of each sample. An aliquot of the same tap water collected on 12 June 2022 was also immediately transferred to a 10 mL polyethylene sample bottle on collection, and sealed, as a control. Prior to analysis, 5 mL aliquots of each of the three samples – tap water, rainwater collector water, and the open bottle water – were transferred to glass Thermo™ vials. Samples were then analysed for oxygen and hydrogen isotopes using a Picarro L2130-i Cavity RingDown Spectrometer (CRDS) at the Bloomsbury Environmental Isotope Facility (BEIF), University College London (UCL), UK. The resulting values were expressed in standard delta units, where:

$$\delta\,‰\,VSMOW = \left(\frac{Rsample}{Rstandard} - 1\right) \times 1000 \qquad \text{Equation 1}$$

and δ is $\delta^{18}O$ or $\delta^{2}H$, and correspondingly, R is $^{18}O/^{16}O$ and $^{2}H/^{1}H$, and the standard VSMOW errors (1s) were determined from repeat measurements of each sample as well as determinations of IAEA standards (Table 2).

At the end of the experimental week, 280 mL of water remained in the open water bottle, indicating that 20 mL (or 6.7%) had evaporated. Within the margin of error of reading from parallax, 300 mL of water remained in the rainwater collector bottle, indicating negligible evaporation. These results align with the stable isotope analysis of these samples (Table 2), which for $\delta^{18}O$ were -2.50 ± 0.10 ‰ for the tapwater control, -2.51 ± 0.11 ‰ for water from the rain collector and -2.13 ± 0.10 ‰ for water from the open bottle. Corresponding values for $\delta^{2}H$ were -10.85 ± 1.01 ‰, -10.94 ± 1.01 ‰ and -8.94 ± 1.01 ‰. In summary, the oxygen- and hydrogen-isotope signatures were almost identical, and well within error, for the tap water and water from the rain collector, whereas the water from the open bottle had undergone significant modification.

The test demonstrated convincingly that the rain collector prevented evaporation, given that the isotopic composition of the water held in the rain collector was effectively identical to the control, whereas the water held in the open bottle had undergone significant evaporation, as indicated both by its isotope composition (Figure 2) and loss of volume.

Table 2: Oxygen- and hydrogen-isotope values for the water samples used to test the rain collector

Sample	δ^{18} O	Internal error	External error		δ^{2} H	Internal error	External error
	‰ VSMOW				‰ VSMOW		
Tapwater control	–2.50	0.02	0.10		–10.85	0.17	1.00
Rain collector	–2.51	0.04	0.10		–10.94	0.15	1.00
Open bottle	–2.13	0.03	0.10		–8.94	0.13	1.00

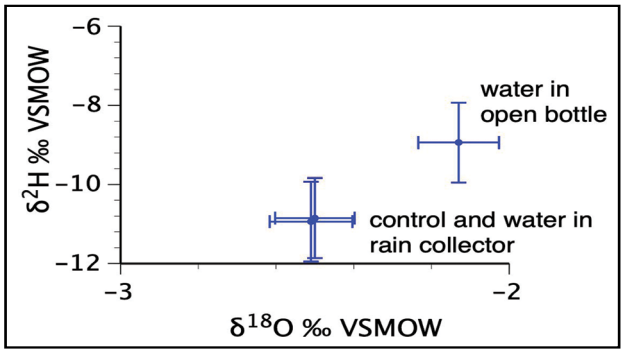


Figure 2: Results of rain gauge test. Errors (1s) are derived from six repeat determinations of each water sample (internal error) as well as determinations of a water standard during the run (external error).

The test was performed for the period of one week during winter when humidity in Johannesburg is at its minimum. It therefore demonstrates the efficacy of the rainwater collector in preventing evaporation over the period for which sampling from these rainwater collectors is intended. Should the design be used for monthly sampling, a testing period of one month would be advised to confirm that evaporation and the effects thereof on fractionation are successfully prevented over a longer period.

Conclusion

We have described the construction of a low-cost, portable precipitation collector that prevents evaporation and is therefore suitable for collection of precipitation samples destined for stable isotope analysis. Our testing has shown that the design prevents evaporation of water kept in the collector for one week, our intended sampling frequency. We have now deployed samplers in three locations in South Africa (Pretoria, Bloemfontein and Cape Town) spanning two of the rainfall seasonality zones, and easterly and westerly derived moisture, and these will be emptied for isotope analysis weekly over at least the year from August 2022.

Additional testing would be required to confirm the suitability of this specific design for preventing evaporation over longer periods, although Gröning et al.[27] reported that water held in their tube dip-in with pressure equilibration sampler underwent minimal evaporative loss over almost one year. Although our rain collector was constructed using materials procured in southern Africa, similar collectors could easily be constructed at low cost in other countries using comparable materials, although careful testing prior to deployment would be advisable.

Acknowledgements

We thank Prof. Anne Fitchett and Mr William Mainganye for access to equipment and assistance in the University of the Witwatersrand Civil and Environmental Engineering Laboratory for the drilling of holes in the component parts of the rainwater collector. We thank Cath D'Alton for the production of the cross-section diagram of the rainwater collector and Anne-Lise Jourdan for undertaking stable isotope analyses. This project was supported by the 2021 WITS-UCL Strategic Partnership seed fund.

Authors' contributions

J.A.H. and J.M.F. contributed equally to conceptualisation, methodology, data collection, sample analysis, data analysis, and funding acquisition. J.A.H. led the writing of this paper, including the initial draft and revisions.

References

1. Good SP, Mallia DV, Lin JC, Bowen GJ. Stable isotope analysis of precipitation samples obtained via crowdsourcing reveals the spatiotemporal evolution of Superstorm Sandy. PLoS ONE. 2014;9(3):e91117. https://doi.org/10.1371/journal.pone.0091117
2. Vystavna Y, Matiatos I, Wassenaar LI. Temperature and precipitation effects on the isotopic composition of global precipitation reveal long-term climate dynamics. Sci Rep. 2021;11:18503. https://doi.org/10.1038/s41598-021-98094-6
3. Jones MD, Leng MJ, Arrowsmith C, Deuchar C, Hodgson J, Dawson T. Local δ18O and δ2H variability in UK rainfall. Hydrol Earth Syst Sci. 2007;4:2403–2423. https://doi.org/10.5194/hessd-4-2403-2007
4. Rozanski K, Araguás-Araguás L, Gonfiantini R. Isotopic patterns in modern global precipitation. In: Swart PK, Lohmann KC, McKenzie J, Savin S, editors. Climate change in continental isotopic records. Geophysical Monograph Series Vol. 78. Washington DC: American Geophysical Union; 1993. p.1–36. https://doi.org/10.1029/gm078p0001
5. Terzer S, Wassenaar LI, Araguás-Araguás LJ, Aggarwal PK. Global isoscapes for δ^{18}O and δ^{2}H in precipitation: Improved prediction using regionalized climatic regression models. Hydrol Earth Syst Sci. 2013;17:4713–4728. https://doi.org/10.5194/hess-17-4713-2013
6. Darling WG, Talbot JC. The O and H stable isotopic composition of fresh waters in the British Isles. 1. Rainfall. Hydrol Earth Syst Sci, 2003;7:163–181. https://doi.org/10.5194/hess-7-183-2003
7. Tian C, Wang L. Stable isotope variations of daily precipitation from 2014–2018 in the central United States. Sci Data. 2019;6:190018. https://doi.org/10.1038/sdata.2019.18
8. Wu H, Zhang X, Xiaoyan L, Li G, Huang Y. Seasonal variations of deuterium and oxygen-18 isotopes and their response to moisture source for precipitation events in the subtropical monsoon region. Hydrol Process. 2015;29:90–102. https://doi.org/10.1002/hyp.10132
9. Yu WS, Yao TD, Tian LD, Ma YM, Naoyuki K, IchiyanagiK, et al. Stable isotope variations in precipitation and moisture trajectories on the western Tibetan Plateau, China. Arct Antarct Alp Res. 2007;39:688–693. https://doi.org/10.1657/1523-0430(07-511)[YU]2.0.CO;2
10. Jeelani G, Deshpande RD, Galkowski M, Rozanski K. Isotopic composition of daily precipitation along the southern foothills of the Himalayas: Impact of marine and continental sources of atmospheric moisture. Atmos Chem Phys. 2018;18:8789–8805. https://doi.org/10.5194/acp-18-8789-2018
11. Friedman I, Harris JM, Smith GI, Johnson CA. Stable isotope composition of waters in the Great Basin, United States 1. Air-mass trajectories. J Geophys Res-Atmos. 2002;107(D19):ACL-14. https://doi.org/10.1029/2001JD000565
12. Aravena R, Suzuki O, Pena H, Grilli A, Pollastri A, Fuenzalida H. Isotopic composition and origin of the precipitation in Northern Chile. Appl Geochem. 1999;14:411–422. https://doi.org/10.1016/S0883-2927(98)00067-5
13. Krklec K, Domínguez-Villar D, Lojen S. The impact of moisture sources on the oxygen isotope composition of precipitation at a continental site in central Europe. J Hydrol. 2018;561:810–821. https://doi.org/10.1016/j.jhydrol.2018.04.045
14. Balagizi CM, Kasereka MM, Cuoco E, Liotta M. Influence of moisture source dynamics and weather patterns on stable isotopes ratios of precipitation in Central-Eastern Africa. Sci Total Environ. 2018;628–629:1058–1078. https://doi.org/10.1016/j.scitotenv.2018.01.284

15. Roffe SJ, Fitchett JM, Curtis CJ. Classifying and mapping rainfall seasonality in South Africa: A review. S Afr Geogr J. 2019;101(2):158–174. https://doi.org/10.1080/03736245.2019.1573151

16. Kruger AC, Nxumalo MP. Historical rainfall trends in South Africa: 1921–2015. Water SA. 2017;43(2):285–297. https://doi.org/10.4314/wsa.v43i2.12

17. Roffe SJ, Fitchett JM, Curtis CJ. Investigating changes in rainfall seasonality across South Africa: 1987–2016. Int J Climatol. 2021;41:E2031–E2050. https://doi.org/10.1002/joc.6830

18. Sousa PM, Blamey RC, Reason CJ, Ramos AM, Trigo RM. The 'Day Zero' Cape Town drought and the poleward migration of moisture corridors. Environ Res Lett. 2018;13(12):124025. https://doi.org/10.1088/1748-9326/aaebc7

19. Leketa K, Abiye T. Investigating stable isotope effects and moisture trajectories for rainfall events in Johannesburg, South Africa. Water SA. 2020;46:429–437. https://doi.org/10.17159/wsa/2020.v46.i3.8653

20. IAEA/WMO. Global network of isotopes in precipitation. The GNIP Database. 2022. Available from: https://nucleus.iaea.org/wiser

21. Harris C, Burgers C, Miller J, Rawoot F. O- and H-isotope record of Cape Town rainfall from 1996 to 2008, and its application to recharge studies of Table Mountain groundwater, South Africa. S Afr J Geol. 2010;113:33–56. https://doi.org/10.2113/gssajg.113.1.33

22. Braun K, Bar-Matthews M, Ayalon A, Zilberman T, Matthews A. Rainfall isotopic variability at the intersection between winter and summer rainfall regimes in coastal South Africa (Mossel Bay, Western Cape Province). S Afr J Geol. 2017;120:323–340. https://doi.org/10.25131/gssajg.120.3.323

23. Durowoju OS, Odiyo JO, Ekosse GIE. Determination of isotopic composition of rainwater to generate local meteoric water line in Thohoyandou, Limpopo Province, South Africa. Water SA. 2019;45(2):183–189. https://doi.org/10.4314/wsa.v45i2.04

24. Kaseke KF, Wang L, Wanke H, Turewicz V, Koeniger P. An analysis of precipitation isotope distributions across Namibia using historical data. PLoS ONE. 2016;11(5):e0154598. https://doi.org/10.1371/journal.pone.0154598

25. Kaseke KF, Wang L, Wanke H, Tian C, Lanning M, Jiao W. Precipitation origins and key drivers of precipitation isotope (^{18}O, ^{2}H, and ^{17}O) compositions over Windhoek. J Geophys Res-Atmos. 2018;123. https://doi.org/10.1029/2018JD028470

26. IAEA/GNIP. Precipitation sampling guide. 2014. Available from: http://www-naweb.iaea.org/napc/ih/documents/other/gnip_manual_v2.02_en_hq.pdf

27. Gröning M, Lutz HO, Roller-Lutz Z, Kralik M, Gourcy L, Pöltenstein L. A simple rain collector preventing water re-evaporation dedicated for $\delta^{18}O$ and $\delta^{2}H$ analysis of cumulative precipitation samples. J Hydrol. 2012;448–449:195–200. https://doi.org/10.1016/j.jhydrol.2012.04.041

AUTHORS:
Houraa Daher[1]
Ben P. Kirtman[1]

AFFILIATION:
[1]Rosenstiel School of Marine and Atmospheric Science, University of Miami, Miami, Florida, USA

CORRESPONDENCE TO:
Houraa Daher

EMAIL:
hdaher@rsmas.miami.edu

DATES:
Received: 06 Apr. 2022
Revised: 10 Mar. 2023
Accepted: 13 Mar. 2023
Published: 08 Aug. 2023

HOW TO CITE:
Daher H, Kirtman BP. Future climate change in the Agulhas system and its associated impact on South African rainfall. S Afr J Sci. 2023;119(7/8), Art. #13733. https://doi.org/10.17159/sajs.2023/13733

ARTICLE INCLUDES:
☒ Peer review
☐ Supplementary material

DATA AVAILABILITY:
☐ Open data set
☒ All data included
☐ On request from author(s)
☐ Not available
☐ Not applicable

EDITORS:
Jennifer Fitchett
Adriaan van der Walt

KEYWORDS:
Agulhas system, South African rainfall, climate change, CCSM4, CMIP5

FUNDING:
US National Oceanic and Atmospheric Administration (NA18OAR4310293, NA15OAR4320064), US National Science Foundation (OCE1419569, OCE1559151), US Department of Energy (DE-SC0019433)

Future climate change in the Agulhas system and its associated impact on South African rainfall

South African climate variability has been linked to changes in both the Agulhas system and external forcing (i.e. CO_2 and ozone). We analysed future climate change in the Agulhas system volume transport and its associated impacts on South Africa's precipitation using the Community Climate System Model version 4 as part of the Coupled Model Intercomparison Project, phase 5. Output from one historical and three future greenhouse gas emission scenarios were examined to project various climate storylines. We found that the Agulhas Current volume transport decreases across all three scenarios and that the current displays a strong baroclinic component with an increase in transport at the surface and decrease at intermediate depths. Agulhas leakage was found to increase with historical emissions. Additionally, an east-west dipole pattern for convective precipitation was found over South Africa, with an increase over the eastern region related to an increase in greenhouse gas emissions and a decrease in the western region linked to the location of Hadley cell edge latitude. Moving into the 21st century, future predictions in regional climate variability are shown to be dependent on the intensity of greenhouse gas emissions and are extremely important for South Africa, a region prone to drought and flooding and home to a large population dependent on rain-fed agriculture.

Significance:

- Future climate variability in the Agulhas system and South African region is heavily dependent on changes in external forcing.
- The Agulhas Current volume transport decreases as the greenhouse gas emissions continue to increase and a strong baroclinic component is found with an increase in transport at the surface and a decrease at the intermediate depths.
- A strong east-west dipole precipitation pattern is found over South Africa with the increase in the eastern region related to the increase in greenhouse gas emissions and the decrease in the western region related to the location of the Hadley cell edge latitude.

Introduction

South Africa is a region prone to precipitation changes with drought and flooding events being a common occurrence.[1,2] Understanding the rainfall variability over South Africa is of great interest as Cape Town was one of the first major cities in the world to nearly run out of water in 2018 and because it is home to a significant population that is dependent on rain-fed agriculture. Studies have identified two regions of rainfall over South Africa: eastern and northern South Africa and southwest South Africa. Precipitation over eastern and northern South Africa primarily occurs in summer and precipitation over southwest South Africa occurs in winter.[2]

The precipitation pattern over eastern and northern South Africa can be explained by tropical cloud bands and associated convection in the region as a result of sea surface temperature (SST) anomalies.[3] These warm SST anomalies come from the Agulhas Current which leads to the advection of moist marine air over the region.[4-9] The Agulhas Current, located off the eastern coast of South Africa, is the largest western boundary current in the southern hemisphere and provides moisture to the atmospheric boundary layer via latent heat fluxes which are projected to increase significantly over western boundary systems[10-14] as the climate system warms. The SST of the Agulhas Current is linked to the El Niño-Southern Oscillation (ENSO), originating in the equatorial eastern Pacific, with El Niño correlated with drier conditions and La Niña with wetter conditions in the eastern and northern region.[3,8,15-22]

Undergoing a mechanism different from that of the precipitation in the eastern and northern region, the precipitation found over southwestern South Africa is driven by cold fronts associated with mid-latitude cyclones formed in the South Atlantic.[2,20,23-26] The precipitation associated with these fronts was found to be linked to the expansion of the Hadley cell edge resulting in a poleward shift of these mid-latitude cyclones and low-pressure systems such that a post-frontal high-pressure system is located above South Africa suppressing the precipitation typically seen over the southwestern region.[27] With the poleward shift of the low-pressure belts, the easterly winds increase and the precipitation is expected to decrease as the rain-bearing storms weaken and deflect poleward.[28] This behaviour, representative of the ozone depletion period, explains the drought experienced in Cape Town in 2018 and suggests that as ozone recovers, wetter conditions and fewer droughts in the region can be expected.[29,30] In addition to the mid-latitude cold fronts, it has been shown that ENSO, SST anomalies, and Agulhas leakage, the inflow of warm and salty water from the Indian Ocean to the Atlantic Ocean, also have an impact on the local precipitation in this region.[2,23,26,31,32]

Using the Community Climate System Model version 4 (CCSM4) simulations from the Coupled Model Intercomparison Project phase 5 (CMIP5) experiments[33], we investigated future climate change projections in

the Agulhas system and regional precipitation variability in South Africa. Agulhas Current and Agulhas leakage transports were calculated and precipitation over South Africa analysed. We compared different emissions scenarios and three different time periods to determine the significance of greenhouse gas and ozone forcing in the region.

Methods

Model and climate scenarios

In this study, we analysed model output from NCAR's CCSM4 coupled climate model.[34] The ocean model used was Parallel Ocean Program, version 2 (POP2) at a 1° horizontal resolution with 60 vertical layers. The atmosphere model in CCSM4 is Community Atmosphere Model, version 4 (CAM4) at a 1° horizontal resolution with 26 vertical layers. The monthly mean outputs were considered in this analysis.

The CCSM4 output used are from the CMIP5 archive as CCSM4 was not included in the more recent CMIP6 data. Data from four different forcing scenarios were considered: historical (pre-2006), representative concentration pathway 2.6 (RCP2.6), RCP4.5, and RCP8.5 (2006–2100). The historical forcing is the 20th-century simulation using a combined anthropogenic and natural forcing.[34] RCP2.6 is the extreme mitigation scenario in which emissions peak at 3.0 W/m^2 in year 2050 and then decrease to 2.6 W/m^2, hence the name RCP2.6. RCP4.5 is the stabilisation scenario in which emissions peak at year 2075 and then remain constant until the end of the century where they are at 4.5 W/m^2. And lastly, RCP8.5 is the scenario with the largest emissions, where the forcing increases over time and is still increasing as it reaches 8.5 W/m^2 by 2100. Following Barnes et al.[35], three different RCP scenarios were considered to allow future climate predictions across a range of possible emission scenarios.

Continuing to follow Barnes et al.[35], the data were split into three time periods: ozone depletion (OD), ozone recovery (OR), and post-ozone recovery (POR). The OD period (1970–2005) is mainly driven by the depletion of the ozone and the increase in greenhouse gas (GHG) emissions. The results from this period will be the same across all three RCPs as data from the historical run are used (i.e. observed estimates of atmospheric gas concentrations). In the OR period (2006–2045), the ozone begins to recover and the GHGs continue to increase, therefore the ozone and GHG forces will start to oppose each other. The changes that are seen during the OD period can be expected to weaken or reverse in the OR period. Lastly, the POR period (2046–2100) will have a recovered ozone, and varying GHG emissions and results from this period are largely driven by GHG emissions. For the RCP2.6 scenario, the ozone recovery is expected to continue to dominate as the GHG emissions decrease resulting in trends similar to those of the OR period. As the emissions increase slightly in RCP4.5, however, the ozone and GHG emissions will offset and trends similar to those of both the OD and OR periods can be expected. For RCP8.5, the final scenario, the emissions are highest and the increase in GHGs will dominate the ozone recovery and trends can be expected to return to those observed during the OD period.

Lagrangian particle tracking

An offline Lagrangian particle tracking tool called Parcels was used to calculate Agulhas leakage in this study. Lagrangian particle tracking is used to ensure that all Agulhas leakage (filaments and Agulhas Rings) is considered and that water crossing into the Atlantic Ocean originated in the Agulhas Current.[36-39] Parcels (Probably a Really Computationally Efficient Lagrangian Simulator), is a Python Lagrangian tracking tool created to track passive and active tracers.[40] Parcels uses linear and nearest-neighbour interpolation in time and space and particles are advected using the Runge-Kutta 4 scheme.

A cross-section of particles is released along 34°S, the latitude of the maximum wind stress curl over the South Indian Ocean and therefore

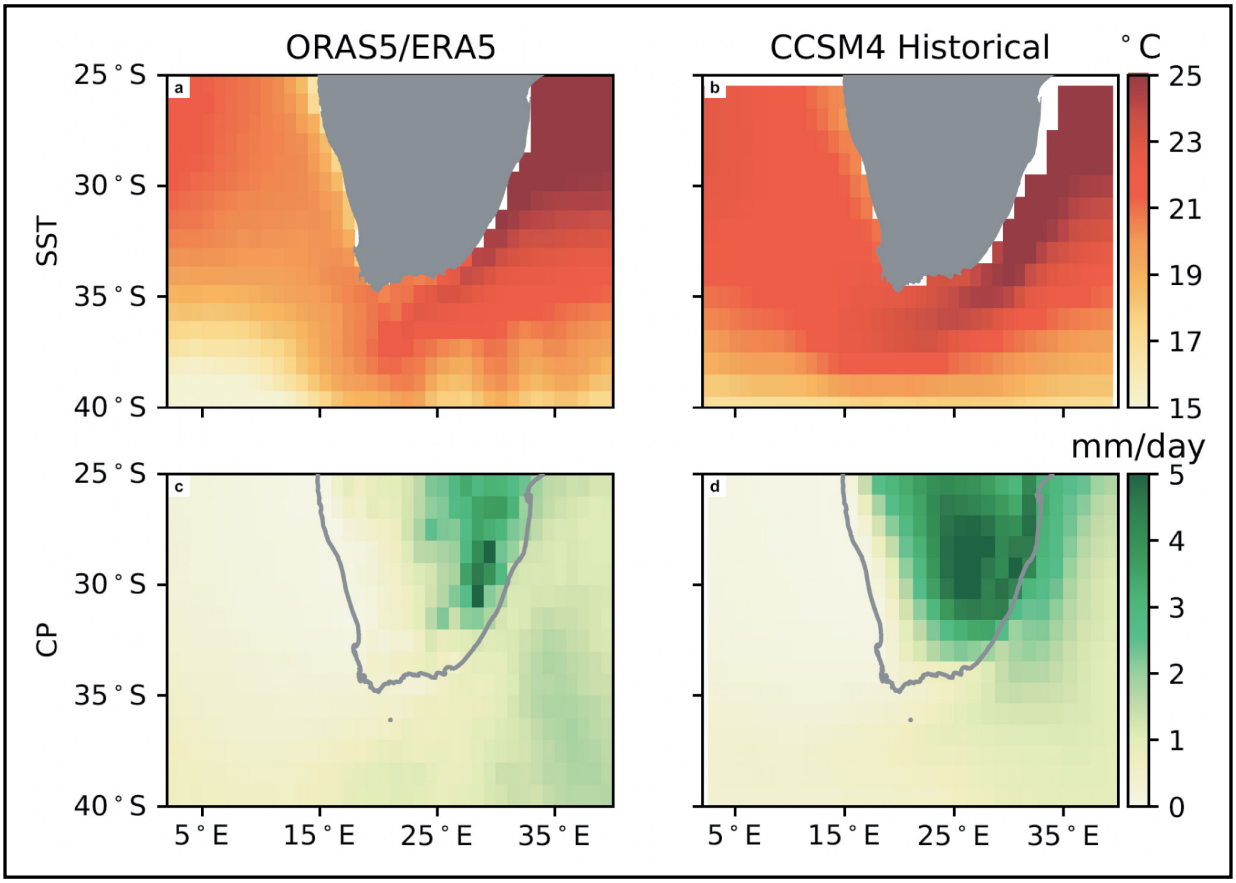


Figure 1: Comparison between reanalysis data (ORAS5/ERA5) from 1979 to 2005 (first column) and CCSM4 historical data from 1970 to 2005 (second column) for (a,b) the mean sea surface temperature (SST) and (c,d) convective precipitation (CP). The reanalysis data have been re-gridded from 0.25°×0.25° to 1°×1° to match the CCSM4 model output resolution to enable meaningful comparison.

the location of maximum Agulhas Current transport. Each particle is assigned a volume transport equal to the initial release velocity multiplied by the grid-cell size. This cross-section of particles is released every month from 1960 to 2100, allowing 10 years of particle circulation before transport estimates are calculated. Although previous studies tend to release more frequently than monthly, Cheng et al.[39] found that the temporal resolution of the velocity fields does not have much of an impact on the estimate of Agulhas leakage in CCSM4, with a difference of approximately 1 Sv (1 Sv = 1×10^6 m^3/s) when comparing the daily fields to the monthly mean fields. A particle is considered Agulhas leakage so long as its initial transport is southward, and it has crossed the GoodHope line an odd number of times. Following previous studies[36,37,39,41,42], the GoodHope line, a hydrographic section separating the Indian Ocean from the Atlantic Ocean[43], is used as the boundary to determine Agulhas leakage. The 130-year Agulhas leakage time series is calculated by summing the volume transport of the leakage particles at their time of final crossing of the GoodHope line.

In order to build on the South African precipitation analysis and Lagrangian experiments found in Cheng et al.[32,39], in this study we also used CCSM4 for consistency despite it underperforming compared to some of the other CMIP5 models. Cheng et al.[32,39], however, used data from the 20th century only, and did not run any future climate analysis. Additionally, the emissions in Cheng et al.[32] were kept constant, and in Cheng et al.[39], only the historical forcing scenario was used. Cheng et al.[39] validated CCSM with satellite and in situ data and found that their results of the Agulhas Current, retroflection, and Agulhas Return Current agreed with what was seen in the AVISO data.

Furthermore, SST and convective precipitation from the CCSM4 historical data used in this study were compared to reanalysis data from 1979 to 2005 (Figure 1). The ORAS5 SST (Figure 1a, 0.25° horizontal resolution re-gridded to 1°) and CCSM4 SST (Figure 1b, 1° resolution) agree with each other in that the strongest temperatures are found over the Agulhas Current as it brings warm water from the equator towards the poles and that there is cooler water found along the western coast of South Africa in the Benguela Current, a region of upwelling. The low resolution of CCSM4, however, does not capture the Agulhas retroflection and Agulhas Return Current that is seen clearly in the reanalysis data. For the convective precipitation, both the ERA5 (Figure 1c, 0.25° horizontal resolution re-gridded to 1°) and CCSM4 (Figure 1d, 1° resolution) data show the strongest precipitation in the eastern region of South Africa. But again, the low resolution of CCSM4 shows a broader, less accurate depiction of South African convective precipitation with the ERA5 data confined more to the east along the coast.

Result analysis

The five-member ensemble mean is shown for each of the CMIP5 forcing scenarios, removing some of the interannual variability. Results are shown for the austral summer, December–February (DJF), where the largest changes are seen due to the lagged response of the SON stratospheric ozone signal to reach the lower troposphere.[29,44] All time series were smoothed using a 10-year moving average filter with a time step of one year as in Barnes et al.[35] The difference in the trends of the unsmoothed and smoothed time series is insignificant. And lastly, all trends and linear regression maps were calculated using linear least-squares regression with all significant results shown within the 95% confidence interval.

Results and discussion

Agulhas Current and Agulhas leakage

The Agulhas Current volume transport is calculated across 34°S and a mean volume transport of 65 Sv is found for the ozone depletion period compared to the observed transport of 77 Sv[45], a reasonable transport estimate for a low-resolution model that does not capture

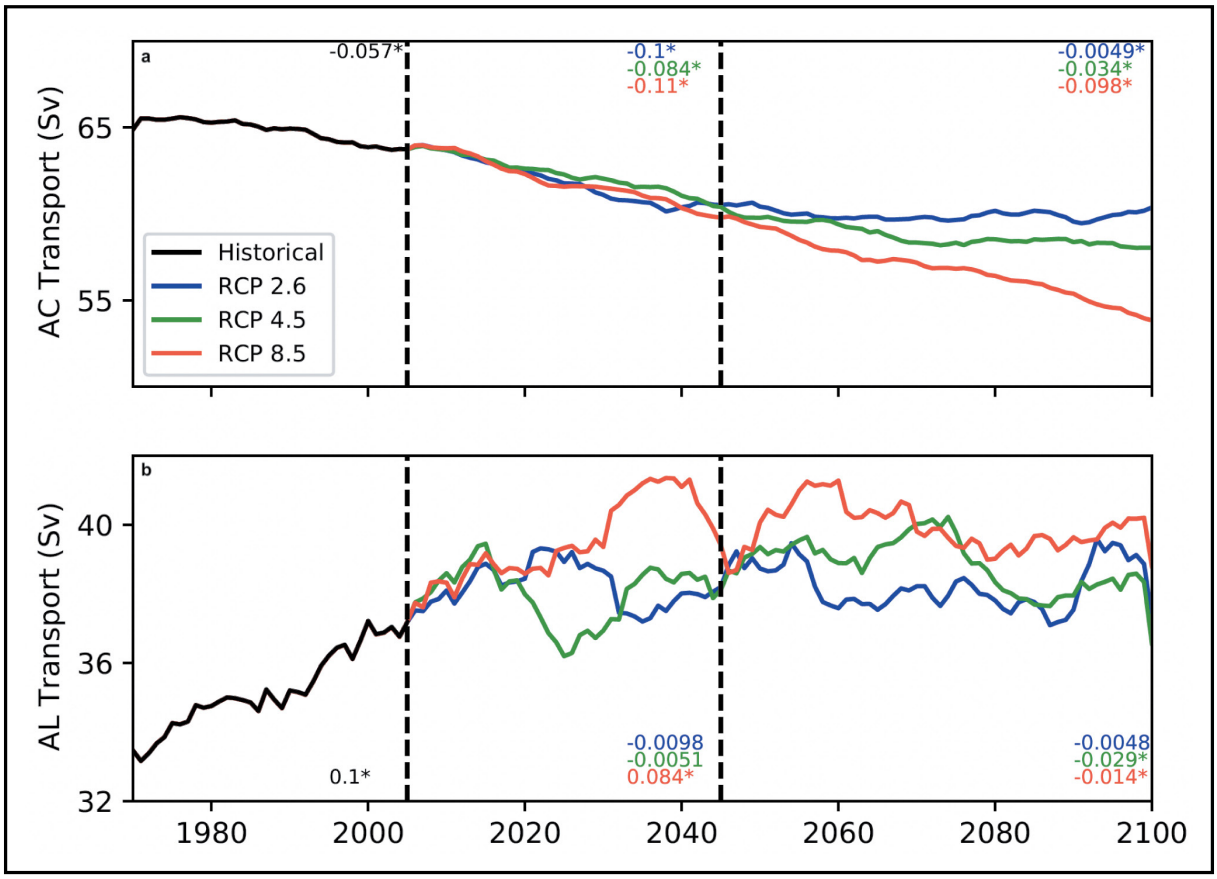


Figure 2: (a) Time series of the Agulhas Current (AC) volume transport and (b) Agulhas leakage (AL) volume transport. Results for historical (black), RCP2.6 (blue), RCP4.5 (green), and RCP8.5 (red) are shown. Trends for each period are shown with those within the 95% confidence interval marked with an asterisk (*).

the mesoscale features of the current. The Agulhas Current volume transport is found to be decreasing through time across all three RCPs (Figure 2a) with a 5 Sv decrease in RCP2.6 and 10 Sv decrease in RCP8.5. Although the transport saturates near 2050, right after the POR period begins, with all three RCPs converging on similar transport values, the downward trend continues to the end of the century in RCP8.5 (Figure 2a, red line). In RCP2.6, however, the transport begins to increase at the tail end of the century (Figure 2a, blue line), demonstrating a slight lag in the response of the ocean to the decrease in emissions in the atmosphere.

The vertical structure of the Agulhas Current is found to have a baroclinic component in the water column (Figure 3), with a speed up found at the surface in the upper layer (> 500 m) and a slowdown at intermediate depths (500–1500 m). For the ozone recovery minus ozone depletion (OR-OD) period, there is not much difference between the three RCPs (Figure 3a–c) with a slight intensification found in the core of the current in the upper 100 m and a weakening seen throughout the rest of the water column resulting in the overall decrease in volume transport. In POR-OR, however, the vertical structure is different across all RCPs with RCP2.6 showing an increase in transport in the upper 500 m (Figure 3d), likely leading to the increase in transport seen at the end of the century (Figure 2a, blue line), and RCP8.5 showing a strong decrease in poleward transport at the intermediate depths (Figure 3f), resulting in the decrease in transport observed (Figure 2a, red line). The baroclinicity found in the Agulhas Current is similar to that found in the Kuroshio Current by Chen et al.[46], who used the historical and RCP4.5 simulations of CMIP5. Chen et al.[46] discuss that the baroclinicity found may be a result of the stronger stratification and downward heat mixing[47] within the vertical water column that leads to a weakening of the subtropical gyre in the lower thermocline[47,48] and therefore a slowdown at greater depths.

The Agulhas leakage volume transport was calculated across all three RCPs (Figure 2b) following the method explained above. A mean transport of 36 Sv was found for the ozone depletion period – a better estimate than the 43 Sv calculated using the same low-resolution model with identical historical forcing[38] and 15 Sv more than the observed transport of 21 Sv.[49] During the ozone depletion period, there is a clear increase in Agulhas leakage transport over time (Figure 2b), in agreement with previous literature showing an increase in Agulhas leakage transport with anthropogenic climate change.[36,50] After the ozone depletion period, however, there is a strong weakening of the trend seen across all three RCPs through the end of the century, with little overall change in transport and with only RCP8.5 showing significant changes in leakage as a result of the extreme increase in emissions. Similar to that seen in the Agulhas Current transports, there is a convergence in leakage transport found near 2050, just as the changes in emissions begin to vary the most across the three RCPs. The Agulhas leakage transport trends are noticeably noisier compared to the Agulhas Current transports (Figure 2a), which were found to be more robust according to the significance test.

Correlation coefficients between the Agulhas Current and Agulhas leakage transport were found to be significant, with values of −0.76, −0.78, and −0.75 for RCP2.6, RCP4.5, and RCP8.5, respectively. This relationship is compared to the results from Van Sebille et al.[37], who used a high-resolution ocean model, and therefore found lower, more accurate values of Agulhas leakage and a correlation coefficient of −0.67. The Agulhas Current and Agulhas leakage are inversely related with a decrease (increase) in Agulhas Current associated with an increase (decrease) in Agulhas leakage. Looking at the volume transport from the ozone depletion period only, a correlation coefficient of −0.92 is found, showing a robust relationship for this period.

South Africa precipitation

A clear east-west dipole pattern of convective precipitation is found in both the OR-OD and POR-OR maps (Figure 4). Similar to the vertical structure of the Agulhas Current (Figure 3), there is not much difference across the three RCPs for convective precipitation in the OR-OD maps (Figure 4a–c) with an increase (decrease) of approximately a quarter of a millimetre per day over the eastern (western) region of South Africa. The

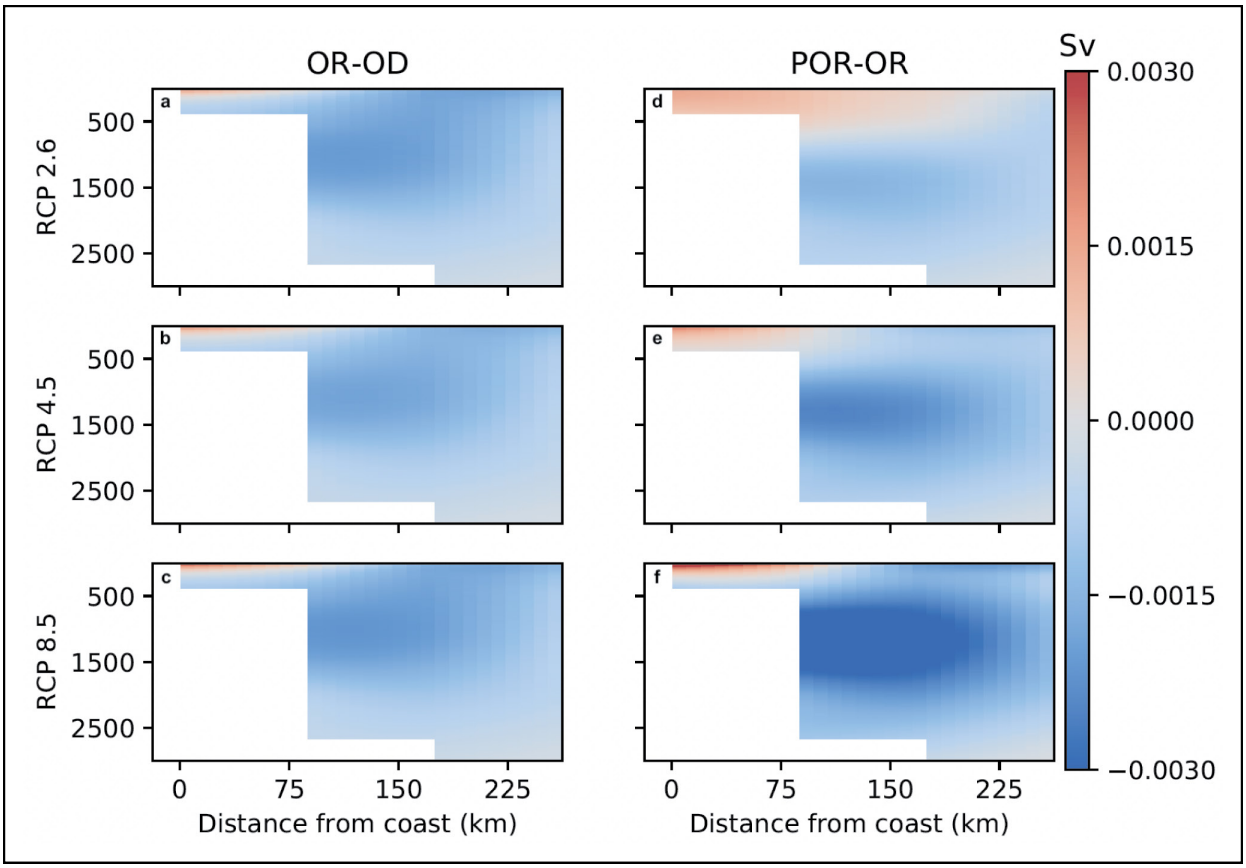


Figure 3: The Agulhas Current cross-sectional volume transport across 34°S showing the differences between the OR-OD periods (a–c) and POR-OR periods (d–f). Results for RCP2.6, RCP4.5, and RCP8.5 are shown in the first, second, and third rows, respectively.

results for POR-OR (Figure 4d–f), however, vary for each RCP scenario, highlighting the importance of the different emission levels. In RCP2.6 (Figure 4d), there is little change in convective precipitation seen in both the eastern and western regions of South Africa. RCP4.5 (Figure 4e) shows an increase of about half a millimetre per day in the eastern region and not much change in the western region. Lastly, RCP8.5 (Figure 4f) shows the most drastic change in convective precipitation in the two regions, with nearly an increase of one millimetre per day in the eastern region, roughly four times the rate seen in OR-OD (Figure 4c), and a decrease of almost half a millimetre per day in the western region, nearly double the amount of OR-OD (Figure 4c).

The total convective precipitation was calculated over the eastern (Figure 4a, blue-green box) and western (Figure 4a, brown box) regions and, overall, it is clear that there is more precipitation over eastern South Africa than western South Africa per day (Figure 5). The convective precipitation rates for the eastern box (EB) are seen to be increasing throughout all three periods (Figure 5; EB curves) with similar values of precipitation observed for all three RCPs until 2050. Then the largest and significant increase is seen during the POR period for RCP8.5 (Figure 5; red line), suggesting once more that the increased emissions in this RCP have a significant impact on regional precipitation in South Africa, whereas with the recovery of the ozone and the stabilisation of GHGs, trends in RCP2.6 and RCP4.5 weaken (Figure 5; blue and green lines). Opposite and weaker trends can be found in the western box (Figure 5; WB curves) with a decrease in convective precipitation seen throughout most of the century with the exception of RCP2.6 in the POR period (Figure 5; blue line) where an increase in precipitation can be found towards the tail end of the century, likely correlating with an equatorward shift of the Hadley cell and storm tracks. The largest decrease is seen in RCP8.5 (Figure 5; red line), especially in the last 20 years of the century as the emissions continue to increase and cause a poleward expansion of the Hadley cell and frontal systems. Additionally, RCP8.5 is the only one of the RCPs to show a significant trend after the historical period ends.

As discussed previously, the convective precipitation over eastern South Africa is known to exhibit a strong correlation with warm SSTs in the Agulhas Current. This relationship likely explains the increased precipitation found over eastern South Africa, especially during the POR period in RCP8.5 when emission levels are highest. The convective precipitation over western South Africa, however, is related to the location of the Hadley cell edge. Additionally, the east-west dipole precipitation pattern found over South Africa is associated with the increase in easterly winds linked to the poleward shift of the frontal systems.

SST and moisture flux

A clear relationship was found between SSTs in the Agulhas system and the moisture advected over eastern South Africa (Figure 6). Consistent with the previous maps, there is not much difference observed in SST and moisture flux (u*q and v*q) for the OR-OD period (Figure 6a–c) as there is very little change in emissions. Throughout the Agulhas system there is warming of 1 °C and moisture is advected downstream of the Agulhas Current and over the Agulhas bank and leakage corridor, forming a cyclonic low-pressure system over South Africa. This loss of heat from the ocean to the atmospheric boundary layer and cyclonic system results in the increase in convective precipitation observed in eastern South Africa (Figure 4a–c) and is consistent with findings in previous studies.[4-9] For RCP2.6 (Figure 4d) in the POR-OR period, however, there is almost no change in SST or moisture flux with there being very weak cyclonic motion as the emissions decrease. RCP4.5 (Figure 4e) shows a result similar to what was seen during the OR-OD period with the emissions stabilising and remaining constant. RCP8.5 (Figure 6f), with the greatest increase in emissions, shows the most extreme response in both SSTs and moisture flux. There is a continued increase in SSTs of nearly 2 °C everywhere and a significant increase in the moisture flux strength resulting in a large air–sea heat/moisture exchange. Therefore the formation of the intense cyclonic low-pressure system observed leads to a considerable increase in convective precipitation (Figure 4f).

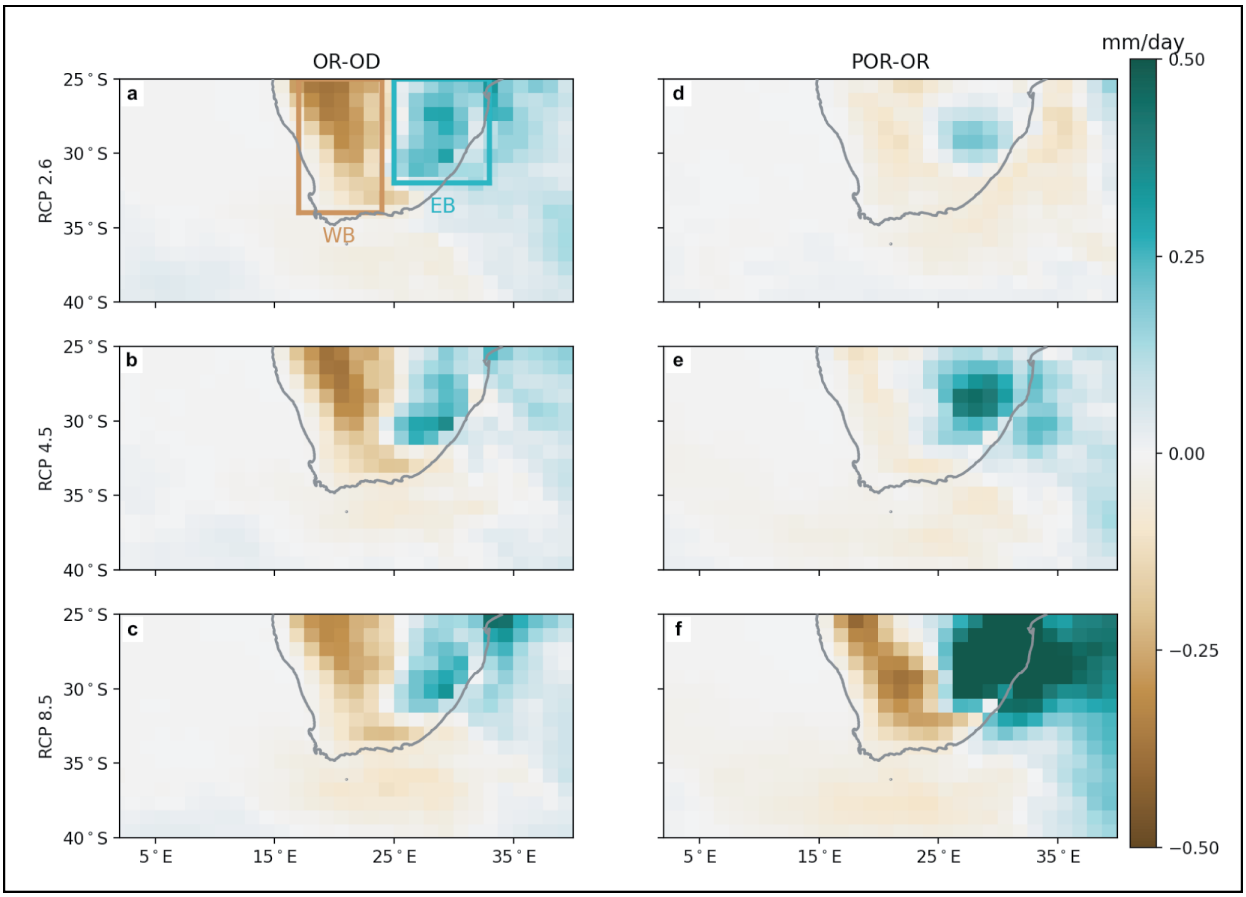


Figure 4: The convective precipitation across 34°S showing the differences between the OR-OD periods (a–c) and POR-OR periods (d–f). Results for RCP2.6, RCP4.5, and RCP8.5 are shown in the first, second, and third rows, respectively. The eastern box (EB) response is shown in blue-green and the western box (WB) in brown.

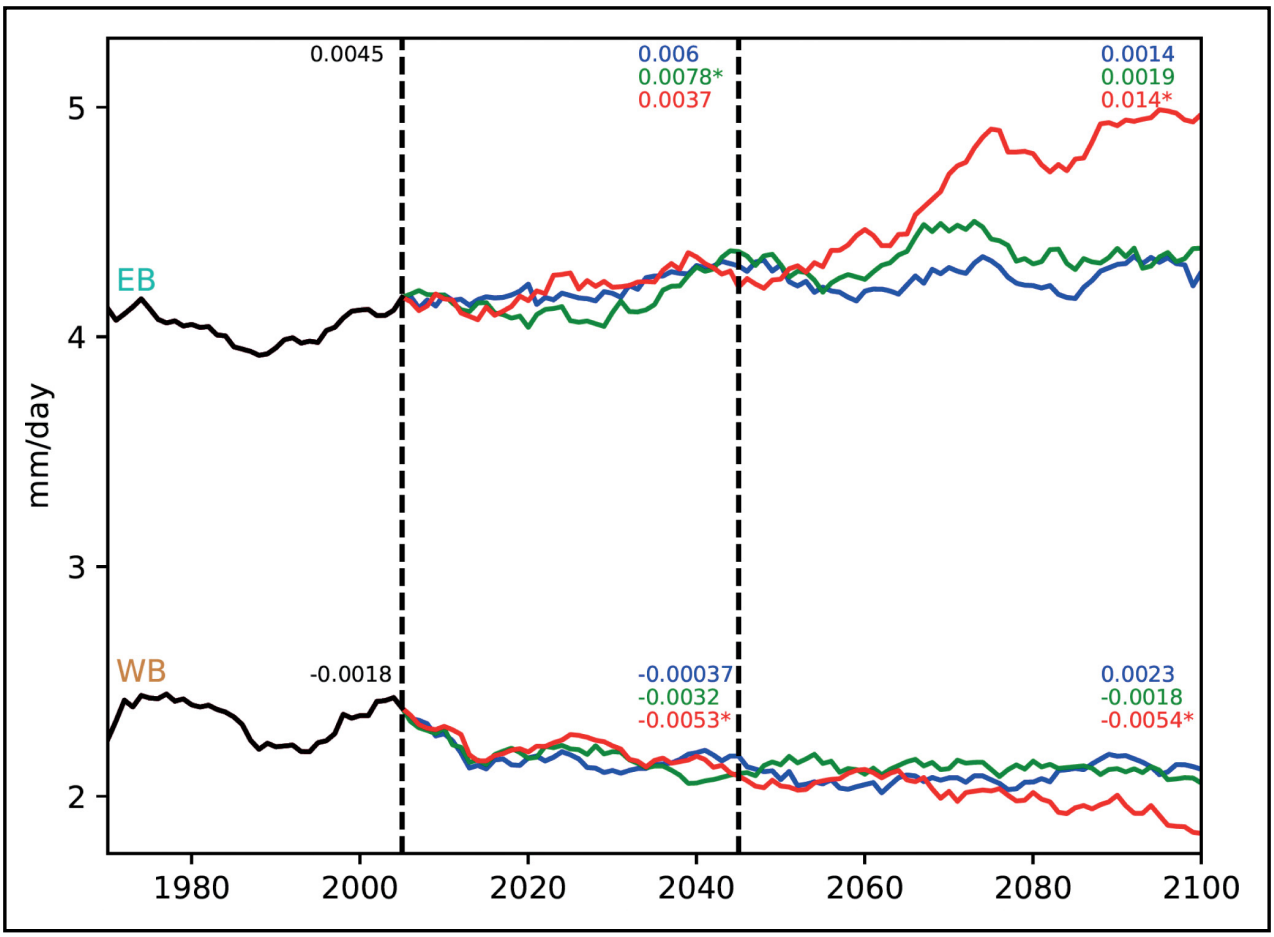


Figure 5: Time series of the mean convective precipitation in the eastern box (EB) and the western box (WB) for historical (black), RCP2.6 (blue), RCP4.5 (green), and RCP8.5 (red).

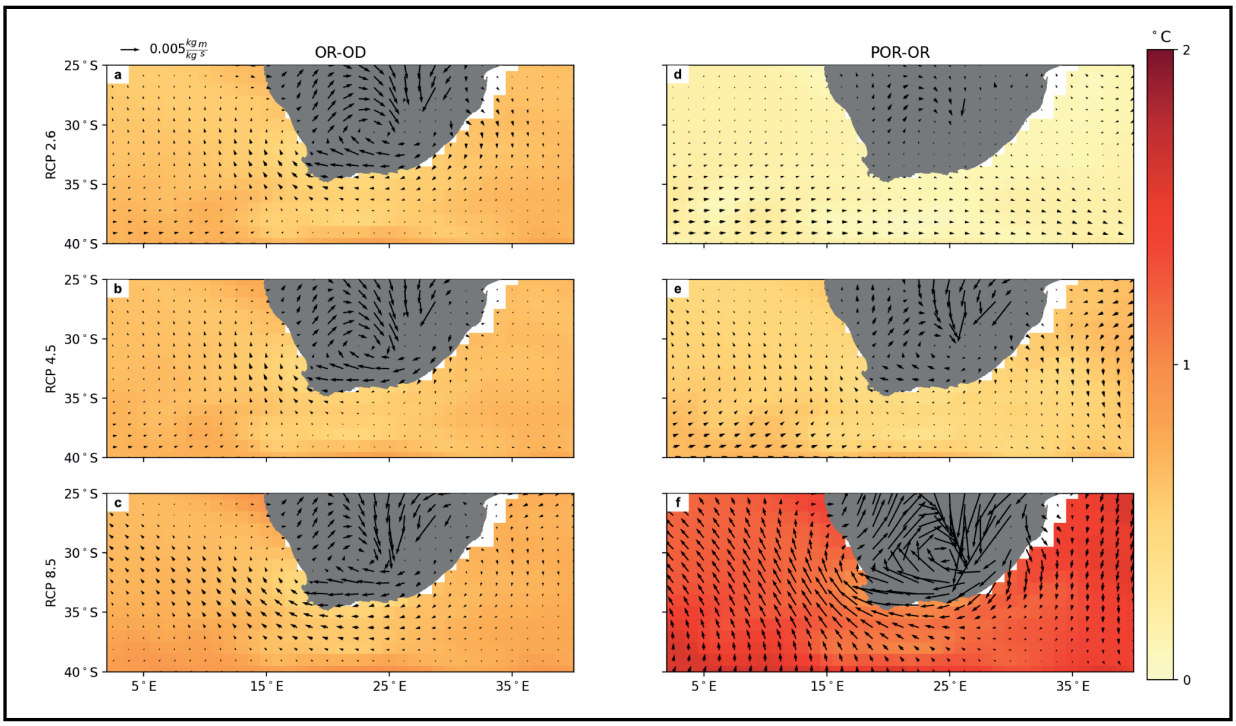


Figure 6: The sea surface temperature (°C) and moisture flux (vectors) showing the differences between the OR-OD periods (a–c) and POR-OR periods (d–f). Results for RCP2.6, RCP4.5, and RCP8.5 are shown in the first, second, and third rows, respectively.

Hadley cell

The latitude of the Hadley cell edge and storm tracks plays a large role in precipitation over western South Africa, with a poleward shift associated with a decrease in precipitation and an equatorward shift linked to an increased rate. The location of the Hadley cell is controlled by both the ozone and GHGs with ozone depletion and an increase in GHGs contributing to the poleward shift and ozone recovery and decrease of GHGs resulting in an equatorward shift. Following previous studies, the Hadley cell edge is defined as the latitude of the zonal mean of the meridional mass stream function where the stream function is equal to zero at 500 hPa.[51-56] The meridional mass stream function, ψ, is calculated by:

$$\psi(\phi, p) = \frac{2\pi a \cos\phi}{g} \int_0^p v \, dp,$$

where ϕ is the latitude, p is pressure, a is the radius of the earth, g is the gravitational acceleration, and v is the meridional wind.

During the OD period, there is a significant poleward expansion of the Hadley cell edge of roughly 1° (Figure 7a, black line). This trend is weakened across all three RCP scenarios as the ozone recovers, with an equatorward shift shown in RCP2.6 that continues into the POR period (Figure 7a, blue line). Similar to previous results, after the mid-century, there is a sharp increase in the equatorward trend for RCP2.6 that is followed by a stabilisation over the last few decades of the century, resulting in little change in the position of the Hadley cell edge. This trend matches that seen for the convective precipitation in the western box for RCP2.6. With the equatorward shift of the Hadley cell edge and the return of mid-latitude cyclones and cold fronts, there is an increase in the associated precipitation in this region. RCP4.5 also shows very little change in the latitudinal position of the Hadley cell edge during this period, indicating once more that the ozone and GHG forcing cancel each other out in this scenario (Figure 7a, green line), in agreement with the western box precipitation for RCP4.5. Lastly, there is a poleward expansion of the Hadley cell edge for RCP8.5 that continues through the end of the century (Figure 7a, red line). Overall, a 2° poleward expansion of the Hadley cell edge is discovered in the RCP8.5 scenario with half of that expansion occurring in the first 35 years during the OD period.

The poleward shift of the Hadley cell edge and high-pressure systems is linked to a suppression of precipitation over western South Africa and the decrease in convective precipitation found. The different trends seen during the POR period across each RCP scenario for the Hadley cell edge latitude were found to be significant and continue to highlight the importance of the different emissions scenarios.

The overall strength and expansion of the Hadley cell varies with each emission scenario(Figure 7b,c). The difference between the late era (2080–2100) and early era (1970–1990) of the meridional mass stream function at all levels in the atmosphere is shown in Figure 7. These periods were selected to show the extreme changes found in the Hadley cell circulation. There is very little difference found in the Hadley cell strength in RCP2.6 (Figure 7b). RCP4.5 (Figure 7c) shows a slight increase and expansion in the overturning circulation, but the greatest change is seen in RCP8.5 once again. There is an increase in the Hadley cell strength in RCP8.5 (Figure 7d) of over three times the circulation seen in the other RCP scenarios, as well as an overall expansion of the cell observed. Previous studies have shown that these patterns typical accompany the poleward shift in Hadley cell edge as the climate warms[51-56], with the expansion caused by an increase in the subtropical static stability which pushes the baroclinic instability zone and frontal systems poleward, therefore leading to the formation of high-pressure systems and less precipitation over western South Africa.[27,52]

Conclusions

Three different emission scenarios (RCP2.6, RCP4.5, and RCP8.5) were examined to understand the influence of varying GHG emissions in future climate prediction. The data were split into three different periods: ozone depletion (1970–2005), ozone recovery (2006–2045), and post-ozone recovery (2046–2100).

Both the Agulhas Current and Agulhas leakage volume transports were calculated as they have been shown to be linked to precipitation changes over South Africa. The Agulhas Current was found to decrease with time across all three RCPs, with the most change seen in RCP8.5. A baroclinic component was seen with a speed up in the upper layer and slowdown in the intermediate layer, resulting in an overall decrease in volume transport. The Agulhas leakage volume transport did not show a clear trend throughout time, but rather a strong increase in the ozone

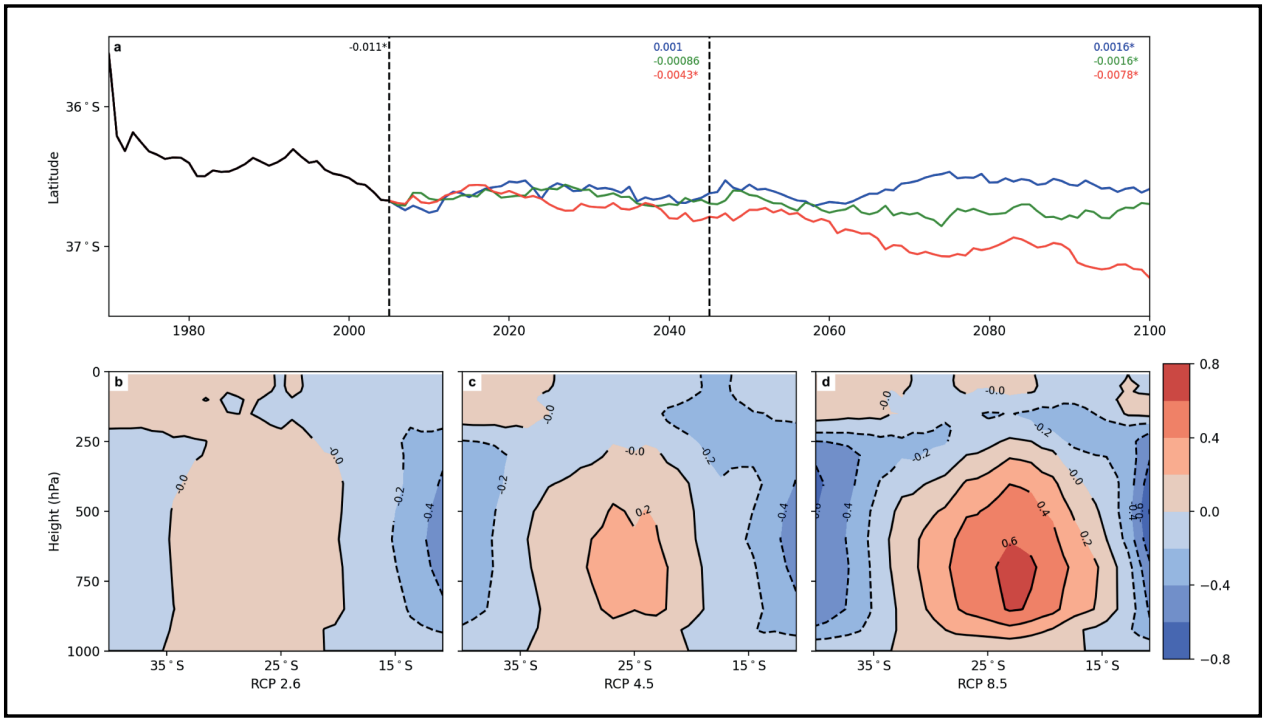


Figure 7: (a) Time series of the Hadley cell edge latitude. (b–d) The difference in the zonal mean of the meridional mass stream function between the last 20 years and first 20 years (2080–2100 minus 1970–1990) for (b) RCP2.6, (c) RCP4.5, and (d) RCP8.5. The contour interval is 0.2 × 1010 kg/s.

depletion period followed by a stabilisation of transport through the end of the century. The increase in the ozone depletion period is associated with anthropogenic climate change.

An increase in convective precipitation was found over the eastern region of South Africa across all RCPs with a significant increase during the post-ozone recovery period of RCP8.5. The precipitation in this region is linked to the warm SSTs in the Agulhas Current and the cyclonic motion found over South Africa, hence the strong trend found during the post-ozone recovery period for RCP8.5 when there is the greatest increase in temperature in the current and loss of heat to the atmosphere, leading to strong moist air being advected over land. The opposite pattern was seen over western South Africa with a decrease in convective precipitation. The precipitation in the western region is related to the position of the Hadley cell, indicating that the external forcing is the main driver behind this change. In RCP2.6, during the post-ozone recovery period, there was an increase in precipitation over the western region likely due to the ozone recovery causing an equatorward shift of the Hadley cell edge and a decrease seen during RCP8.5 due to the poleward shift and strengthening of meridional circulation caused by increased GHGs.

To summarise, when the GHG emissions are not increasing at an alarming rate (RCP2.6), it is found that the ozone recovery dictates the climate response in the Agulhas system and South Africa. But when the GHGs are increasing significantly throughout the 21st century (RCP8.5), the ozone recovery signal is overpowered by the GHG forcing and trends similar to those of the ozone depletion period are witnessed. Future climate change prediction in the Agulhas system and South Africa precipitation will be heavily dependent on the severity of the GHG emission levels moving forward.

Acknowledgements

All CCSM4 data from the CMIP5 archive can be accessed via the Climate Data Gateway at NCAR (https://www.earthsystemgrid.org/). We acknowledge the support from the National Oceanic and Atmospheric Administration (NA18OAR4310293, NA15OAR4320064), the National Science Foundation (OCE1419569, OCE1559151), and the Department of Energy (DE-SC0019433). B.P.K. is the William R. Middelthon Chair of Earth Sciences and is grateful for the associated support.

Competing interests

We have no competing interests to declare.

Authors' contributions

H.D.: Conceptualisation, methodology, data collection, data analysis, validation, writing – the initial draft. B.P.K.: Conceptualisation, methodology, writing – revisions, student supervision, funding acquisition.

References

1. Reason CJ, Allan RJ, Lindesay JA, Ansell TJ. ENSO and climatic signals across the Indian Ocean basin in the global context: Part I, Interannual composite patterns. Int J Climatol. 2000;20(11):1285–1327. https://doi.org/10.1002/1097-0088(200009)20:11<1285::AID-JOC536>3.0.CO;2-R
2. Reason CJ, Rouault M, Melice JL, Jagadheesha D. Interannual winter rainfall variability in SW South Africa and large scale ocean–atmosphere interactions. Meteorol Atmos Phys. 2002;80(1):19–29. https://doi.org/10.1007/s007030200011
3. Reason CJ, Rouault M. ENSO-like decadal variability and South African rainfall. Geophys Res Lett. 2002;29(13):16-1–16-4. https://doi.org/10.1029/2002GL014663
4. Reason CJ. Evidence for the influence of the Agulhas current on regional atmospheric circulation patterns. J Clim. 2001;14(12):2769–2778. https://doi.org/10.1175/1520-0442(2001)014<2769:EFTIOT>2.0.CO;2
5. Reason CJ. Subtropical Indian Ocean SST dipole events and Southern African rainfall. Geophys Res Lett. 2001;28(11):2225–2227. https://doi.org/10.1029/2000GL012735
6. Rouault M, White SA, Reason CJ, Lutjeharms JR, Jobard I. Ocean–atmosphere interaction in the Agulhas Current region and a South African extreme weather event. Weather Forecast. 2002;17(4):655–669. https://doi.org/10.1175/1520-0434(2002)017<0655:OAIITA>2.0.CO;2
7. Blamey RC, Reason CJ. Numerical simulation of a mesoscale convective system over the east coast of South Africa. Tellus A: Dyn Meteorol Oceanogr. 2009;61(1):17–34. http://doi.org/10.3402/tellusa.v61i1.15529
8. Jury MR. Passive suppression of South African rainfall by the Agulhas current. Earth Interact. 2015;19(13):1–4. http://doi.org/10.1175/ei-d-15-0017.1
9. Nkwinkwa Njouodo AS, Koseki S, Keenlyside N, Rouault M. Atmospheric signature of the Agulhas current. Geophys Res Lett. 2018;45(10):5185–5193. https://doi.org/10.1029/2018GL077042
10. Jury M, Walker N. Marine boundary layer modification across the edge of the Agulhas current. J Geophys Res Oceans. 1988; 93(C1):647–654. https://doi.org/10.1029/JC093iC01p00647
11. Walker ND, Mey RD. Ocean/atmosphere heat fluxes within the Agulhas retroflection region. J Geophys Res Oceans. 1988; 93(C12):15473–15483. https://doi.org/10.1029/JC093iC12p15473
12. Walker ND. Links between South African summer rainfall and temperature variability of the Agulhas and Benguela current systems. J Geophys Res Oceans. 1990; 95(C3):3297–3319. https://doi.org/10.1029/JC095iC03p03297
13. Rouault M, Lee-Thorp AM, Lutjeharms JR. The atmospheric boundary layer above the Agulhas current during alongcurrent winds. J Phys Oceanogr. 2000;30(1):40–50. https://doi.org/10.1175/1520-0485(2000)030<0040:TABLAT>2.0.CO;2
14. Rouault M, Reason CJ, Lutjeharms JR, Beljaars AC. Underestimation of latent and sensible heat fluxes above the Agulhas current in NCEP and ECMWF analyses. J Clim. 2003;16(4):776–782. https://doi.org/10.1175/1520-0442(2003)016<0776:UOLASH>2.0.CO;2
15. Lindesay JA. South African rainfall, the Southern Oscillation and a Southern Hemisphere semi-annual cycle. J Climatol. 1988;8(1):17–30. https://doi.org/10.1002/joc.3370080103
16. Rouault M, Richard Y. Spatial extension and intensity of droughts since 1922 in South Africa. Water SA. 2003;29:489–500. http://doi.org/10.4314/wsa.v29i4.5057
17. Rouault M, Richard Y. Intensity and spatial extent of droughts in Southern Africa. Geophys Res Lett. 2005;32(15). https://doi.org/10.1029/2005GL022436
18. Fauchereau N, Pohl B, Reason CJ, Rouault M, Richard Y. Recurrent daily OLR patterns in the Southern Africa/Southwest Indian Ocean region, implications for South African rainfall and teleconnections. Clim Dyn. 2009;32(4):575–591. https://doi.org/10.1007/s00382-008-0426-2
19. Rouault M, Penven P, Pohl B. Warming in the Agulhas current system since the 1980's. Geophys Res Lett. 2009;36(12). https://doi.org/10.1029/2009GL037987
20. Rouault M, Pohl B, Penven P. Coastal oceanic climate change and variability from 1982 to 2009 around South Africa. Afr J Mar Sci. 2010;32(2):237–246. https://doi.org/10.2989/1814232X.2010.501563
21. Hoell A, Funk C, Magadzire T, Zinke J, Husak G. El Niño–Southern Oscillation diversity and Southern Africa teleconnections during austral summer. Clim Dyn. 2015;45(5):1583–1599. https://doi.org/10.1007/s00382-014-2414-z
22. Putrasahan D, Kirtman BP, Beal LM. Modulation of SST interannual variability in the Agulhas leakage region associated with ENSO. J Clim. 2016;29(19):7089–7102. https://doi.org/10.1175/JCLI-D-15-0172.1
23. Reason CJ, Jagadheesha D. Relationships between South Atlantic SST variability and atmospheric circulation over the South African region during austral winter. J Clim. 2005;18(16):3339–3355. https://doi.org/10.1175/JCLI3474.1
24. Singleton AT, Reason CJ. Numerical simulations of a severe rainfall event over the Eastern Cape coast of South Africa: Sensitivity to sea surface temperature and topography. Tellus A: Dyn Meteorol Oceanogr. 2006;58(3):335–367. https://doi.org/10.1111/j.1600-0870.2006.00180.x
25. Singleton AT, Reason CJ. Variability in the characteristics of cut-off low pressure systems over subtropical southern Africa. Int J Climatol. 2007;27(3):295–310. https://doi.org/10.1002/joc.1399
26. Philippon N, Rouault M, Richard Y, Favre A. The influence of ENSO on winter rainfall in South Africa. Int J Climatol. 2012;32(15):2333–2347. https://doi.org/10.1002/joc.3403
27. Burls NJ, Blamey RC, Cash BA, Swenson ET, Fahad AA, Bopape MJ, et al. The Cape Town "Day Zero" drought and Hadley cell expansion. npj Clim Atmos Sci. 2019;2(1):1–8. https://doi.org/10.1038/s41612-019-0084-6

28. Jury MR. Marine climate change over the eastern Agulhas Bank of South Africa. Ocean Sci. 2020;16(6):1529–1544.

29. Polvani LM, Waugh DW, Correa GJ, Son SW. Stratospheric ozone depletion: The main driver of twentieth-century atmospheric circulation changes in the Southern Hemisphere. J Clim. 2011;24(3):795–812. https://doi.org/10.1175/2010JCLI3772.1

30. Thompson DW, Solomon S, Kushner PJ, England MH, Grise KM, Karoly DJ. Signatures of the Antarctic ozone hole in Southern Hemisphere surface climate change. Nat Geosci. 2011;4(11):741–749. https://doi.org/10.1038/ngeo1296

31. Reason CJ, Rouault M. Links between the Antarctic oscillation and winter rainfall over western South Africa. Geophys Res Lett. 2005;32(7). https://doi.org/10.1029/2005GL022419

32. Cheng Y, Beal LM, Kirtman BP, Putrasahan D. Interannual Agulhas leakage variability and its regional climate imprints. J Clim. 2018;31(24):10105–10121. https://doi.org/10.1175/JCLI-D-17-0647.1

33. Taylor KE, Stouffer RJ, Meehl GA. An overview of CMIP5 and the experiment design. Bull Am Meteorol Soc. 2012;93(4):485–498. https://doi.org/10.1175/BAMS-D-11-00094.1

34. Gent PR, Danabasoglu G, Donner LJ, Holland MM, Hunke EC, Jayne SR, et al. The community climate system model version 4. J Clim. 2011;24(19):4973–4991. https://doi.org/10.1175/2011JCLI4083.1

35. Barnes EA, Barnes NW, Polvani LM. Delayed Southern Hemisphere climate change induced by stratospheric ozone recovery, as projected by the CMIP5 models. J Clim. 2014;27(2):852–867. https://doi.org/10.1175/JCLI-D-13-00246.1

36. Biastoch A, Böning CW, Schwarzkopf FU, Lutjeharms JR. Increase in Agulhas leakage due to poleward shift of Southern Hemisphere westerlies. Nature. 2009;462(7272):495–498. https://doi.org/10.1038/nature08519

37. Van Sebille E, Biastoch A, Van Leeuwen PJ, De Ruijter WP. A weaker Agulhas current leads to more Agulhas leakage. Geophys Res Lett. 2009;36(3). https://doi.org/10.1029/2008GL036614

38. Weijer W, Sloyan BM, Maltrud ME, Jeffery N, Hecht MW, Hartin CA, et al. The Southern Ocean and its climate in CCSM4. J Clim. 2012;25(8):2652–2675. https://doi.org/10.1175/JCLI-D-11-00302.1

39. Cheng Y, Putrasahan D, Beal L, Kirtman B. Quantifying Agulhas leakage in a high-resolution climate model. J Clim. 2016;29(19):6881–6892. https://doi.org/10.1175/JCLI-D-15-0568.1

40. Lange M, Van Sebille E. Parcels v0. 9: Prototyping a Lagrangian ocean analysis framework for the petascale age. Geosci Model Dev. 2017;10(11):4175–4186. https://doi.org/10.5194/gmd-10-4175-2017

41. Van Sebille E, Barron CN, Biastoch A, Van Leeuwen PJ, Vossepoel FC, De Ruijter WP. Relating Agulhas leakage to the Agulhas current retroflection location. Ocean Sci. 2009;5(4):511–521. https://doi.org/10.5194/os-5-511-2009

42. Ragoasha N, Herbette S, Cambon G, Veitch J, Reason C, Roy C. Lagrangian pathways in the Southern Benguela upwelling system. J Mar Syst. 2019;195:50–66. https://doi.org/10.1016/j.jmarsys.2019.03.008

43. Ansorge IJ, Speich S, Lutjeharms JR, Goni GJ, Rautenbach CD, Froneman PW, et al. Monitoring the oceanic flow between Africa and Antarctica: Report of the first GoodHope cruise: Research in action. S Afr J Sci. 2005;101(1):29–35.

44. Thompson DW, Solomon S. Interpretation of recent Southern Hemisphere climate change. Science. 2002;296(5569):895–899. https://doi.org/10.1126/science.1069270

45. Beal LM, Elipot S, Houk A, Leber GM. Capturing the transport variability of a western boundary jet: Results from the Agulhas current time-series experiment (ACT). J Phys Oceanogr. 2015;45(5):1302–1324. https://doi.org/10.1175/JPO-D-14-0119.1

46. Chen C, Wang G, Xie SP, Liu W. Why does global warming weaken the Gulf Stream but intensify the Kuroshio? J Clim. 2019;32(21):7437–7451. https://doi.org/10.1175/JCLI-D-18-0895.1

47. Wang G, Xie SP, Huang RX, Chen C. Robust warming pattern of global subtropical oceans and its mechanism. J Clim. 2015;28(21):8574–8584. https://doi.org/10.1175/JCLI-D-14-00809.1

48. Zhang X, Church JA, Platten SM, Monselesan D. Projection of subtropical gyre circulation and associated sea level changes in the Pacific based on CMIP3 climate models. Clim Dyn. 2014;43(1):131–144. https://doi.org/10.1007/s00382-013-1902-x

49. Daher H, Beal LM, Schwarzkopf FU. A new improved estimation of Agulhas leakage using observations and simulations of Lagrangian floats and drifters. J Geophys Res Oceans. 2020;125(4). https://doi.org/10.1029/2019JC015753

50. Biastoch A, Böning CW. Anthropogenic impact on Agulhas leakage. Geophys Res Lett. 2013;40(6):1138–1143. https://doi.org/10.1002/grl.50243

51. Hu Y, Fu Q. Observed poleward expansion of the Hadley circulation since 1979. Atmos Chem Phys. 2007;7(19):5229–5236. https://doi.org/10.5194/acp-7-5229-2007

52. Lu J, Vecchi GA, Reichler T. Expansion of the Hadley cell under global warming. Geophys Res Lett. 2007;34(6). https://doi.org/10.1029/2006GL028443

53. Johanson CM, Fu Q. Hadley cell widening: Model simulations versus observations. J Clim. 2009;22(10):2713–2725. https://doi.org/10.1175/2008JCLI2620.1

54. Polvani LM, Previdi M, Deser C. Large cancellation, due to ozone recovery, of future Southern Hemisphere atmospheric circulation trends. Geophys Res Lett. 2011;38(4). https://doi.org/10.1029/2011GL046712

55. Min SK, Son SW. Multimodel attribution of the Southern Hemisphere Hadley cell widening: Major role of ozone depletion. J Geophys Res Atmos. 2013;118(7):3007–3015. https://doi.org/10.1002/jgrd.50232

56. Choi J, Son SW, Lu J, Min SK. Further observational evidence of Hadley cell widening in the Southern Hemisphere. Geophys Res Lett. 2014;41(7):2590–2597. https://doi.org/10.1002/2014GL059426

AUTHORS:
Buster P. Mogonong[1,3]
Jolene T. Fisher[1]
David Furniss[1]
Debbie Jewitt[1,2]

AFFILIATIONS:
[1]Centre for African Ecology, School of Animal, Plant and Environmental Sciences, University of the Witwatersrand, Johannesburg, South Africa
[2]Conservation Research and Assessment, Scientific Services, Ezemvelo KZN Wildlife, Pietermaritzburg, South Africa
[3]South African Environmental Observation Network (SAEON), Arid Lands Node, Kimberley, South Africa

CORRESPONDENCE TO:
Buster Mogonong

EMAIL:
b.p.mogonong@gmail.com

DATES:
Received: 13 Apr. 2021
Revised: 06 Feb. 2023
Accepted: 17 Apr. 2023
Published: 08 Aug. 2023

HOW TO CITE:
Mogonong BP, Fisher JT, Furniss D, Jewitt D. Land cover change in marginalised landscapes of South Africa (1984–2014): Insights into the influence of socio-economic and political factors. S Afr J Sci. 2023;119(7/8), Art. #10709. https://doi.org/10.17159/sajs.2023/10709

ARTICLE INCLUDES:
☒ Peer review
☒ Supplementary material

DATA AVAILABILITY:
☐ Open data set
☐ All data included
☒ On request from author(s)
☐ Not available
☐ Not applicable

EDITOR:
Sydney Moyo

KEYWORDS:
rural landscapes, grassland, persistence, systematic and intensity analysis, uThukela Municipal District

FUNDING:
South African National Research Foundation (99289; SFH180514329246)

Land cover change in marginalised landscapes of South Africa (1984–2014): Insights into the influence of socio-economic and political factors

Rural landscapes in South Africa experience high conversion rates due to intense land use; however, the changes are site specific and depend on the socio-economic and political history of the area. Land cover change (LCC) was assessed in response to socio-economic and political factors in uThukela Municipal District, KwaZulu-Natal, using Landsat imagery from 1984 to 2014, while making comparisons to other studies in South Africa. Socio-economic/political data were used to gain insights into the observed LCC patterns. Land cover was classified using a random forest classifier, and accuracies ranging from 87% to 92% were achieved. Systematic and intensity analysis methods were used to describe patterns, rates, and transitions of LCC in Imbabazane (ILM) and Okhahlamba (OLM) local municipalities. The results showed a reduced rate of change intensity from 3.4% to 0.9% in ILM and from 3.1% to 1.1% in OLM between 1984 and 2014. Grassland was persistent, covering over 70% in both local municipalities between 1984 and 2014. Although persistent, grassland experienced respective losses of 3.7% and 14.3% in both observation periods in ILM and of 10.2% and 13.3% in OLM. During the analysis period, settlements and cropland gained actively in both local municipalities. The changes represent a degree of population, local authority, and people's perception as influencers of land use and LCC. It is therefore argued that socio-economic and political changes can potentially influence land use and LCC; however, natural ecosystems can persist under those conditions, and this requires more research efforts.

Significance:

This study contributes towards a growing knowledge and understanding of land cover change studies in marginalised landscapes in South Africa. The findings enforce the notion that natural vegetation systems can be altered by human-induced land use such as expansion of settlement and commercial agricultural. We show that in recent times there has been a decline in the overall rate of land cover conversion, and a high persistence of grassland amid global change, although the quality of the vegetation needs further research. We argue that the changes observed in marginalised landscapes are potentially driven by socio-economic and political dynamics.

Introduction

In marginalised rural areas, provisioning ecosystem services are crucial to the livelihoods of local communities.[1,2] Natural resource extraction and use result in land use and land cover change over time. Studies have documented the impacts of human-induced land use and land cover change in various landscapes around the world.[3-12] Various authors have concluded that intense natural resource harvesting, such as fuelwood collection, can cause an ecological crisis.[13,14] Although these conclusions were made, no reports of ecological crises have been made in South African landscapes due to the persistence of natural resources resulting from regulated use by local communities. For example, woody vegetation was persistent in Bushbuckridge, and this was attributed to coppicing and fast regeneration of harvested trees[15,16], while in Zimbabwe, human response to resource scarcity was the driver of ecosystem resilience[17]. Although the ecosystems show resilience, ecosystem service loss remains a concern in rural areas amidst intense land use and land cover change.

Land use and land cover change are regarded as major drivers of biodiversity and ecosystem loss in the world[18] and lead to changes in biogeochemical cycles, increased climatic variability[19] and ecosystem services loss. In rural communities where reliance and demand on natural resources is high, socio-economic and political factors play a role in maintaining key ecosystem services such as provisioning of fuel wood.[19,20,21] Ecosystem services loss can lead to a change in livelihood strategies and force local communities to shift towards non-land-based strategies. This shift is defined by Cumming et al.[22] as a green-red loop. Green-loop systems depend on ecosystem services from the environment, while red-loop systems obtain basic needs from the markets.[22] Land cover change over time can be used to infer ecosystem service loss due to natural land cover loss.[23,24]

There are multiple causes of land cover change in Africa (including South Africa), such as climate and land use change; however, socio-political and economic dynamics are also crucial drivers for shaping pattern and processes of land cover change given the colonial history of the continent.[4,25] Indigenous communities across the continent have been marginalised, and rural populations, therefore, depend on natural resources to sustain and diversify their livelihoods.[1,26,27] In South Africa, marginalised communities were forced onto communal lands, commonly known as Bantustans. Bantustans were a result of the *Bantu Act of 1951*, which saw different ethnic groups separated into ten homelands in South Africa.[28] Bantustan areas were, and still are, less productive and less developed than surrounding landscapes,but are relied on heavily by local people for non-timber forest products to support livelihoods. These landscapes are thus subjected to intense resource extraction resulting in rapid land use change from natural vegetation to anthropogenic land covers[8,9,29-32], which may lead to ecological processes such as land degradation, habitat fragmentation and biodiversity loss. To a greater extent, areas outside Bantustans – such as

commercial farms, urban spaces and mining areas – also experience rapid land use and land cover change. Land use change also results in loss of ecosystem services, leading to a socio-ecological system transition from a green-loop to red-loop. Long-term land cover change studies provide insights that can help to unpack socio-ecological system feedback mechanisms (e.g. green-loop or red-loop[22]) in rural areas.

Few studies have assessed land cover change and its drivers in KwaZulu-Natal. From the few studies that were conducted, the analysis was done over a broader spatial and short temporal scale[33], or the areas of interest were outside the former Bantustan area[10]. Moreover, the influence of socio-economic and political factors on land use and cover change have been understudied in uThukela Municipal District. The landscapes in uThukela Municipal District (KwaZulu-Natal) are a mosaic of natural and anthropogenic land cover. Parts of uThukela Municipal District are within the boundaries of the former KwaZulu Bantustan, resulting in a diverse landscape of social, political, and ecological systems.[28] uThukela Municipal District provides a good opportunity to assess patterns of land cover change in a landscape that has a diverse range of land use activities and socio-economic and political history. The information obtained from land cover change analysis will expand on the existing knowledge of human-induced land cover change patterns in rural South Africa. Furthermore, land cover change analysis will provide insights on how people shape the landscape around them, which is important to integrate into landscape ecology and socio-ecological systems framework.

This study was therefore aimed at understanding the influence of socio-economic and political changes on the patterns of land cover change in uThukela Municipal District, KwaZulu-Natal, South Africa. This was achieved by mapping land cover over a 30-year period in uThukela Municipal District and the use of socio-economic data. Using systematic and intensity analysis, the following were determined: (1) the dominant land cover in the landscape over time, (2) the intensity of change over time, (3) the intensity of gain/loss per land cover class and (4) the transition of categories in the landscape. Furthermore, the observed patterns of land cover change were described based on people's perception of the landscape around them and how changes in socio-economic and political factors influence land cover change and the function of the landscape to provide ecosystem services in marginalised landscapes.

Materials and methods

Study area

The study was conducted in uThukela Municipal District, KwaZulu-Natal, South Africa (Figure 1). Two local municipalities (i.e. Okhahlamba (OLM) and Imbabazane (ILM)) were selected as case studies for land cover change analysis. The two local municipalities consist of areas that are inside and outside of the boundaries of the Bantustan (KwaZulu), resulting in a complex socio-ecological system. The two local municipalities further house various land use activities such as large-scale crop farming, agroforestry, and protected areas, which result in a heterogeneous landscape. Okhahlamba and Imbabazane local municipalities are both Drakensberg, and sub-escarpment grassland dominated landscapes, with little woody vegetation. The two grassland bioregions receive an annual precipitation of 732 mm and 763 mm, respectively.[34]

Imbabazane local municipality was merged with uMtshezi local municipality in August 2016 to form Inkosi Langalibalele local municipality. However, for this study, Imbabazane local municipality boundaries were used to match the 2014 observation point of this study (before merging). Okhahlamba and Imbabazane local municipalities had population sizes of 132 068 and 113 073, respectively, in 2011.[35] Unemployment remains high at 43.4% and 48.6% in OLM and ILM, respectively.[35] Black people dominate these areas, with Zulu being the primary spoken language. Land tenure is administered by the Ingonyama Trust and local authorities govern the use of the landscapes.

Remote sensing data

Image acquisition

Landsat imagery was used to map land cover in OLM and ILM. Winter (August) images were used as they have low atmospheric noise (such as clouds) that may obscure the earth surface. The images were obtained from the USGS website[36], with geometric corrections already performed on these images. Images for path 168 and row 80 were obtained as they cover the study areas in one scene. Three observation points were chosen – 1984, 1991 (Landsat 4-5 TM), and 2014 (Landsat 8 OLI) – and images that correspond to the observation points were obtained. The three observation points were chosen to assess land cover in the early stages of municipal development, giving a baseline land cover (1984), development that might have been impacted by apartheid laws (1991), and development post-apartheid (2014). While

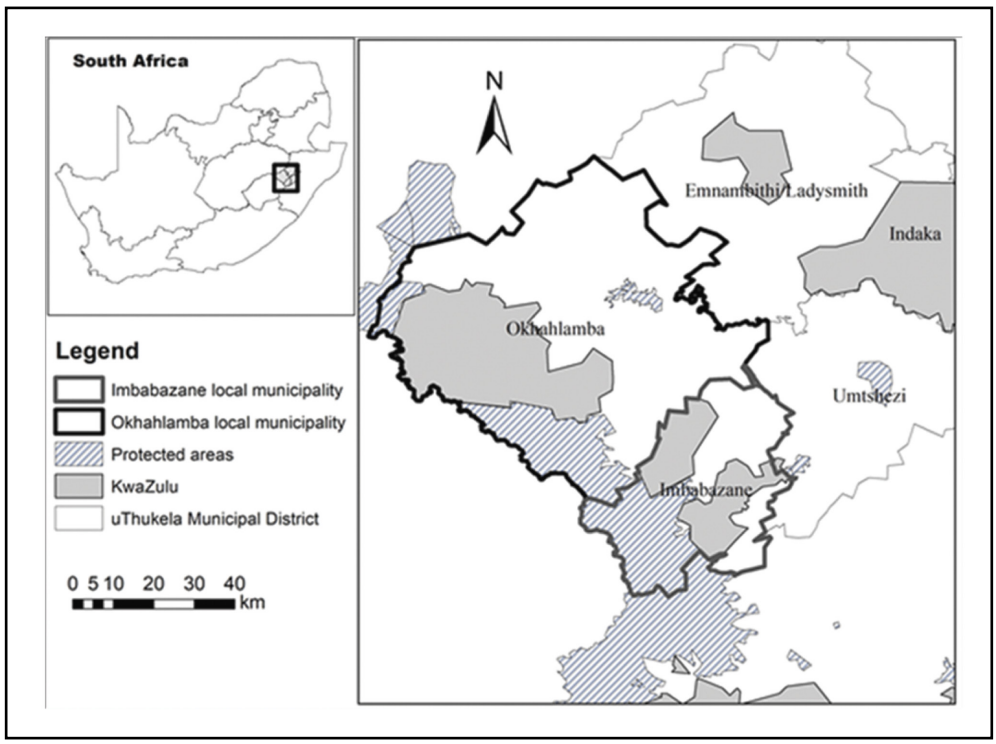


Figure 1: The location of Okhahlamba and Imbabazane local municipalities, and KwaZulu former apartheid homeland in KwaZulu-Natal, South Africa.

the socio-economic and political factors would not have been the only drivers of land cover change, factors such as topography and climate would have remained constant, with the predominant driving influences of change being socio-economic/political. The selected observation points allow for analysing the influence of socio-economic and political factors on land use/land cover change in marginalised landscapes of South Africa.

Image pre-processing

Although the acquired images had negligible cloud interference, image pre-processing of the Landsat images was conducted prior to classification. Radiometric correction followed by atmospheric correction (using the FLAASH tool) was done in ENVI v5.3 to remove all the noise that might have resulted due to atmospheric interference from clouds.

Land cover classification

The classification system by Thompson[37] was revised and used to define land cover categories in this study (Table 1). Random forest classification was used to classify the pixels on the image using R studio v1.1.456 (Raster, rgdal, caret, e1071 and RandomForest packages were used). A minimum of 50 points per class were used to train the random forest classifier model. The output pixelated classified image was then scripted into QGIS v2.18 to classify the different land cover categories by pixel value. Any misclassified pixels were manually assigned to the correct land cover type.

Accuracy assessment was conducted using a combination of aerial photographs, Google Earth, and ground truthing during site visits in January 2019. A minimum of 50 points per land cover category were used for accuracy assessment. The overall accuracy of the classified images ranged from 87% to 90% (Supplementary table A1). User and producer accuracy assessment was also calculated for the individual classes using contingency tables (Supplementary tables A2 and A3).

Land cover change analysis

Protected areas were removed from the analysed images of the two local municipalities in uThukela Municipal District to avoid skewing the data towards high persistence of natural land cover categories. The area removed covered 515 km^2 (13%) in OLM and 405 km^2 (29%) in ILM.

Land cover change was analysed using transition matrix and intensity analysis which uses statistical methods to depict the patterns of land cover change over time.[38,39] The detailed description of these methods is explained in the mentioned references and thus will not be repeated; however, the equations and cross-tabulation template used in this paper are attached as supplementary information (Appendix B). Intensity analysis considers the size of each class and accounts for the number of years in each time interval; thus, the gap between observation points does not affect the analysis process. This method further analyses the data in three levels. The first level is the interval level, which analyses the overall change in the landscape against a uniform change. The change can be categorised as fast or slow compared to the uniform change. The second level is the category level, which assesses the intensity of gain and loss for each land cover category. Gain or loss intensity can be categorised as active or dormant against a uniform intensity. The third level is the transitional level, which analyses the transition of each land cover category from one class to the other and depicts whether a category was targeted or avoided during the changes in the landscape for a particular period. The intensity analysis was performed using the IntensityAnalysis03 Excel software.[40]

Table 1: Land cover classification system derived from the South African National Land Cover Classification System

Land cover type	Brief description
Shrubland	Woodlands that are disturbed, some evidence of fuelwood harvest sites and high coppice growth
Plantation	Planted forests for wood production/timber
Grassland	Areas dominated by grasses and other herbaceous vegetation
Fallow	Abandoned fields, characterised by secondary grassland, including crop fields that are left unplanted for regeneration purposes
Other	Land cover that is not classified in any of the selected land cover types, e.g. roads
Cropland	Cultivated fields, including commercial and subsistence farms
Settlement	Structures associated with human habitation, i.e. homesteads and yards
Bare	Patches of land that are not covered by vegetation
Water body	River, dam and any other water body

Adopted and modified from Thompson[37]

Socio-economic and political data

Socio-economic and political data were used to further the understanding of factors influencing patterns in land cover change. Socio-economic data were obtained from Statistics South Africa (Stats SA) and household interviews were used to gain insight into people's perception of the changes in the case study municipalities. Data on population sizes were used to account for the population change of the entire municipality. Household surveys in the form of questionnaires (open-ended questions) were used to collect socio-economic data during site visits (January 2019). A total of 30 households (n=30) were interviewed per municipality, resulting in a total of 60 participant households. The interviews were targeted at people who had resided in the area for at least 30 years to provide a general overview of the land use and land cover change over time. A referral approach was used, in which a member of the tribal authority was appointed to assist in identifying the participants in the area. Even though a referral approach may be biased towards people with whom the appointed member has good relations, the interviews conducted still reflected the perceptions of the local people residing in the area. The questions asked revolved around people's perception of the landscape integrity (degraded or not) and the role of the local authority regarding land use and tenure.Ethical clearance for the study was granted by the University of the Witwatersrand's Human Research Ethics Committee (reference number H180623).

The analysis of the household surveys data was done using a thematic approach in which responses were grouped into themes. This approach was used to have numerical proportion of response that can be visually presented in the form of a graph. The analysis of the data was performed in Microsoft Excel v.2112.

Results

Land cover

The land cover maps generated for ILM and OLM showed that grassland was the dominant land cover category (Figures 2 and 3), covering 75.5%, 88.7% and 79.5% over the analysis periods in ILM (Figure 4a). In OLM, grassland covered 84.8%, 83.6% and 78.4% over the analysis period (Figure 4b). Commercial farming represented by irrigated croplands was prominent in OLM (2.4% in 2014) compared to ILM (0.4% in 2014); however, the category showed an increasing trend in percentage cover in both local municipalities (Figure 4).

Systematic pattern of land cover change

The landscape in ILM experienced a total change of 23.7% between 1984 and 1991 and 21.4% between 1991 and 2014 (Table 2). Swapping contributed a total of 9.1% to the total change between 1984 and 1991 and 12.4% between 1991 and 2014. Between 1984 and 1991, cropland and settlement transformed 0.3% and 0.1% of grassland, respectively; however, grassland persistence (62.3%) remained high (Supplementary table C1). Between 1991 and 2014, 1.9% and 0.3% of grassland was transformed to settlement and cropland, respectively, indicating settlement and agricultural expansion (Supplementary table C2). In OLM, a total change of 21.7% (1984–1991) and 24.5% (1991–2014) was experienced (Table 3). Swapping contributed 18.2% between 1984 and 1991, and 17.3% between 1991 and 2014. Between 1984 and 1991, 0.2% of grassland was transformed to settlement and 0.6% to cropland. Grassland persistence (69.9%) remained high during this period (Supplementary table C3).

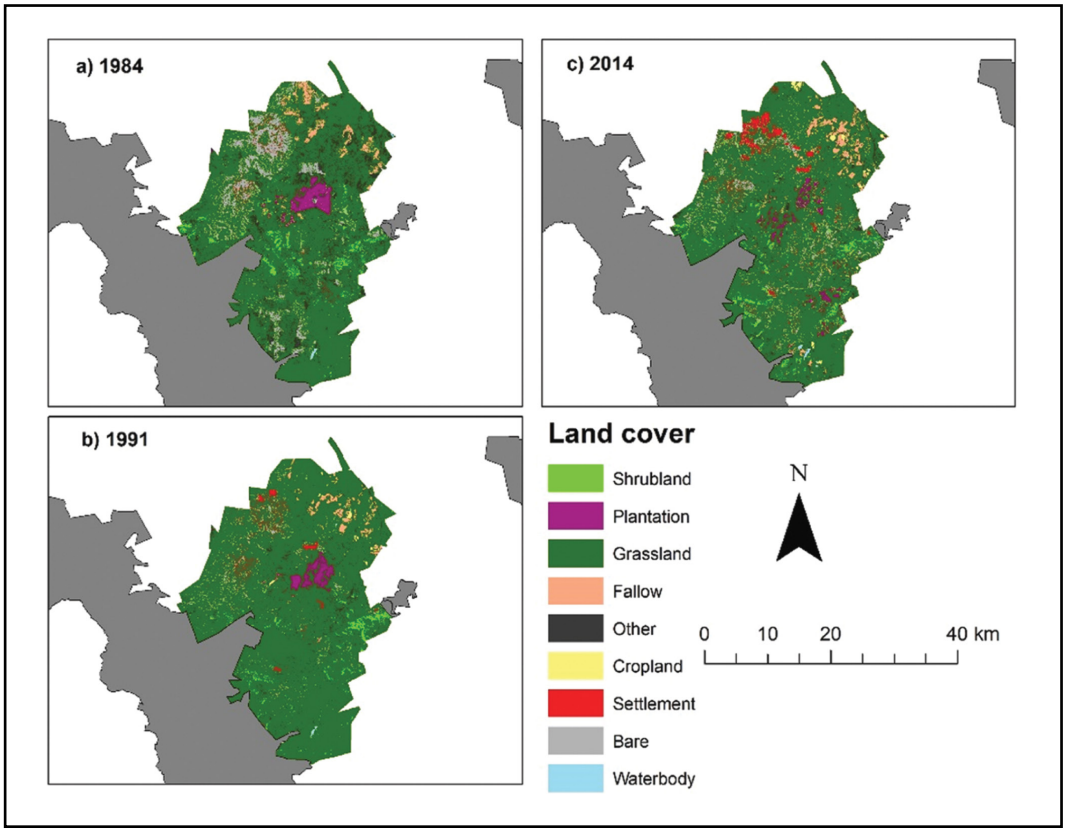


Figure 2: Land cover maps for the three observation time points in Imbabazane local municipality, KwaZulu-Natal, South Africa. The solid grey areas on the map show the boundaries of the area that is covered by protected areas.

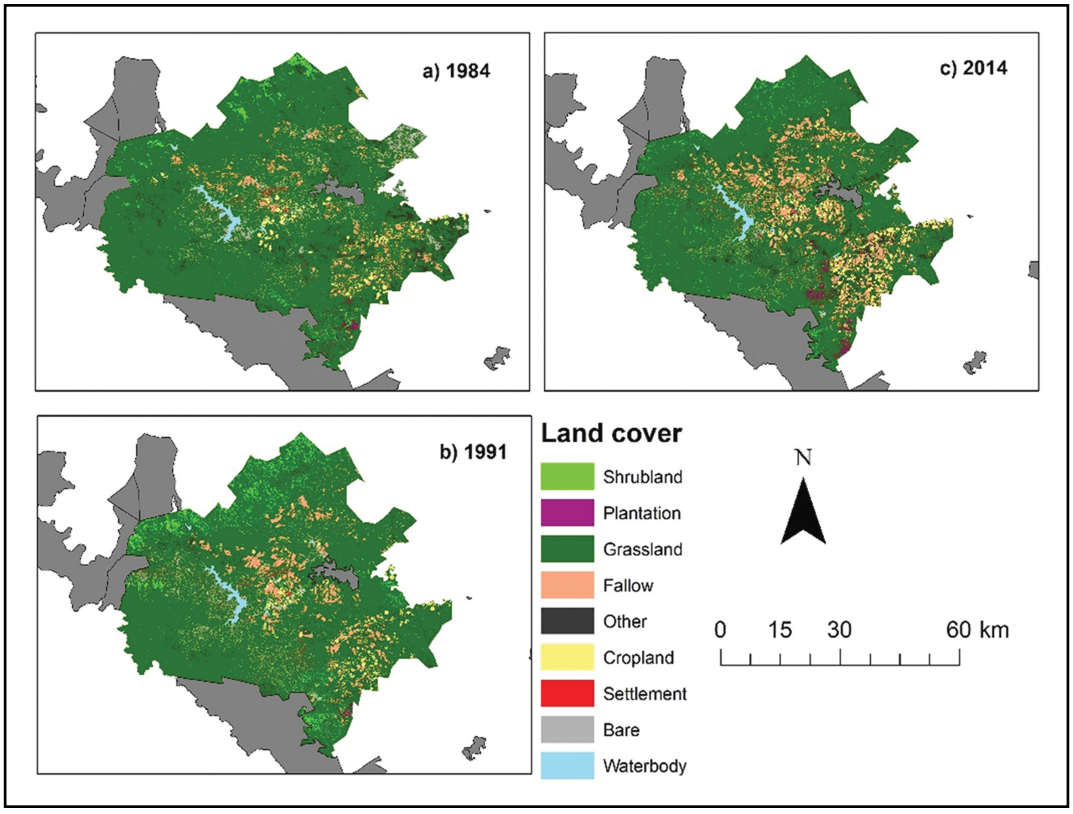


Figure 3: Land cover maps for the three observation time points in Okhahlamba local municipality, KwaZulu-Natal, South Africa. The solid grey areas on the map show the boundaries of the area that is covered by protected areas.

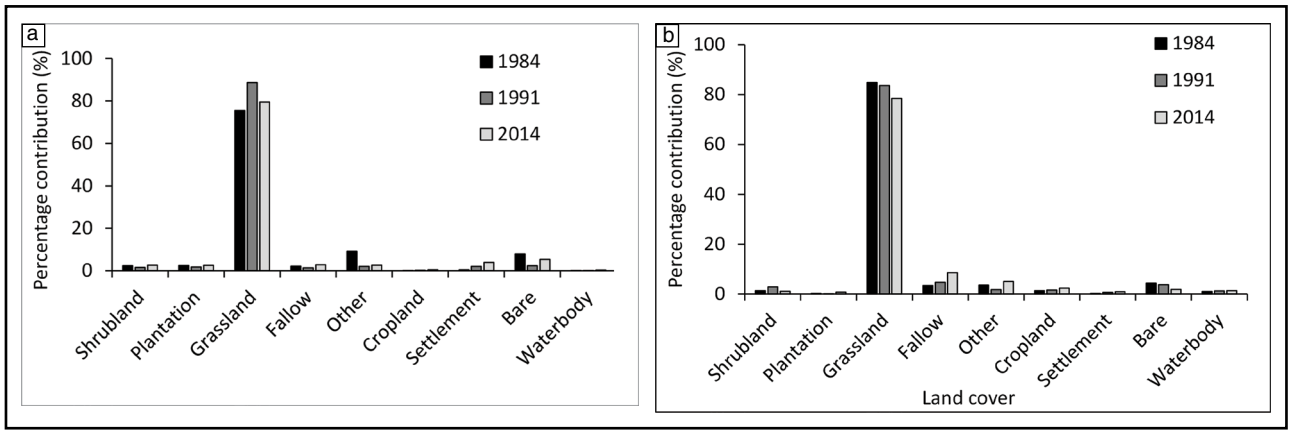


Figure 4: Percentage contribution of each land cover category in the landscapes of (a) Imbabazane and (b) Okhahlamba local municipalities, KwaZulu-Natal, South Africa.

Table 2: Summary table showing the total changes (%) in Imbabazane local municipality landscape between 1984 and 1991 and between 1991 and 2014

	1984–1991					1991–2014				
Land cover	**Gain**	**Loss**	**Total change**	**Swap**	**Absolute value of net change**	**Gain**	**Loss**	**Total change**	**Swap**	**Absolute value of net change**
Shrubland	1.,0	1.8	2.8	2.0	0.8	2.1	1.0	3.0	1.9	1.1
Plantation	0.4	1.1	1.6	0.9	0.7	2.0	1.2	3.1	2.3	0.8
Grassland	17.6	3.7	21.3	7.4	13.9	5.3	14.3	19.5	10.6	9.0
Fallow	1.0	1.8	2.8	2.1	0.7	2.2	0.7	2.9	1.4	1.5
Other	1.6	8.5	10.1	3.1	7.0	2.5	1.8	4.3	3.7	0.6
Cropland	0.2	0.0	0.2	0.0	0.2	0.3	0.1	0.5	0.2	0.2
Settlement	0.7	0.2	0.8	0.3	0.5	2.1	0.5	2.6	1.1	1.5
Bare	1.1	6.5	7.6	2.2	5.5	4.8	1.8	6.5	3.5	3.0
Water body	0.1	0.0	0.1	0.1	0.0	0.2	0.0	0.3	0.1	0.2
Total	**23.7**	**23.7**	**23.7**	**9.1**	**14.6**	**21.4**	**21.4**	**21.4**	**12.4**	**9.0**

Table 3: Summary table showing the total changes (%) in Okhahlamba local municipality landscape between 1984 and 1991 and between 1991 and 2014

	1984–1991					1991–2014				
Land cover	**Gain**	**Loss**	**Total change**	**Swap**	**Absolute value of net change**	**Gain**	**Loss**	**Total change**	**Swap**	**Absolute value of net change**
Shrubland	2.4	0.8	3.2	1.6	1.6	0.8	2.6	3.4	1.5	1.9
Plantation	0.0	0.2	0.2	0.1	0.1	0.6	0.1	0.7	0.1	0.6
Grassland	9.2	10.2	19.4	18.5	0.9	8.0	13.3	21.3	16.0	5.3
Fallow	3.7	2.3	6.1	4.7	1.4	6.1	2.2	8.3	4.4	3.9
Other	1.5	3.4	5.0	3.1	1.9	4.8	1.4	6.2	2.9	3.3
Cropland	1.0	0.8	1.8	1.5	0.2	1.9	1.0	2.9	2.0	0.9
Settlement	0.2	0.1	0.3	0.3	0.1	0.5	0.2	0.7	0.3	0.3
Bare	3.1	3.7	6.8	6.3	0.6	1.4	3.3	4.7	6.6	-1.9
Water body	0.4	0.2	0.6	0.3	0.2	0.4	0.3	0.7	0.6	0.1
Total	**21.7**	**21.7**	**21.7**	**18.2**	**3.5**	**24.5**	**24.5**	**24.5**	**17.3**	**7.2**

Between 1991 and 2014, cropland transformed 1.3% of grassland while settlement transformed 0.3% of this land cover (Supplementary table C4).

Intensity analysis

Interval level

In ILM, the uniform change was 1.5% and the calculated intensity of change at interval level was 3.4% (1984–1991) and 0.9% (1991–2014), indicating a fast change between 1984 and 1991 (Figure 5a). In OLM, the uniform change was also 1.5% and calculated intensity of change was 3.1% (1984–1991) and 1.1% (1991–2014), indicating a fast change between 1984 and 1991 (Figure 5b).

Category level

The category level analysis showed similar patterns of gains and loss for both local municipalities. The uniform intensity of gain/loss was 3.4% (1984–1991) and 0.9% (1991–2014) in ILM. All land cover categories, except for grassland, gained actively for both time periods (Figure 6). Settlement and grassland loss was dormant in both time periods, while the other land cover categories showed active loss. Cropland gain/loss intensity was active in both time periods, even though this land cover category covered a small portion of the landscape.

In OLM, the uniform intensity of gain/loss was 3.1% (1984–1991) and 1.1% (1991–2014). Grassland gain/loss was dormant for both time periods against the uniform intensity (Figure 7). Settlement gain/loss was active against the uniform intensity for both time periods (Figure 7b). Cropland gain/loss was active for both time periods. The other land cover categories' gain/loss intensity was active against the uniform intensity for both time periods, except for 'Water body' which was dormant between 1984 and 1991.

Transition level

The analysis in ILM showed that grassland was mainly transitioned to cropland and fallow between 1984 and 1991 (Supplementary table D1), and cropland only between 1991 and 2014. Settlement encroached into bare areas during this period, while cropland encroached into fallow and 'other'. Cropland was changed to fallow during both time periods, indicating large-scale rotational farming amongst different crop fields.

In OLM, grassland was mainly transitioned to bare, shrubland, fallow and 'other' between 1984 and 1991 (Supplementary table D2). Cropland encroached into fallow, bare, 'other' and plantation for the same period, while fallow replaced cropland. Between 1991 and 2014, cropland was replaced by fallow, while fallow also transitioned to cropland during the same period. Settlement targeted bare during this period.

Socio-economic and political factors

There was a general perception that the function of the landscapes in the two local municipalities are degraded, with 77% and 70% of the participants perceiving the landscape as degraded in ILM and OLM, respectively (Figure 8). There was also a perceived degradation of power of the local authorities in both local municipalities under focus. Over 60% of the locals perceived that people no longer respect the chief (Figure 9). Moreover, the respondents expressed that locals do not follow cultural practices like they used to before 1994 (Supplementary tables E1–E2).

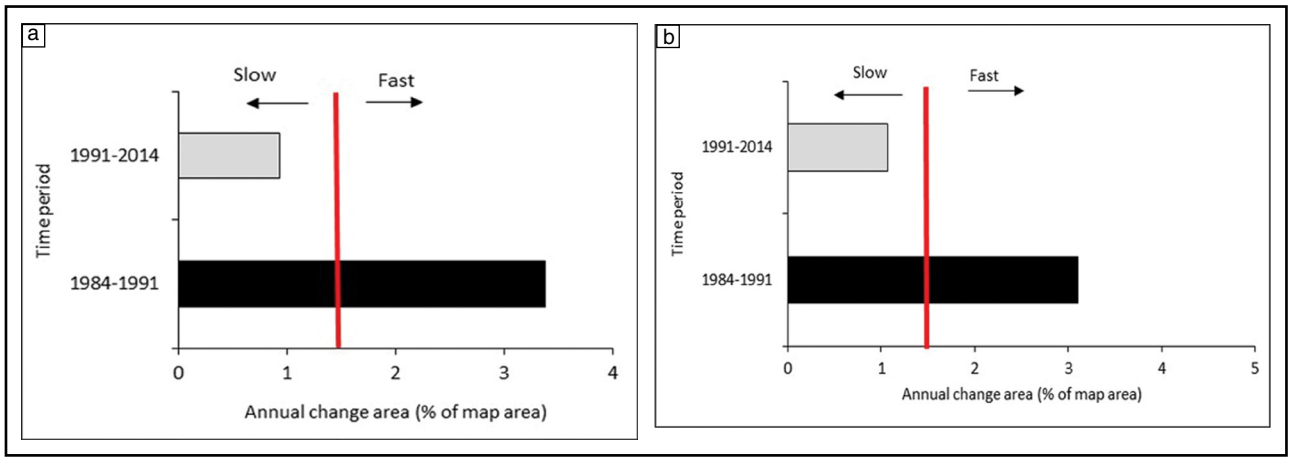


Figure 5: Annual change area for (a) Imbabazane local municipality and (b) Okhahlamba local municipality. The red line in the graphs shows the uniform rate of change, and if the bar falls left of the red line, the change is slow, while if the bar is to the right of the line, change is fast.

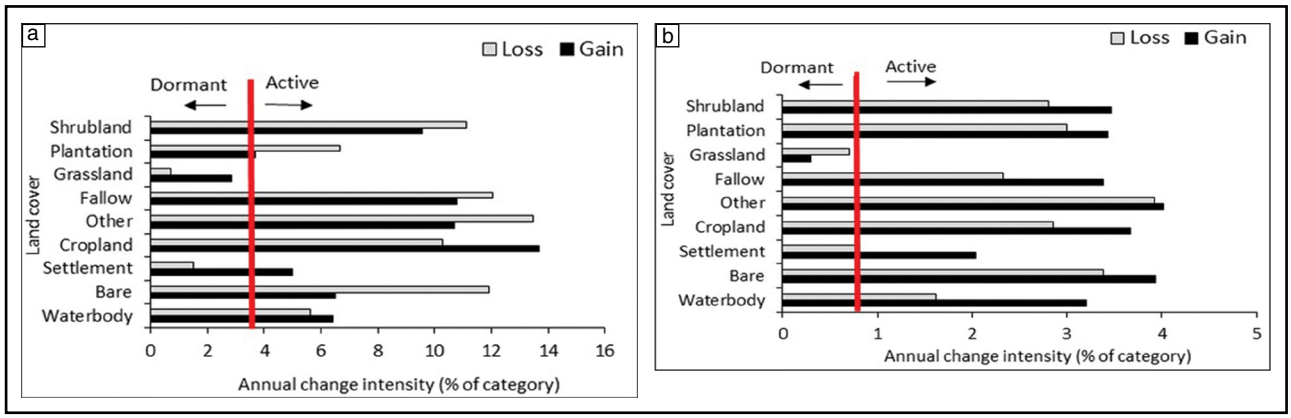


Figure 6: Annual change intensity per category for (a) 1984–1991 and (b) 1991–2014 in Imbabazane local municipality, KwaZulu-Natal, South Africa. The red line in the graphs shows the uniform intensity, and if the bar falls left of the red line, the intensity is dormant, while if the bar is to the right of the line, the intensity is active.

The population numbers in both OLM and ILM declined during the census period (Figure 9), indicating movement of people out of the local municipalities. Overall population declines of 4.7% and 2% were observed in ILM and OLM, respectively, between 2001 and 2016. Although the population numbers declined, the locals perceived the number of homesteads to have increased (Supplementary tables E1–E2).

Discussion

We anticipated a substantial reduction in natural vegetation, together with an increased rate of conversion from natural to anthropogenic land cover over the study period. However, the results of our study were contrary to expectations in a marginalised landscape. The annual change area was reduced by 2.5% and 2% in ILM and OLM, respectively, during the analysis period, which is similar to trends observed in other parts of the world[41], including KwaZulu-Natal[33]. We expected the rate of change

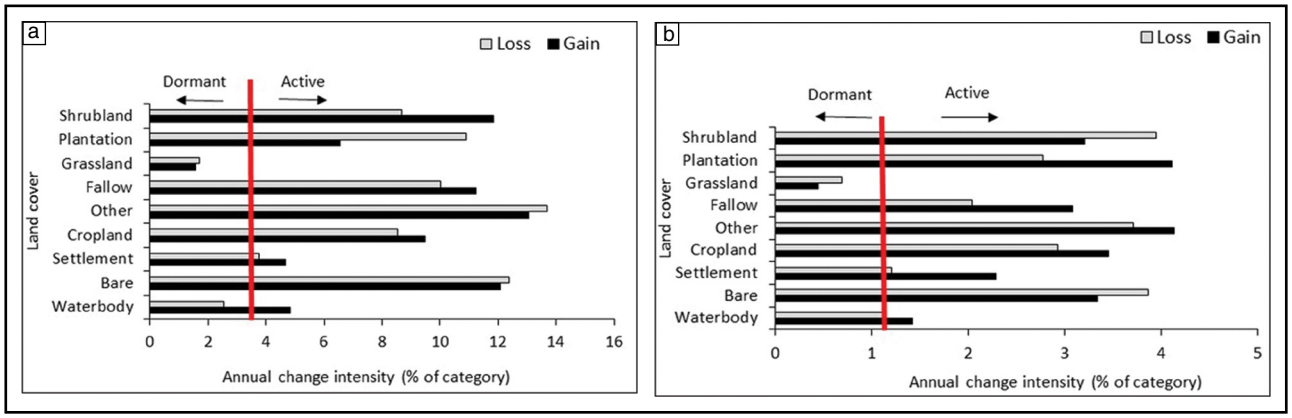


Figure 7: Percentage intensity of gain and loss for each land cover type for the two observation periods in Okhahlamba local municipality, KwaZulu-Natal, South Africa. The red line in the graphs shows the uniform intensity, and if the bar falls left of the red line, the intensity is dormant, while if the bar is to the right of the line, the intensity is active.

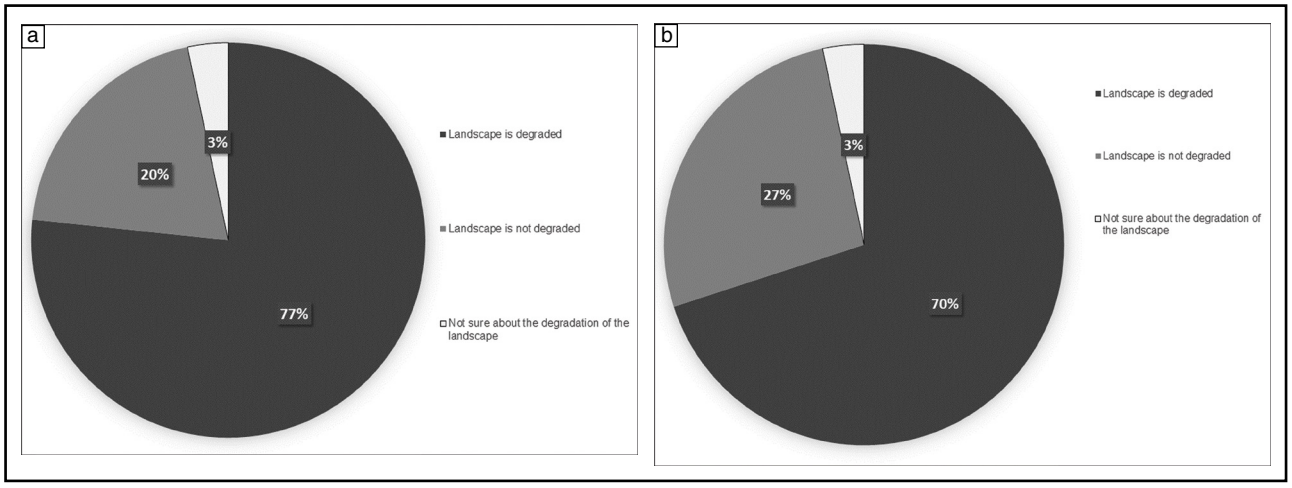


Figure 8: People's perceptions of the quality of the landscape around them. The majority of the 30 participants perceived the landscape to be degraded in (a) Imbabazane and (b) Okhahlamba local municipalities.

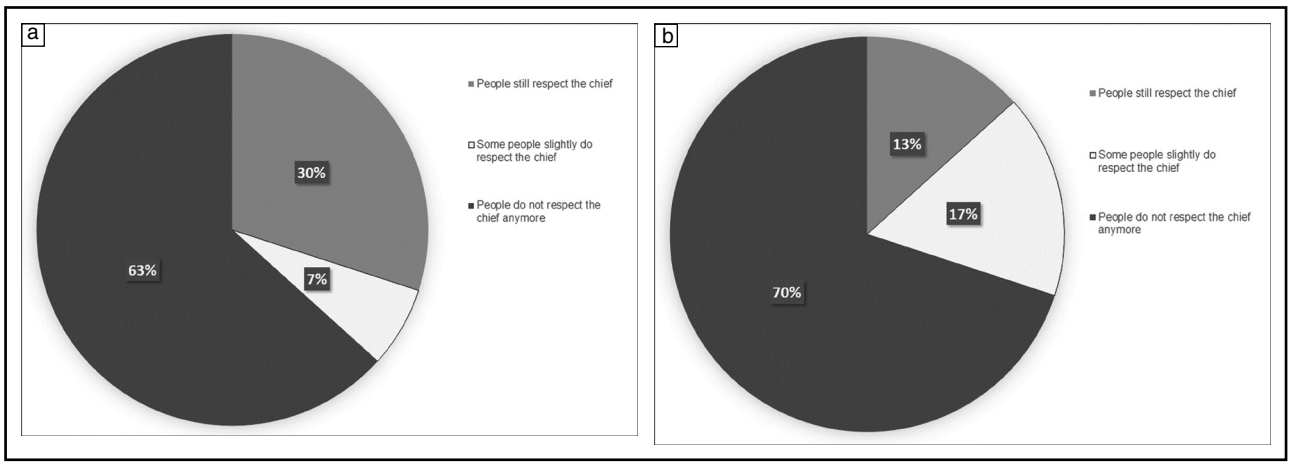


Figure 9: Perception of locals ($n = 30$) on whether people in recent times still respect the local authorities (i.e. chief) in (a) Imbabazane and (b) Okhahlamba local municipalities.

to increase post-apartheid, given that people have the freedom to use the landscapes around them, together with introduction of governmental development programmes such as the Integrated Development Plan. The observed changes show high persistence of grassland, indicating its ability to withstand human impacts. One of the reasons that could be driving the persistence of grassland is the reduced use of the landscapes by villagers, given that there is a general perception that the landscapes in the two local municipalities are degraded (Figure 8). Moreover, the movement of people (outmigration) to the cities has resulted in a reduced population (Figure 10), which coincides with the reduced rate of land cover conversion during the analysis period.

Local authorities such as the chief and headman were previously responsible for managing land tenure and use in rural communities in South Africa.[7,27] These structures are still present; however, there is a perceived degradation of the local authority's power over land use in the studied municipalities (Figure 9). This notion is observed in other rural areas such as Bushbuckridge in South Africa.[7] Reduced influence on land tenure and use by the local authorities has implications for land cover change because locals have a level of freedom to use the landscapes around them in a way that benefits them. For example, harvesting of non-timber forest products could be high and unregulated, and thus leading to loss of natural ecosystems, as seen by the reduced grassland cover in our study area between 1991 and 2014.

The overall land cover change showed an increase in the spatial extent of cropland and settlement in both local municipalities. Settlement and cropland encroached into grassland patches; however, grassland remained persistent during the analysis period. Overall, grassland net change was higher in the first period compared to the second period in ILM, characterised by a high swapping phenomenon and a loss of 3.7% and 14.3% in both observation periods, respectively. Similar to ILM, grassland had a high swapping phenomenon in OLM and a loss of 10.2% and 13.3% in both observation periods, respectively. Although the loss of grassland was high in the recent time periods, the rate in which the conversion was taking place was dormant in both local municipalities. The slow rate of change could partly be explained by the swapping phenomenon, which represents a spatial change in grassland cover where a loss in one location results in a gain in a different location in the landscape, creating a steady loss rate. Encroachment by settlement and cropland on grassland has important ecological implications for the local municipalities in uThukela Municipal District because it can lead to ecological crisis, such as habitat and biodiversity loss.[33,42] Apart from biodiversity loss, disturbing grassland can destabilise the ecosystem's composition and function. Grasslands provide important ecosystem services that are vital for maintaining the primary function of the ecosystem. These services include providing habitats for flora and fauna and regulating atmospheric carbon through carbon sequestration.[43] The results of our study show high spatial persistence of grassland ecosystems in uThukela with over 60% of grassland being persistent throughout the study period (Supplementary tables D1–D4). The quality of these patches was not assessed in this study and thus we cannot comment on the quality, species richness and composition in these patches because the regrowth of secondary grassland occurred in fallow lands; however, local people in the communities perceive the landscapes to have become degraded. Secondary grasslands are associated with low forb species richness and are slower to return to the natural state.[44] Future research is needed to determine how long it takes for secondary grassland to return to natural grassland to understand the impact of land cover change on grassland biodiversity.

Grasslands play a vital role in the livelihoods of rural communities by providing important natural resources necessary to sustain rural populations. Provision of wild fruits, vegetables, grazing land, and medicinal plants are some of the key natural resources obtained from grasslands. Local people rely on natural resources such as non-timber forest products to diversify their livelihoods and generate income to alleviate poverty.[1,2,45] Thus, transforming grasslands can hinder rural livelihood diversification by reducing areas where resources can be harvested in ILM and OLM, forcing people into a 'red loop' where products are bought in the market instead of being harvested.[22,46] The observed reduction in the rate of change post-apartheid potentially indicates that locals are less reliant on land-based strategies to sustain their livelihoods in rural South Africa in recent times. Ragie et al.[47] showed that there is a shift towards off-farm income generation in Bushbuckridge, South Africa, suggesting that rural communities substitute practices such as farming with other income-generating activities such as social grants, remittances, and wage labour. Moreover, competition between uses of communal land for grazing and cultivation could limit the extent of expansion of cropland, which can be a source of local conflict between farm owners and local land users.

The results of the land cover change over the 30-year period show an overall increase in human-induced land cover categories. These results are similar to those of other studies in communal landscapes whereby human-induced land cover categories such as settlement have expanded over time[8,9], resulting in reduced and fragmented natural vegetation near built-up areas. Although the drivers of change were not assessed, the results in this study show that settlement and cropland contributed to grassland loss in OLM and ILM. Built-up, commercial agriculture and mining activities have been observed to be the drivers of natural vegetation loss in KwaZulu-Natal.[33] In other parts of South Africa such as the Kruger to Canyons Biosphere Reserve, intact natural vegetation declined by 6.8% whereas settlement increased by 39.7% between 1993 and 2006.[8] In Bushbuckridge municipality (found in Gazankulu and Lebowa former homeland), woodland declined while settlement and cropland increased between 1965 to 2009.[48] Land cover associated with green space (natural vegetation) declined over a 22-year period in an urban settlement space in eThekwini municipality, KwaZulu-Natal.[49] These findings concur with our results, indicating that anthropogenic activities shape the landscapes in rural and urban areas, leading to natural vegetation loss and land cover change.

Cropland and fallow swapping occurred due to large-scale rotational farming and field abandonment. Field abandonment is one of the most debated processes in South Africa, especially in former Bantustan areas.[50]

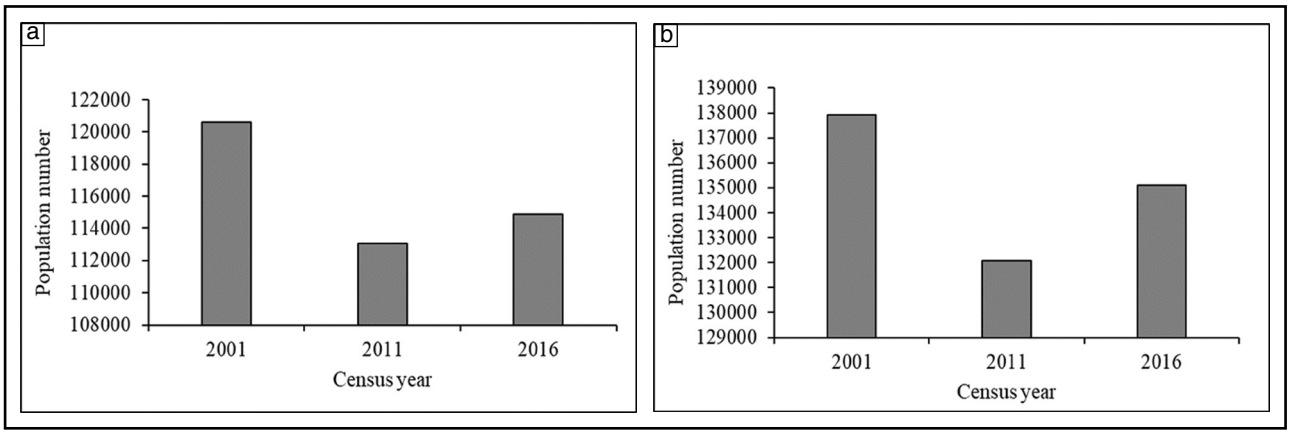


Figure 10: Population numbers in (a) Imbabazane and (b) Okhahlamba local municipalities as per Statistics South Africa (Stats SA) census. An overall decline in population numbers was observed for both local municipalities.

At fine scales such as village level, subsistence farming has drastically declined over time in South African Bantustans.[30-32] Factors that drive cropland abandonment at village scale are similar across South African Bantustans and they include lack of human resources, limited equipment to farm, infertile soil and drought.[30-32] This is no different for the villages in OLM and ILM; however, locals in uThukela Municipal District are shifting towards home gardens rather than owning a big crop field. Shifting towards gardens has also been observed in Transkei and Ciskei former homelands in South Africa.[32] The analysis of the Landsat imagery in this study does not capture this transition because of its resolution, which is a limitation when attempting to understand fine scale patterns of change. This emphasises the importance of considering scale, site visits and conducting semi-structured interviews when assessing land use and land cover change.

The results of this study expand on the existing knowledge of land use and land cover change in South Africa, indicating that the patterns of land use and land cover change are similar across all marginalised landscapes in the country despite the differences in geographic locations. The rate of change was higher in the first time period (i.e.1984–1991) in which total changes of 23.7% and 21.7% of the landscapes were observed in ILM and OLM, respectively. The first time period coincides with the early establishment of the municipalities during the apartheid era in the province and one would expect that some form of development would take place. The drivers of change were not assessed in this study, and thus future studies should seek to understand the drivers of land cover change such as climate, settlement expansion, socio-economic and political factors and mine area expansion.

Conclusion

Landscape transformation was higher in the observation period between 1984 and 1991, but low for the observation period 1991–2014. This can be attributed to rapid development in the early developmental stages of the municipalities, with land use change becoming steady and stable in recent times. Grassland was observed to be persistent in both landscapes, regardless of a loss in both local municipalities. The loss of grassland was dormant throughout the study period, partly due to high swapping phenomenon. Commercial agriculture and settlements encroach in the matrix of these patches. We cannot conclude the drivers of change and quality of the grassland because it was not assessed; however, secondary grassland regrew in previously farmed areas and the landscapes were perceived to be degraded by the local communities. Future research is needed to assess the drivers of land cover change and time span of the transition between secondary and natural grassland in previously farmed areas to understand the influence of socio-economic and political changes on land use and land cover change and its impact on biodiversity and ecosystem services in rural areas.

Acknowledgements

We thank Mr Bheki Radebe for providing expert knowledge of the landscapes in uThukela Municipal District during the ground truthing exercise. We also thank the reviewers for their constructive feedback on the manuscript. This project was funded by the South African National Research Foundation (NRF) Thuthuka Grant (99289) and DAAD-NRF In-Country Master's and Doctoral Scholarships Programme (reference no. SFH180514329246).

Competing interests

We have no competing interests to declare.

Authors' contributions

B.P.M.: Conceptualisation, methodology, software, validation, formal analysis, investigation, data curation, writing – initial draft, visualisation, project administration, funding acquisition. D.J.: Conceptualisation, methodology, validation, writing – review and editing, supervision. J.T.F.: Conceptualisation, methodology, validation, resources, writing – review and editing, supervision, funding acquisition. D.F.: Validation, writing – review and editing. All authors read and agreed to submission of the final version of the manuscript.

References

1. Shackleton C, Shackleton S. The importance of non-timber forest products in rural livelihood security and as safety nets: A review of evidence from South Africa. S Afr J Sci. 2004;100(11):658–664.
2. Shackleton S, Delang CO, Angelsen A. From subsistence to safety nets and cash income: Exploring the diverse values of non-timber forest products for livelihoods and poverty alleviation. In: Non-timber forest products in the global context. Springer; 2011. p. 55–81. https://doi.org/10.1007/978-3-642-17983-9_3
3. Gibbs HK, Ruesch AS, Achard F, Clayton MK, Holmgren P, Ramankutty N, et al. Tropical forests were the primary sources of new agricultural land in the 1980s and 1990s. Proc Natl Acad Sci USA. 2010;107(38):16732–16737. https://doi.org/10.1073/pnas.0910275107
4. Reid RS, Kruska RL, Muthui N, Taye A, Wotton S, Wilson CJ, et al. Land-use and land-cover dynamics in response to changes in climatic, biological and socio-political forces: The case of southwestern Ethiopia. Landsc Ecol. 2000;15(4):339–355. https://doi.org/10.1023/A:1008177712995
5. Mapedza E, Wright J, Fawcett R. An investigation of land cover change in Mafungautsi Forest, Zimbabwe, using GIS and participatory mapping. Appl Geogr. 2003;23(1):1–21. https://doi.org/10.1016/S0143-6228(02)00070-X
6. Kiage L, Liu KB, Walker N, Lam N, Huh O. Recent land-cover/use change associated with land degradation in the Lake Baringo catchment, Kenya, East Africa: Evidence from Landsat TM and ETM+. Int J Remote Sens. 2007;28(19):4285–4309. https://doi.org/10.1080/01431160701241753
7. Kirkland T, Hunter LM, Twine W. "The bush is no more": Insights on institutional change and natural resource availability in rural South Africa. Soc Nat Resour. 2007;20(4):337–350. https://doi.org/10.1080/08941920601161353
8. Coetzer KL, Erasmus BF, Witkowski E, Bachoo AK. Land-cover change in the Kruger to Canyons Biosphere Reserve (1993-2006): A first step towards creating a conservation plan for the subregion. S Afr J Sci. 2010;106(7–8), Art. #221. https://doi.org/10.4102/sajs.v106i7/8.221
9. Coetzer KL, Erasmus BF, Witkowski ET, Reyers B. The race for space: Tracking land-cover transformation in a socio-ecological landscape, South Africa. Environ Manage. 2013;52(3):595–611. https://doi.org/10.1007/s00267-013-0094-9
10. Burgoyne C, Kelso C, Ahmed F. Human activity and vegetation change around Mkuze Game Reserve, South Africa. S Afr Geogr J. 2016;98(2):217–234. https://doi.org/10.1080/03736245.2015.1028978
11. Pullanikkatil D, Palamuleni LG, Ruhiiga TM. Land use/land cover change and implications for ecosystems services in the Likangala River Catchment, Malawi. Phys Chem Earth. 2016;93:96–103. https://doi.org/10.1016/j.pce.2016.03.002
12. Pullanikkatil D, Palamuleni L, Ruhiiga T. Assessment of land use change in Likangala River catchment, Malawi: A remote sensing and DPSIR approach. Appl Geogr. 2016;71:9–23. https://doi.org/10.1016/j.apgeog.2016.04.005
13. Dovie DB, Witkowski E, Shackleton CM. The fuelwood crisis in southern Africa-relating fuelwood use to livelihoods in a rural village. GeoJournal. 2004;60(2):123–133. https://doi.org/10.1023/B:GEJO.0000033597.34013.9f
14. Wessels KJ, Colgan M, Erasmus BFN, Asner G, Twine W, Mathieu R, et al. Unsustainable fuelwood extraction from South African savannas. Environ Res Lett. 2013;8(1):014007. https://doi.org/10.1088/1748-9326/8/1/014007
15. Twine WC, Holdo RM. Fuelwood sustainability revisited: Integrating size structure and resprouting into a spatially realistic fuelshed model. J Appl Ecol. 2016;53(6):1766–1776. https://doi.org/10.1111/1365-2664.12713
16. Swemmer AM, Mashele M, Ndhlovu PD. Evidence for ecological sustainability of fuelwood harvesting at a rural village in South Africa. Reg Environ Change. 2019;19(2):403–413. https://doi.org/10.1007/s10113-018-1402-y
17. Vermeulen S, Campbell B, Mangono J. Shifting patterns of fuel and wood use by households in rural Zimbabwe. Energy Environ. 2000;11(3):233–254. https://doi.org/10.1260/0958305001500112
18. Millennium Ecosystem Assessment. Ecosystems and human well-being: Health synthesis. Geneva: World Health Organization; 2005.
19. Feddema JJ, Oleson KW, Bonan GB, Mearns LO, Buja LE, Meehl GA, et al. The importance of land-cover change in simulating future climates. Science. 2005;310(5754):1674–1678. https://doi.org/10.1126/science.1118160

20. Berkes F, Colding J, Folke C. Rediscovery of traditional ecological knowledge as adaptive management. Ecol Appl. 2000;10(5):1251–1262. https://doi.org/10.1890/1051-0761(2000)010[1251:ROTEKA]2.0.CO;2

21. Cumming GS. Spatial resilience: Integrating landscape ecology, resilience, and sustainability. Landsc Ecol. 2011;26(7): 899–909. https://doi.org/10.1007/s10980-011-9623-1

22. Cumming GS, Buerkert A, Hoffmann EM, Schlecht E, Von Cramon-Taubadel S, Tscharntke T. Implications of agricultural transitions and urbanization for ecosystem services. Nature. 2014;515(7525):50–57. https://doi.org/10.1038/nature13945

23. Wu K, Ye X, Qi Z, Zhang H. Impacts of land use/land cover change and socioeconomic development on regional ecosystem services: The case of fast-growing Hangzhou metropolitan area, China. Cities. 2013;31:276–284. https://doi.org/10.1016/j.cities.2012.08.003

24. Karki S, Thandar AM, Uddin K, Tun S, Aye WM, Aryal K, et al. Impact of land use land cover change on ecosystem services: A comparative analysis on observed data and people's perception in Inle Lake, Myanmar. Environ Syst Res. 2018;7(1):25. https://doi.org/10.1186/s40068-018-0128-7

25. McCusker B. Land use and cover change as an indicator of transformation on recently redistributed farms in Limpopo Province, South Africa. Hum Ecol. 2004;32(1):49–75. https://doi.org/10.1023/B:HUEC.0000015220.22795.27

26. De Boer WF, Baquete DS. Natural resource use, crop damage and attitudes of rural people in the vicinity of the Maputo Elephant Reserve, Mozambique. Environ Conserv. 1998;25(3):208–218. https://doi.org/10.1017/S0376892998000265

27. Twine W. Socio-economic transitions influence vegetation change in the communal rangelands of the South African lowveld. Afr J Range Forage Sci. 2005;22(2):93–99. https://doi.org/10.2989/10220110509485866

28. Ramutsindela M. Resilient geographies: Land, boundaries and the consolidation of the former bantustans in post-1994 South Africa. Geogr J. 2007;173(1):43–55. https://doi.org/10.1111/j.1475-4959.2007.00230.x

29. Giannecchini M, Twine W, Vogel C. Land-cover change and human-environment interactions in a rural cultural landscape in South Africa. Geogr J. 2007;173(1):26–42. https://doi.org/10.1111/j.1475-4959.2007.00227.x

30. Blair D, Shackleton CM, Mograbi PJ. Cropland abandonment in South African smallholder communal lands: Land cover change (1950-2010) and farmer perceptions of contributing factors. Land. 2018;7(4):121. https://doi.org/10.3390/land7040121

31. Hebinck P, Mtati N, Shackleton C. More than just fields: Reframing deagrarianisation in landscapes and livelihoods. J Rural Stud. 2018;61:323–334. https://doi.org/10.1016/j.jrurstud.2018.01.004

32. Shackleton C, Mograbi P, Drimie S, Fay D, Hebinck P, Hoffman M, et al. Deactivation of field cultivation in communal areas of South Africa: Patterns, drivers and socio-economic and ecological consequences. Land Use Policy. 2019;82:686–699. https://doi.org/10.1016/j.landusepol.2019.01.009

33. Jewitt D, Goodman PS, Erasmus BF, O'Connor TG, Witkowski ET. Systematic land-cover change in KwaZulu-Natal, South Africa: Implications for biodiversity. S Afr J Sci. 2015;111 (9–10), Art. #2015-0019. https://doi.org/10.17159/sajs.2015/20150019

34. Mucina L, Rutherford M. The vegetation of South Africa, Lesotho and Swaziland. Strelitzia 19. Pretoria: South African National Biodiversity Institute; 2006.

35. Statistics South Africa (Stats SA). Statistics by place [webpage on the Internet]. c2011 [cited 2019 Feb 09]. Available from: https://www.statssa.gov.za/?page_id=964

36. USGS. Viewer [webpage on the Internet]. No date [cited 2018 Mar 05]. Available from: https://earthexplorer.usgs.gov/

37. Thompson M. A standard land-cover classification scheme for remote-sensing applications in South Africa. S Afr J Sci. 1996;92(1):34–42.

38. Pontius Jr RG, Shusas E, McEachern M. Detecting important categorical land changes while accounting for persistence. Agric Ecosyst Environ. 2004;101(2–3):251–268. https://doi.org/10.1016/j.agee.2003.09.008

39. Aldwaik SZ, Pontius Jr RG. Intensity analysis to unify measurements of size and stationarity of land changes by interval, category, and transition. Landsc Urban Plan. 2012;106(1):103–114. https://doi.org/10.1016/j.landurbplan.2012.02.010

40. Aldwaik SZ, Pontius Jr RG. IntensityAnalysis03 [software program]; 2009.

41. Minaei M, Kainz W. Land cover change dynamics based on intensity analysis in gorganrood watershed, Iran. J Agric Sci Technol. 2018;20(5):965–978.

42. Sala OE, Chapin FS, Armesto JJ, Berlow E, Bloomfield J, Dirzo R, et al. Global biodiversity scenarios for the year 2100. Science. 2000;287(5459):1770–1774. https://doi.org/10.1126/science.287.5459.1770

43. Bengtsson J, Bullock JM, Egoh B, Everson C, Everson T, O'Connor T, et al. Grasslands – more important for ecosystem services than you might think. Ecosphere. 2019;10(2):e02582. https://doi.org/10.1002/ecs2.2582

44. Zaloumis NP, Bond WJ. Grassland restoration after afforestation: No direction home? Austral Ecol. 2011;36(4):357–366. https://doi.org/10.1111/j.1442-9993.2010.02158.x

45. Shackleton CM, Shackleton SE. Household wealth status and natural resource use in the Kat River valley, South Africa. Ecol Econ. 2006;57(2):306–317. https://doi.org/10.1016/j.ecolecon.2005.04.011

46. Hamann M, Biggs R, Reyers B. Mapping social-ecological systems: Identifying 'green-loop' and 'red-loop' dynamics based on characteristic bundles of ecosystem service use. Glob Environ Change. 2015;34:218–226. https://doi.org/10.1016/j.gloenvcha.2015.07.008

47. Ragie FH, Olivier DW, Hunter LM, Erasmus BF, Vogel C, Collinson M, et al. A portfolio perspective of rural livelihoods in Bushbuckridge, South Africa. S Afr J Sci. 2020;116 (9–10), Art. #7522. https://doi.org/10.17159/sajs.2020/7522

48. Matsika R. The spatio-temporal dynamics of woody biomass supply and demand in response to human utilisation in an African savanna woodland [PhD thesis]. Johannesburg: University of the Witwatersrand; 2012.

49. Otunga C, Odindi J, Mutanga O. Land use land cover change in the fringe of eThekwini Municipality: Implications for urban green spaces using remote sensing. S Afr J Geomat. 2014;3(2):145–162. https://doi.org/10.4314/sajg.v3i2.3

50. Mkhongi FA, Musakwa W. Trajectories of deagrarianization in South Africa – Past, current and emerging trends: A bibliometric analysis and systematic review. Geogr Environ Sustain. 2022;3(4):325–333. https://doi.org/10.1016/j.geosus.2022.10.003