Landscape and Urban Planning 233 (2023) 104707

Contents lists available at ScienceDirect

Landscape and Urban Planning

journal homepage: www.elsevier.com/locate/landurbplan

Review Article

Status of urban ecology in Africa: A systematic review

Adewale G. Awoyemi [a,b,c,*], Juan Diego Ibáñez-Álamo [a,*]

[a] Department of Zoology, Faculty of Sciences, University of Granada, Granada, Spain
[b] Forest Center, International Institute of Tropical Agriculture, Ibadan, Nigeria
[c] A.P. Leventis Ornithological Research Institute, Jos, Nigeria

HIGHLIGHTS

- African urban ecology is understudied.
- There are important geographic, ecological and scientific biases.
- Urban ecology is significantly more studied in wealthier African countries.
- More urbanized areas (now or in the future) are not the main focus of study.
- We need to redirect our priorities regarding urban ecology in Africa.

ARTICLE INFO

Keywords:
Africa
Biodiversity
Literature search
Research effort
Review
Urban ecology

ABSTRACT

Urbanization is an extreme human activity and is expanding worldwide, consequently increasing the attention of scientists across research areas of urban ecology. Recent studies have warned of the lack of information from certain regions, particularly Africa, which is rapidly urbanizing. Thus, we did a detailed literature search to determine the state of knowledge in African urban ecology in the last century. We found 795 relevant papers from where data were collected and tested to understand geographic and ecological mismatches in research effort, allowing us to identify important knowledge gaps (e.g., taxonomy and scientific fields). We also tested the effect of current and future urbanization intensity, human population density, size and conservation status of ecoregions and Gross Domestic Product (GDP) on research effort. Our results suggest a low turnout of papers and a dearth of knowledge about African urban ecology. Studies were conducted in 72% of African countries, with South Africa alone accounting for almost 40% of all published papers. The studies were either conducted at the city (55%) or local/country (34%) level, suggesting the lack of transnational research collaboration. Interestingly, only country GDP and the size and conservation status of ecoregions significantly predicted the number of publications, suggesting that research effort is driven by economic reasons and the relevance of conservation in African urban ecology. We need to account for these biases to advance our understanding of the impacts of urbanization on African biodiversity.

1. Introduction

Rapidly expanding urbanization is a major threat to nature worldwide, leading to the reduction of biodiversity and alteration of species interactions and ecosystem services (Gaston, 2010; McDonald, Kareiva, & Forman, 2008; McKinney, 2006; United Nations, 2016). The impacts of urbanization could be even worse in the near future due to the geometric progression of human population. According to the United Nations (2019), the global human population density will increase from 60 humans/km^2 in 2020 to 78 humans/km^2 in 2050, while the global urban land cover will increase from 824,200 km^2 to 1,145,698 km^2 during the same period (Angel et al., 2011). Thus, research on urban ecology is imperative to achieve sustainable development, allowing for the understanding of ecological processes in urban areas and providing necessary data for urban planning, landscape design, policy formulation and biodiversity conservation (Corbyn, 2010; Moragues-Faus & Carroll, 2018).

Given the availability of various definitions of urban ecology, we follow the scientific proposition that incorporates the 'interaction of organisms, built structures and the physical environment where people

* Corresponding authors at: Avda. Fuentenueva s/n, 18071 Granada, Spain (A.G. Awoyemi).
E-mail addresses: Awoyemi@correo.ugr.es (A.G. Awoyemi), jia@ugr.es (J.D. Ibáñez-Álamo).

https://doi.org/10.1016/j.landurbplan.2023.104707
Received 2 February 2022; Received in revised form 23 January 2023; Accepted 23 January 2023
Available online 31 January 2023

are concentrated' (Forman, 2014). Due to the transformative potential of urbanization, the concept of social and ecological integration (inclusiveness) has been proposed to enhance biodiversity in urban areas (e.g., Haase et al., 2017). For instance, Ferketic et al., (2010) demonstrated the usefulness of inclusiveness in promoting conservation justice in Cape Town (South Africa), thereby influencing the ecology of the city, and an understanding of such a nexus is useful to design resilient and sustainable urban areas (Childers et al., 2015; Grimm et al., 2008).

The globally recognized multi-disciplinary fields and the embedded scientific topics in urban ecology have attracted increasing attention from researchers (e.g., Anderson et al., 2013; Cilliers et al., 2013; Girma et al., 2019). However, several papers have highlighted important knowledge gaps across regions, taxa and scientific topics (e.g., Magle et al., 2012; Tóth et al., 2020; van der Walt et al., 2015). Probably, one of the most important mismatches between urban ecology research effort and the urbanization process is the lack of knowledge on the topic from the most rapidly urbanizing continents of South America, Asia and Africa (Ibáñez-Álamo et al., 2017; Seto et al., 2012; Shackleton et al., 2021). As identified in these studies, geographic biases impede the full comprehension of the real impacts of urbanization on nature. Future studies conducted in appropriate areas will therefore be useful to determine ameliorative strategies needed to promote the co-existence of humans with nature, thereby enhancing urban habitats and the associated biodiversity, which is in line with the 11th Sustainable Development Goal of the United Nations (2021).

Literature reviews provide an opportunity for summarizing the state of evidence-based knowledge applied in many fields (e.g., Ibáñez-Álamo et al., 2017; Magle et al., 2012). Broadly, this involves the incorporation of published literature in any given field (Garousi et al., 2019). However, the generalization and application of findings from literature reviews in decision-making have been a subject for debate, mainly due to transparency, objectivity, repeatability and credibility (Sánchez-Tójar et al., 2020). Since traditional approaches to literature reviews are prone to errors (Grant & Booth, 2009), rigorous methodological approaches have been developed and applied more recently in the field of urban ecology (e.g., Cilliers et al., 2018; Kendal et al., 2020; Ibáñez-Álamo et al., 2017), allowing for an important advancement in our understanding of the effect of urban areas on organisms.

In the present study, we conducted a systematic literature review to determine trends in urban ecological research conducted in Africa. Relative to other regions such as Asia, Europe and North America (Forman, 2016; Lin & Grimm, 2015; Magle et al., 2012; Wu et al., 2014), there have been few attempts aimed at synthesizing the state of knowledge in African urban ecology (e.g., Cilliers et al., 2013; Shackleton et al., 2017; Lindley et al., 2018; du Toit et al., 2018). Our aims were to (i) analyze the current status of research effort on urban ecology in this continent, (ii) identify research gaps (geographic, taxonomic and ecological) and (iii) provide recommendations and insights on future prospects. Additionally, (iv) we investigated the potential association of urban ecology research effort with some factors previously associated with the number of scientific publications. On the one hand, we tested whether the number of publications in the field (i.e., urban ecology) per country could be influenced by human population density, economic wealth, as well as the current or future urbanization prospects. Given the positive association between human population density and the degree of urbanization (e.g., Gao & O'Neill, 2021; Qizhi et al., 2016), we would expect that countries with high human population density would hold the majority of studies in urban ecology. Furthermore, if urban ecology research effort is driven by the intensity of urbanization, based on the scientific reasoning of geographic focus areas of particular interest, we could predict a positive association of the number of publications on this topic in those countries currently more urbanized or with the highest rate of urban expansion (i.e., future urbanization). Although the relationship between urbanization and economic growth is often contested (e.g., Chen et al., 2014; Moomaw & Shatter, 1996), we would expect that wealthier countries (i.e., higher Gross Domestic Product –GDP–) are those concentrating the majority of urban ecological studies as increased funding positively influences publication rates (Man et al., 2004). On the other hand, we also tested whether the number of publications in the field could be influenced by the conservation status and size of African ecoregions. Previous reviews have pointed out the positive association between the conservation status of study sites and research effort (e.g., de Lima et al., 2011). Thus, if research effort is based on conservation-oriented reasons, we would expect that threatened ecoregions will be more studied. In addition, since smaller areas generally support lower species richness (see Rantalainen et al., 2005), we would expect that larger ecoregions will provide more study opportunities for researchers specializing in different species and scientific topics, and will therefore be more studied. Considering the marked differences between Global North and Global South urban settings (Shackleton et al., 2021), we acknowledge that there could be other factors (e.g., climate severity, colonial history or high diversity in human-nature interactions) shaping the urban ecology research effort in Africa, which is considered part of the Global South. However, we did not include them because of the difficulty of extracting such information and to avoid over-parameterization of models. Findings of this study will provide additional information about African urban landscapes that should generate interest among researchers, conservation practitioners and policy-makers.

2. Methods

2.1. Bibliographic search and paper screening

We performed a literature search in Web of Science, Google Scholar and Scopus on 8 March 2021 using different combinations of 89 relevant keywords within the article titles, abstracts and keywords, covering the period 1920–2020. The search string containing research focus (23 keywords; e.g., ecology, biodiversity and wellbeing) and urban terms (5 keywords; e.g., urban, city and town) were matched with region (Africa and country name). We performed independent searches for each of the 58 countries and autonomous territories in the continent. A detailed description of these search terms, and the relevant Web of Science categories (41) and Scopus study fields (10) selected can be found in Table S1. The relevance of the use of such comprehensive keywords has been demonstrated by previous studies (e.g., Raji & Downs, 2021; Roy et al., 2012; Tan & bin Abdul Hamid, 2014).

We then uploaded all detected papers on Rayyan (https://www.rayyan.ai/) for screening. Rayyan is a web-based App that uses a semi-automation process to screen paper's preliminary pages with a high degree of precision (Olofsson et al., 2017; Ouzzani et al., 2016). Its adaptability and many functions allow the detection of duplicates, verification, collaboration and decisions in systematic reviews (Abreha, 2019; de Keijzer et al., 2016). In the present study, both authors independently performed the paper selection process by activating the "blind function" in Rayyan and reached a consensus thereafter.

Our selection process followed the Preferred Reporting Items for Systematic Reviews and meta-analyses (PRISMA Statement) (Abreha, 2019; Moher et al., 2009), which is presented in Fig. 1. Based on article titles and abstracts, we first excluded duplicates, non-African studies and investigations carried out outside urban settings. We also excluded papers on human diseases, climate change, pollution and agriculture when they were exclusively focused on clear different disciplines, such as malaria studies exclusively focused on the medical science (e.g., Kigozi et al., 2020) or agricultural papers investigating different crop varieties without any socio-ecological, biodiversity or human dimensions focus (e.g., Kent et al., 2001). Several systematic reviews already exist on these disciplines (e.g., Fayiga et al., 2018; Hulme et al., 2001; Orsini et al., 2013). The remaining articles were then screened and those that met the following criteria were retained for data extraction: (1) urban landscape, ecological and sociological studies, (2) journal articles

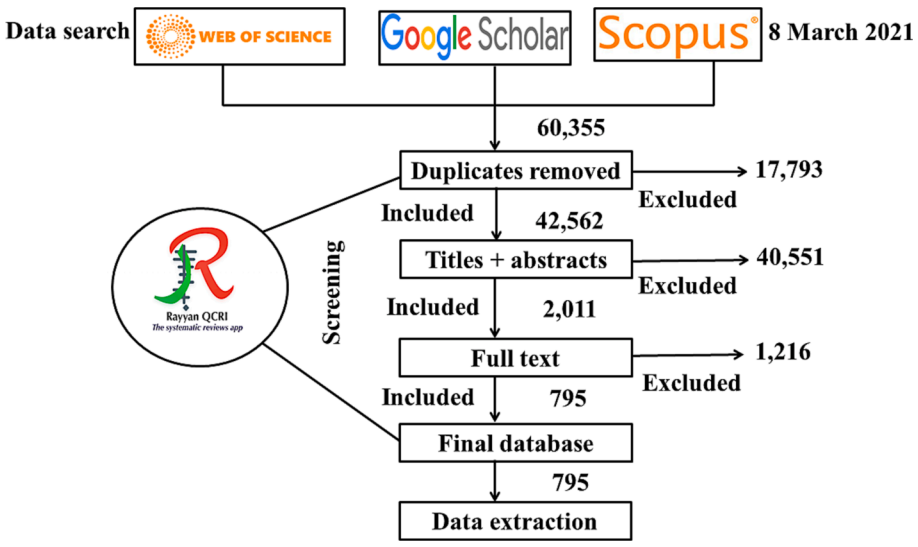


Fig. 1. PRISMA flow diagram for determining the state of urban ecology in Africa using the Rayyan Software.

published in English, (3) peer-reviewed as a first step towards quality control (Beninde et al., 2015; Raji & Downs, 2021), and (4) biodiversity conservation studies (including pet animals and introduced species).

2.2. Data extraction and categorization

We extracted the following data from each included paper: title, year of publication, journal, country of study and study sites. We then classified each paper based on type (field study, review or perspective) and scale, which included city (conducted in a single city), local (involving more than one city in a country), regional (involving more than one African country) and global (involving more than the African continent). Further, we followed the classification of Magle et al. (2012) to allocate each paper to one of the following scientific fields, including animal behavior, community ecology, conservation, human dimensions, human-wildlife conflict, landscape ecology, population ecology, wildlife disease and wildlife management. For taxonomic studies, we extracted information on the kingdoms and classes of focal species based on the classification of the Global Biodiversity Information Facility (GBIF) (GBIF, 2021; accessed May 2022).

With the exception of reviews and perspectives, we obtained the coordinates of all 1405 African study sites included in the selected papers by using Google Earth. This ensured conformity and completion given that the coordinates of some sites were either not originally provided in the papers or were presented in different formats. We then obtained information on all terrestrial ecoregions found in Africa from the World Wildlife Fund for Nature (WWF: Olson et al., 2001). Further data on the ecoregions, including size, conservation status and the biome they are located in, were also collected (Burgess et al., 2004). In addition, we obtained data on urbanization intensity and urban land cover (2015) across the continent, as well as the total population (2015) and total land area of each studied country from Africapolis (OECD/SWAC, 2020; accessed 9th June 2021). Urban land cover was used as a proxy for country urbanization intensity, while the total population was divided by the total land area to obtain the population density of each country. We then overlaid the study sites across ecoregions and urbanization intensity, as well as urbanization intensity across ecoregions, using QGIS (version 3.24 Tisler). Africapolis is the single most important and comprehensive geospatial database on cities and urbanization dynamics in Africa, which incorporates data on demography, satellite and aerial imagery and other cartographic sources (OECD/SWAC, 2020). To investigate urbanization prospect based on the urban land cover, data on the average annual rate of change of the percentage urban expansion by country (2015–2050) were integrated (United Nations, 2018). The Gross Domestic Product (GDP 2020; US$) of each studied country was also extracted from the National Accounts Section of the United Nations Statistics Division (accessed 6th May 2022).

2.3. Statistical analyses

All analyses were carried out using R Version 1.4.1717 (R Core Team, 2016). We performed descriptive statistics using the number of published urban ecological studies to determine temporal and spatial trends in urban ecological knowledge across years, countries, study scales, scientific fields, journals, and taxonomic kingdoms and classes.

We first used the number of published urban ecological studies (hereafter: research effort) per country as the response variable to test the effect of urbanization intensity, urbanization prospect, human population density and GDP using general linear models (LM). We used the "performance" package to check for multi-collinearity among the independent variables (Bernat-Ponce et al., 2021; Lüdeck et al., 2021) and tested the normality (Shapiro & Wilk, 1965) of the dependent variable ($p < 0.05$). The independent variables had low correlation (Variance Inflation Factor < 5) and, consequently, were all included in the models, but research effort was log-transformed to obtain reasonably normally distributed residuals from final models, and models that did not violate LM assumptions when examined visually as diagnostic plots (Crawley, 2013). Using the stepwise backward selection method (Crawley, 2013), variables with the highest p values were removed and the procedure repeated until the best model was selected as the one with the lowest Akaike Information Criterion value (Burnham & Anderson, 2002). Statistical significance was set at p value < 0.05. We also conducted a sensitivity analysis (Moher et al., 2009) due to the disproportionate weight of South African studies in our database, causing outliers. Of the overall 710 field studies that mentioned the 42 African countries represented here, 313 (44 %) were from South Africa. The second model therefore incorporated the same variables as the first but without South African papers.

Secondly, we tested for mismatches in the distribution of research effort across ecoregions. Note that this information could not be combined with the one collected at the country level and thus requires for an additional model to be tested. Given that research effort was not normally distributed ($p < 0.05$) even after log-transformation, we built a separate model using Poisson Logistic Regression to test if the size and conservation status of ecoregions (factor: Critical, Endangered, Vulnerable, Relatively Stable or Relatively Intact) influence research effort. We

then conducted a Tukey post-hoc test for a pairwise comparison across the different categories of conservation status using the package "emmeans" (Manley et al., 2015; Yvoz et al., 2020).

3. Results

Our search string detected a total of 60,355 papers out of which 17,793 duplicates were removed. The output of the remaining processes of Rayyan screening led to the retention of 795 papers considered in this review (Fig. 1). Out of them, 691 (87 %) were field studies, 90 (11 %) reviews and 14 (2 %) perspectives, all of which were published in 377 journals (Table S2). The first urban ecology studies focused on Africa date back from the 1970s (Okpala, 1978; Hugo, 1979), but the publication rate on the topic was slow (<10 papers/year) until 2006 when an exponential growth started, culminating in 126 papers published in 2020 (Fig. 2). From a geographical point of view, we found studies from 72 % of the countries that make up the African continent (42 out of 58 countries and autonomous territories; Fig. 3). However, a single country (South Africa) published 4 out of every 10 papers on the topic (N = 313), with the highly-urbanized and biodiversity-rich countries of tropical regions of the continent recording little (<40 papers; e.g., Democratic Republic of the Congo and Kenya) or even no urban studies (e.g., Angola and Liberia; Figs. 3 and 4) for the period of study (1920–2020). Furthermore, papers found in our literature search showed that most urban ecological research in Africa (89 %) was performed within countries, either focused on a single city (N = 434; 55 %) or conducted locally (N = 270; 34 %). We identified very few international research as only 4 % of the studies were carried out regionally (i.e., including more than one African country; N = 29) and only 8 % were coordinated at a global scale (i.e., including data from other continents too; N = 62).

The result of the LM analysis for all countries shows that research effort significantly increased with higher GDP, but not according to any other predictors (Table 2; Fig. 5). Contrary to our expectation, countries with higher human density and current or future urbanization prospects (up to 2050) have not been more studied (Table 1). In contrast, wealthier African countries have significantly investigated more on urban ecology (Table 1; Fig. 5). The same significant pattern was found for the sensitivity analysis (i.e., when South Africa was removed; Table S3).

Regarding ecoregions, we found information from 75 out of the 119 ecologically relevant regions in Africa (Fig. 6a-b; Table S4). This implies 37 % of ecoregions without a single urban ecology study. The research effort at this respect is not homogeneously distributed and varies considerably depending on the biome (Table 2). Furthermore, 22 out of the 44 African ecoregions without urban ecology studies are classified as threatened (Table S4) (Burgess et al., 2004). The Poisson Logistic Regression shows that research effort significantly increased in larger and more threatened ecoregions (Table 3). Urban areas in critical, endangered and vulnerable ecoregions have been more intensively studied (Fig. 7).

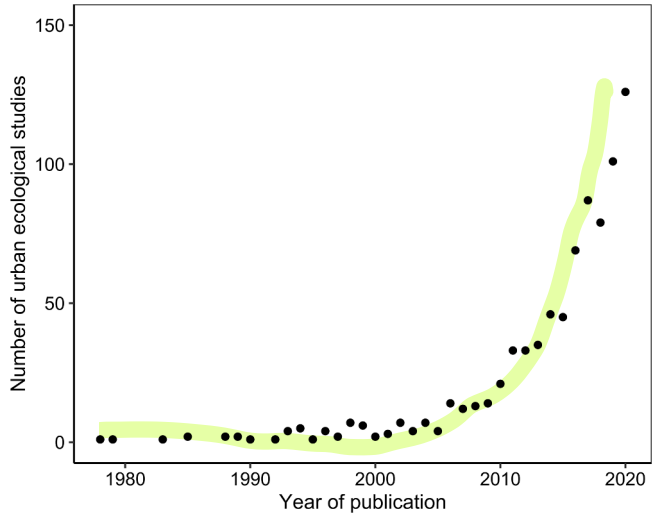


Fig. 2. Urban ecology research effort (number of urban ecological studies) across years.

Our review also showed important taxonomic biases in the study of urban ecology in Africa. We found information on studies focusing on seven kingdoms, with Animalia and Plantae being the most studied so far (Fig. 8). This result also highlights our limited understanding of other organisms, including Archaea, Bacteria, Chromista, Fungi and Protozoa, which when combined accounted only for 5 % of the studies. The number of studied classes was considerably higher in Animalia (27) than Plantae (9), with Aves (N = 138; 34 %) and Mammalia (N = 95; 23 %) accounting for the majority of studied animal groups (Fig. 9). Regarding plants, the most commonly studied classes were Magnoliopsida (N = 253; 66 %) and Liliopsida (N = 94; 24 %).

From a more conceptual point of view, we found variation in research effort among scientific fields (Fig. 10). The main focus of urban ecology in Africa seems to be applied studies given that conservation and human dimensions studies were the two most commonly investigated fields, with 41 % of all papers falling into these two categories. The scientific fields of wildlife management, wildlife disease and human-wildlife conflict were the least studied, accounting for merely 6 % of the total publications represented in this review. Our data showed that pattern approaches (e.g., Population, Community or Landscape Ecology) are more common than mechanistic studies (e.g., Animal Behavior) in Africa (Fig. 10). The first animal behaviour studies were published in the early 1990s, investigating insects (Paillette et al., 1993) and birds (Van Zyl, 1994). But the focus on this discipline has considerably increased since 2015, with 64 % of all Africa urban ecology studies on animal behavior published after this year (Table S2). Despite this increasing interest, there is still an important taxonomic bias, and only 44 % of the 27 animal classes were represented in animal behaviour studies, including Mammalia (38), Aves (47), Reptilia (7), Amphibia (6), Insecta (5), Gastropoda (2), Actinopterygii (2), Arachnida (1), Clitellata (1), Entognatha (1), Malacostraca (1) and Sarcopterygii (1).

4. Discussion

4.1. Spatio-temporal patterns in knowledge

Our literature search shows almost 800 urban ecology papers for the entire African continent. According to a recent review investigating the top 20 countries publishing on urban ecology (Shackleton et al., 2021), this number is lower than the number of publications from medium-sized European countries, such as Germany (2,479) or Spain (1,864), and much lower than the research effort identified for the United States (12,728), China (6,655) or Australia (2,900). This suggests that urban ecology research in Africa is still considerably low compared to other regions of the World (e.g., Europe, North America, Asia or Australia), matching previous findings that already indicated the African continent was the least studied regarding urban ecology (e.g., Magle et al., 2012 stated that Africa accounted for 2.8 % of published papers on urban wildlife ecology in 2010). It is interesting to note that despite the exponential growth in research effort during the last 15 years, mimicking the global trend on the topic (Lin & Grimm, 2015), Africa has not increased its relative contribution to the field like other regions (e.g., Asia) that were also underrepresented a decade ago (Magle et al., 2012; Wu et al., 2014; Shackleton et al., 2021). The overall number of urban ecology papers in Africa does not seem to be associated with a delayed start in the discipline. Our review shows that African urban ecology started at the end of 1970s around the same time that this discipline started in other regions of the World (McDonnel, 2011; Wu, Xiang, & Zhao, 2014). We cannot be completely sure that there have not been

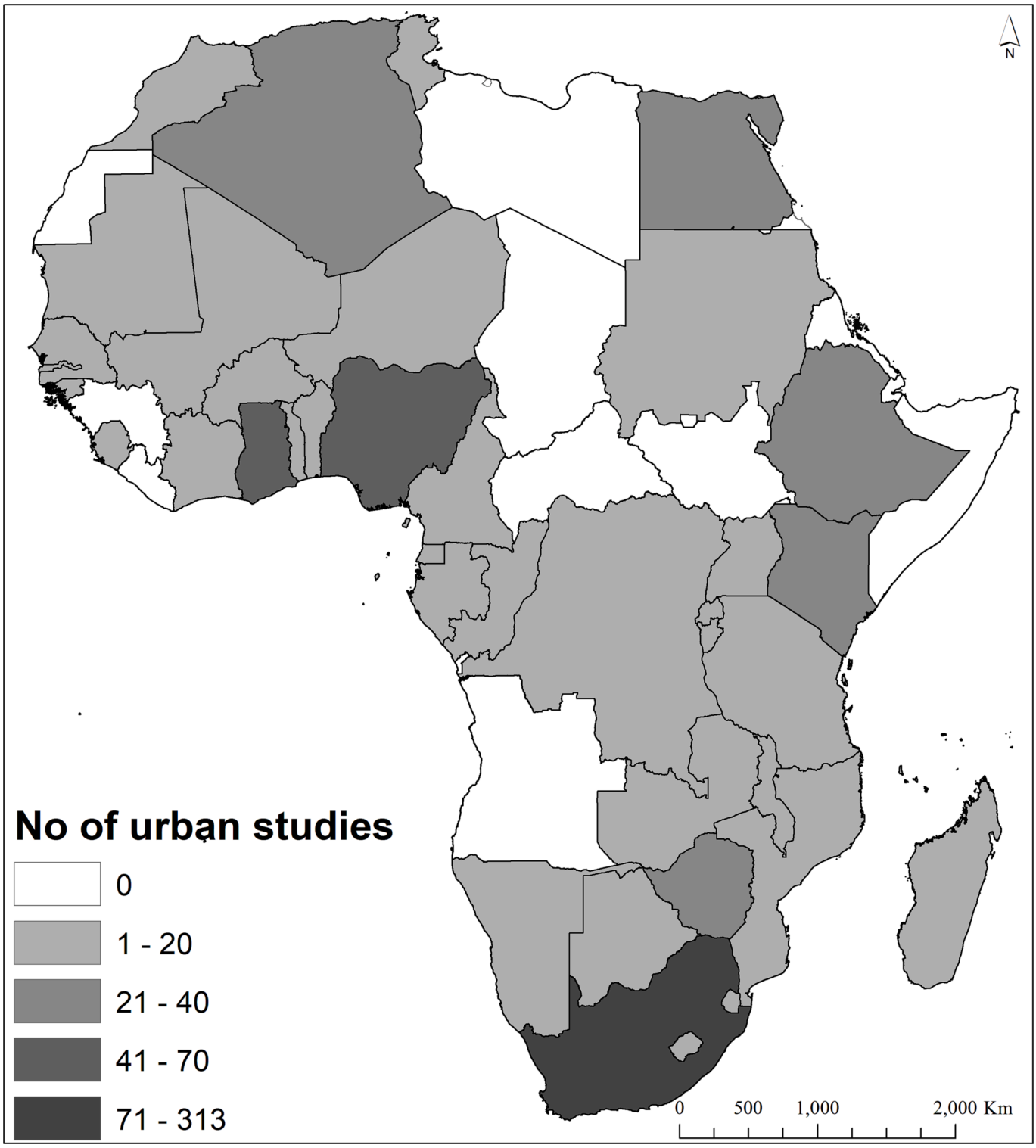


Fig. 3. The distribution of urban ecological studies across African countries.

earlier publications in non-English languages, but probably the first African paper explicitly mentioning the concept of urban ecology corresponded to Okpala's study (1978). This pioneering investigation focused on socio-economic aspects from Lagos (Nigeria), already highlighting the potential conflict of trying to apply European or American urban ecology theory to the African case, an argument that is still valid within the Global North and Global South framework (Shackleton et al., 2021). The current underrepresentation of African urban ecology is particularly worrying as most African urban settings are considered as clear representatives of the Global South urban settings, integrating particular biophysical and socio-economic contexts (Shackleton et al., 2021). Thus, the lack of knowledge at this respect impedes us to complement our understanding of urban ecology, which is based on the more traditional Global North perspective.

There could be other different reasons explaining the low number of publications from Africa. The lack of local capacity/experts in the field is

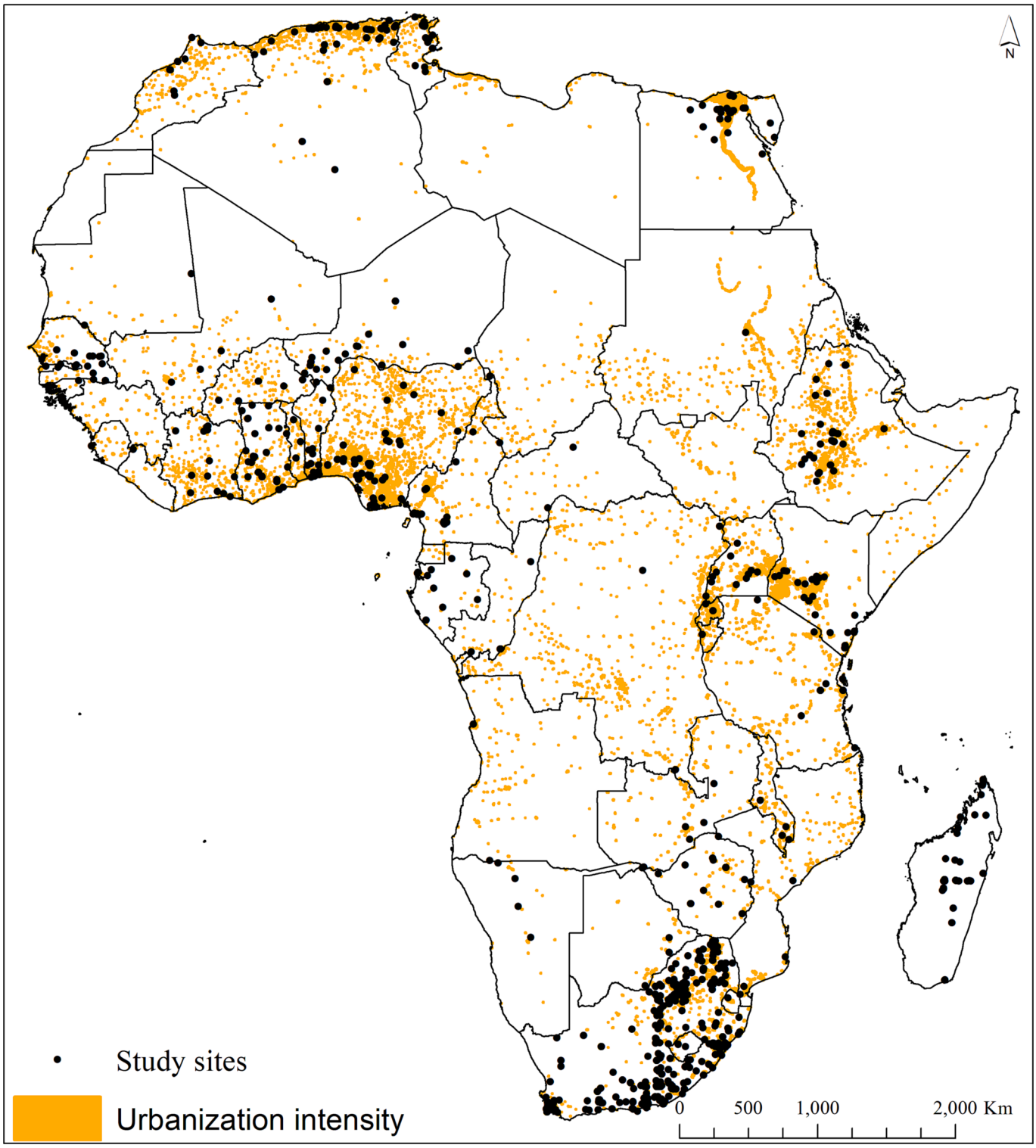


Fig. 4. The distribution of urban ecological study sites superimposed on urbanization intensity.

one of them. This factor has been previously highlighted as a key difference between the Global North and Global South urban settings that could influence the lower level of urban ecology research effort in the latter (Shackleton et al., 2021). According to the UNESCO's database for the period 2015–2020 (UNESCO, 2020; accessed 30 Oct 2022), the number of researchers per million of inhabitants in Northern (732.4) and, particularly, Sub-Saharan Africa (97.4), is considerably lower than in other regions of the planet, such as North America (4,544.8), Europe (3,010.4) or Oceania (3,510.5). This low ratio of skilled people has been demonstrated to influence research effort in Africa regarding other fields such as ornithology (Cresswell, 2018). Therefore, we encourage funding bodies to finance the education of local urban ecologists and researchers to overcome this potential restriction. Another potential reason explaining the low research effort is partially linked to the previous one: the lack of investment in Research and Development (R&D) in Africa compared to other continents. Despite the African Union aims at

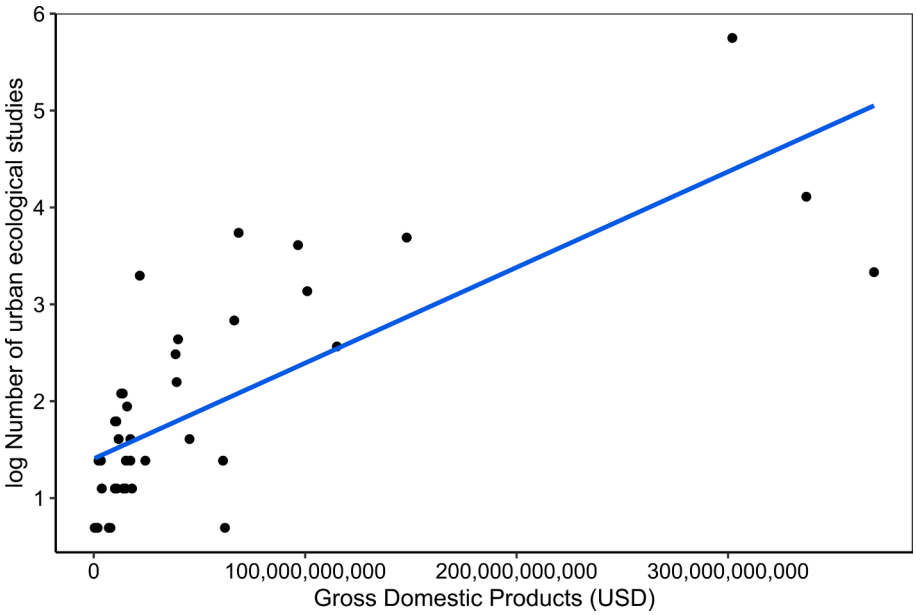


Fig. 5. Relationship between urban ecology research effort (number of urban ecological studies) across all countries and Gross Domestic Products (USD). Note that the y-axis is on a logarithmic scale and that there are several overlapping points.

Table 1
Results of a GLM exploring the predictors of the number of urban ecological studies published across all countries. The number of urban studies +1 was log-transformed to achieve a normal distribution of residuals. The last model (F_{40} = 51.9, P < 0.001; AIC = 100.57) incorporated only the significant variable and had an adjusted R^2 = 0.55.

	Estimate	SE	*t-value*	*p-value*
Intercept	1.41E + 00	1.38E-01	10.22	<0.001
Gross Domestic Product	9.88E-12	1.37E-12	7.203	<0.001
Rejected variables				
Urbanization intensity	1.07E-01	1.20E-01	0.892	0.378
Human population density	−9.24E-04	9.68E-04	−0.955	0.346
Urbanization prospect	5.19E-02	3.69E-02	1.4	0.167

Table 2
Urban ecology research effort (i.e., studied ecoregion/total ecoregion %) across African biomes and ecoregions.

Biome	Total ecoregion	Studied ecoregion	Research effort (%)
Temperate Coniferous Forests	1	1	100
Mangroves	5	4	80
Tropical and Subtropical Moist Broadleaf Forests	30	23	77
Mediterranean Forests, Woodlands, and Scrub	7	5	71
Tropical and Subtropical Grasslands, Savannas, Shrublands, and Woodlands	24	16	67
Montane Grasslands and Shrublands	16	10	63
Flooded Grasslands and Savannas	10	6	60
Deserts and Xeric Shrublands	23	9	39
Tropical and Subtropical Dry Broadleaf Forests	3	1	33

reaching to the 1 % of GDP invested in R&D (United Nations. Economic Commission for Africa 2018), current data indicate that it is 0.64 % and 0.34 % for northern and sub-Saharan Africa, respectively. This is quite far from the values of North American, European or Eastern Asian countries that reached a mean of 2.6 % in 2020. Matching the target proposed by the African Union will certainly help to increase the focus on multiple topics, including urban ecology. However, there are ways to improve knowledge on urban ecology in Africa even without the need of large economic investments. For example, the use of available databases, such as the various atlas projects, which have been successfully implemented in the continent (Botts et al., 2011; Lee & Nel, 2020). Other repositories, such as the Global Biodiversity Information Facility, laboratories, herbaria and museums in and outside of Africa are also useful tools to advance our understanding of the ecology of African urban areas and biodiversity as some recent studies have already shown (e.g., Cohen et al., 2021; Fishpool & Collar, 2018). This approach could also be implemented in collaboration with inhabitants of African urban areas through citizen science projects (e.g., iNaturalist or the Southern African Bird Atlas Project) that can serve to improve information on certain urban questions (e.g., animal distribution) as well as promote the connection between citizens and nature (Reynolds et al., 2021). Engaging citizens could also be instrumental to help increase the urban governance in the Global South, including Africa (Shackleton et al., 2021), and ultimately promote additional support for urban ecology studies in this continent.

Our review also shows that research effort is not homogeneously distributed within the African continent. From a political point of view, there is an important variation among African countries in their urban ecology research effort. One single country (South Africa) stands out as it is responsible for almost 40 % of published papers on the topic. This is so despite only representing 4 % of African territory and 1.02 % of all urban areas in the region (OECD/SWAC, 2020). This high rate of urban ecology publications matches previous information indicating that South Africa is very active in the field at the global level (Shackleton et al., 2021). This does not seem to depend on its number of researchers per million of inhabitants (411.6) or its R&D investment (0.62 % of GDP), which is lower than the mean for Northern Africa (UNESCO, 2020), an area that not even combining all its countries reaches half the number of papers published in South Africa. This country started publishing urban ecology papers at the earliest stages in Africa (Hugo, 1979), so it is possible that this long-term publication period is behind its uniqueness. Another possibility could be that several South African cities (e.g., Cape Town and Durban) are located in biodiversity hotspots of global importance (Cilliers & Siebert, 2012). Alternatively, given that

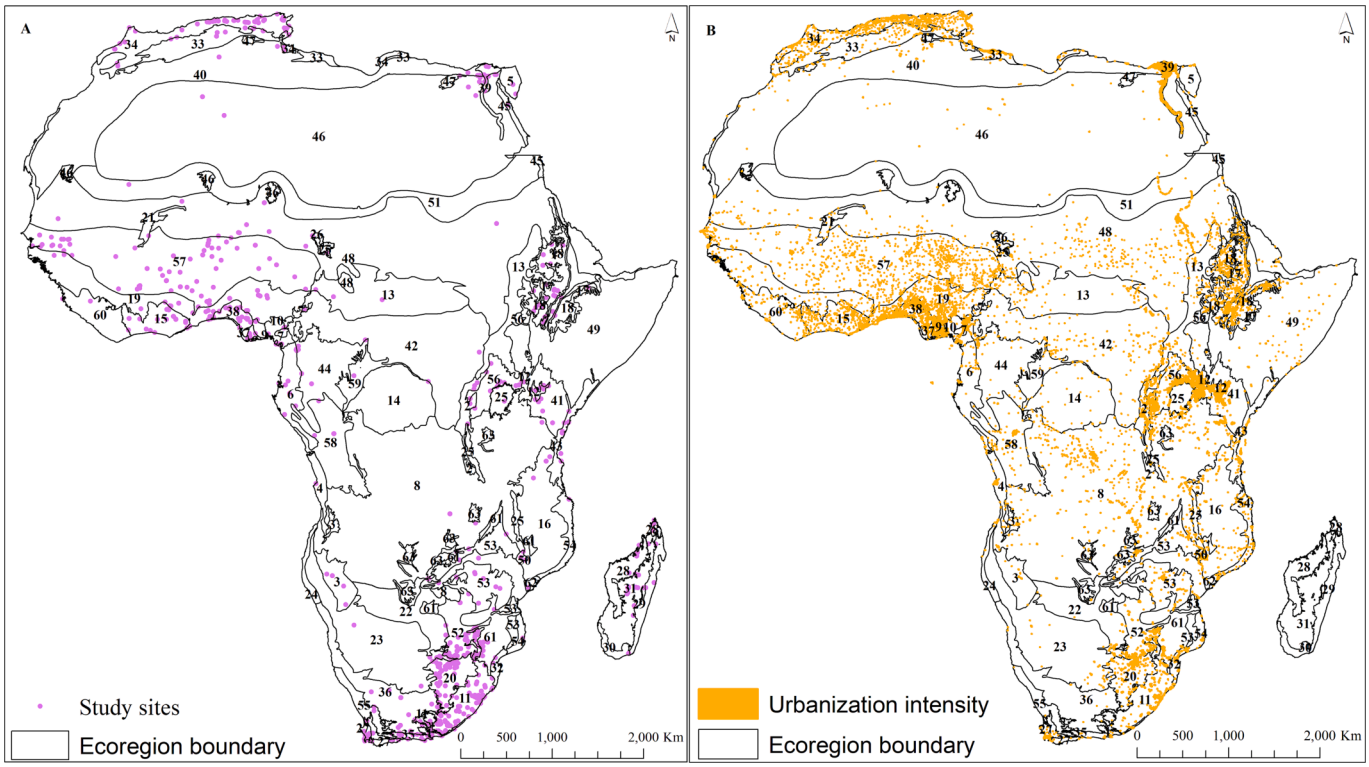


Fig. 6. Map of the African terrestrial ecoregions showing the distribution of urban ecological study sites (a) and urbanization intensity (b). The maps were simplified to facilitate interpretation. Thus, we retain outlines of relatively large ecoregions >10,000 km^2 and those including study sites. However, the names of all ecoregions, their corresponding numbers in the map and additional details (e.g., size) are included in Table S4.

Table 3
Results of a Poisson Logistic Regression exploring the relationship between the number of published urban studies and the conservation status and size of ecoregions. Conservation status is a factor with 5 levels (Critical, Endangered, Relatively Intact, Relatively Stable, Vulnerable) and size is a continuous variable. Critical has been set as the intercept in the model.

	Estimate	SE	z-value	*p-value*
Intercept	2.99E + 00	4.50E-02	66.467	**<0.001**
Endangered	2.44E-01	6.46E-02	3.782	**<0.001**
Relatively Intact	−2.33E + 00	2.13E-01	−10.971	**<0.001**
Relatively Stable	−1.13E + 00	9.02E-02	−12.524	**<0.001**
Vulnerable	−2.62E-02	1.10E-01	−0.239	0.811
Size	5.45E-07	4.69E-08	11.609	**<0.001**

Global North urban principles do not always apply to Global South urban areas (Okpala, 1978; Shackleton et al., 2021), there could be a special interest by funders and/or researchers from this country to acquire first-hand knowledge of direct application to South-African urban settings. For instance, some universities from this country (e.g., Witwatersrand) have strategically focused on global change research, including urban ecology (Scholes et al., 2013) or have developed specific institutes for the study of 'urbanism from an African perspective' (e.g., The African Centre for Cities, from the University of Cape Town; <https://www.africancentreforcities.net/about/acc-at-uct/>). Independently of the reasons for this important outlier, urban ecology research effort varies considerably within African countries. We identified that 28 % of these countries did not publish a single urban ecology study and thus, they completely depend on urban knowledge obtained elsewhere that sometimes might not be really useful for their local situations.

Our analyses show that the number of publications per country on the topic is not associated with current or future urbanization. This result contradicts our initial prediction; however, it could be well understood from a Global South perspective. African countries, like other countries from this group, have several particularities compared to those from the Global North (Shackleton et al., 2021). One of them is the extremely high urbanization rate. Africa is the continent of the World with the most intense urbanization (Cohen, 2006; Seto et al., 2012), with many African countries experiencing urbanization rates above 4 % (e.g., Mali, Nigeria, Angola or Mozambique), an order of magnitude higher than those from other regions of the planet (World Bank, 2021). This factor leads to unplanned urbanization (Zhang, 2016) and compromises sustainable urban development in the continent by impeding the implementation of ecologically-sound practices (Cohen, 2006) and hence potentially explaining the mismatch between urbanization and urban ecology research effort.

Furthermore, we found that the human population density of a country was not significantly associated with the number of publications on urban ecology either. The reasons for this lack of association could be the same as explained before for the current and future urbanization prospects as these are positively correlated with human population density (e.g., Gao & O'Neill, 2021; Qizhi et al., 2016). However, this predictor could also be associated with other potential factors that might prevent investing resources and effort in investigating about urban ecology. For example, there is an increase in people living in extreme poverty in Africa, with more than half of the urban population living in slums and informal settlements (World Cities Report, 2016). Highly populated areas also require a higher infrastructure investment, which is particularly needed in Africa (Zhang, 2016). Thus, socio-economic priorities combined with an insufficient capacity of urban governance (Zhang, 2016; Shackleton et al., 2021) could prevent finding the initially expected effect of human population density. Considering all these results and factors, particularly the uncoupled distribution between urban ecology knowledge and future urban prospects, we would recommend local authorities, funding bodies and researchers to make an effort in the study of the areas that soon will be transformed into urban landscapes.

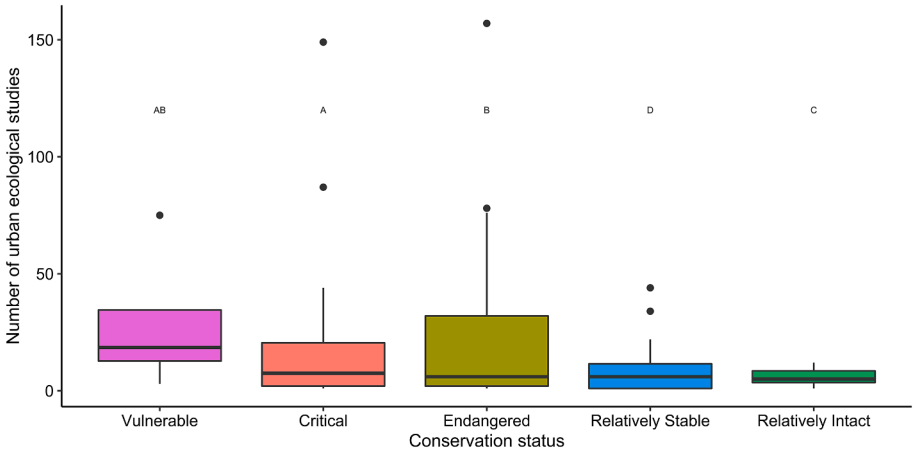


Fig. 7. Urban ecology research effort (number of urban ecology studies) across the conservation categories of ecoregions. Box-plots show median, quartiles, 5- and 95- percentiles and extreme values. Different letters indicate significant differences ($P < 0.01$) between conservation status according to Tukey post-hoc tests using the package "emmeans" (Manley et al., 2015; Yvoz et al., 2020).

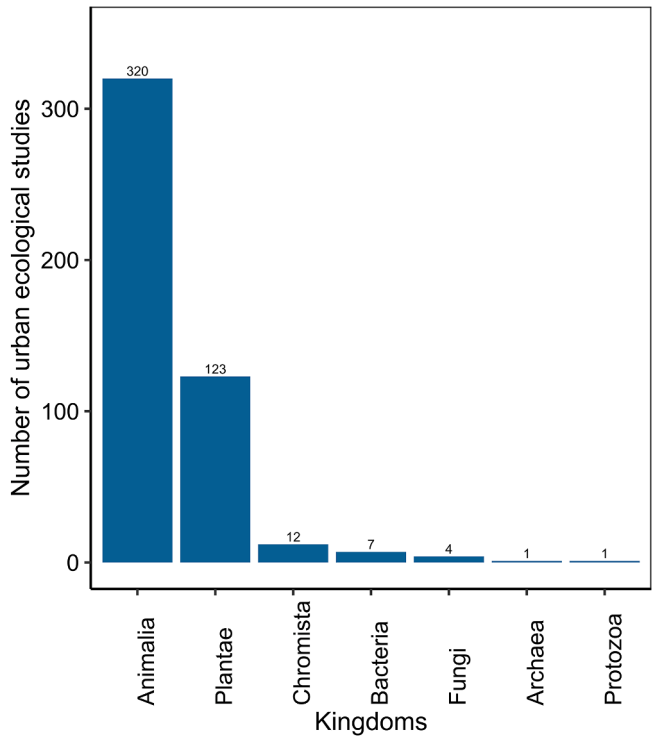


Fig. 8. Urban ecology research effort (number of urban ecological studies) across taxonomic kingdoms.

This is particularly important in the tropical African belt given that it will concentrate the greatest urban expansion in the future (Seto et al., 2012), but also holds the largest biodiversity of the continent (Cazzolla Gatti et al., 2015).

Interestingly, our results indicate that the number of published urban ecological studies depended on economic factors (i.e., GDP). This association has been found in other cross-sectional (e.g., (Doi & Takahara, 2016; Fisher et al., 2011) and longitudinal studies (Vinkler, 2008). This economic indicator is in addition significantly associated with a higher rate of influential publications within their subject area (Bornmann et al., 2014). However, other investigations showed that R&D investment rather than per capita GDP is positively associated with research productivity in different continents (Meo et al., 2013, 2014). It is possible that GDP is a better predictor of R&D in Africa than in other regions, thus potentially explaining the obtained finding. This influence of economic factors on urban ecology research effort is crucial given the link between cities and economic wealth (Zhang, 2016), which could lead us to think that as urbanization progresses in Africa, the better their economies will be and consequently more research on urban ecology could be made. This scenario seems unlikely as this association between economic and urban growth is decoupled in the African continent (Cohen, 2004), which does not warranty this increasing research effort in the future. Other factors not considered in our analyses could also explain the country-wide variation in urban ecology research. For example, political instability could play an important role for the lack of studies on the topic in certain countries such as Western Sahara, South Sudan or Libya.

The fact that the majority of published studies were conducted locally within a single city or country (e.g., Koricho et al., 2020; Lindley et al., 2018; Muleya & Campbell, 2020) suggests the need for investigation of local/national cases for the application of specific solutions. However, it also highlights the lack of transnational collaboration among African countries. This low level of international research both within Africa and with countries from other continents is particularly important considering that: (1) it impedes the generalization of findings at the continental and global scale, and (2) reduces the number of substantive contributions to scientific progress (Bornmann et al., 2014). Therefore, we recommend funders and researchers alike to strengthen or promote the creation of new international networks or institutes on African urban ecology as well as encourage urban ecologists of the continent to participate in other global actions, networks (e.g., the Urban Biodiversity Research Coordination Network) or societies (e.g., Society for Urban Ecology) that are already running.

The geographic variation in research effort could also be linked to conservation aspects. Conservation research in Africa is particularly relevant and prolific in the global context (Doi & Takahara, 2016). There are still some controversies on whether conservation status is significantly and positively associated with research effort at the species level (e.g., Brooke et al., 2014; Ducatez & Lefebvre, 2014; Ibáñez-Álamo et al., 2017), but countries with a higher level of environmental protection activity investigate more in ecology (Doi & Takahara, 2016). Our results match this finding given that urban ecology research effort is significantly associated with the conservation status of African ecoregions. The ecologically relevant regions belonging to the most threatened categories (Critical, Endangered and Vulnerable) showed the highest number of publications on the topic. This is logical considering the previously described restricted R&D investment in Africa that would divert the current available resources towards areas of conservation concern. Despite this, we found that about half (50 %) of African

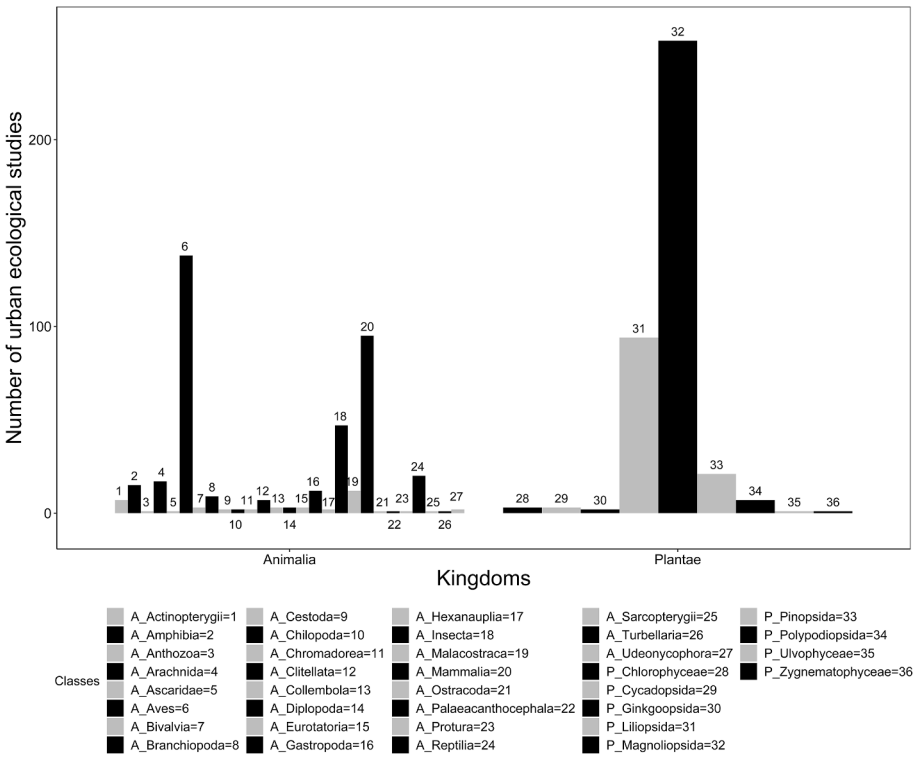


Fig. 9. Urban ecology research effort (number of urban ecological studies) per class of the two most studied kingdoms (Animals and Plants). Each number on/below the bars corresponds with the number and class in the legend.

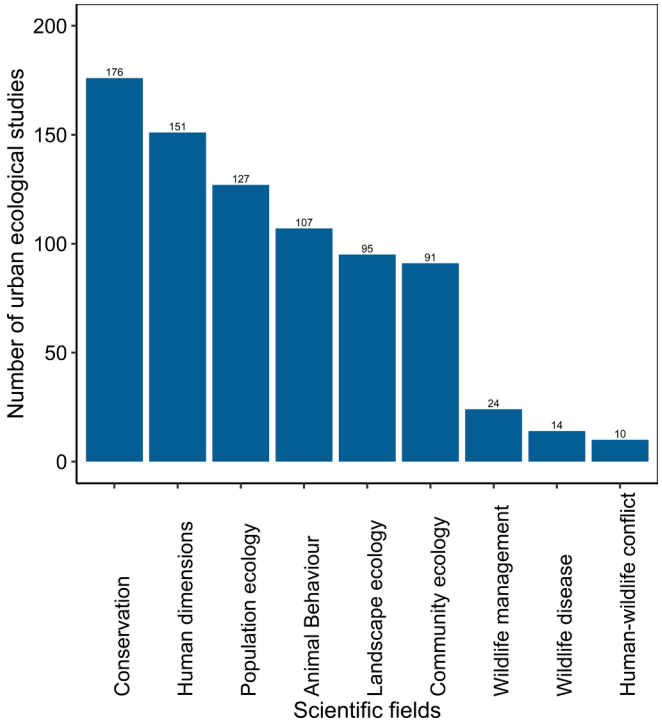


Fig. 10. Urban ecology research effort (number of urban ecological studies) across scientific fields.

ecoregions without a single published study on the topic are classified as threatened, and urbanization is considered a leading threat in the area (Burgess et al., 2004), suggesting the need for additional studies to determine the ecological effects of urbanization and propose suitable conservation actions. On the other side, the significant effect of ecoregion size fitted our initial expectations as larger ecoregions would support higher biodiversity levels (Rantalainen et al., 2005) and consequently a higher likelihood of being investigated. As larger and more threatened ecoregions were significantly more studied in the continent, there is a need to expend greater research effort on smaller and relatively stable ecoregions (e.g., East African Montane Moorlands and Lake Chad Flooded Savanna), which are more likely to suffer unnoticed fragmentation from urbanization and other anthropogenic land-use changes as also indicated by previous studies (e.g., Beyer, Venter, Grantham, & Watson, 2020; Burgess, Hales, Ricketts, & Dinerstein, 2006; McDonald et al., 2008). Particularly surprising is the lack of studies from the majority (77 %) of ecoregions from the Tropical and Subtropical Dry Broadleaf Forests biome. These ecoregions mainly correspond with large areas of Madagascar, a megadiverse country (<https://www.biodiversitya-z.org/content/megadiverse-countries> accessed 30 October 2022) with the lowest percentage of urban land-cover in the whole continent (0.04 %; OECD/SWAC, 2020). In contrast, other forested biomes are quite well represented, which makes sense considering that forests, especially those from Western Africa, support higher biodiversity and endangered species, thus promoting a more intense ecological research effort (Doi & Takahara, 2016).

4.2. *Gaps in knowledge according to taxonomy and scientific fields*

Our review also offers interesting information on the current methodological and conceptual orientation of urban ecological research in Africa. From a methodological point of view, we found an important taxonomic bias in the study of urban ecology in Africa similar to those previously reported (e.g., Callaghan et al., 2020; Shwartz et al., 2014). This taxonomic bias has a strong effect in our urban ecology knowledge given that the impact of urbanization varies considerably depending on the type of organisms considered (McKinney 2008; Paul & Meyer, 2001). Our literature search offered studies focused on organisms belonging to seven kingdoms, although the majority of urban ecology research used either animals or plants as model systems. This result highlights our

limited understanding of other organisms in the African urban context, including Archaea, Bacteria, Chromista, Fungi and Protozoa, which should be prioritized for future studies. This is justified by current literature highlighting their relevance in natural environments (e.g., Epp Schmidt et al., 2019; Kartzinel et al., 2019; Thompson et al., 2017). The uneven distribution of urban ecology research effort went down to lower taxonomic levels (e.g., classes). Among animals, birds and mammals were the two most studied groups. The publication bias towards these two classes in urban ecology is not restricted to Africa alone (Donaldson et al., 2017; Shwartz et al., 2014), and has also been identified in other study fields such as conservation biology (Lawler et al., 2006) and invasion ecology (Pyšek et al., 2008). Several reasons have been proposed to explain this bias for birds and mammals, such as body size (Brodie 2009) or conservation status of focal species (Donaldson et al., 2017). Regarding plants, flowering plants (Magnoliopsida and Liliopsida) dominate urban ecology research effort in Africa, replicating the patterns found by other research effort studies on plants (Richardson & Rejmanek, 2011; Stranga & Katsanevakis, 2021). In contrast with plants, with the richly diverse Magnoliopsida (Tracheophyta) relatively well studied (Cilliers & Bredenkampl, 1999; Moussa et al., 2020; van der Walt et al., 2015), the most diverse animal group of Arthropoda is clearly underrepresented in urban ecology and calling for additional scientific attention (reviewed here; McIntyre, 2000).

Urban ecology research effort in Africa also varied in terms of scientific disciplines. Conservation was the most studied scientific field. This result is in agreement with previous findings already highlighting the relevance of Africa in the study of environmental conservation and ecology (Doi & Takahara, 2016), and matches also with our initial result that indicates preference for ecoregions of conservation concern. Interestingly, a handful of such African conservation studies diagnosed different socio-environmental issues in urban areas and developed useful frameworks or plans for promoting nature conservation and sustainable urban development in the continent (e.g., Boon et al., 2016; Cilliers et al., 2004; Goosen & Cilliers, 2020; Rebelo et al., 2011). While these findings imply the availability of data that could be useful for promoting conservation actions, they are mostly restricted to South Africa. For an effective implementation of conservation actions, more studies are needed from unrepresented areas as they may help to discover local issues such as environmental injustice (Ernstson, 2013). The human dimension field is well-represented within African urban ecological research, which points to the relevance of multifaceted approaches in Africa, particularly regarding ecosystem services that complements conservation or ecological studies (e.g., population ecology or animal behavior). For instance, the majority of human dimension studies in our review indicate that people in African urban areas appreciate the socio-ecological services (Dipeolu et al., 2020; Rogerson & Rogerson, 2020) and economic benefits provided by urban biodiversity (Babalola et al., 2013; King & Shackleton, 2020). In a study by Popoola and Ajewole (2002), most Nigerian respondents were even willing to support the conservation of urban nature through personal funds. The conservation of urban biodiversity is tightly linked to public support (Miller & Hobbs, 2002), and thus, human dimension studies could be useful educational tools to reconcile urban development and nature preservation in the continent (McDuff, 2000). In addition, unlike in other regions where the important roles of urban biodiversity in enhancing ecosystem services and human well-being have been well-documented (Brown & Grant, 2005; Dallimer et al., 2012; O'Sullivan et al., 2017), this interplay is much more complex in the African case (Wangai et al., 2016) usually not considering the ecosystem disservices that could be of critical importance in areas of the Global South (Davoren & Shackleton, 2021). In general, ecosystem services in Africa have been poorly studied (du Toit et al., 2018), although there is a clear effort in recent years to overcome this important gap (e.g., Dobbs et al., 2021; Escobedo, 2021; Shackleton et al., 2021; Wangai et al., 2016), including the evaluation of how different frameworks are applied to African urban settings (Lindley et al., 2018).

We identified that many urban ecology papers focused on Africa used pattern approaches either at the species or community level. Several reviews on urban ecology or specific aspects of urban ecology (e.g., urban ornithology) have also found similar results at the global level (Magle et al., 2012; Marzluff, 2016; Wu et al., 2014). As we have stated before, Africa is understudied in urban ecology, and we lack many basic information on even the presence/absence of certain organisms in cities of this continent. Some of the studies in these categories describe new species (e.g., Malonza et al., 2016; Smales et al., 2017), provide information on potentially problematic organisms (e.g., invasive species; Bigirimana et al., 2011; Hima et al., 2019) or provide much needed information on the distribution of organisms in African urban settings (e.g., Moussa et al., 2020; Muchayi et al., 2017). But some of these articles also used applied approaches by integrating human-nature interaction aspects. For example, Chamberlain et al. (2019) found evidence supporting the luxury effect in South Africa. This effect states that there is a positive correlation between wealth and biodiversity, and thus relates to environmental injustice issues (Reynolds et al., 2021). These pattern-approach studies that also consider applied aspects and the particularities of Global South urban areas are excellent examples on how we can advance in our understanding of African urban ecology. Some researchers have highlighted the lack of urban ecology mechanistic studies in countries of the Global South compared to those from the Global North (Marzluff, 2016). Mechanistic studies would, for example, include animal behavior papers that could explain the observed patterns (e.g., feeding behavior explaining the presence of certain animals in cities). Africa has produced quite a lot of animal behavior studies centered in urban areas but most of them were observational (e.g., McPherson et al., 2016; Widdows & Downs, 2016), with only a handful of experimental manipulations (Cronk & Pillay, 2018; Patterson et al., 2016) that are much more powerful to identify cause-effect associations. Future studies should try to put more emphasis on experimental manipulations to fill in this important gap in our urban ecology knowledge.

Landscape ecology is still not as well studied as in other regions regarding urban areas (Magle et al., 2012; Wu et al., 2014), but it offers unique opportunities for the development of this field in Africa. On the one hand, landscape ecology studies in our database extensively utilized the Geographic Information System (GIS) for estimating land cover and habitat heterogeneity (e.g., Benza et al., 2016; Kowe et al., 2020). The use of GIS techniques could enhance better coverage of study sites (e.g., conflicting/dangerous/remote areas), helping to complete the missing geographic areas in urban ecology research detected in our review. These techniques require highly qualified personnel but provide useful information at minimal time and cost (Langat et al., 2019), thus, offering a good opportunity for capacity building in the continent while considering the economic restrictions in R&D of the region (see above). On the other hand, landscape ecology is an integrative discipline merging geospatial patterns, ecological and socio-economic processes and ecosystem services/disservices, thus favoring the interdisciplinary collaborations between sociologists, ecologists and geographers among others (Wu et al., 2014), thereby facilitating the establishment of much needed interdisciplinary collaborations in African urban ecology. For all these reasons, we expect that the field of urban landscape ecology will continue to increase as it has happened at the global scale (Magle et al., 2012).

5. Conclusions

This review shows that research effort on urban ecology is still low in Africa, with the exception of South Africa, particularly in the highly urbanized and biodiversity-rich areas of the continent. This continent is an important representative of the Global South, and thus the lack of information on the topic is an important impediment to try to overcome the traditional Global North perspective on urban ecology (Shackleton et al., 2021). In addition, the information presented here could be crucial to achieve the 11th Sustainable Development Goal in the rapidly

urbanizing African continent (Cobbinah et al., 2015). Urban areas, if well-planned, can still provide substantial benefits for biodiversity, act as hotspots and habitat corridors for some threatened species (Ives et al., 2016; Kumdet et al., 2021) and serve important socio-ecological (Dipeolu et al., 2020; Rogerson & Rogerson, 2020) and economic benefits (Babalola et al., 2013; King & Shackleton, 2020). To our knowledge, this is the first general literature review of urban ecological studies for the entire African continent that follows rigorous, verifiable and repeatable methodological approaches recommended in recent times (Ibáñez-Álamo et al., 2017; Magle et al., 2012; Moher et al., 2009; Sánchez-Tójar et al., 2020). Previous methodologically-similar reviews of African urban ecology, though interesting and useful, either focused mainly on socio-ecological systems (e.g., Cilliers, 2019; Lindley et al., 2018) or specific aspects of African urban biodiversity (e.g., Güneralp et al., 2018; Roets et al., 2019; Trimble & van Aarde, 2014). The low research effort in African urban ecology seems to point to socio-economic factors such as the low level of skilled people and reduced investment in R&D typical from this continent (e.g., Cresswell, 2018). We believe that this situation could be partially reverted if African countries follow the African Union recommendation of investing 1 % of their GDP in R&D, although other socio-economic needs (e.g., infrastructure, security, health issues) could make this change very difficult (Zhang, 2016).

Economic factors (GDP) rather than other urban indicators (e.g., urbanization intensity, human population density) are also crucial to explain urban ecology research effort within the continent. South Africa congregates many of the papers on the topic, while there are 16 African countries without urban ecology studies, providing clear targets for future investigations. The South African case could be useful to identify specific aspects that could be reproduced in other neighboring countries to try to boost urban ecology research. Thus, studies comparing different urban ecology aspects between South Africa and other African countries would be particularly interesting at this respect. In addition, it is especially worrisome the uncoupled nature between future urbanization prospects and urban ecology knowledge as local authorities will not count with valuable information to take scientifically-based actions. This lack of information has already been suggested as an important impediment to achieve sustainable urban development in Africa (Cobbinah et al., 2015; Patel et al., 2017).

In addition, greater research effort is expended on larger and threatened ecoregions. Threatened sites and species are usually prioritized for conservation actions (Brooks et al., 2006), and could influence research effort (e.g., de Lima et al., 2011). However, relatively stable ecoregions could suffer unnoticed effects of urbanization, which could be detrimental to certain biodiversity that may suffer regional extinction before being identified. This pattern has been previously reported in Africa (Ahrends et al., 2011), and could even be more severe in the future given the mismatches in the allocation of research effort across regions. This research bias towards threatened areas is partially linked to the fact that conservation studies dominate the urban ecology literature produced in the African continent. Our literature search also indicated that African urban ecology research is multidimensional with an important contribution to human dimension studies including those on ecosystem services and disservices. These studies have increased in recent years providing much needed information for the urban settings of this continent and ultimately helping to improve our understanding of the complex urban environment in which many different components interact (e.g., sociological, ecological, economical…).

6. Recommendations and future prospects

We argue that for African urban ecology to provide more useful information for decision-making and promote sustainable development, future research should try to overcome the detected geographic, taxonomic and ecological biases. To help in this endeavor, we provide a list of the articles reviewed here as well as the journals of publication, where key stakeholders or researchers could obtain relevant data on the topic (Table S2).

Based on our review, we propose the following recommendations to promote urban ecology research in this continent: (1) strengthening collaboration and networking among researchers across regions and countries, as previously suggested in a more general context (McPhearson et al., 2016). This will allow for larger scale studies that will provide an additional and complementary perspective to city/local studies that tackle more specific problems. (2) Helping the education of local experts on urban ecological studies can be also instrumental to overcome some of the previously described publication biases on the topic (Shackleton et al., 2021). (3) Engaging with the citizenship through citizen science projects. This will allow the acquisition of additional scientific information at the same time as it promotes a better urban governance through participation of urban inhabitants. (4) Use of low-cost techniques like GIS or available databases (e.g., museums) to maximize the scientific outcome considering the economic restrictions of the region. We hope that this review will help to re-orientate our research effort on the topic and fill in some important knowledge gaps highlighted here to grant a balanced strategy between urban development and nature conservation in this unique continent.

CRediT authorship contribution statement

Adewale G. Awoyemi: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Resources, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Juan Diego Ibáñez-Álamo:** Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

The A.G. Leventis and APLORI Foundations funded this study as part of the PhD project awarded to A.G. Awoyemi. JDI-A has been funded by the Spanish Ministry of Science and Innovation (PID2019-107423GA-I00/SRA (State Research Agency/10.13039/501100011033), while the open access charge for this paper was funded by the University of Granada/CBUA. We are also grateful to D. Alcaraz, S. Karau and T. Alabi, who kindly helped with the use of QGIS, as well as O. Olubodun, who provided the PDFs of some studies from various African archives. Two anonymous reviewers and the editor provided useful comments that improved an earlier version of this manuscript. The donor played no part in the conduct of this study.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.landurbplan.2023.104707.

References

Abreha, S. K. (2019). Model-based cost-effectiveness analysis of external beam radiation therapy for the treatment of localized prostate cancer: A systematic review. *Cost Effectiveness and Resource Allocation, 17*(1), 1–12. https://doi.org/10.1186/s12962-019-0178-3

Ahrends, A., Burgess, N. D., Gereau, R. E., Marchant, R., Bulling, M. T., Lovett, J. C., Platts, P. J., Wilkins Kindemba, V., Owen, N., Fanning, E., & Rahbek, C. (2011). Funding begets biodiversity. *Diversity and Distributions, 17*(2), 191–200. https://doi.org/10.1111/j.1472-4642.2010.00737.x

Anderson, P. M. L., Brown-Luthango, M., Cartwright, A., Farouk, I., & Smit, W. (2013). Brokering communities of knowledge and practice: Reflections on the African Centre for Cities' CityLab programme. *Cities, 32*, 1–10. https://doi.org/10.1016/j.cities.2013.02.002

Angel, S., Parent, J., Civco, D. L., Blei, A., & Potere, D. (2011). The dimensions of global urban expansion: Estimates and projections for all countries, 2000–2050. *Progress in Planning, 75*(2), 53–107. https://doi.org/10.1016/j.progress.2011.04.001

Dipeolu, A. A., Aluko, R. O., Omoniyi, S. S., & Oyinloye, A. M. (2020). Assessment of residents' socio-demographic factors associated with visit to green infrastructure facilities in Lagos Metropolis, Nigeria. *International Journal of Built Environment and Sustainability, 7*(2), 45–55. https://doi.org/10.11113/ijbes.v7.n2.489

Babalola, F. D., Borokini, T. I., Onefeli, A. O., & Muchie, M. (2013). Socio-economic contributions of an indigenous tree in urban areas of Southwest Nigeria. *African Journal of Science, Technology, Innovation and Development, 5*(6), 479–489. https://doi.org/10.1080/20421338.2013.820449

Beninde, J., Veith, M., & Hochkirch, A. (2015). Biodiversity in cities needs space: A meta-analysis of factors determining intra-urban biodiversity variation. *Ecology Letters, 18*(6), 581–592. https://doi.org/10.1111/ele.12427

Benza, M., Weeks, J. R., Stow, D. A., López-Carr, D., & Clarke, K. C. (2016). A pattern-based definition of urban context using remote sensing and GIS. *Remote Sensing of Environment, 183*, 250–264. https://doi.org/10.1016/j.rse.2016.06.011

Bernat-Ponce, E., Ferrer, D., Gil-Delgado, J. A., & López-Iborra, G. M. (2021). Effect of replacing surface with underground rubbish containers on urban House Sparrows *Passer domesticus*. *Urban Ecosystems*. https://doi.org/10.1007/s11252-021-01138-y

Beyer, H. L., Venter, O., Grantham, H. S., & Watson, J. E. M. (2020). Substantial losses in ecoregion intactness highlight urgency of globally coordinated action. *Conservation Letters, 13*, e12692.

Bigirimana, J., Bogaert, J., de Canniere Charles, C., Lejoly, J., & Parmentier, I. (2011). Alien plant species dominate the vegetation in a city of Sub-Saharan Africa. *Landscape and Urban Planning, 100*(3), 251–267. https://doi.org/10.1016/j.landurbplan.2010.12.012

Boon, R., Cockburn, J., Douwes, E., Govender, N., Ground, L., McLean, C., Roberts, D., Rouget, M., & Slotow, R. (2016). Managing a threatened savanna ecosystem (KwaZuluNatal Sandstone Sourveld) in an urban biodiversity hotspot: Durban, South Africa. *Bothalia, 46*(2). https://doi.org/10.4102/abc.v46i2.2112

Bornmann, L., Stefaner, M., de Moya Anegón, F., & Mutz, R. (2014). What is the effect of country-specific characteristics on the research performance of scientific institutions? Using multi-level statistical models to rank and map universities and research-focused institutions worldwide. *Journal of Informetrics, 8*(3), 581–593. https://doi.org/10.1016/j.joi.2014.04.008

Botts, E. A., Erasmus, B. F. N., & Alexander, G. J. (2011). Geographic sampling bias in the South African Frog Atlas Project: Implications for conservation planning. *Biodiversity and Conservation, 20*(1), 119–139. https://doi.org/10.1007/s10531-010-9950-6

Brodie, J. F. (2009). Is research effort allocated efficiently for Conservation? Felidae as a global case study. *Biodiversity and Conservation, 18*(11), 2927–2939. https://doi.org/10.1007/s10531-009-9617-3

Brooke, Z. M., Bielby, J., Nambiar, K., & Carbone, C. (2014). Correlates of research effort in carnivores: Body size, range size and diet matter. *PLoS ONE, 9*(4). https://doi.org/10.1371/journal.pone.0093195

Brooks, T. M., Mittermeier, R. A., da Fonseca, G. A. B., Gerlach, J., Hoffmann, M., Lamoreux, J. F., Mittermeier, C. G., Pilgrim, J. D., & Rodrigues, A. S. L. (2006). Global biodiversity conservation priorities. *Science, 313*(5783), 58–61. https://www.science.org.

Brown, C., & Grant, M. (2005). Biodiversity and human health: What role for nature in healthy urban planning? *Built Environment, 31*(4), 326–338.

Burgess, N. D., Hales, J. D. A., Ricketts, T. H., & Dinerstein, E. (2006). Factoring species, non-species values and threats into biodiversity prioritisation across the ecoregions of Africa and its islands. *Biological Conservation, 127*(4), 383–401. https://doi.org/10.1016/j.biocon.2005.08.018

Burgess, N., D'Amico Hales, J., Underwood, E., Dinerstein, E., Olson, D., Itoua, I., Schipper, J., Ricketts, T., & Newman, K. (2004). *Terrestrial eco-regions of Africa and Madagascar: A conservation assessment*. Washington, DC: Island Press. https://www.researchgate.net/publication/292588815.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach* (2nd Ed.). Springer.

Callaghan, C. T., Ozeroff, I., Hitchcock, C., & Chandler, M. (2020). Capitalizing on opportunistic citizen science data to monitor urban biodiversity: A multi-taxa framework. *Biological Conservation, 251*. https://doi.org/10.1016/j.biocon.2020.108753

Cazzolla Gatti, R., Castaldi, S., Lindsell, J. A., Coomes, D. A., Marchetti, M., Maesano, M., di Paola, A., Paparella, F., & Valentini, R. (2015). The impact of selective logging and clearcutting on forest structure, tree diversity and above-ground biomass of African tropical forests. *Ecological Research, 30*(1), 119–132. https://doi.org/10.1007/s11284-014-1217-3

Chamberlain, D. E., Henry, D. A. W., Reynolds, C., Caprio, E., & Amar, A. (2019). The relationship between wealth and biodiversity: A test of the Luxury Effect on bird species richness in the developing world. *Global Change Biology, 25*(9), 3045–3055. https://doi.org/10.1111/gcb.14682

Chen, M., Huang, Y., Tang, Z., Lu, D., Liu, H., & Ma, L. (2014). The provincial pattern of the relationship between urbanization and economic development in China. *Journal of Geographical Sciences, 24*(1), 33–45. https://doi.org/10.1007/s11442-014-1071-9

Childers, D. L., Cadenasso, M. L., Morgan Grove, J., Marshall, V., McGrath, B., & Pickett, S. T. A. (2015). An ecology for cities: A transformational nexus of design and ecology to advance climate change resilience and urban sustainability. *Sustainability, 7*(4), 3774–3791. https://doi.org/10.3390/su7043774

Cilliers, E. J. (2019). Reflecting on green infrastructure and spatial planning in Africa: The complexities, perceptions, and way forward. *Sustainability, 11*(2). https://doi.org/10.3390/su11020455

Cilliers, S., Cilliers, J., Lubbe, R., & Siebert, S. (2013). Ecosystem services of urban green spaces in African countries-perspectives and challenges. *Urban Ecosystems, 16*(4), 681–702. https://doi.org/10.1007/s11252-012-0254-3

Cilliers, S. S., & Bredenkampl, G. J. (1999). Analysis of the spontaneous vegetation of intensively managed urban open spaces in the Potchefstroom Municipal Area, North West Province, South Africa. *South African Journal of Botany, 65*(1), 59–68.

Cilliers, S. S., Müller, N., & Drewes, E. (2004). Overview on urban nature conservation: Situation in the western-grassland biome of South Africa. *Urban Forestry and Urban Greening, 3*(1), 49–62. https://doi.org/10.1016/j.ufug.2004.04.003

Cilliers, S. S., & Siebert, S. J. (2012). Urban ecology in Cape Town: South African comparisons and reflections. *Ecology and Society, 17*(3). https://doi.org/10.5751/ES-05146-170333

Cilliers, S. S., Siebert, S. J., du Toit, M. J., Barthel, S., Mishra, S., Cornelius, S. F., & Davoren, E. (2018). Garden ecosystem services of Sub-Saharan Africa and the role of health clinic gardens as social-ecological systems. *Landscape and Urban Planning, 180*, 294–307. https://doi.org/10.1016/j.landurbplan.2017.01.011

Cobbinah, P. B., Erdiaw-Kwasie, M. O., & Amoateng, P. (2015). Africa's urbanisation: Implications for sustainable development. *Cities, 47*, 62–72. https://doi.org/10.1016/j.cities.2015.03.013

Cohen, B. (2004). Urban growth in developing countries: A review of current trends and a caution regarding existing forecasts. *World Development, 32*(1), 23–51. https://doi.org/10.1016/j.worlddev.2003.04.008

Cohen, B. (2006). Urbanization in developing countries: Current trends, future projections, and key challenges for sustainability. *Technology in Society, 28*(1–2), 63–80. https://doi.org/10.1016/j.techsoc.2005.10.005

Cohen, C., Collar, N. J., Dagne, A., Fishpool, L. D. C., Marsden, S. J., Spottiswoode, C. N., & Wotton, S. R. (2021). Status of Taita Falcon Falco fasciinucha in Ethiopia and the identification problem posed by African Hobby F. cuvierii. *ABC Bulletin, 28* (2), 225–233.

Corbyn, Z. (2010). Urban ecosystems research joins mainstream ecology. *Nature, 467* (September), 153. https://doi.org/10.1038/news.2010.359

Crawley, M. J. (2013). *The R book* (2nd Ed.). John Wiley & Sons Ltd.

Cresswell, W. (2018). The continuing lack of ornithological research capacity in almost all of West Africa. *Ostrich, 89*(2), 123–129. https://doi.org/10.2989/00306525.2017.1388301

Cronk, N. E., & Pillay, N. (2018). Food choice and feeding on carrion in two African mongoose species in an urban environment. *Acta Ethologica, 21*(2), 127–136. https://doi.org/10.1007/s10211-018-0291-x

Dallimer, M., Irvine, K. N., Skinner, A. M. J., Davies, Z. G., Rouquette, J. R., Maltby, L. L., Warren, P. H., Armsworth, P. R., & Gaston, K. J. (2012). Biodiversity and the feel-good factor: Understanding associations between self-reported human well-being and species richness. *BioScience, 62*(1), 47–55. https://doi.org/10.1525/bio.2012.62.1.9

Davoren, E., & Shackleton, C. M. (2021). Urban ecosystem disservices in the global South. In C.M. Shackleton, S.S. Cilliers, E. Davoren, & M.J. du Toit (Eds.), *Urban ecology in the global South, cities and nature* (pp. 265–292). Springer Nature.

de Keijzer, C., Gascon, M., Nieuwenhuijsen, M. J., & Dadvand, P. (2016). Long-term green space exposure and cognition across the life course: A systematic review. *Current Environmental Health Reports, 3*(4), 468–477. https://doi.org/10.1007/s40572-016-0116-x

de Lima, R. F., Bird, J. P., & Barlow, J. (2011). Research effort allocation and the conservation of restricted-range island bird species. *Biological Conservation, 144*(1), 627–632. https://doi.org/10.1016/j.biocon.2010.10.021

Dobbs, C., Vasquez, A., Olave, P., & Olave, M. (2021). Cultural urban ecosystem services. In C.M. Shackleton, S.S. Cilliers, E. Davoren, & M.J. du Toit (Eds.), *Urban ecology in the global South, cities and nature* (pp. 245–264). Springer Nature.

Doi, H., & Takahara, T. (2016). Global patterns of conservation research importance in different countries of the world. *PeerJ, 4*(e2173). https://doi.org/10.7717/peerj.2173

Donaldson, M. R., Burnett, N. J., Braun, D. C., Suski, C. D., Hinch, S. G., Cooke, S. J., & Kerr, J. T. (2017). Taxonomic bias and international biodiversity conservation research. *FACETS, 1*(1), 105–113. https://doi.org/10.1139/facets-2016-0011

du Toit, M. J., Cilliers, S. S., Dallimer, M., Goddard, M., Guenat, S., & Cornelius, S. F. (2018). Urban green infrastructure and ecosystem services in sub-Saharan Africa. *Landscape and Urban Planning, 180*, 249–261. https://doi.org/10.1016/j.landurbplan.2018.06.001

Ducatez, S., & Lefebvre, L. (2014). Patterns of research effort in birds. *PLoS ONE, 9*(2). https://doi.org/10.1371/journal.pone.0089955

Epp Schmidt, D. J., Kotze, D. J., Hornung, E., Setälä, H., Yesilonis, I., Szlavecz, K., Dombos, M., Pouyat, R., Cilliers, S., Tóth, Z., & Yarwood, S. (2019). Metagenomics reveals bacterial and archaeal adaptation to urban land-use: N catabolism, methanogenesis, and nutrient acquisition. *Frontiers in Microbiology, 10*. https://doi.org/10.3389/fmicb.2019.02330

Ernstson, H. (2013). The social production of ecosystem services: A framework for studying environmental justice and ecological complexity in urbanized landscapes. *Landscape and Urban Planning, 109*(1), 7–17. https://doi.org/10.1016/j.landurbplan.2012.10.005

Escobedo, F. J. (2021). Understanding urban regulating ecosystem services in the global South. In C.M. Shackleton, S.S. Cilliers, E. Davoren, & M.J. du Toit (Eds.), *Urban ecology in the global South, cities and nature* (pp. 227–244). Springer Nature.

Fayiga, A. O., Ipinmoroti, M. O., & Chirenje, T. (2018). Environmental pollution in Africa. In *Environment, development and sustainability*, 20 (1), 41–73. Springer Netherlands. https://doi.org/10.1007/s10668-016-9894-4.

Ferketic, J. S., Latimer, A. M., & Silander, J. A. (2010). Conservation justice in metropolitan Cape Town: A study at the Macassar Dunes Conservation Area. *Biological Conservation, 143*(5), 1168–1174. https://doi.org/10.1016/j.biocon.2010.02.024

Fisher, R., Radford, B. T., Knowlton, N., Brainard, R. E., Michaelis, F. B., & Caley, M. J. (2011). Global mismatch between research effort and conservation needs of tropical coral reefs. *Conservation Letters, 4*, 64–72.

Fishpool, L. D. C., & Collar, N. J. (2018). On the validity of the Golden-naped Weaver Ploceus aureonucha and the Yellow-legged Weaver *P. flavipes*, with comments on forest "nuthatch-weavers." *ABC Bulletin, 25* (2), 159–179.

Forman, R. T. T. (2014). *Urban ecology: Science of cities* (1st Ed.). Cambridge University Press.

Forman, R. T. T. (2016). Urban ecology principles: Are urban ecology and natural area ecology really different? *Landscape Ecology, 31*(8), 1653–1662. https://doi.org/10.1007/s10980-016-0424-4

Gao, J., & O'Neill, B. (2021). Different spatiotemporal patterns in global human population and built-up land. *Earth's Future, 9*(8). https://doi.org/10.1029/2020EF001920

Garousi, V., Felderer, M., & Mäntylä, M. V. (2019). Guidelines for including grey literature and conducting multivocal literature reviews in software engineering. *Information and Software Technology, 106*, 101–121. https://doi.org/10.1016/j.infsof.2018.09.006

Gaston, K. (2010). *Urban ecology*. Cambridge: Cambridge University Press.

GBIF: The global biodiversity information facility. (2021). *What is GBIF? Available from* https://www.gbif.org/what-is-gbif *[22—28 December 2021]*.

Girma, Y., Terefe, H., & Pauleit, S. (2019). Urban green spaces use and management in rapidly urbanizing countries: The case of emerging towns of Oromia special zone surrounding Finfinne, Ethiopia. *Urban Forestry and Urban Greening, 43*. https://doi.org/10.1016/j.ufug.2019.05.019

Goosen, Z., & Cilliers, E. J. (2020). Enhancing social sustainability through the planning of third places: A theory-based framework. *Social Indicators Research, 150*(3), 835–866. https://doi.org/10.1007/s11205-020-02350-7

Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal, 26*(2), 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Grimm, N. B., Faeth, S. H., Golubiewski, N. E., Redman, C. L., Wu, J., Bai, X., & Briggs, J. M. (2008). Global change and the ecology of cities. *Science, 319*(5864), 756–760. https://www.science.org.

Güneralp, B., Lwasa, S., Masundire, H., Parnell, S., & Seto, K. C. (2018). Urbanization in Africa: Challenges and opportunities for conservation. *Environmental Research Letters, 13*(1), Article 015002. https://doi.org/10.1088/1748-9326/aa94fe

Haase, D., Kabisch, S., Haase, A., Andersson, E., Banzhaf, E., Baró, F., Brenck, M., Fischer, L. K., Frantzeskaki, N., Kabisch, N., Krellenberg, K., Kremer, P., Kronenberg, J., Larondelle, N., Mathey, J., Pauleit, S., Ring, I., Rink, D., Schwarz, N., & Wolff, M. (2017). Greening cities – To be socially inclusive? About the alleged paradox of society and ecology in cities. *Habitat International, 64*, 41–48. https://doi.org/10.1016/j.habitatint.2017.04.005

Hima, K., Houémenou, G., Badou, S., Garba, M., Dossou, H. J., Etougbétché, J., Gauthier, P., Artige, E., Fossati-Gaschignard, O., Gagaré, S., Dobigny, G., & Dalecky, A. (2019). Native and invasive small mammals in urban habitats along the commercial axis connecting Benin and Niger, West Africa. *Diversity, 11*(12), 1–20. https://doi.org/10.3390/d11120238

Hugo, M. L. (1979). A note on ethnicity and outdoor recreation: A South African Experience. *South African Journal of Sociology, 19*, 62–70.

Hulme, M., Doherty, R., Ngara, T., New, M., & Lister, D. (2001). African climate change: 1900–2100. *Climate Research, 17*, 145–168. http://www.cru.uea.ac.uk.

Ibáñez-Álamo, J. D., Rubio, E., Benedetti, Y., & Morelli, F. (2017). Global loss of avian evolutionary uniqueness in urban areas. *Global Change Biology, 23*(8), 2990–2998. https://doi.org/10.1111/gcb.13567

Ives, C. D., Lentini, P. E., Threlfall, C. G., Ikin, K., Shanahan, D. F., Garrard, G. E., Bekessy, S. A., Fuller, R. A., Mumaw, L., Rayner, L., Rowe, R., Valentine, L., & Kendal, D. (2016). Cities are hotspots for threatened species. *Global Ecology and Biogeography, 25*, 117–126.

Kartzinel, T.R., Hsing, J. C., Musili, P. M., Brown, B. R. P., & Pringle, R. M. (2019). Covariation of diet and gut microbiome in African megafauna. *Proceedings of the National Academy of Sciences, 116* (47), 23588–23593. https://doi.org/10.5061/dryad.c119gm5.

Kendal, D., Egerer, M., Byrne, J. A., Jones, P. J., Marsh, P., Threlfall, C. G., Allegretto, G., Kaplan, H., Nguyen, H. K. D., Pearson, S., Wright, A., & Flies, E. J. (2020). City-size bias in knowledge on the effects of urban nature on people and biodiversity. *Environmental Research Letters, 15*(12). https://doi.org/10.1088/1748-9326/abc5e4

Kent, R., Johnson, D. E., & Becker, M. (2001). The influences of cropping system on weed communities of rice in Côte d'Ivoire, West Africa. *Agriculture, Ecosystems and Environment, 87*, 299–307. https://doi.org/10.1016/S0167-8809(01)00153-0

Kigozi, S. P., Kigozi, R. N., Sebuguzi, C. M., Cano, J., Rutazaana, D., Opigo, J., Bousema, T., Yeka, A., Gasasira, A., Sartorius, B., & Pullan, R. L. (2020). Spatial-temporal patterns of malaria incidence in Uganda using HMIS data from 2015 to 2019. *BMC Public Health, 20*, 1913. https://doi.org/10.1186/s12889-020-10007-w

King, A., & Shackleton, C. M. (2020). Maintenance of public and private urban green infrastructure provides significant employment in Eastern Cape towns, South Africa. *Urban Forestry and Urban Greening, 54*. https://doi.org/10.1016/j.ufug.2020.126740

Koricho, H. H., Seboka, A. D., & Song, S. (2020). Assessment of the structure, diversity, and composition of woody species of urban forests of Adama city, Central Ethiopia. *Arboricultural Journal*. https://doi.org/10.1080/03071375.2020.1798702.

Kowe, P., Mutanga, O., Odindi, J., & Dube, T. (2020). A quantitative framework for analysing long term spatial clustering and vegetation fragmentation in an urban landscape using multi-temporal landsat data. *International Journal of Applied Earth Observation and Geoinformation, 88*(102057). https://doi.org/10.1016/j.jag.2020.102057

Kumdet, P. S., Ivande, S. T., & Dami, F. D. (2021). Key drivers of avifauna in greenspace of institutional campuses in a state in Western Africa. *Urban Forestry and Urban Greening, 61*(March), Article 127092. https://doi.org/10.1016/j.ufug.2021.127092

Langat, P. K., Kumar, L., & Koech, R. (2019). Monitoring river channel dynamics using remote sensing and GIS techniques. *Geomorphology, 325*, 92–102. https://doi.org/10.1016/j.geomorph.2018.10.007

Lawler, J. J., Aukema, J. E., Grant, J. B., Halpern, B. S., Kareiva, P., Nelson, C. R., Ohleth, K., Olden, J. D., Schlaepfer, M. A., Silliman, B. R., & Zaradic, P. (2006). Conservation science: A 20-year report card. *Frontiers in Ecology and the Environment, 4*(9), 473–480. https://doi.org/10.1890/1540-9295(2006)4[473:CSAYRC]2.0.CO;2

Lin, T., & Grimm, N. B. (2015). Comparative study of urban ecology development in the U.S. and China: Opportunity and challenge. *Urban Ecosystems, 18*(2), 599–611. https://doi.org/10.1007/s11252-014-0413-9

Lindley, S., Pauleit, S., Yeshitela, K., Cilliers, S., & Shackleton, C. (2018). Rethinking urban green infrastructure and ecosystem services from the perspective of sub-Saharan African cities. *Landscape and Urban Planning, 180*, 328–338. https://doi.org/10.1016/j.landurbplan.2018.08.016

Lüdeck, D., Makowski, M. S., Ben-Shachar, D., Patil, I., Waggoner, P., Wiernik, B. M., Arel-Bundock, V., & Jullum, M. (2021). *Assessment of regression models performance. Accessed 10 Jan 2022*.

Magle, S. B., Hunt, V. M., Vernon, M., & Crooks, K. R. (2012). Urban wildlife research: Past, present, and future. *Biological Conservation, 155*, 23–32. https://doi.org/10.1016/j.biocon.2012.06.018

Malonza, P. K., Granthon, C., & Williams, D. A. (2016). A new species of dwarf gecko in the genus Lygodactylus (squamata: Gekkonidae) from central Kenya. *Zootaxa, 4061*(4), 418–428. https://doi.org/10.11646/zootaxa.4061.4.6

Man, M. P., Weinkauf, J. G., Tsang, M., & Sin, D. D. (2004). Why do some countries publish more than others? An international comparison of research funding, English proficiency and publication output in highly ranked general medical journals. *European Journal of Epidemiology, 19*, 811–817.

Manley, C. B., Rakocinski, C. F., Lee, P. G., & Blaylock, R. B. (2015). Feeding frequency mediates aggression and cannibalism in larval hatchery-reared spotted seatrout, *Cynoscion nebulosus*. *Aquaculture, 437*, 155–160. https://doi.org/10.1016/j.aquaculture.2014.11.012

Marzluff, J. M. (2016). A decadal review of urban ornithology and a prospectus for the future. *Ibis, 159*, 1–13. https://doi.org/10.1111/ibi.12430

McDonald, R. I., Kareiva, P., & Forman, R. T. T. (2008). The implications of current and future urbanization for global protected areas and biodiversity conservation. *Biological Conservation, 141*(6), 1695–1703. https://doi.org/10.1016/j.biocon.2008.04.025

McDonnel, M. (2011). The history of urban ecology: An ecologist perspective. *Urban Ecology: Patterns, Processes and Applications, 9*, 5–13.

McDuff, M. (2000). Thirty years of environmental education in Africa: The role of the wildlife clubs of Kenya. *Environmental Education Research, 6*(4), 383–396. https://doi.org/10.1080/713664697

McIntyre, N. E. (2000). Ecology of urban arthropods: A review and a call to action. *Annals of the Entomological Society of America, 93*(4), 825–835. https://academic.oup.com/aesa/article/93/4/825/22659.

McKinney, M. L. (2006). Urbanization as a major cause of biotic homogenization. *Biological Conservation, 127*(3), 247–260. https://doi.org/10.1016/j.biocon.2005.09.005

McKinney, M. L. (2008). Effects of urbanization on species richness: A review of plants and animals. *Urban Ecosystems, 11*, 161–176.

McPhearson, T., Pickett, S. T. A., Grimm, N. B., Niemelä, J., Alberti, M., Elmqvist, T., Weber, C., Haase, D., Breuste, J., & Qureshi, S. (2016). Advancing urban ecology toward a science of cities. *BioScience, 66*(3), 198–212. https://doi.org/10.1093/biosci/biw002

McPherson, S. C., Brown, M., & Downs, C. T. (2016). Crowned eagle nest sites in an urban landscape: Requirements of a large eagle in the Durban Metropolitan Open Space System. *Landscape and Urban Planning, 146*, 43–50. https://doi.org/10.1016/j.landurbplan.2015.10.004

Meo, S. A., Al Masri, A. A., Usmani, A. M., Memon, A. N., & Zaidi, S. Z. (2013). Impact of GDP, spending on R&D, number of universities and scientific journals on research publications among Asian countries. *PLoS ONE, 8*(6), e66449. https://doi.org/10.1371/journal.pone.0066449

Meo, S. A., Usmani, A. M., & Meo, A. (2014). Impact of R&D expenditures on research publications, patents and high-tech exports among European countries. *Europe Review for Medical and Pharmacological Science, 18*, 1–9.

Miller, J. R., & Hobbs, R. J. (2002). Conservation where people live and work. *Conservation Biology, 16*(2).

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & P.g.. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *BMJ, 339* (7716), 332–336. https://doi.org/10.1136/bmj.b2535

Moomaw, R. L., & Shatter, A. M. (1996). Urbanization and economic development: A bias toward large cities? *Journal of Urban Economics, 40*(21), 13–37.

Moragues-Faus, A., & Carroll, B. (2018). Reshaping urban political ecologies: An analysis of policy trajectories to deliver food security. *Food Security, 10*(6), 1337–1351. https://doi.org/10.1007/s12571-018-0855-7

Moussa, S., Kuyah, S., Kyereh, B., Tougiani, A., & Mahamane, S. (2020). Diversity and structure of urban forests of Sahel cities in Niger. *Urban Ecosystems, 23*(4), 851–864. https://doi.org/10.1007/s11252-020-00984-6

Muchayi, G. K., Gandiwa, E., & Muboko, N. (2017). Composition and structure of woody vegetation in an urban environment in northern Zimbabwe. *Tropical Ecology, 58*(2), 347–356. www.tropecol.com.

Muleya, N., & Campbell, M. (2020). A multisensory approach to measure public space quality in the city of Bulawayo, Zimbabwe. *Town and Regional Planning, 76*, 56–71. https://doi.org/10.18820/2415-0495/trp76i1.5

OECD/SWAC. (2020). Africapolis (database), www.africapolis.org (accessed 11 February 2020).

Okpala, D. C. I. (1978). Urban ecology and residential location theories: Application in Nigeria's socio-cultural milieu. *Socio-Economic Planning Sciences, 12*, 177–183.

Olofsson, H., Brolund, A., Hellberg, C., Silverstein, R., Stenström, K., Österberg, M., & Dagerhamn, J. (2017). Can abstract screening workload be reduced using text mining? User experiences of the tool Rayyan. *Research Synthesis Methods, 8*(3), 275–280. https://doi.org/10.1002/jrsm.1237

Olson, D. M., Dinerstein, E., Wikramanayake, E. D., Burgess, N. D., Powell, G. V. N., Underwood, E. C., D'Amico, J. A., Itoua, I., Strand, H. E., Morrison, J. C., Loucks, C. J., Allnutt, T. F., Ricketts, T. H., Kura, Y., Lamoreux, J. F., Wettengel, W. W., Hedao, P., & Ka, K. R. (2001). Terrestrial ecoregions of the world: A new map of life on Earth. *Bioscience, 51*(11), 933–938.

Orsini, F., Kahane, R., Nono-Womdim, R., & Gianquinto, G. (2013). Urban agriculture in the developing world: A review. *Agronomy for Sustainable Development, 33*(4), 695–720. https://doi.org/10.1007/s13593-013-0143-z

O'Sullivan, O. S., Holt, A. R., Warren, P. H., & Evans, K. L. (2017). Optimising UK urban road verge contributions to biodiversity and ecosystem services with cost-effective management. *Journal of Environmental Management, 191*, 162–171. https://doi.org/10.1016/j.jenvman.2016.12.062

Ouzzani, M., Hammady, H., Fedorowicz, Z., & Elmagarmid, A. (2016). Rayyan-a web and mobile app for systematic reviews. *Systematic Reviews, 5*(1), 1–10. https://doi.org/10.1186/s13643-016-0384-4

Paillette, M., Capy, P., Pla, E., & David, J. R. (1993). Reproductive isolation through homogamy in 2 sympatric habitat races of *Drosophila melanosgaster*. *Revue D Ecologie-La Terre Et La Vie, 48*(2), 229–238.

Patel, Z., Greyling, S., Simon, D., Arfvidsson, H., Moodley, N., Primo, N., & Wright, C. (2017). Local responses to global sustainability agendas: Learning from experimenting with the urban sustainable development goal in Cape Town. *Sustainability Science, 12*(5), 785–797. https://doi.org/10.1007/s11625-017-0500-y

Patterson, L., Kalle, R., & Downs, C. (2016). Predation of artificial bird nests in suburban gardens of KwaZulu-Natal, South Africa. *Urban Ecosystems, 19*(2), 615–630. https://doi.org/10.1007/s11252-016-0526-4

Paul, M. J., & Meyer, J. L. (2001). Streams in the urban landscape. *Annual Review of Ecology and Systematics, 32*, 333–365.

Popoola, L., & Ajewole, O. (2002). Willingness to pay for rehabilitation of Ibadan urban environment through reforestation projects. *International Journal of Sustainable Development and World Ecology, 9*(3), 256–268. https://doi.org/10.1080/13504500209470121

Pyšek, P., Richardson, D. M., Pergl, J., Jarošík, V., Sixtová, Z., & Weber, E. (2008). Geographical and taxonomic biases in invasion ecology. *Trends in Ecology and Evolution, 23*(5), 237–244. https://doi.org/10.1016/j.tree.2008.02.002

Qizhi, M., Long, Y., Wu, K., Ying, L., & Kang, W. (2016). Spatio-temporal changes of population density and urbanization pattern in China (2000–2010). *China City Planning Review, 25*(4), 8–14. https://www.researchgate.net/publication/311707586.

R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Raji, I. A., & Downs, C. T. (2021). Ficus-frugivore interactions, especially in areas of land-use change, in Africa: A systematic review. *Acta Oecologica, 113*(103774). https://doi.org/10.1016/j.actao.2021.103774

Rantalainen, M. L., Fritze, H., Haimi, J., Pennanen, T., & Setala, H. (2005). Species richness and food web structure of soil decomposer community as affected by the size of habitat fragment and habitat corridors. *Global Change Biology, 11*, 1614–1627. https://doi.org/10.1111/j.1365-2486.2005.00999.x

Rebelo, A. G., Holmes, P. M., Dorse, C., & Wood, J. (2011). Impacts of urbanization in a biodiversity hotspot: Conservation challenges in Metropolitan Cape Town. *South African Journal of Botany, 77*(1), 20–35. https://doi.org/10.1016/j.sajb.2010.04.006

Reynolds, C., Byrne, M. J., Chamberlain, D. E., Howes, C. G., Seymour, C.L., Sumasgutner, P., & Taylor, P. J. (2021). Urban animal diversity in the global South. In C.M. Shackleton, S.S. Cilliers, E. Davoren, & M.J. du Toit (Eds.), *Urban ecology in the global South, cities and nature* (pp. 169–202). Springer Nature.

Richardson, D. M., & Rejmanek, M. (2011). Trees and shrubs as invasive alien species - a global review. *Diversity and Distributions, 17*, 788–809.

Roets, F., Allison, J. D., & Basson, R. J. (2019). Recent records of fruit chafers (Scarabaeidae: Cetoniinae: Cetoniini) in the Southwestern Cape Region of South Africa suggest that range expansions were facilitated by human-mediated jump-dispersal and pre-adaptation to transformed landscapes. *African Entomology, 27*(1), 135–145. https://doi.org/10.4001/003.027.0135

Rogerson, C. M., & Rogerson, J. M. (2020). Municipal caravan parks in South Africa: Geography, management and planning. *Bulletin of Geography. Socio-Economic Series, 49*(49), 109–124. https://doi.org/10.2478/bog-2020-0027

Roy, S., Byrne, J., & Pickering, C. (2012). A systematic quantitative review of urban tree benefits, costs, and assessment methods across cities in different climatic zones. *Urban Forestry and Urban Greening, 11*(4), 351–363. https://doi.org/10.1016/j.ufug.2012.06.006

Sánchez-Tójar, A., Lagisz, M., Moran, N. P., Nakagawa, S., Noble, D. W. A., & Reinhold, K. (2020). The jury is still out regarding the generality of adaptive 'transgenerational' effects. *Ecology Letters, 23*(11), 1715–1718. https://doi.org/10.1111/ele.13479

Scholes, M. C., Crouch, A., Erasmus, B. F. N., Schwaibold, U., Vogel, C. H., & Thormeyer, T. (2013). Initiatives to enhance African Environmental Science. *Environmental Development, 5*(1), 164–168. https://doi.org/10.1016/j.envdev.2012.11.007

Seto, K. C., Güneralp, B., & Hutyra, L. R. (2012). Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proceedings of the National Academy of Sciences of the United States of America, 109*(40), 16083–16088. https://doi.org/10.1073/pnas.1211658109

Shackleton, C. M., Cilliers, S. S., du Toit, M. J., & Davoren, E. (2021). The need for an urban ecology of the global South. In C.M. Shackleton, S.S. Cilliers, E. Davoren, & M. J. du Toit (Eds.), *Urban ecology in the global South* (pp. 1–26). Springer Nature.

Shackleton, C. M., Hurley, P. T., Dahlberg, A. C., Emery, M. R., & Nagendra, H. (2017). Urban foraging: A ubiquitous human practice overlooked by urban planners, policy, and research. *Sustainability, 9*(10). https://doi.org/10.3390/su9101884

Shapiro, S. S., & Wilk, M. B. (1965). *An analysis of variance test for normality (complete samples)* (Vol. 52, Issue 3). https://www.jstor.org/stable/2333709.

Shwartz, A., Turbé, A., Julliard, R., Simon, L., & Prévot, A. C. (2014). Outstanding challenges for urban conservation research and action. *Global Environmental Change, 28*(1), 39–49. https://doi.org/10.1016/j.gloenvcha.2014.06.002

Smales, L. R., Halajian, A., Mokgawa, M. P., & Luus-Powell, W. J. (2017). A new species of Centrorhynchus Lühe, 1911 (Acanthocephala: Centrorhynchidae) from the lizard buzzard *Kaupifalco monogrammicus* (Temminck) (Aves: Accipitridae) in South Africa. *Systematic Parasitology, 94*(3), 423–430. https://doi.org/10.1007/s11230-017-9710-z

Stranga, Y., & Katsanevakis, S. (2021). Eight years of bioinvasions records: Patterns and trends in alien and cryptogenic species records. *Management of Biological Invasions, 12*(2), 221–239. https://doi.org/10.3391/mbi.2021.12.2.01

Tan, P. Y., & bin Abdul Hamid, A.R. (2014). Urban ecological research in Singapore and its relevance to the advancement of urban ecology and sustainability. *Landscape and Urban Planning, 125*, 271–289. https://doi.org/10.1016/j.landurbplan.2014.01.019

Thompson, L. R., Sanders, J. G., McDonald, D., Amir, A., Ladau, J., Locey, K. J., Prill, R. J., Tripathi, A., Gibbons, S. M., Ackermann, G., Navas-Molina, J. A., Janssen, S., Kopylova, E., Vázquez-Baeza, Y., González, A., Morton, J. T., Mirarab, S., Xu, Z. Z., Jiang, L., ….. Knight, R. & The earth microbiome Project consortium (2017). A communal catalogue reveals Earth's multiscale microbial diversity. *Nature, 551*(7681), 457–463. https://doi.org/10.1038/nature24621.

Tóth, Z., Szlavecz, K., Epp Schmidt, D. J., Hornung, E., Setälä, H., Yesilonis, I. D., Kotze, D. J., Dombos, M., Pouyat, R., Mishra, S., Cilliers, S., Yarwood, S., & Csuzdi, C. (2020). Earthworm assemblages in urban habitats across biogeographical regions. *Applied Soil Ecology, 151*. https://doi.org/10.1016/j.apsoil.2020.103530

Trimble, M. J., & van Aarde, R. J. (2014). Supporting conservation with biodiversity research in sub-Saharan Africa' s human-modified landscapes. *Biodiversity Conservation, 23*, 2345–2369. https://doi.org/10.1007/s10531-014-0716-4

Lee, A. T. K., & Nel, H. (2020). BirdLasser: The influence of a mobile app on a citizen science project. *African Zoology, 55*(2), 155–160. https://doi.org/10.1080/15627020.2020.1717376

UNESCO. (2020). *UNESCO Institute for statistics. Accessed 30 October 2022.*

United Nations. (2016). *Urbanization and development: Emerging futures. World Cities Report 2016*. Nairobi: United Nations.

United Nations. (2021). The Sustainable Development Goals Report: Make cities and human settlements inclusive, safe, resilient and sustainable.

United Nations Department of Economic and Social Affairs (Population Division). (2018). *World Urbanization Prospects: The 2018 Revision, Online Edition.*

United Nations Department of Economic and Social Affairs (Population Division). (2019). *World Population Prospects 2019, Online Edition. Rev. 1.*

United Nations. Economic Commission for Africa (2018). Towards Achieving the African Union's recommendation of expenditure of 1% of GDP on research and development. Addis Ababa. © UN. ECA. https://hdl.handle.net/10855/24306.

van der Walt, L., Cilliers, S. S., du Toit, M. J., & Kellner, K. (2015). Conservation of fragmented grasslands as part of the urban green infrastructure: How important are species diversity, functional diversity and landscape functionality? *Urban Ecosystems, 18*(1), 87–113. https://doi.org/10.1007/s11252-014-0393-9

van Zyl, A. J. (1994). The influence of the environment on breeding success of a suburban population of Crested Barbets *Trachyphonus vaillantii*. *Ostrich, 65*(3–4), 291–296.

Vinkler, P. (2008). Correlation between the structure of scientific research, scientometric indicators and GDP in EU and non-EU countries. *Scientometrics, 74*(2), 237–254. https://doi.org/10.1007/s11192-008-0215-z

Wangai, P. W., Burkhard, B., & Müller, F. (2016). A review of studies on ecosystem services in Africa. *International Journal of Sustainable Built Environment, 5*(2), 225–245. https://doi.org/10.1016/j.ijsbe.2016.08.005

Widdows, C. D., & Downs, C. T. (2016). Urban roost temperatures of large-spotted-genets: The effect of anthropogenic structures. *Journal of Thermal Biology, 57*, 66–71. https://doi.org/10.1016/j.jtherbio.2016.02.004

World Bank. (2021). *Urban population growth. Accessed 30 October 2022.*

World Cities Report. (2016). *Urbanization and development: Emerging futures.*

Wu, J., Xiang, W.-N., & Zhao, J. (2014). Urban ecology in China: Historical developments and future directions. *Landscape and Urban Planning, 125*, 222–233. https://doi.org/10.1016/j.landurbplan.2014.02.010

Yvoz, S., Cordeau, S., Zuccolo, C., & Petit, S. (2020). Crop type and within-field location as sources of intraspecific variations in the phenology and the production of floral and fruit resources by weeds. *Agriculture, Ecosystems and Environment, 302*. https://doi.org/10.1016/j.agee.2020.107082

Zhang, X. Q. (2016). The trends, promises and challenges of urbanisation in the world. *Habitat International, 54*, 241–252. https://doi.org/10.1016/j.habitatint.2015.11.018

Mr. Adewale G. AWOYEMI is a conservation biologist, who currently heads the Forest Center at the International Institute of Tropical Agriculture in Ibadan, Nigeria. In March 2021, he was awarded as one of the Top 100 Young African Conservation Leaders. He is simultaneously pursuing his PhD at the University of Granada (Spain) to investigate the impacts of urbanization on avian health and biodiversity in tropical environments under the supervision of Prof. Juan Diego IBÁÑEZ-ÁLAMO. Mr. Awoyemi is a registered member of the West African Ornithological Society, and African and Ibadan Bird Clubs.

Prof. Juan Diego IBÁÑEZ-ÁLAMO is a Senior Lecturer at the University of Granada. His research line combines avian behavioral ecology, eco-physiology and community ecology and has published over 60 papers in top journals. His research is particularly oriented to investigate the impact of urbanization on birds, but has also investigated other animal interactions (e.g., brood parasites, predator-prey, host-microbe). He is currently focused on exploring how urbanization is affecting avian biodiversity worldwide, its effects on avian physiology and behavior, and human-bird interactions in cities. He has received several Spanish Awards for his research and popularization of science activities.

SSM - Qualitative Research in Health 3 (2023) 100244

Contents lists available at ScienceDirect

SSM - Qualitative Research in Health

journal homepage: www.journals.elsevier.com/ssm-qualitative-research-in-health

Experiences with COVID-19 case investigation and contact tracing: A qualitative analysis☆

Nickolas DeLuca [a,*], Elise Caruso [a], Reena Gupta [b], Charlene Kemmerer [b], Rebecca Coughlin [b], Olivia Chan [b], Divya Vohra [b], John E. Oeltmann [a], Melanie M. Taylor [a], Patrick K. Moonan [a], Phoebe G. Thorpe [a], Penny S. Loosier [a], CDC COVID-19 Case Investigation and Contact Tracing Task Forcea, Geraldine Haile [b]

[a] U.S. Centers for Disease Control and Prevention, COVID-19 Response Team, Atlanta, GA, USA
[b] Mathematica, Cambridge, MA, USA

ARTICLE INFO

Keywords:
COVID-19
Case investigation
Contact tracing
Isolation
Quarantine
Health equity

ABSTRACT

Case investigation and contact tracing (CI/CT) is a critical part of the public health response to COVID-19. Individuals' experiences with CI/CT for COVID-19 varied based on geographic location, changes in knowledge and guidelines, access to testing and vaccination, as well as demographic characteristics including age, race, ethnicity, income, and political ideology. In this paper, we explore the experiences and behaviors of adults with positive SARS-CoV-2 test results, or who were exposed to a person with COVID-19, to understand their knowledge, motivations, and facilitators and barriers to their actions. We conducted focus groups and one-on-one interviews with 94 cases and 90 contacts from across the United States. We found that participants were concerned about infecting or exposing others, which motivated them to isolate or quarantine, notify contacts, and get tested. Although most cases and contacts were not contacted by CI/CT professionals, those who were reported a positive experience and received helpful information. Many cases and contacts reported seeking information from family, friends, health care providers, as well as television news and Internet sources. Although participants reported similar perspectives and experiences across demographic characteristics, some highlighted inequities in receiving COVID-19 information and resources.

1. Background and purpose

Case investigation and contact tracing (CI/CT) are longstanding public health measures used to mitigate the spread of infectious diseases and were a critical part of the public health response to COVID-19 (Centers for Disease Control and Prevention [CDC], 2022a; World Health Organization (WHO), 2021). *Case investigation* involves interviewing someone who has confirmed or suspected COVID-19 to learn whom they may have exposed to the virus, eliciting their contacts, counseling them to monitor their symptoms, and recommending isolation (i.e., staying home and away from others). *Contact tracing* is the subsequent process of notifying close contacts of their potential exposure, referring them to testing, counseling them to monitor their symptoms, and encouraging them to quarantine (i.e., staying home and away from others after potential exposure) (CDC, 2022a).

During the COVID-19 pandemic, health departments faced challenges and lacked resources to provide CI/CT to all cases and contacts, especially during surges of high COVID-19 incidence (Lash et al., 2021). The success of CI/CT also depends on the participation of individuals, which can be influenced by factors including potential stigma and lack of trust in government (Lash et al., 2021). Research has also revealed that individual knowledge of and responses to CI/CT varied based on demographic characteristics, such as age, race, ethnicity, income, and political ideology (McClain & Rainie, 2020). Moreover, CI/CT may have also been impacted based on access to COVID-19 testing. For example, although increased availability and use of self-tests beginning in 2021 let more people learn of their COVID-19 infection outside of a clinical or laboratory setting, these results were not routinely reported to health

☆ **Additional Contributions**: Holly Matulewicz, Melanie Au, Jason Markesich, Tricia Higgins (Mathematica), Michelle Fiscus, Elinor Higgins, Clare Cartheuser (National Academy for State Health Policy), David Lintern, Kelly Bell (IPSOS), Hugo S. Puerto (University of Washington).

* Corresponding author. US Centers for Disease Control and Prevention, 1600 Clifton Road NE, Atlanta, GA, 30333, USA.
E-mail address: nkdeluca@cdc.gov (N. DeLuca).

https://doi.org/10.1016/j.ssmqr.2023.100244
Received 28 October 2022; Received in revised form 27 February 2023; Accepted 1 March 2023
Available online 3 March 2023

departments, removing a key step in the CI/CT process (Rader et al., 2022).

Additionally, evolving guidance for testing, isolation, quarantine, and vaccination may have prompted individuals to alter their behaviors. For example, in December 2021, the Centers for Disease Control and Prevention (CDC) shortened the recommended duration of isolation for individuals with COVID-19 from 10 days to at least 5 days, with the recommendation to wear a well-fitting mask for 10 full days around others (CDC, 2021). The duration of quarantine and definition of a contact evolved over the course of the pandemic and were based on vaccination status and history of prior infection. For example, in December 2020, some jurisdictions implemented options to shorten quarantine for eligible people to 7 or 10 days from the recommended 14 days (CDC, 2020). By December 2021, CDC recommended that contacts who were not up to date with COVID-19 vaccines quarantine for at least 5 days, and all contacts, regardless of their vaccination status, get tested and wear a well-fitting mask around others for 10 days (CDC, 2021).

2. Methods

This study is part of a larger project that surveyed a nationally representative sample of the United States (U.S.) population who tested positive for SARS-CoV-2 (case) or were exposed to someone with COVID-19 (contact) to learn about their subsequent behaviors to prevent COVID-19 transmission (Oeltmann et al., working paper). At the time of the survey, respondents were routed as cases if they self-reported ever testing positive for COVID-19, regardless of close contact with someone who had COVID-19; respondents were routed as contacts if they only reported close contact with someone who had COVID-19 and had not tested positive for COVID-19. The survey sample was drawn from the Ipsos KnowledgePanel (Ipsos, 2022), a probability-based, web-based panel that provides a representative sampling frame for noninstitutionalized adults (aged >18 years) residing in the U.S. A question was appended to the survey asking participants if they would be willing to participate in a focus group discussion (FGD) related to their COVID-19 positive result or exposure (Oeltmann et al., working paper).

We offered eligibility questions to 22,514 KnowledgePanel participants. Of 15,923 participants who completed these questions, 9269 were eligible and participated in the survey. The survey completion rate was 70.7% (15,923/22,514). Of 9109 survey participants who responded to the question about FGD participation, 3208 (35.2%) expressed interest and 164 participated in a FGD. Thirteen participants did not respond to the question about FGD participation. Individuals who participated in a FGD were similar to those who did not regarding previous isolation or quarantine. For populations harder to reach through the survey, we partnered with a national community-based organization to recruit 7 additional participants.

We conducted FGDs and interviews with a subset of survey participants and a supplement of adult community members to explore the experiences and behaviors of cases and contacts to understand their knowledge, motivations, and facilitators and barriers to their actions. We explored differences by race, ethnicity, socioeconomic status, and political ideology to answer two research questions:

1. What were individuals' experiences after learning they had COVID-19 or had been exposed to COVID-19?
2. What factors affected individuals' experiences and behaviors and why?

2.1. Participants

Participants included cases and contacts representing people from a range of racial, ethnic, political, and socioeconomic groups (Table 1). Participants confirmed their positive SARS-CoV-2 test result or exposure; age; Hispanic, Latino, or Spanish origin; race; highest level of education; and political ideology.

Table 1
Focus group and interview participant demographics (N = 184).

	%	N
COVID-19 Status		
Case	51%	94
Contact	49%	90
Gender		
Male	50%	92
Female	50%	92
Age		
18-39	36%	65
40-59	38%	71
60+	26%	48
Race/Ethnicity		
American Indian or Alaskan Native, non-Hispanic	5%	10
Asian and Native Hawaiian or Pacific Islander, non-Hispanic	14%	25
Black or African American, non-Hispanic	28%	51
Hispanic	26%	48
White, non-Hispanic	27%	50
Education		
Less than high school	9%	17
High school	20%	37
Some college	27%	49
Bachelor's degree or higher	44%	81
Political ideology[a]		
Liberal	21%	38
Middle of the road	32%	59
Conservative	38%	70
No response	9%	17
Contacted by CI/CT professional		
Yes	26%	48
No	73%	135
No response	1%	1

[a] Participants were classified as Liberal if they chose extremely liberal, liberal, or slightly liberal; Middle of the road if they chose moderate/middle of the road; and Conservative if they selected slightly conservative, conservative, and extremely conservative.

Participants were scheduled for a FGD that aligned with their COVID-19 status, race, ethnicity, educational level, and political ideology. Focus groups ranged from 2 to 10 participants. To supplement smaller FGDs, we conducted 6 interviews.

2.2. Virtual focus groups and interviews

Focus group guides were developed and pretested prior to data collection. Following the pretests, moderators debriefed to adjust questions and probes, ordering of questions, and length of guides. We adapted the focus group guides for use in the one-on-one interviews.

Virtual FGDs and interviews were recorded using secure video teleconferencing and conducted in English and Spanish. A lead moderator facilitated each FGD, and a support moderator took notes. One interviewer facilitated each interview. All moderators and interviewers participated in a 2-h training to employ an effective, equitable, and trauma-informed approach that valued all participants' viewpoints (Mathematica, 2022). The training focused on being empathetic to experiences and feelings and avoiding re-traumatization associated with COVID-19.

All participants received documentation that described the purpose of the study, the voluntary nature of participation, and what to expect during each FGD or interview. This documentation was reviewed at the start of each group, and consent was obtained from each participant. Each participant received the equivalent of $75 for participating.

Thirty-three 60-min virtual, synchronous FGDs were conducted with 178 participants from March 17, 2022, to April 29, 2022. In-depth interviews were conducted with 6 participants during May 5–9, 2022. All FGDs and interviews were recorded, transcribed, and translated (if conducted in Spanish) for coding and analysis.

2.3. Analysis

Codebooks were developed based on the structured guides and themes that emerged during implementation. To ensure interrater reliability, a team of five coders applied codes to early transcripts to resolve any discrepancies and differences and met regularly to discuss emerging themes. We used NVivo software (Version 12, March 2020) for coding transcripts and used framework matrices to create analytic summaries of coded text. We used the Framework Method to analyze data based on demographic characteristics to compare summaries and identify areas of dissent (Gale et al., 2013).

2.4. Ethical considerations

Participation was voluntary, and all participants were provided privacy and confidentiality protections. This study was reviewed by CDC and deemed not to be research as defined in 45 CFR 46.102(l) (U.S. Department of Health and Human Services, Title 45 Code of Federal Regulations 46, Protection of Human Subjects).

3. Results

3.1. Reactions to COVID-19

Upon learning of their positive SARS-CoV-2 test result, cases noted that they were most concerned about overall community health and getting others sick, particularly those with preexisting conditions that may place them at higher risk for severe outcomes from COVID-19, and others, such as the elderly, or children who may not have been eligible for vaccination at the time.

> "I just didn't want to further spread and infect people or harm people. I felt like you kind of have a personal responsibility to do what you can for the betterment of everyone." (Case)

> "My grandma lives with us. And I work at a restaurant. I deal with a lot of elderly people, and when I got home, I was running a fever. At that point, I just already called off of work just to be safe. But I just didn't want to infect anybody else." (Case)

Some cases expressed concerns with the uncertainty of the disease and their health outcome (getting very sick, intubated, or dying), or missing work and losing wages. A few noted that they had no concerns about their health because they heard the symptoms were mild. Cases who were vaccinated generally expressed less worry about becoming very sick or dying following their positive test result than those who were not vaccinated.

Public sentiment and perceptions of COVID-19, particularly a sense of shame and fear, was a challenge for many cases. For example, one participant reported feeling "demoralized" when her roommates would monitor her to ensure she isolated.

3.2. Reactions to exposure to COVID-19

Most contacts were notified of COVID-19 exposure by a family member, friend, or someone at their workplace or child's school. Few contacts were reached by formal contact tracing professionals, and even fewer learned of their exposure solely from a contact tracer. Many contacts were worried about their personal risk after being notified of exposure. Some experienced hospitalization or loss of a loved one, which made them concerned about their own disease severity. This seemed particularly true among those with underlying health conditions or those who were immunocompromised.

> "I have COPD, I'm on a breathing treatment four times a day and an inhaler two times a day… I was concerned about winding up on a ventilator." (Contact)

Like cases, contacts were also concerned about potentially exposing family members, friends, and coworkers. Contacts were especially concerned if they lived in multigenerational households with elderly family members at increased risk for severe outcomes from COVID-19 or young children who were not eligible for vaccination.

> "Everyone started to get worried because my mother had just gotten out of the hospital. She was in the hospital for gallbladder surgery right at Thanksgiving, so she was the last person that needs to get COVID at this point. We all said, 'OK, we're going to monitor ourselves. We're just going to stay home and see what happens and if we start getting symptoms.'" (Contact)

Some contacts described negative emotions as a result of notification of exposure. One participant, who was on vacation when their spouse tested positive, said they felt "terror" about what would happen next—they worried their spouse would be stuck in the location they were visiting. Another used the word "trauma" to describe what it was like to be exposed while not having the ability to take time off work:

> "That's something else nobody ever talked about, all the other trauma, like from what the kids went through to the psychological stuff we went through to everything else we went through. When people worried about bringing COVID home because I had to go to work." (Contact)

Several contacts felt that their race or ethnicity influenced their reaction to learning of exposure to COVID-19. For example, an Asian person who was a contact expressed concern about the stigma associated with spreading the virus to others:

> "There's not a lot of Asian people [here]. And there was like kind of this stigma, so I didn't want to be sick and spreading it. And kind of contributing to the stereotype that oh, it's an Asian virus or a China virus." (Contact)

Some Black or African American persons who were contacts shared that some members of their community have underlying health conditions that increase their risk for poor outcomes from COVID-19, and as a result, their community continues to follow mitigation strategies:

> "Overall, it's just a more generalized fear within the community about because they said preexisting conditions make you more prone to have fatalities with COVID… I think we're more risk-adverse than other races and that may be one of the reasons why we would continue to wear a mask." (Contact)

3.3. Information sources and educational resources

Cases reported that their test results were accompanied with information about isolation. For many, this information was included with an email or phone call notifying them of the test result. Some also reported being contacted by their doctor's office or a public health professional after their positive result to provide information about isolating, how to manage symptoms, and when to seek medical care. The completeness of information participants received was inconsistent. For example, some received information about managing symptoms without any guidance about isolation or received information about isolation without any guidance about managing symptoms.

Cases who were contacted by case investigators were generally satisfied with their experience. They appreciated calls that provided detailed information and reassurance.

> "They told me the little they knew about it, what I had to do. The lady was very—she cheered me up. She said, 'Nothing will happen. Just stay at home, quarantine the whole family, be up to date with all your symptoms. And if things get worse, go to [the] emergency room, but don't panic. Everything's going to be OK.'"(Case)

Others noted that they were contacted long after they tested positive, when the information was too late to be useful for them. One participant perceived the interaction as "intrusive":

> "I actually was called well after I had tested positive. My [isolation] was almost [over] and I did get a call from contact and trace services… And they did ask for contact information, and I thought it was very intrusive." (Case)

Participants reported trusting the people delivering the information because it came from a medical professional or case investigator.

> "Yes, I did trust what they said because maybe they had more information than we did, because they didn't want us to go out so that we wouldn't get others sick." (Case)

Cases reported receiving information that was detailed, timely, and actionable as the most helpful. For those who were not satisfied with their interaction with public health professionals, several noted that information not different from what they already knew. Some mentioned wanting to better understand the course of the disease, what variant they had, and what symptoms they should expect.

> "I think it was hard, because the information was so—or the virus was so new, and the way people reacted was so new. They didn't really have anything, but I wish they would have told me more about day ten, eleven, twelve, you're going to just get really better really fast. Or more like, 'These are the symptoms you can expect.' Or 'This is when you maybe should come into the hospital.' They really didn't say any of that. It was more kind of, 'Yes. You're sick. Go away. Stay away.'" (Case)

Some Black or African American persons who were cases perceived inequities in information offered to White communities as compared to communities of color, particularly around information and access to monoclonal antibodies, vaccination, and testing.

> "I live in a middle-class neighborhood. It's not like I don't have access. There was nowhere, no signs around to say, hey, go to your local CVS and get this testing done. Or go here to get the vaccine even later on. No matter what, communities of color are always left behind. And so even though it's on the news, communities of color are less vaccinated or less tested. If you walked around, there were not as many signs as there was in the White community right next door where the sign said, Go here, go there and stand there. And it's open all day. That's what made me sad." (Case)

Among contacts who were reached by contact tracers or a health professional, most were given information on quarantining, testing, symptoms, and additional actions to take (e.g., wearing a mask and social distancing). Among those who were not reached by contact tracers or health professionals, most reported being notified of their potential exposure by a family member, friend or someone at their workplace or child's school. Because of this informal notification process, many did not receive detailed information about COVID-19.

Contacts who were notified of their potential exposure by a contact tracer generally described the experience as positive and noted that the information was comprehensive.

> "I thought that was helpful to confirm all that information. And it was also helpful the way she kind of approached it and… making sure I had all my questions answered, and next steps if I needed access to some kind of resource to get me food while I was isolated or something like that, like all those things, were pretty helpful." (Contact)

Other contacts notified by contact tracers expressed dissatisfaction.

> "I think they were trying to just go through the motions of what they're supposed to say and do because of the job, because a lot of information that was given to me, again, wasn't relevant by that point because [my child] was already allowed to go back to school. We had already self-isolated at home. We were washing hands regularly. I was making and preparing food to take to her in her bedroom using only disposable utensils, and she was only coming out to go to the bathroom. After she'd go bathroom, I was cleaning everything off and all of that. All the things they were telling me to do, we had already done." (Contact)

3.4. Reliance on independent information gathering

Both cases and contacts reported seeking information about COVID-19 from family, friends, health care providers, news channels, and websites.

> "My dad actually helped with a lot of the research. I think we mostly just tried to look for the CDC sites, things that would be very reliable information." (Case)

Some indicated that they felt they did not need to seek out additional information because so much information had already been circulating.

> "I didn't get a call until well into feeling really sick. And by that time, I was already in quarantine. I already knew the protocols. It's not so much that I looked up what the protocols were. I think just all of 2020 had been so in your face everywhere, that you already knew what to do." (Case)

Participants acknowledged that some sources could be hard to trust and reported looking for consistency across sources or following what was either the most detailed or strictest guideline.

> "I tried to stay away from more opinion. I took personal anecdotes with a grain of salt … it was a little tricky at first, because there was a lot of back and forth on you should do this or no, now you should do that. So that was a little tricky as well, a little frustrating. I want to say nobody had a good handle on it at a global scale, so it is understandable that maybe there was some variation with the reports that were coming out even from the CDC." (Contact)

Many participants also noted that information and guidelines often seemed confusing.

> "I got the positive result via email and then it was just trying to kind of figure out so what do I do now? Frankly, the CDC guidance was not very clear. If there was some flowchart that was just yes or no and then it directs you to, OK, then do this, it would have been a lot simpler." (Case)

Among Spanish-speaking participants, some noted not receiving sufficient information in their preferred language because resources were only available in English.

> "I understand English, but obviously not 100 percent, the basic only, and when I did the test the people that were there, they only spoke English, and I wanted to ask questions, like when are the results going to be ready and this and that, do I have to come back and get it [or] take another test. I couldn't get those answers back then." (Contact, translated from Spanish)

3.5. Experiences with COVID-19 testing

Cases reported getting tested for COVID-19 because they were experiencing symptoms, knew they had been exposed, or both. Some reported getting tested due to an employer mandate. Access and availability of testing varied over the course of the pandemic. Those who contracted COVID-19 earlier in the pandemic encountered limited availability of tests.

Several contacts expressed concern with the availability and costs of tests.

"I really didn't want to take a test away from somebody that might have needed it more than us. I don't think that they were in great supply. I just didn't see an urgent need for us to take a test when we had no symptoms, and we weren't going anywhere." (Contact)

Many participants reported that the wait time for scheduling testing was long, and it was hard to get tested. Many also explained that because the wait for test results could be long, quarantine was over by the time the results were returned.

"The biggest thing was just trying to find a place where we could get tested, and then when we did find them to find out that for days they were completely booked out. So, that was the most difficult part. But once we got scheduled, it was really fast, really easy." (Contact)

Several contacts exposed during times of surges expressed that it would have been helpful to have access to more home tests.

Other participants noted that they would get tested now because home tests became available for free through an online ordering system for all Americans in January 2022.

"I feel great that we have those tests that the government provided that you could get through the USPS. So I feel like now anybody who wants to get tested, like you just can… it's so different." (Contact)

Many participants reported that the testing center notified them of their results directly through a phone call, email, text, or an electronic health record patient portal. A few were notified by phone calls from a health department or other health agency. Some described conflicting test results or taking multiple tests before testing positive.

"I got tested a couple times, some of them were negative. And then finally we found one that was positive because even though it said we didn't have it, I could tell because the symptoms that we did." (Case)

Many contacts were tested after exposure, with some noting that they waited at least three days to test after exposure.

"I was told not to take the test too early, that if you take the test too early, you'll end up with a false negative, especially with the home test. So if you think you've been exposed, you should really wait three to five days." (Contact)

Most contacts got tested to confirm that they did not have COVID-19, with several noting that they were required by their employer to get tested after exposure.

"I was notified through my supervisor at work that a close colleague that I'd worked with had tested positive. This was 2020, so it was like a similar two-week quarantine and then I had to do a negative test to return to the office." (Contact)

Some contacts decided not to test because they did not have any symptoms. In many instances, they also did not have quick and easy access to tests.

"I just figured if I wasn't feeling sick or anything like that, within the five days, I figured I was OK." (Contact)

One contact reported not getting tested because they were vaccinated.

"Since I was already vaccinated, my daughter asked me if I wanted to take the COVID test, and I told her no. I told her that I was already vaccinated with two doses." (Contact)

Almost all contacts who did not get tested said they would have done so if they had symptoms.

"If I would've started feeling any of the symptoms, then I would've done it. I just didn't see a point in waiting in a line or wasting time." (Contact)

Some Latino persons who were contacts shared that they were less likely to seek health care due to their upbringing unless their conditions were severe. These contacts were reluctant to seek testing, especially if it was not affordable.

"My family is pretty hardheaded when it comes to sickness or disease, especially myself. I can't remember the last time I've been to the doctor. Unless like a body part's falling off, I won't go. We're taught to work hard and work through anything that's trying to stop you." (Contact)

3.6. Experiences with isolation

Most cases reported isolating, with duration ranging from 2 days to 2 months, with 10 days reported most. Primary factors influencing the duration of isolation included length and severity of symptoms, recommendations by public health officials or providers, and employer requirements. Most cases began isolation after their positive test; a few began upon the onset of symptoms before they tested positive. Most noted that they would follow the same isolation behavior if they tested positive for COVID-19 again. Among those who tested positive earlier in the pandemic, many noted that they would follow updated guidance and only isolate for 5 days instead of 10. A few participants who identified as politically conservative noted that they would not take isolation as seriously, would live normally, or would treat a new COVID-19 infection like the common cold.

3.7. Motivators and facilitators of isolation

Many cases reported isolating because they did not want to infect others, especially elderly or immunocompromised people, and potentially cause hospitalization or death. Several cited the social responsibility to protect others from COVID-19 as a motivator to isolate.

"I feel like we have a social responsibility to take care of others because even though you may not have gotten COVID as extreme, you don't know that the person next to you could possibly get… This whole pandemic wasn't just about yourself and what you wanted to do, it's about how do we take care of each other during this crisis that no one wanted or predicted." (Case)

Several cases reported isolating for more pragmatic reasons, reporting they felt too sick to go anywhere or because their jobs required isolation.

"I didn't even want to get out of bed let alone go anywhere. And then it was more like another two to three weeks just to get my strength back." (Case)

Having a place to isolate was an important facilitator of isolation for many cases. Some were able to stay in their own room, and others already lived alone. Many cited having necessary supplies at home as a crucial factor in not having to leave home. In addition, for many, others within the household were able to shop and bring food and resources to the rooms in which cases were isolating.

"I was pretty much segregated to my room and the nearest bathroom, my sister would—she'd bring me food… there was a cooler just outside my room that she'd put waters and drinks in and other foods for me." (Case)

Many cases also had family, friends, and neighbors outside the household who could drop off supplies, or access to delivery services.

"It was nice to have those little services that will come drop the food at your door, and then, you don't have to worry about going out or giving anybody—contracting everybody." (Case)

Participants reported that it was helpful to have social support during the isolation period through virtual check-ins with family and friends or

through their faith community. Feelings of loneliness or despair while in isolation were reported, and participants described the mental health toll of being away from family and friends.

> "Not being able to visit my mother who was in a nursing home at the time, and she also had COVID. And then she was hospitalized, and I couldn't go visit her, so that was the hardest part for me. But then one day I tested negative, and I was able to do that, and then she passed away. It was hard." (Case)

Supportive work policies aided isolation. A few cases reported that their colleagues were understanding and provided coverage at work. Many were able to work from home or take paid time off or leave.

> "It was easier for me to recover being that I was able to work from home and I had the option to quarantine. I think that it could be really detrimental to somebody who has a job that requires in-person and manual labor." (Case)

> "My work has a 14-day policy where I have to stay home, so that kind of made me stay home." (Case)

Conversely, others reported unsupportive work situations.

> "I lost pay from work, because by the time I contracted the virus our company stopped enforcing the COVID relief policy… it was about a month after the government stopped enforcing it for the companies as well. So, I lost out on the money on that." (Case)

3.8. Challenges of isolation

Several challenges impeded the ability to isolate. Some cases reported leaving isolation to get groceries and other supplies, but most tried to stay away from others through curbside pickup or drive-through services. Some, particularly those who reported having to care for dependents or elderly parents, found it difficult to isolate or chose not to isolate from others in their household. Parents noted that it was especially challenging to isolate with young children.

> "I have a special needs nephew… I couldn't stay in my room because I'm the only one that cares for him and my grandson." (Case)

Living arrangements (e.g., space size, not having separate bedrooms or bathrooms) also proved to be a barrier for some.

> "We lived in a two-bedroom apartment. I didn't have my own bedroom, my own room, so I try to limit as much exposure to my grandma as possible." (Case)

Additionally, financial constraints—including the cost of delivery services, lost wages from time off from work, and the lack of paid leave—made isolation challenging for some.

> "I got beat up by my doctor. She said, 'Why did you get on a plane?' [I said] 'You going to pay my hotel fee?' Because I wasn't at home. I wasn't in a place where I could stay… I needed to get back home. I didn't have finances to stay where I was for a whole other fourteen days." (Case)

Many cases described modifications to isolation that included leaving their home but still staying away from others, like taking walks, walking a pet, hiking, going for drives, and using drive-through and curbside pickup services to get fresh air and tackle boredom.

3.9. Experiences with quarantine

Most contacts reported staying home after learning they were exposed, but the length of time varied. Some noted staying home only until receiving a negative test result, which they used to justify ending quarantine. When asked what they would do about quarantining if notified of a new exposure, many said they would likely do the same unless they learned of updated guidance.

3.10. Motivators and facilitators of quarantine

When asked why they quarantined, almost all contacts expressed wanting to be careful and not infect others.

> "I didn't want to spread it. I didn't want to go into an office space or anywhere, and then find out that I had passed it on to someone not knowing everybody's underlying health conditions." (Contact)

> "I didn't want people to go through the situation that my friend was going through. Maybe because of your stubbornness or your negligence someone might have gotten sick. What I didn't want was to be a risk factor for other people." (Contact)

Among contacts who quarantined, many noted that they never had to leave home. These participants discussed their reliance on family and friends or delivery services to get food and other necessities.

> "I do a lot of home delivery stuff anyway. It was just convenient, just put in the food order and have it delivered at a certain time or whatever." (Contact)

Supportive work policies aided in the ability to quarantine for some, like the experience of cases.

> "One thing is it's part of our company's policy to be quarantined as soon as you know you've been exposed. If you have symptoms and you test positive then yes, you follow the quarantine, whatever period that was." (Contacts)

A popular sentiment among contacts was that it was not difficult to stay home because it is what they were doing anyway.

> "Well, for me it was easy to stay home because I'm more of a homebody. All I need is a good book, turn on the TV for a distraction, or background noise and I'm good." (Contact)

Some American Indian or Alaska Native persons who were contacts noted that support they received from their tribal community helped them practice mitigation strategies and quarantine.

> "I'm part of a Native American tribe… and in our community we have a lot of information that we pass around, and we help each other out. We were given stuff for disinfecting and cleaning and masks, and anything that we would need, we were provided from our tribe. And during quarantine, they helped us a lot too." (Contact)

3.11. Challenges with quarantine

Several contacts said they left quarantine at least once to do routine tasks or run errands, including grocery shopping, curbside pickup, paying bills, or walking the dog.

> "[I left quarantine] to buy essentials like water and juice and fresh milk and that kind of thing." (Contact)

Like cases, contacts described trying to be safe around others when leaving the house.

> "I [did] go to the grocery store. I stayed ten feet [away] and then fully masked." (Contact)

Other reasons for leaving quarantine included getting tested or vaccinated.

> "I left within the first couple of days to get the COVID test and then I stayed at home for a week." (Contact)

Many contacts expressed loneliness during quarantine. Others noted that it was most challenging to cancel plans that were already made or not see family, especially if they had a relative who was ill.

> "For me it was difficult because, during that time, my dad had passed away. He passed away in Mexico when I was in quarantine. I couldn't go. That's something that, personally, just was very difficult for me because I couldn't say goodbye to my dad." (Contact)

Contacts who identified as essential employees, particularly those in the health care field, described experiencing multiple exposures. Most of these contacts reported having employers who required a negative test result before they could return to work. Some expressed pressure to continue working regardless of exposure.

> "I had to go to work first time because work wouldn't give me the time off. I was extra careful to keep more than the six feet." (Contact)

Contacts also commented on the impact of changing guidance. In the beginning, employers were more likely to make them quarantine until they had a negative test result. As the pandemic continued, some reported their employers recommended quarantine based on the nature of the exposure, if the contact was wearing a mask, and vaccination status.

Finally, a few contacts did not think their exposure was serious enough to modify their behaviors, especially if they did not have symptoms.

> "Even when I got the notification, I felt fine. No fever, no headache. No loss of smell. No nothing. So I just kept doing my regular thing." (Contact)

3.12. Identification and notification of contacts

Cases reported limited interaction with case investigators from public health departments. Although some were called, some reported that the focus of the call was related to symptoms, testing, and isolation, and not case investigation. Some were asked if they notified their close contacts and were reminded to do so. Few were asked for close contacts' information or were offered help with notifying close contacts. Despite not being asked, some volunteered names and exposure information to case investigators.

Some cases felt comfortable providing contact information to case investigators. A few referred to the case investigation process as an easy and positive experience. Others offered locations that they visited because they did not know contacts from public places.

Among cases who were not contacted by case investigators, many said that they would have provided close contact information if they had been contacted.

> "I have no problem sharing this information. I feel like people have concerns about privacy and things like that. But when you're dealing with a situation like this, you have to put a lot of the stuff that we normally would have concerns about to the side and just do the right thing." (Case)

Cases who did not or would not feel comfortable providing close contact information cited privacy concerns.

> "I would feel like that would be a breach of privacy. Because if someone just came to you and said, 'Hey… a close person you worked with tested positive,' without you being informed first, wouldn't you be wondering, 'How did [that person] get that information?'" (Case)

Several cases did not provide information about contacts because they preferred notifying their close contacts. A few thought it would be more personal to do their own outreach:

> "Getting a call from a county official for contact tracing just seems a little bit more intense than me being, 'Hey, this happened'… I felt guilty so I wanted to be like 'I'm sorry for exposing you and I felt really bad about it.'" (Case)

Cases largely felt comfortable notifying their contacts. Most did so out of concern for others, to prevent transmission, and to allow contacts to quickly take appropriate action. Many reported that their contacts were understanding, sympathetic, and appreciative after they were notified. Some reported that when they notified contacts, the contacts were not concerned about disease severity, especially among participants who identified as having conservative political ideology.

4. Discussion

In this study, we examined experiences and behaviors of COVID-19 cases and contacts. We sought to understand motivations, facilitators, and barriers to isolation, quarantine, testing, and notifying contacts to enhance further guidance and messaging to improve adoption of public health prevention measures.

4.1. Motivators and facilitators for isolation and quarantine

Participants were motivated to follow public health guidance around isolation and quarantine for their own health and to avoid spreading COVID-19 to others. Messaging that emphasizes the importance of protecting others may be an important motivator, consistent with recommendations from other studies on effective messaging about COVID-19 (Bokemper et al., 2022; Luttrell & Petty, 2021). Supportive work policies, social support from others, and linkage to services that provided food and other necessities were important facilitators for isolating and quarantining. It is critical for public health efforts to incorporate assessment and referrals for social support services (Lash et al., 2021).

4.2. Challenges staying away from others within the household

Although people largely were able to stay away from others outside the household, for some it was challenging to stay away from others within the household. This was especially difficult for parents and caregivers, as well as spouses and partners. Given the evidence of substantial household transmission of COVID-19 (Baker et al., 2022; Lewis et al., 2021), this is an area to emphasize for future messaging. If isolation within the household is not feasible, it is important for public health measures to provide alternative options and to encourage other mitigation measures including mask use, being outside when possible, increasing ventilation, and social distancing (CDC, 2022a).

4.3. Financial implications for isolation and quarantine

The ability to continue work was an important factor influencing isolation and quarantine. Participants who were able to work remotely while in isolation or quarantine reported few inconveniences and concern about income disruption. Participants required to be in-person for their jobs or who did not have adequate sick leave were concerned about loss of income and job security. Some participants reported being directed back to work by their employer before they were able to complete the recommended isolation or quarantine period. Public health efforts that provide supportive sick leave policies and other health programs may promote adherence to quarantine and isolation. For example, in the U.S., states that gained access to two weeks of paid sick leave for workers through the Families First Coronavirus Response Act (FFCRA) reported about 400 fewer confirmed cases per state per day (Pichler et al., 2020).

4.4. Harm reduction strategies during isolation and quarantine

Although participants largely understood the importance of staying away from others, many reported modifications to isolation or

quarantine that included leaving the household while wearing a mask and maintaining distance from others. Public health officials could acknowledge and recommend potential harm reduction strategies that minimize risk of transmission. Guidance on safer activities that people may do if they must modify isolation and quarantine guidance may be helpful. For example, emphasizing behaviors that are done alone, outside, while practicing social distance, or with mask use.

4.5. Testing for COVID-19

Participants reported varying degrees of access to COVID-19 testing throughout the pandemic. During times of limited availability, many participants expressed frustration with the inability to obtain tests and expressed gratitude during times of greater accessibility. Many participants reported willingness and intentions to use home tests if they thought they had an exposure or symptoms of COVID-19. It is important to ensure continued access to at-home tests as the COVID-19 pandemic continues to facilitate rapid isolation.

There was some confusion about the optimal time to test and which tests were preferred. There was a mix of contacts who sought testing after their exposure and others who opted not to test unless they developed COVID-19 symptoms. Given that asymptomatic COVID-19 cases occur, it is essential that people exposed to COVID-19 get tested and rapidly notify their close contacts if they receive a positive result. Public health messaging that is clear about changes to testing guidance, the efficacy of tests, and incubation periods for the virus variants that cause COVID-19 can help to ensure that people are testing at appropriate time intervals and continuing testing after exposure, even if they do not have symptoms.

Participants who received information about next steps after their test result felt it was helpful, although the amount and timing of information varied. It is important to have actionable instructions, including how to connect with a public health professional if needed, and educational resources, timely and consistently accompanying COVID-19 test results.

4.6. Health department interactions

Very few participants reported being contacted by health department officials. Those who were, expressed that they wanted more information and wanted to be contacted sooner, as they reported being contacted long after their test or exposure. This may be due to the difficulty of conducting CI/CT during times of high transmission (Lash et al., 2021). However, these participants largely reported a positive experience and felt that the information received complemented the information they had already learned. This suggests support for CI/CT and underscores the importance of reaching cases and contacts in a timely manner. Although CDC has moved away from CI/CT for everyone, CI/CT is still recommended for priority populations and settings and should be done as timely and thoroughly as possible (CDC, 2022b).

4.7. Identification and notification of contacts

Cases generally understood the importance of informing their close contacts of potential exposure to COVID-19. Although few cases reported formal interaction with CI/CT officials, those who did seemed willing to share their contacts. A few expressed hesitancy with providing names of their contacts, citing mistrust of government officials or perceived invasion of privacy. Overall, participants felt more comfortable notifying their own contacts. Reducing the time to notify contacts can significantly reduce COVID-19 transmission by facilitating rapid quarantine (Jeon et al., 2022). However, cases who choose to personally notify contacts may not do so quickly or comprehensively, especially if they do not fully understand the breadth of their potential contacts. It is recommended that public health messaging include clear information on who is considered a close contact, and public health agencies could develop resources to assist individuals to promptly inform their own contacts.

4.8. Risk perception and health information sources

COVID-19 vaccination status played an important role in participants' perception of potential disease severity. Although many still took a COVID-19 diagnosis seriously, they expressed feeling safer knowing they had been vaccinated. It is recommended that public health officials continue to encourage vaccination and vaccine booster efforts to ensure people are up to date with COVID-19 vaccinations.

Given that most participants did not have formal interaction with CI/CT officials, guidance and communication from trusted public health authorities are continuously needed. Public health agencies should prioritize timely and consistent communication of actionable health messages, following best practices in crisis and emergency risk communication (Centers for Disease Control and Prevention (CDC), 2018). These messages are amplified by the media, other health professionals, and the public to create norms around expected behavior (Kasperson et al., 1988). Participants reported sources of information that included medical professionals, the media, friends, family, the workplace, peers, and websites. Consistent creative methods are needed to disseminate public health information about COVID-19.

4.9. Political ideology and behavior

Previous research has found that political ideology can influence perceptions and actions related to CI/CT for COVID-19 (McClain and Rainie, 2020). However, we discovered that individuals of different political ideologies generally reported similar actions following a positive COVID-19 test result or exposure. An exception is that a few individuals who identified as conservative reported they did not consider COVID-19 a serious disease and would be less likely to take prevention measures if they test positive or were exposed in the future, although that sentiment was not widely expressed. Public health messaging that reaches individuals across the political spectrum to ensure that guidance and information encourage appropriate actions is recommended.

4.10. Health equity considerations

Health equity concerns were raised by participants across groups and topics. It is important to consider equity for the study of the reach of CI/CT services. People from some racial and ethnic groups experienced disproportionately high rates of hospitalization and mortality from COVID-19, stressing the need for CI/CT to be executed in ways that are appropriate to these communities (CDC, 2022b). For instance, participants reported lack of availability of information in their preferred language and observed disparities in availability of testing and vaccination. Some Asian persons who were participants experienced and/or feared racially-based stigma and discrimination due to the COVID-19 pandemic (Dhanani & Franz, 2020). This may add to challenges to health seeking behavior, trust in government, and public health communication for populations who may be experiencing discrimination or racism as the result of stigma associated with COVID-19. Several Black or African American persons who were participants expressed perceived institutional and structural racism about inequitable access to information, testing, and treatments, and the perception of the need to make extra efforts to ensure that they obtained services and support that they needed. These findings add to the growing body of literature describing the need for a more equitable approach to the COVID-19 response (Romano et al., 2021; Wang et al., 2020).

4.11. Limitations and conclusion

This assessment has limitations, including a nonrandomly drawn sample and self-selection bias. Individuals who chose to complete a survey and participate in a focus group or interview about COVID-19 may have greater interest and/or knowledge of the topic than the general public. Although great effort was made to ensure inclusion of a diverse

set of participants, the final sample was not representative of the U.S. population. The FGDs were only conducted in English and Spanish and therefore exclude the perspectives of non-English or non-Spanish speaking individuals. Social desirability may have influenced responses, as individuals may have over- or under-reported their behaviors based on the opinions shared by others in the focus group, as well as recall bias and acquiescence bias. Differences in actions taken by participants may also be attributed to evolving guidelines by CDC and other public health authorities. We explored differences by race, ethnicity, education, and political ideology and report on relevant findings. We did not explore differences by gender, age, region, or urbanicity, which would be interesting to assess in future research.

Our assessment highlighted specific motivators, facilitators, and barriers to the adoption of COVID-19 mitigation measures. Effective communications require reminding individuals that COVID-19 is highly contagious, can cause serious illness, and may be deadly for some individuals. People can protect themselves and their loved ones by following isolation and quarantine guidance, getting tested, and promptly notifying contacts. These findings contribute to the body of evidence describing CI/CT for COVID-19. To our knowledge, this is the first study of such a large sample in qualitative research to explore behaviors regarding COVID-19 mitigation measures. This work goes beyond describing what people did and provides context as to why people did or did not adhere to public health guidance. Future public health messages, services, programs, and interventions for COVID-19 need to address the barriers, facilitators, and motivators identified in this research to enhance adoption and adherence to public health guidance. Lessons from this research highlight the importance of a robust public health infrastructure and increased capacity to effectively respond to future public health emergencies.

Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

The work was supported by the Centers for Disease Control and Prevention under contract number 75D30121C11647.

References

Baker, J. M., Nakayama, J. Y., O'Hegarty, M., McGowan, A., Teran, R. A., Bart, S. M., … Tate, J. E. (2022). SARS-CoV-2 B. 1.1. 529 (Omicron) variant transmission within households—four US jurisdictions. *Morbidity and Mortality Weekly Report, 71*(9), 341. November 2021–February 2022.

Bokemper, S. E., Huber, G. A., James, E. K., Gerber, A. S., & Omer, S. B. (2022). Testing persuasive messaging to encourage COVID-19 risk reduction. *PLoS One, 17*(3), Article e0264782.

CDC. (2020). Science brief: Options to reduce quarantine for contacts of persons with SARS-COV-2 infection using symptom monitoring and diagnostic testing. Retrieved from https://public4.pagefreezer.com/browse/CDC%20Covid%20Pages/11-05-2022T12:30/https:/www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/scientific-brief-options-to-reduce-quarantine.html. (Accessed 8 July 2022).

CDC. (2021). *CDC updates and shortens recommended isolation and quarantine period for general population*. Retrieved from https://www.cdc.gov/media/releases/2021/s1227-isolation-quarantine-guidance.html. (Accessed 6 June 2022).

CDC. (2022a). *Case investigation and contact tracing: Part of a multipronged approach to fight the COVID-19 pandemic*. Retrieved from https://www.cdc.gov/coronavirus/2019-ncov/php/principles-contact-tracing.html. (Accessed 1 June 2022).

CDC. (2022b). *Contact tracing for COVID-19*. Retrieved from https://www.cdc.gov/coronavirus/2019-ncov/php/contact-tracing/contact-tracing-plan/contact-tracing.html. (Accessed 11 July 2022).

Centers for Disease Control and Prevention (CDC). (2018). Crisis and emergency risk communication (CERC). Retrieved from https://emergency.cdc.gov/cerc/index.asp. (Accessed 8 July 2022).

Dhanani, L. Y., & Franz, B. (2020). Unexpected public health consequences of the COVID-19 pandemic: A national survey examining anti-Asian attitudes in the USA. *International Journal of Public Health, 65*(6), 747–754.

Gale, N. K., Heath, G., Cameron, E., Rashid, S., & Redwood, S. (2013). Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Medical Research Methodology, 13*(1), 1–8.

Ipsos. (2022). *KnowledgePanel—a methodological overview*. Retrieved from https://www.ipsos.com/sites/default/files/ipsosknowledgepanelmethodology.pdf. (Accessed 7 July 2022).

Jeon, S., Rainisch, G., Lash, R. R., Moonan, P. K., Oeltmann, J. E., Greening, B., … Meltzer, M. I. (2022). Estimates of cases and hospitalizations averted by COVID-19 case investigation and contact tracing in 14 health jurisdictions in the United States. *Journal of Public Health Management and Practice, 28*(1), 16–24.

Kasperson, R. E., Renn, O., Slovic, P., Brown, H. S., Emel, J., Goble, R., … Ratick, S. (1988). The social amplification of risk: A conceptual framework. *Risk Analysis, 8*(2), 177–187.

Lash, R. R., Moonan, P. K., Byers, B. L., Bonacci, R. A., Bonner, K. E., Donahue, M., … Weisbeck, A. (2021). COVID-19 case investigation and contact tracing in the US, 2020. *JAMA Network Open, 4*(6), Article e2115850. e2115850.

Lewis, N. M., Chu, V. T., Ye, D., Conners, E. E., Gharpure, R., Laws, R. L., … Kirking, H. L. (2021). Household transmission of severe acute respiratory syndrome coronavirus-2 in the United States. *Clinical Infectious Diseases, 73*(7), e1805–e1813.

Luttrell, A., & Petty, R. E. (2021). Evaluations of self-focused versus other-focused arguments for social distancing: An extension of moral matching effects. *Social Psychological and Personality Science, 12*(6), 946–954.

Mathematica. COVID-19 After Action Review Toolkit. https://www.mathematica.org/features/covid-19-after-action-review-toolkit. Accessed July 6, 2022.

McClain, C., & Rainie, L. (2020). The challenges of contact tracing as U.S. battles COVID-19. Retrieved from https://www.pewresearch.org/internet/2020/10/30/the-challenges-of-contact-tracing-as-u-s-battles-covid-19/. (Accessed 1 June 2022).

Oeltmann, J.E., Vohra, D., Matulewicz, H.H., DeLuca, N., Smith, J.P., Couzens, C., ... Moonan, P.K. *working paper*.

Pichler, S., Wen, K., & Ziebarth, N. R. (2020). COVID-19 emergency sick leave has helped flatten the curve in the United States: Study examines the impact of emergency sick leave on the spread of COVID-19. *Health Affairs, 39*(12), 2197–2204.

Rader, B., Gertz, A., Iuliano, A. D., Gilmer, M., Wronski, L., Astley, C. M., … Brownstein, J. S. (2022). Use of at-home COVID-19 tests—United States, August 23, 2021–March 12, 2022. *Morbidity and Mortality Weekly Report, 71*(13), 489.

Romano, S. D., Blackstock, A. J., Taylor, E. V., Felix, S. E. B., Adjei, S., Singleton, C. M., … Boehmer, T. K. (2021). Trends in racial and ethnic disparities in COVID-19 hospitalizations, by region—United States, March–December 2020. *Morbidity and Mortality Weekly Report, 70*(15), 560.

Wang, M. L., Behrman, P., Dulin, A., Baskin, M. L., Buscemi, J., Alcaraz, K. I., … Fitzgibbon, M. (2020). Addressing inequities in COVID-19 morbidity and mortality: Research and policy recommendations. *Translational Behavioral Medicine, 10*(3), 516–519.

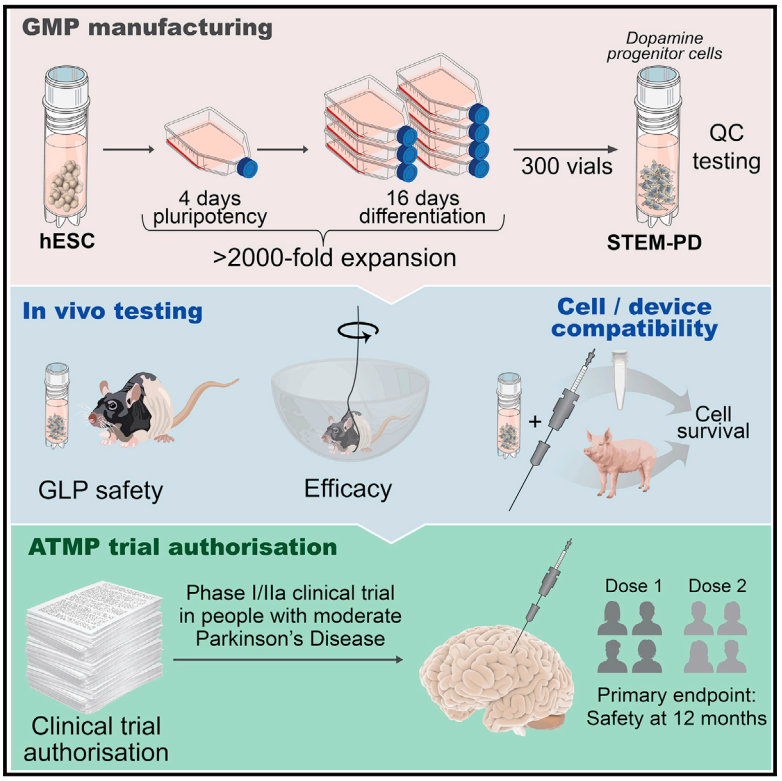
World Health Organization (WHO). (2021). Operational guide for engaging communities in contact tracing. Retrieved from https://www.who.int/publications/i/item/WHO-2019-nCoV-Contact_tracing-Community_engagement-2021.1-eng. (Accessed 1 June 2022).

Cell Stem Cell

Clinical and Translational Report

Preclinical quality, safety, and efficacy of a human embryonic stem cell-derived product for the treatment of Parkinson's disease, STEM-PD

Graphical abstract



Highlights

- GMP manufacturing of cryopreserved dopamine progenitor cell product STEM-PD from hESCs
- STEM-PD shows no adverse effects or tumor formation in rats up to 9 months
- STEM-PD grafts innervate correct brain regions and reverse motor deficits in rat PD model
- A first-in-human ATMP trial with 8 patients is approved and initiated in Sweden

Authors

Agnete Kirkeby, Jenny Nelander, Deirdre B. Hoban, ..., Gesine Paul, Roger A. Barker, Malin Parmar

Correspondence

agnete.kirkeby@med.lu.se (A.K.), malin.parmar@med.lu.se (M.P.)

In brief

Kirkeby et al. report in this issue on the manufacturing and preclinical testing of a pluripotent stem cell product for replacement of lost dopamine neurons in the brains of Parkinson's disease patients. A clinical trial using this stem cell transplantation product has been approved and initiated within the EU.

Kirkeby et al., 2023, Cell Stem Cell *30*, 1299–1314
October 5, 2023
https://doi.org/10.1016/j.stem.2023.08.014

CellPress

Cell Stem Cell

CellPress
OPEN ACCESS

Clinical and Translational Report

Preclinical quality, safety, and efficacy of a human embryonic stem cell-derived product for the treatment of Parkinson's disease, STEM-PD

Agnete Kirkeby,[1,2,*] Jenny Nelander,[3] Deirdre B. Hoban,[3] Nina Rogelius,[3] Hjálmar Bjartmarz,[4] Novo Nordisk Cell Therapy R&D,[6] Petter Storm,[3] Alessandro Fiorenzano,[3] Andrew F. Adler,[3] Shelby Vale,[3] Janitha Mudannayake,[3] Yu Zhang,[1,3] Tiago Cardoso,[3] Bengt Mattsson,[3] Anne M. Landau,[7] Andreas N. Glud,[8] Jens C. Sørensen,[8] Thea P. Lillethorup,[7] Mark Lowdell,[9] Carla Carvalho,[9] Owen Bain,[9] Trinette van Vliet,[5] Olle Lindvall,[10] Anders Björklund,[3] Bronwen Harry,[11] Emma Cutting,[11] Håkan Widner,[5] Gesine Paul,[5,12] Roger A. Barker,[11,13] and Malin Parmar[3,14,*]

[1]Wallenberg Neuroscience Center, Wallenberg Center for Molecular Medicine and Lund Stem Cell Center, Department of Experimental Medical Science, Lund University, 221 84 Lund, Sweden
[2]Novo Nordisk Foundation Center for Stem Cell Medicine (reNEW) and Department of Neuroscience, University of Copenhagen, 2200 Copenhagen, Denmark
[3]Wallenberg Neuroscience Center, MultiPark and Lund Stem Cell Center, Department of Experimental Medical Science, Lund University, 221 84 Lund, Sweden
[4]Department of Neurosurgery, Skåne University Hospital, 221 85 Lund, Sweden
[5]Department of Neurology, Skåne University Hospital, 221 85 Lund, Sweden
[6]Cell Therapy R&D, Novo Nordisk A/S, 2760 Måløv, Denmark
[7]Department of Nuclear Medicine & PET-Center and Translational Neuropsychiatry Unit, Department of Clinical Medicine, Aarhus University, 8200 Aarhus N, Denmark
[8]Center for Experimental Neuroscience (CENSE), Department of Neurosurgery, Department of Clinical Medicine, Aarhus University Hospital, 8200 Aarhus N, Denmark
[9]Centre for Cell, Gene and Tissue Therapeutics, Royal Free NHS Foundation Trust, Royal Free Hospital, London NW3 2QG, UK
[10]Lund Stem Cell Center and Department of Clinical Sciences Lund, Lund University, 221 84 Lund, Sweden
[11]Department of Clinical Neurosciences, University of Cambridge, Cambridge CB2 0PY, UK
[12]Wallenberg Neuroscience Center, Wallenberg Center for Molecular Medicine, Department of Clinical Sciences, Lund University, 221 84 Lund, Sweden
[13]Wellcome-MRC Cambridge Stem Cell Institute, Cambridge CB2 0AW, UK
[14]Lead contact
*Correspondence: agnete.kirkeby@med.lu.se (A.K.), malin.parmar@med.lu.se (M.P.)
https://doi.org/10.1016/j.stem.2023.08.014

SUMMARY

Cell replacement therapies for Parkinson's disease (PD) based on transplantation of pluripotent stem cell-derived dopaminergic neurons are now entering clinical trials. Here, we present quality, safety, and efficacy data supporting the first-in-human STEM-PD phase I/IIa clinical trial along with the trial design. The STEM-PD product was manufactured under GMP and quality tested *in vitro* and *in vivo* to meet regulatory requirements. Importantly, no adverse effects were observed upon testing of the product in a 39-week rat GLP safety study for toxicity, tumorigenicity, and biodistribution, and a non-GLP efficacy study confirmed that the transplanted cells mediated full functional recovery in a pre-clinical rat model of PD. We further observed highly comparable efficacy results between two different GMP batches, verifying that the product can be serially manufactured. A fully in vivo-tested batch of STEM-PD is now being used in a clinical trial of 8 patients with moderate PD, initiated in 2022.

INTRODUCTION

Parkinson's disease (PD) is a common neurodegenerative disorder characterized by the loss of midbrain dopamine (DA) neurons, predominantly of the A9 subtype located in the substantia nigra pars compacta (SNc). These neurons send projections to the forebrain, where DA is released and functions to modulate key aspects of movement control and some aspects of cognition.[1] In PD, the loss of dopaminergic (DAergic) innervation to the striatum, especially the putamen, leads to motor impairments such as bradykinesia and rigidity.[2] There are currently no disease-modifying treatments, and medical management is mainly focused on controlling the motor symptoms using drugs that act on the DAergic system, such as levodopa (L-DOPA) or

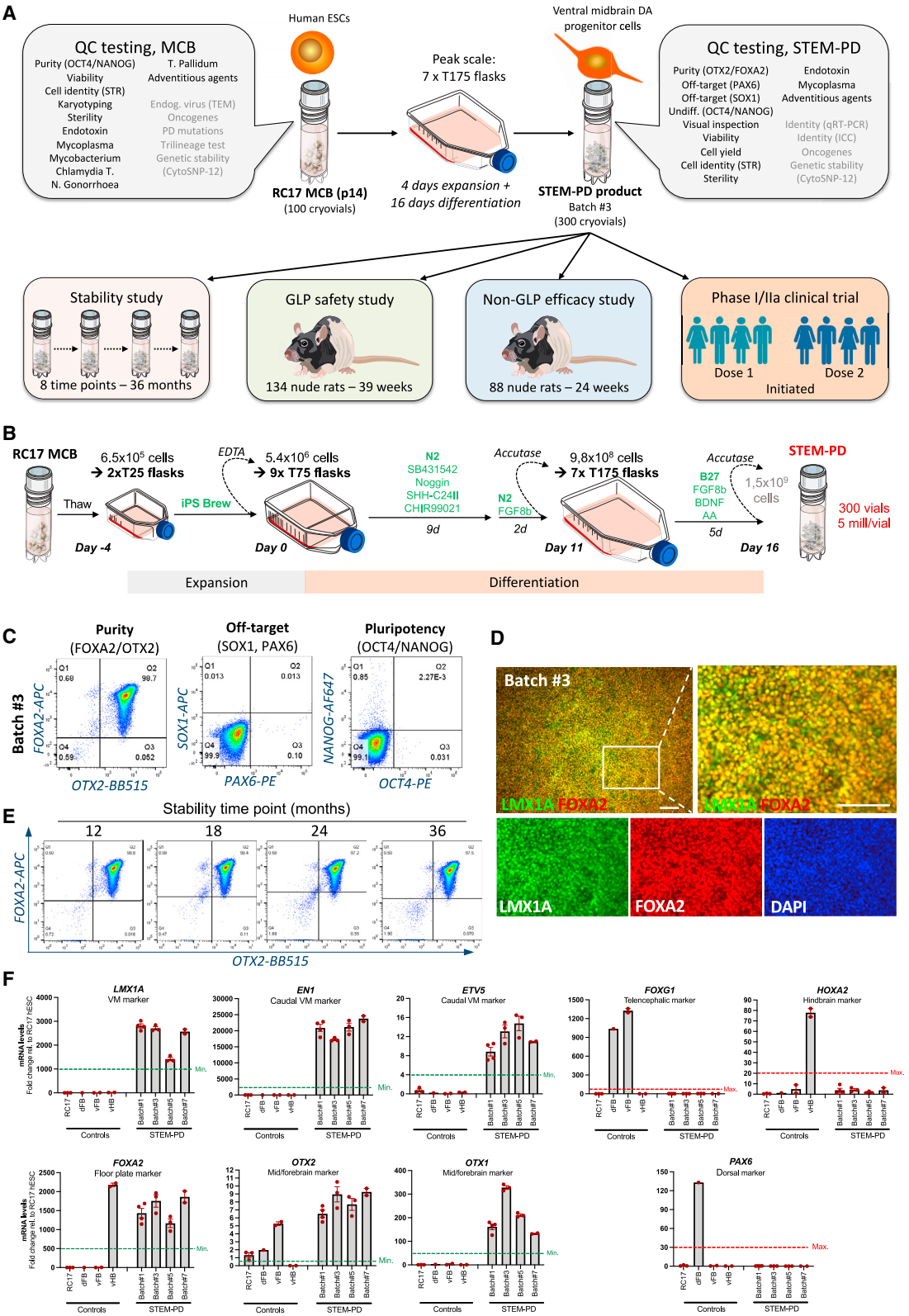


Figure 1. Manufacturing and quality control of STEM-PD

(A) Overview of manufacturing and testing of RC17 MCB and STEM-PD product. Quality control (QC) tests shown in black font were defined as release criteria and conducted under GMP according to European Pharmacopoeia (E.P.) standards, where applicable. Tests listed in gray font were conducted for information only.

(legend continued on next page)

DA agonists. These drugs provide significant clinical benefit in early disease stages, but with time, they lose efficacy and give rise to adverse effects including neuropsychiatric complications, L-DOPA-induced dyskinesias (LIDs), and exacerbation of non-motor aspects of PD such as postural hypotension.[3]

Based on a large number of preclinical studies in animal models of PD, DA cell replacement therapies are being pursued as a potential attractive alternative to oral DA medications.[4] The concept behind this is that physiological and targeted release and reuptake of DA provided by implanted DA neurons would lead to restoration of motor function, while avoiding the loss of efficacy and prominent side effects associated with long-term oral DA medications. Previous studies using DA neurons sourced from the ventral midbrain (VM) of aborted fetuses have demonstrated that intracerebrally grafted human fetal-derived DA neurons can survive long-term in the Parkinsonian brain with physiological release of DA and symptomatic benefit for decades.[5–11] However, the results for patients in receipt of fetal VM transplants have been highly variable, and two double blind placebo-controlled trials failed to reach their primary endpoints, while side effects in the form of graft-induced dyskinesias (GIDs) were observed in some patients.[7,12]

The reasons for these variable trial outcomes have been extensively discussed elsewhere,[13] and led us to undertake another transplant trial using human fetal VM tissue—TRANSEURO (NCT01898390). The TRANSEURO trial initiated in 2012 was designed with the aim to reduce the variation in outcome seen in the previous trials and included more stringent patient selection, cell preparation, handling and dosing, surgical technique, and immune suppression.[14] However, major logistical problems relating to tissue availability meant that recruitment to the TRANSEURO trial had to be terminated before the originally planned number of patients could be grafted.[14] Transplantation of human fetal tissue is further complicated by the fact that each patient receives a different composition and quantity of cells determined by the number and gestational age of the donated tissue, and there is limited possibility to perform quality testing of the cells prior to grafting.

Developing a DA cell replacement product that can be standardized, tested, and manufactured at a large scale will not only allow for adequate control of cell quality and dose but will also enable the therapy to be available to a globally impactful number of patients if successful. We, and others, have therefore focused our efforts toward the derivation of human midbrain DA neurons from pluripotent stem cells through directed differentiation.[15–18] In preclinical studies, we have shown that such differentiated human embryonic stem cell (hESC)-derived DA neurons can fully reverse motor impairments in experimental models of PD and that the neurons function en par with human fetal DA neurons.[19] Efforts to adapt such protocols to clinical standards[20–22] have already led to initiation of clinical trials in Japan and the US[23,24] as well as a single-case medical intervention.[25] Our team has now initiated a first-in-human clinical trial in people with moderate PD in Europe using the recently developed hESC-derived DAergic cell product STEM-PD.[26]

The STEM-PD product was generated using a good manufacturing practice (GMP)-compliant protocol for DA progenitor cell differentiation with scalable manufacturing and includes a method for cryopreservation of the progenitors to facilitate reproducibility, increase safety, and enable shipment of cells between centers.[27] This has enabled us to perform extensive quality control and *in vivo* assessment of the safety and efficacy of the exact same batch of cells that is now being delivered to patients in the STEM-PD clinical trial (EudraCT 2021-001366-38, NCT05635409). This trial is a single arm, first-in-human, phase I/IIa multicenter dose escalation trial to assess the safety and tolerability of hESC-derived DA progenitor cells grafted into the putamen of patients with moderate PD, including secondary efficacy and imaging endpoints. We present here the preclinical data for STEM-PD, some supportive data from research-grade cells, and the clinical trial design forming the basis for the regulatory approval of STEM-PD—a stem-cell-based trial for intracerebral grafting to receive regulatory approval in Europe under the EU advanced therapeutic medicinal product (ATMP) regulation.

RESULTS

Manufacturing and quality control of STEM-PD

A master cell bank (MCB) of 100 hESC cryovials was manufactured from the GMP-grade cell line RC17 obtained under license from Roslin Cells (Edinburgh, UK). This MCB was subjected to a battery of quality tests according to regulatory requirements including confirmation of pluripotency marker expression by flow cytometry (75.3% OCT3/4^{+}NANOG^{+}) and assessment of genetic integrity and stability by karyotyping, CytoSNP array and sequencing of oncogenes and PD-associated genes (Figure 1A; Tables S1 and S2). The MCB was subsequently used for manufacturing of the STEM-PD product (see STAR Methods).[27]

The differentiation protocol involves an initial patterning phase of 9 days using dual SMAD inhibition together with regionalizing growth factors. This includes Noggin and SB431542 for efficient neuralisation[28] combined with the Glycogen Synthase Kinase 3 (GSK3) inhibitor CHIR99021 for posteriorizing toward midbrain fates and sonic hedgehog (SHH) for ventralization to obtain VM progenitor cells (Figure 1B; STAR Methods).[27] At day 9, fibroblast growth factor 8b (FGF8b) expressed at the midbrain-hindbrain boundary was added to fine-tune patterning toward a caudal VM fate, which we have shown to be associated with successful graft outcome in a preclinical rat model of PD.[17]

(B) Detailed overview of the STEM-PD GMP manufacturing process.
(C) Results from flow cytometric release analysis of batch #3 for purity markers FOXA2 and OTX2 as well as off-target markers SOX1 and PAX6 and the pluripotency markers OCT4 and NANOG.
(D) Confirmation of VM identity in STEM-PD batch #3 by ICC. Scale bars, 100 μm.
(E) Stability assessment of STEM-PD batch #3 by flow cytometry for up to 36 months after cryopreservation.
(F) Quantitative real-time PCR analysis for VM and non-VM markers in 4 GMP STEM-PD batches compared with positive and negative non-VM neural progenitor control populations generated from RC17 hESCs. Minimum and maximum fold change cutoff levels for each gene were applied.[27] Data are represented as mean ± SEM. dFB, dorsal forebrain; vFB, ventral forebrain; vHB, ventral hindbrain.

Brain-derived neurotrophic factor (BDNF) and ascorbic acid were added at day 11 to enable progression of the cells toward post-mitotic neuronal progenitor cells, and the cells were harvested and cryopreserved at day 16 (Figure 1B). Each STEM-PD manufacturing run of 20 days resulted in a >2,300-fold expansion in cell number from the initially seeded MCB cells and was conducted at a peak scale of only 7× T175 flasks, generating a total of 300 vials (5×10^6 cells/vial) of the STEM-PD product (Figure 1B).

The cryopreserved STEM-PD product was quality tested for cell line identity, viability, yield, sterility, endotoxins, viruses, and unknown adventitious agents according to EU ATMP guidelines (Figure 1A; Table 1). We present here the results of the full *in vitro* batch release testing from 4 STEM-PD GMP batches, including the batch that is being used in our ongoing STEM-PD clinical trial (batch #3). This batch complied with release specifications and was subsequently used in all our *in vivo* safety and efficacy testing (see full list of specifications and test results in Table 1, and *in vivo* testing results in Figures 2, 3, 4, and 5). Correct differentiation toward VM progenitor fate was assessed by flow cytometry for the purity of FOXA2 and OTX2 double-positive cells (release specification of $\geq$70% FOXA2$^+$/OTX2$^+$) and by the relative absence of off-target non-VM neural cells (release specification of $\leq$2% SOX1$^+$ and $\leq$2% PAX6$^+$ cells, Figure 1C; Table 1). Batch #3 showed a purity of 98.7% FOXA2$^+$/OTX2$^+$ cells, and through stability studies, we demonstrated that the purity was not affected by long-term storage in vapor-phase liquid nitrogen (Figure 1E). The absence of any residual undifferentiated pluripotent cells in the product was confirmed through flow cytometry for OCT3/4 and NANOG (release specification of $\leq$ 0.1% OCT3/4$^+$/NANOG$^+$) (Figure 1C; Table 1). Correct VM identity was further confirmed by immunocytochemistry (ICC) for the VM markers LMX1A and FOXA2 and by quantitative real-time PCR of VM (*EN1*, *ETV5*, *FOXA2*, *LMX1A*, *OTX1*, and *OTX2*) and non-VM (*PAX6*, *FOXG1*, and *HOXA2*) neural markers (Figures 1D and 1F). Genetic stability and the potential presence of any oncogenic mutations was assessed with the same assays as for the MCB (Figure 1A; Table S3).

GLP safety testing in nude rats shows no adverse effects and no biodistribution outside the transplant region

To assess the safety profile of STEM-PD cells prior to use in the clinical trial, we performed a 39-week toxicity, tumorigenicity, and biodistribution study of STEM-PD batch #3 in immunodeficient athymic nude rats under full GLP conditions. This study involved a total of 52 animals administered with vehicle and 70 animals administered with STEM-PD cells at the maximum feasible dose (MFD) through stereotactic implantation unilaterally into the striatum. The MFD was defined as the maximum dose which could feasibly be administered to the rat striatum without graft core necrosis or displacement of the striatal parenchymal boundaries and was established as 700,000 cells distributed over 4 deposits (1.5 μL per deposit) in two tracts. Animals were sacrificed at 4, 26, or 39 weeks post-grafting for assessment of histopathology, organ weight, hematology, blood chemistry, and biodistribution. The 39-week group was further assessed for any ophthalmological adverse events as well as for central nervous system (CNS)-associated signs of abnormalities in motor function or behavior through modified Irwin testing (Figure 2A). A satellite group of 12 animals implanted with 700,000 undifferentiated pluripotent RC17 cells served as a positive control for teratoma formation. At necropsy, the brain was collected from all animals, and 2/3 of the most lateral non-transplanted hemisphere was collected for biodistribution analysis, whereas the remaining part of the brain was used for general histopathological assessment at 7 different rostro-caudal levels according to Bolon sectioning practice.[29] The tissue block containing the graft was additionally used for collection of 20 sections covering the graft area (Figure 2B).

The safety study found no treatment-related effects on mortality, body weight, body temperature, organ weight, food consumption, ophthalmological findings, or on the Irwin assessment of behavior. Some non-adverse inter-group variations in hematology and blood chemistry were observed (not shown), as well as the appearance of peripheral tumors in both treated and untreated animals that were unrelated to the graft, but likely attributable to the immunodeficient nature of the nude rat model (Table S4). Premature euthanasia for welfare reasons occurred in 5.8% of animals in the vehicle group (3 out of 52 animals) and 11.4% of animals in the STEM-PD treated group (8 out of 70 animals). None of the premature terminations were attributed to the STEM-PD treatment (Table S4).

Histopathological assessment of the transplanted hemisphere described findings that would normally be expected from a cell implantation procedure, including needle/trephination tracks and the presence of graft tissue (Figures 2C and 2D). Xenografts with a human neural identity were detected by immunohistochemistry (IHC) for human neural cell adhesion molecule (hNCAM) in 19 of 20 and 39 of 40 STEM-PD-treated animals at 26 and 39 weeks post-transplantation, respectively (Figures 2E and 2F). Only very few scattered Ki67-positive proliferative cells were observed in the graft at 26 and 39 weeks in 47% and 46% of the animals, respectively (Figure 2G). Together with the observation of similar graft volume at 26 and 39 weeks, this indicated that the size of the grafts had stabilized by 26 weeks post-transplantation. In contrast, transplantation of undifferentiated RC17 hESCs resulted in massive teratoma formation (Figure 2H), and 2 out of 12 animals in this group had to be euthanized for welfare reasons before the 12-week endpoint (Table S4). Although not required for regulatory approval, we have subsequently assessed graft proliferation over time in more detail through collection of single nucleus sequencing data from 22 individual animals grafted with 4 different research-grade batches of cells produced with same protocol as STEM-PD. Bioinformatic scoring of cell-cycle phase showed that approximately 0.5% of the cells in the grafts were detected as cycling at 3 months post-transplantation, and this decreased over time. By 12 months, the % cycling cells and % Ki67$^+$ cells were both <0.1%, (Figure S1).

Biodistribution analysis was performed on the transplanted rats by using a high-sensitivity quantitative PCR for human Alu Y elements. Analysis of 20 different organ tissues from a total of 20 vehicle treated and 27 STEM-PD-treated rats at 4 and 26 weeks confirmed that no human cells could be detected outside of the transplanted brain hemisphere (Table S5). Therefore, the remaining animals sacrificed at 39 weeks were assessed only for biodistribution within the CNS (i.e., in the contralateral hemisphere, the cerebrospinal fluid [CSF], and three segments of the spinal cord). Out of 113 samples analyzed,

Table 1. STEM-PD release specification and results

Test parameter	Method of analysis	Specification	Result from 4 GMP batches	Result, batch #3	PASS/FAIL
Purity/identity	flow cytometry: FOXA2 and OTX2	≥70% of the population is positive for both markers	87%–99%	98.7%	PASS
Off-target cells	flow cytometry: SOX1	≤2% of the population is positive	<0.5%	<0.5%	PASS
Off-target cells	flow cytometry: PAX6	≤2% of the population is positive	<0.5%	<0.5%	PASS
Undifferentiated cells	flow cytometry: OCT3/4 and NANOG	≤0.1% of the population is positive for both markers	<0.1%	<0.1%	PASS
Viability	NucleoCounter NC-200	≥70% viability	78%–83%	83%	PASS
Yield	NucleoCounter NC-200	$\geq 3 \times 10^6$ viable cells per vial	$3–4 \times 10^6$	4×10^6	PASS
Appearance	visual assessment	clear to cloudy, colorless to pink dispersion, and free of foreign particles	complies	complies	PASS
Cell line ID	STR DNA amplification	the cell line identity must be identified as the MCB	complies	complies	PASS
Sterility	BacTEC FX40 and BacT/ALERT 3D rapid microbial detection system in accordance to Ph. Eur. 2.6.27	no growth	no growth	no growth	PASS
Endotoxin	LAL test using turbidimetric method, according to Ph. Eur. 2.6.14.	<5 EU/mL	<0.1 EU/mL	<0.1 EU/mL	PASS
Mycoplasma	real-time PCR in accordance with Ph. Eur. 2.6.7	mycoplasma not detected	negative	negative	PASS
Adventitious agent testing	*in vitro* assay (28 days) on MRC-5, Vero and HeLa cells, in accordance with ICH Q5A. Porcine virus testing in PPK and Vero cells; bovine virus testing in Vero and BT cell lines, according to CHMP guidelines and Ph. Eur. requirements; evaluation of reverse transcriptase activity by ultracentrifugation and QFPERT assay	virus not detected	no virus detected	no virus detected	PASS

STEM-PD cells were tested for release according to the specifications in the table. Results from the GMP batches are shown in the last two columns. BT, bovine turbinate; CHMP, Committee for Medicinal Products for Human Use, European Medicines Agency; ICH, International Council for Harmonization; LAL, limulus amebocyte lysate; MCB, master cell bank; PPT, primary porcine kidney; QFPERT, quantitative fluorescent product enhanced reverse transcriptase.

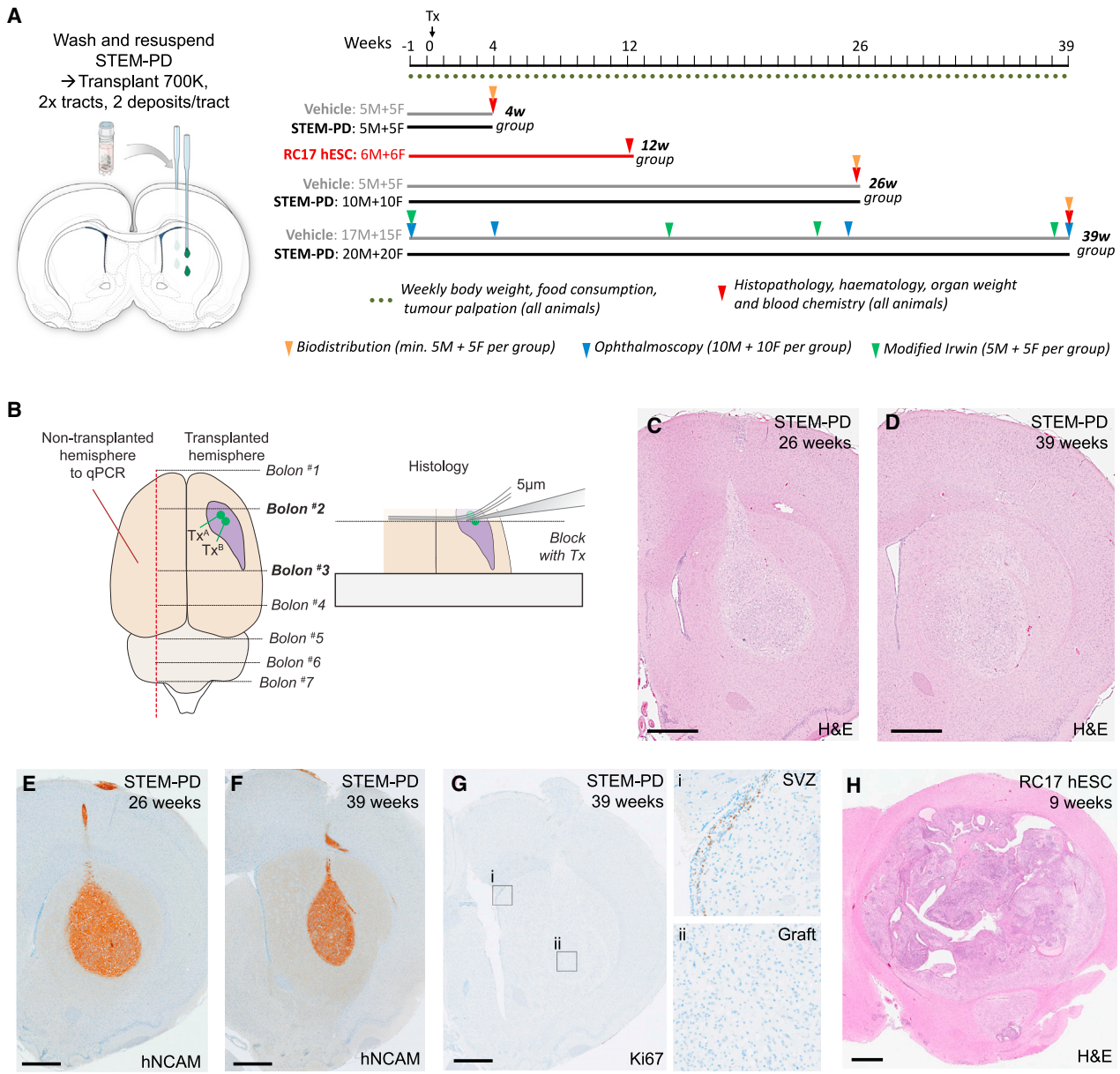


Figure 2. GLP safety study for toxicity, tumorigenicity, and biodistribution

(A) Overview of the GLP safety study design. The rats were divided into 3 groups, receiving either vehicle, an MFD of 700,000 STEM-PD cells, or 700,000 RC17 hESC. The cells were distributed over 4 deposits into one striatum (2 deposits per needle tract), and the animals were assessed at several time points.

(B) The rat brains were dissected to allocate 2/3 of the non-transplanted hemisphere for qPCR and the remaining part for paraffin-embedding and histopathology with Bolon sectioning.

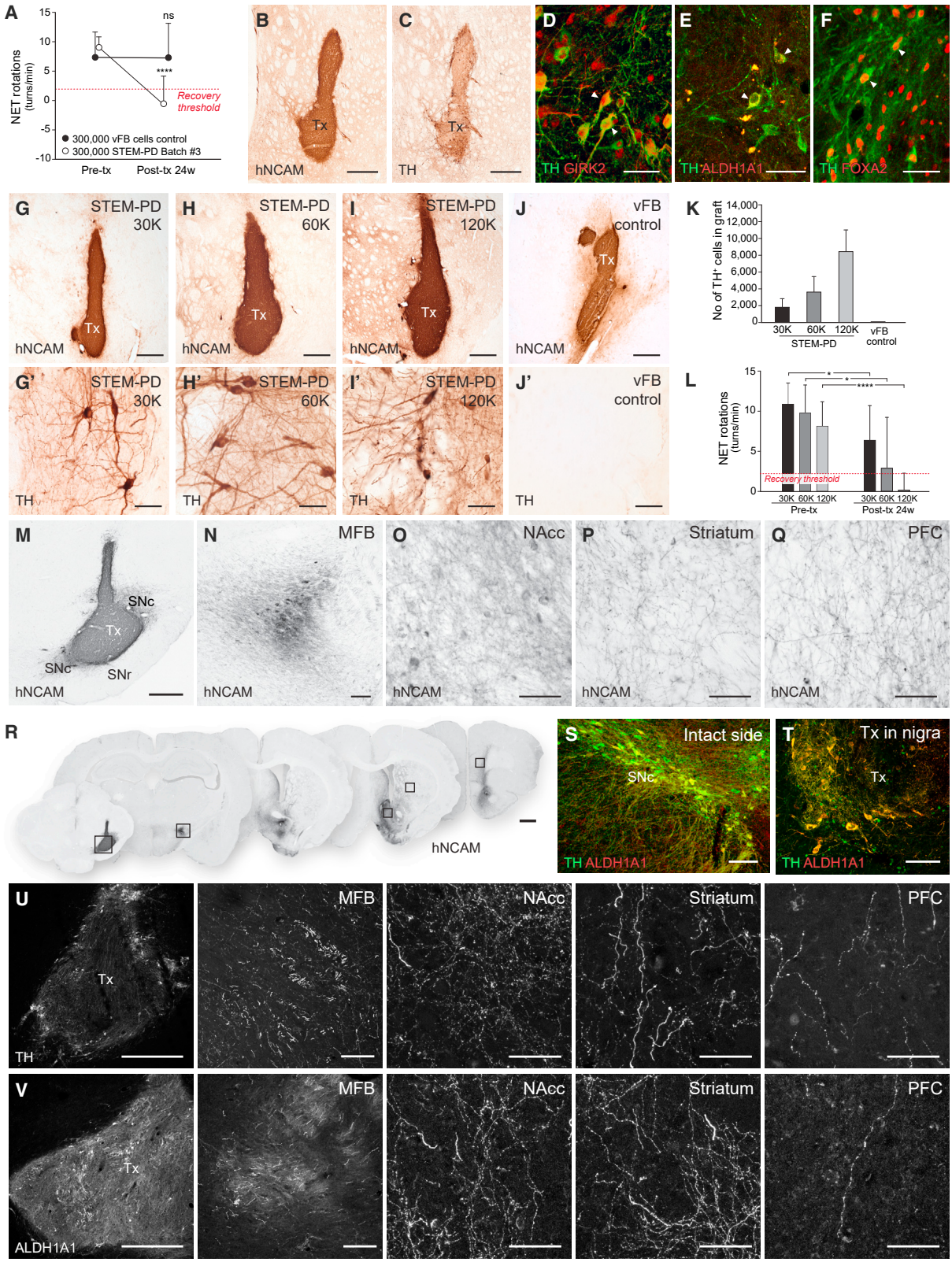
(C–F) Staining by hematoxylin and eosin (H&E) was used to assess general tissue histopathology (C and D), and staining for hNCAM was used to confirm neural identity of the transplanted cells (E and F).

(G) Staining for Ki67 revealed few or no Ki67+ cells in the graft tissue (Gi), as opposed to the neighboring subventricular zone (SVZ, Gii), which showed clear immunoreactivity for Ki67.

(H) H&E brain section from an early sacrificed animal from group 3 (transplanted with 700,000 RC17 hESCs), showing a large teratoma structure at 9 weeks post-transplantation. See Tables S4 and S5 for further results from the study. M, male; F, female; Tx, transplant. All scale bars, 1 mm.

only 1 CSF sample from a STEM-PD-treated animal at 39 weeks showed a human DNA signal above the limit of detection (LoD), although below the limit of quantification (LoQ) (Table S5). Together, the GLP safety study of STEM-PD batch #3 showed that STEM-PD administered at the MFD was well tolerated with no adverse in-life or histopathological findings and with no detectable biodistribution of STEM-PD cells outside of the transplanted brain hemisphere.

CellPress
OPEN ACCESS



(legend on next page)

STEM-PD grafts are rich in DA neurons, innervate the host brain, and reverse amphetamine-induced rotations in rats

Efficacy, potency, and outgrowth capacity of STEM-PD batch #3 was evaluated in the unilateral 6-hydroxydopamine (6-OHDA)-lesioned model of PD (Table S6) with immunodeficient nude rats as recipients. To confirm functional efficacy, 300,000 STEM-PD cells were transplanted to the denervated striatum (group 1 in Table S6), and motor function was evaluated using the amphetamine-induced rotation test prior to transplant surgery (baseline) and at 24 weeks post-transplantation. The mean value for the pre-transplantation rotational scores in this group of animals was 9.0 ± 1.9 turns/min and −0.54 ± 4.58 turns/min at 24 weeks post-transplantation ($p < 0.0001$ for pre-Tx vs. 24 weeks post-Tx in paired t test, Figure 3A). Furthermore, all surviving animals with correct graft placement in group 1 (n = 16) reduced from their pre-transplantation score, and 14 out of 16 displayed rotation scores below 2 turns/min at 24 weeks. No recovery was observed in the lesioned animals transplanted with non-DA ventral forebrain (vFB) control cells (Figure 3A).

At 24 weeks post-transplantation, the rats were perfused, and the brain tissue processed for histology. The transplants were detected by hNCAM and Tyrosine Hydroxylase (TH) (Figures 3B and 3C). The TH^+ neurons co-expressed mature mesencephalic DAergic (mesDA) markers typical of A9 SN neurons such as GIRK2 (KCNJ6), ALDH1A1, and FOXA2 (Figures 3D–3F). Quantification of the total number of TH^+ neurons in the grafts showed that the mean (±SD) content of TH^+ cells was 8,504 ± 4,397 cells/graft (Table S7). Hence, the yield of TH^+ neurons from STEM-PD batch #3 was determined to be 2,835 TH^+ neurons per 1×10^5 transplanted cells, which is in the range of what we have previously observed.[17] Finally, since it has been hypothesized that GIDs observed in some patients receiving fetal human VM transplants may relate to the presence of contaminating serotonergic neurons in the grafts,[7,12,30,31] we performed IHC for 5-hydroxytryptamine (5-HT). No 5-HT^+ cells were detected in any of the 16 analyzed grafts in group 1 (Figure S2).

To investigate dose dependency, we included 3 additional groups (groups 2–4, Table S6), transplanted with lower numbers of STEM-PD cells (i.e., 30,000, 60,000, and 120,000 cells, respectively). To avoid confounding effects of graft size and density of cell suspension on cell maturation, the STEM-PD cells in these experiments were diluted with non-DA carrier cells of vFB identity to retain the same density of cell preparation used for transplantation and same total cell number as in group 1. hNCAM immunostaining confirmed graft viability in all three groups (Figures 3G–3J). Immunostaining for TH (Figures 3G′–3J′), confirmed a dose-response correlation between the number of TH^+ cells and the increasing dose of STEM-PD (Figures 3G′–3J′ and 3K). As a control, we analyzed transplants of vFB carrier cells alone (group 7, Table S6), and confirmed that these cells did not give rise to any TH^+ neurons (Figures 3J and 3J′). We found that 120,000 STEM-PD cells diluted with vFB cells produced almost the same amount of DA neurons as 300,000 STEM-PD cells transplanted alone. This indicates that co-transplantation with non-DA progenitors may potentially be supportive for DA neuron survival or maturation—something that was not expected and would be relevant to investigate further experimentally for future product improvement.

Prior to grafting, the animals in groups 2–4 (dosed with 30K, 60K, and 120K STEM-PD, respectively) showed a mean rotational score of 11.05 ± 2.53, 9.82 ± 3.46, and 8.21 ± 3.01 turns/min, respectively (Figure 3L; Table S7). At 24 weeks post-transplantation, the respective rotation scores for these groups were 6.49 ± 4.27, 2.97 ± 6.32, and 0.15 ± 2.12, thus showing significant reductions in mean rotational scores for all three dosing groups ($p = 0.011$, $p = 0.026$, and $p < 0.0001$, respectively; Figure 3L) but only full recovery in animals dosed with 120K STEM-PD cells. From the low-dose groups (30K and 60K STÈM-PD), we identified two rats that had grafts with <2,000 TH^+ cells (rat #21 with 912 TH^+ cells and rat #36 with 1,720 TH^+ cells, Table S7) while still showing complete behavioral recovery with rotation scores <2 at 24 weeks after dosing. This suggests that even a small number of STEM-PD DA cells can support recovery of motor asymmetry in rats as has been reported previously for fetal cells[32] and research-grade ESCs.[19]

Another important criterion for clinical efficacy is the extent of DAergic innervation into the host brain parenchyma. To assess the growth capacity of graft-derived axons, we transplanted 20 6-OHDA-lesioned nude rats with STEM-PD cells into the midbrain (group 5, Table S7). At 24 weeks post-transplantation, hNCAM staining confirmed the presence of a graft in the substantia nigra (SN) in all 19 surviving animals (Figures 3M

Figure 3. Efficacy testing of STEM-PD batch #3

Efficacy of STEM-PD batch #3 was tested in a long-term *in vivo* study in nude rats.

(A) Recovery of amphetamine-induced rotational assessment pre-transplantation and 24 weeks post-transplantation. One way ANOVA: $F(2,22) = 60.97$, ****$p < 0.0001$ for pre-Tx vs. 24 weeks post-Tx in paired t test, n = 16 for STEM-PD and n = 4 for vFB, ns: non-significant.

(B and C) (B) The presence of human neural grafts was confirmed by hNCAM-positive immunohistochemical staining and (C) DAergic content was confirmed with TH staining.

(D–F) High magnification images of cells confirming their mature DA phenotype and midbrain identity by co-labeling of TH with GIRK2, ALDH1A1, and FOXA2.

(G–J) Transplants of increasing numbers of STEM-PD cells, diluted with non-DA vFB cells (J) to reach a total of 300,000 grafted cells per animal in all groups. (G′–I′) Staining for TH confirmed DAergic content of the grafts in animals dosed with STEM-PD but not with vFB carrier cells alone (J′).

(K and L) (K) TH counts and (L) amphetamine-induced rotational assessment for animals dosed with 30,000, 60,000, and 120,000 STEM-PD cells at 24 weeks post-transplantation, *$p < 0.05$, ****$p < 0.0001$ by paired t test.

(M–Q) (M) hNCAM staining of intranigral grafts and their innervation from the graft core into the (N) MFB, (O) NAcc, (P) striatum, and (Q) PFC.

(R) Whole-brain visualization of hNCAM staining from an intranigral graft. Boxes show locations for magnified images in (M)–(Q).

(S and T) Staining for A9 DA markers TH and ALDH1A1 in endogenous neurons in the intact SNc (S), and in an intranigral STEM-PD transplant (T).

(U and V) TH and ALDH1A1 staining in animals with intranigral STEM-PD grafts, showing A9-like innervation to known DA target structures. All graphs are showing mean ± SD. MFB, medial forebrain bundle; NAcc, nucleus accumbens; PFC, prefrontal cortex; SNc, substantia nigra pars compacta; Tx, transplant; vFB, ventral forebrain. Scale bars: 50 μm in (D)–(F), (G′)–(J′), (O)–(Q), (U), and (V) (NAcc, Striatum, PFC); 100 μm in (N) and (S)–(V) (MFB); 500 μm in (B), (C), (G)–(J), (M), (U), and (V) (Tx); and 1 mm in (R). See Table S6 for details on study outline.

CellPress
OPEN ACCESS

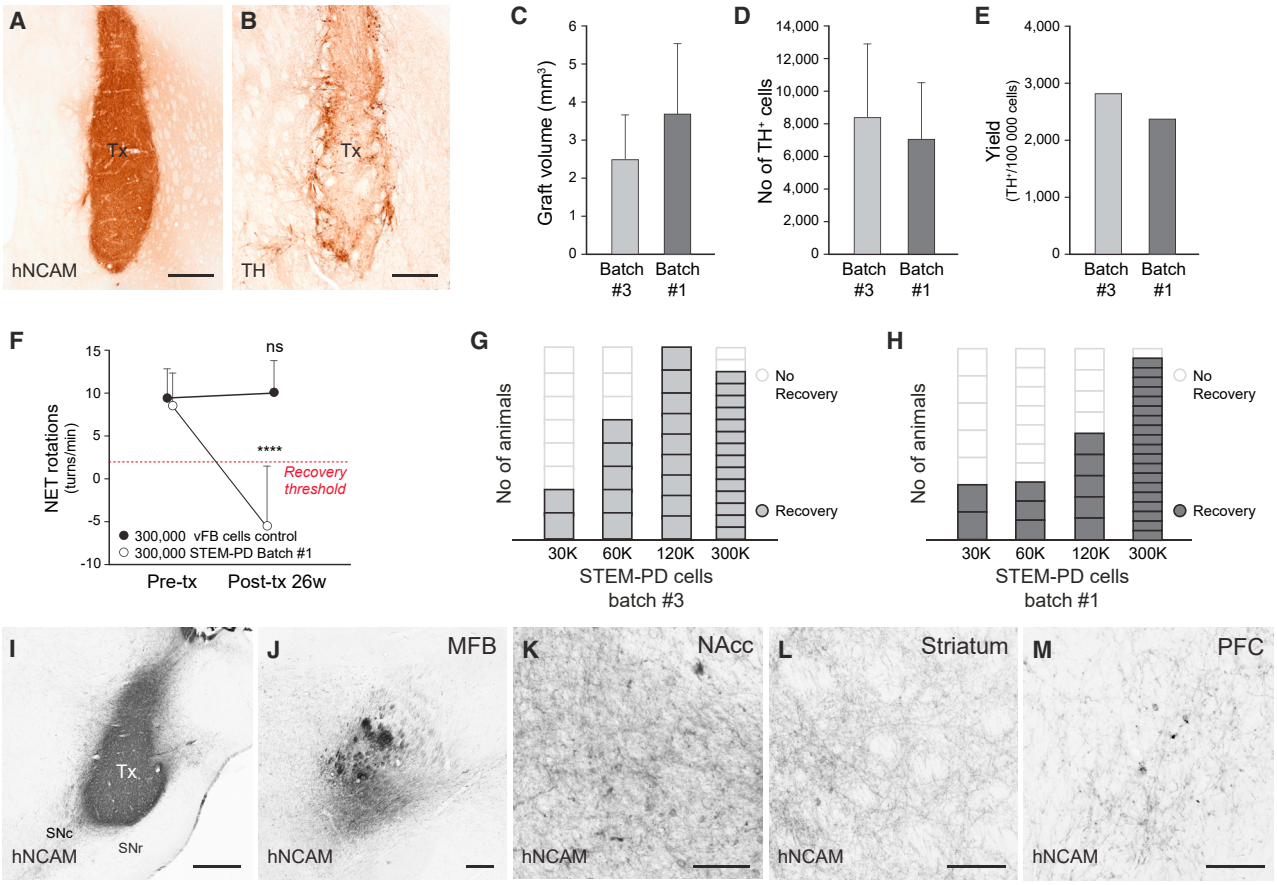


Figure 4. Batch-to-batch comparability on *in vivo* efficacy

Long-term efficacy testing in nude rats was performed with one additional batch of STEM-PD (batch #1), and the results were compared with batch #3.

(A and B) (A) hNCAM and (B) TH staining of batch #1. Scale bars, 500 μm.

(C) Quantification of graft volume in batch #1 and #3 based on HuNu staining.

(D and E) (D) Quantification of TH^+ neurons in batch #1 and #3, and (E) yield of TH^+ neurons per 100,000 grafted cells.

(F) Amphetamine-induced rotational assessment pre-transplantation and 26 weeks post-transplantation for animals dosed with batch #1 ($t = 9.083$, $^{****}p < 0.0001$ by paired t test, $n = 20$).

(G and H) Panels showing the number of recovered animals (i.e., with post-transplantation rotational scores <2) out of total assessed animals in each group transplanted with 30,000, 60,000, and 120,000 cells, respectively, from batch #3 (G) and batch #1 (H).

(I–L) (I) Grafts placed in the midbrain (scale bars, 500 μm; SNc, substantia nigra pars compacta; SNr, substantia nigra reticulata) were confirmed to innervate through the medial forebrain bundle (MFB) (J, scale bars, 100 μm); nucleus accumbens (NAcc) (J), striatum (K), and prefrontal cortex (PFC) (L). Scale bars: 100 μm in (K)–(M). All graphs are showing mean ± SD.

and 3R). hNCAM-positive fibers could be traced in large numbers along the nigrostriatal pathway (NSP) and the medial forebrain bundle (MFB; 19/19 animals) toward the A9 target area in the striatum (15/19 animals), as well as the A10 target areas in nucleus accumbens (NAcc, 19/19 animals) and prefrontal cortex (PFC; 19/19 animals), extending approximately 7–8 mm from the graft core, which is at a scale relevant for re-innervation of the human putamen in a clinical setting (Figures 3M–3R). Co-staining of TH and ALDH1A1 confirmed midbrain DAergic identity of the grafted cells, which were of a similar appearance to the endogenous nigral neurons and further confirmed prominent A9-like DAergic innervation of the DA-depleted dorsolateral striatum (Figures 3S–3V).

Taken together, these data show that STEM-PD cells survive transplantation long-term and mature into DA neurons with the capacity for long-distance fiber outgrowth and the ability to mediate full functional recovery in the DA-lesioned recipient animal.

STEM-PD shows robust and reproducible *in vivo* efficacy in batch-to-batch comparison

To assess batch-to-batch comparability with respect to graft outcome and *in vivo* efficacy, we performed a second full efficacy study with 80 animals using a different batch of GMP-manufactured STEM-PD cells (batch #1), which had a $FOXA2^+$/$OTX2^+$ purity of 86.5% and complied with all release criteria, thereby showing characteristics similar to batch #3 (Figure S3; Table 1). Likewise, batch #1 gave rise to DA-rich grafts with a similar appearance to batch #3 as determined by hNCAM and TH immunostaining (Figures 4A and 4B vs. Figures 3B and 3C), and yielding a similar graft volume and number of TH^+ neurons (Figures 4C–4E). Moreover, batch #1 also mediated complete

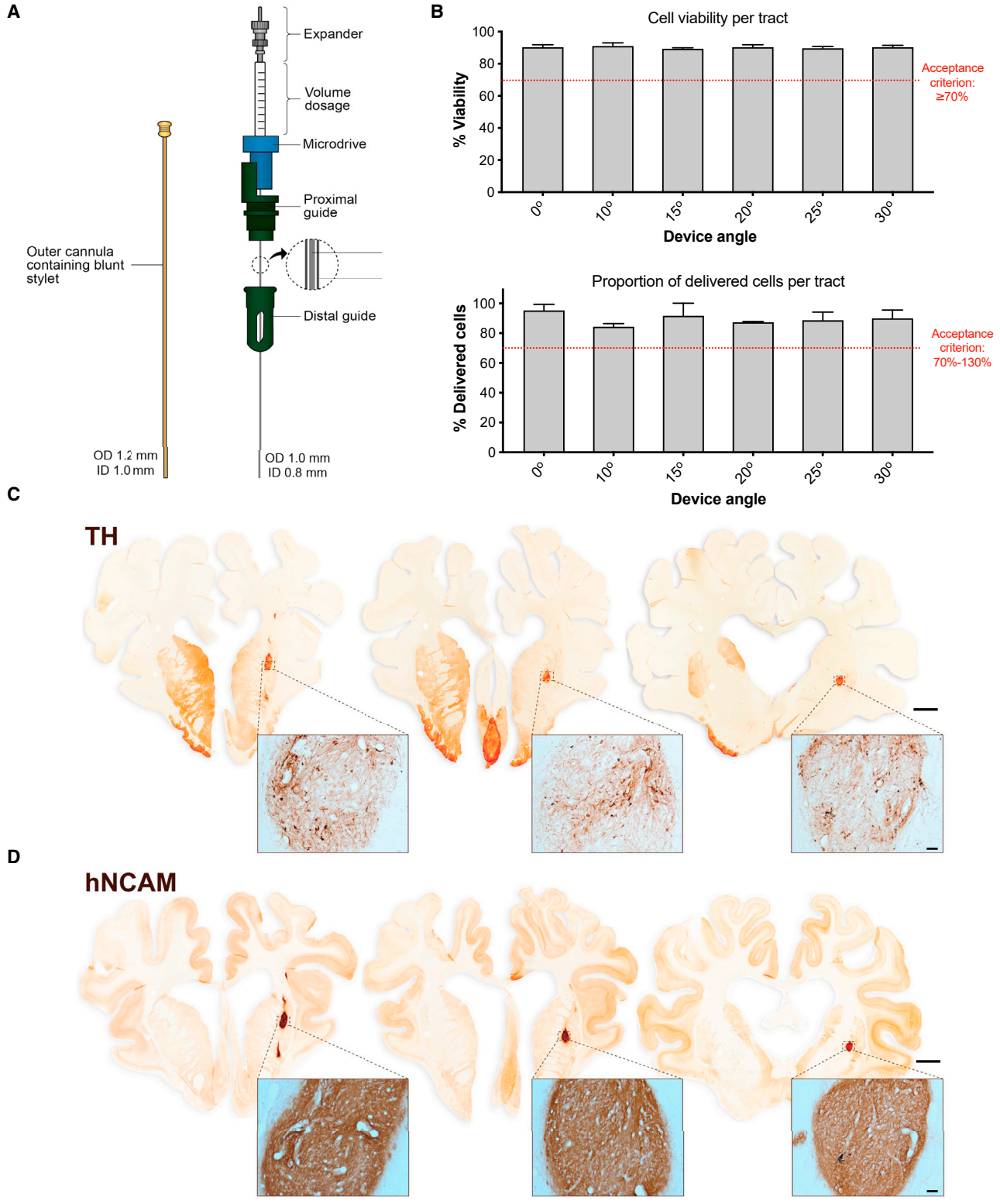


Figure 5. Testing of R-L device compatibility with STEM-PD cells *in vitro* and in minipig

The in-house manufactured Rehncrona-Legradi (R-L) device was tested with STEM-PD cells *in vitro* and in a pig transplantation study for compatibility and feasibility.

(A) Schematic drawing of the R-L device design with an inner and outer cannula and stepwise microdrive at the top to allow accurate serial delivery of up to 8 × 2.5 μL deposits in each tract.

(legend continued on next page)

functional recovery in the 6-OHDA rat model of PD at 26 weeks post-transplantation (Figure 4F) with a dose dependency similar to batch #3 (Figures 4G and 4H). When transplanted to the SN, this batch was similarly able to regrow fibers along the NSP and MFB (in 8/8 animals with correctly placed grafts) to the NAcc (8/8 animals), striatum (6/8 animals), and PFC (in 6/8 animals, Figures 4I–4M).

STEM-PD cells survive long term in the pig brain upon transplantation using the clinical delivery device

For the clinical trial, all surgeries will be performed in one site (Skåne University Hospital, Sweden), and the cells delivered using the Rehncrona-Legradi (R-L) device (Figure 5A) that has successfully been used in the clinic for human fetal VM tissue implants.[5,33–35] We tested the compatibility of this device to deliver viable STEM-PD cells *in vitro* and confirmed that the viability after passing through the device ranged from 89% to 92% (Figure 5B). We further showed that the number of delivered viable cells per simulated tract ranged from 82% to 99% of the expected pre-loaded number of cells at all tested angles of the device position (0°–30°).

The R-L device is too large for use in the rat brain, so to assess the feasibility and compatibility of this device for delivery of STEM-PD cells in a large brain, we conducted a small transplantation study in a single 6-OHDA lesioned Göttingen minipig immunosuppressed with daily PO administration of tacrolimus. Using the R-L device, a total of 1.25 × 10^6 STEM-PD cells (batch #3) were implanted into the right putamen, distributed over 10 deposits of 2.5 μL each. Histological assessment of the pig brain 11 months post-transplantation showed the presence of a surviving human graft in the putamen, rich in TH^+ DA neurons (Figures 5C and 5D). This confirmed that the R-L device is compatible with delivery of STEM-PD cells to the correct target structures in a large brain and that the implanted cells survive long term and mature to DA neurons.

Clinical dose calculation and implantation approach

To set the clinical dose for the STEM-PD trial, we were guided by experiences gained from previous human fetal tissue transplantation trials in PD patients where several post-mortem studies have reported the number of surviving DA neurons in patients with clinically effective grafts (see Table S8). Based on these data, a minimal dose for clinical recovery was estimated to be approximately 1 × 10^5 human TH^+ neurons per grafted putamen, which is the dose we have selected as the starting dose in the STEM-PD trial (dose 1). To determine how many STEM-PD cells should be transplanted to achieve 1 × 10^5 mature TH^+ neurons in the grafts after transplantation, we made use of the DA yield obtained from the STEM-PD batch #3 efficacy study (i.e., 2,835 TH^+ neurons derived per 1 × 10^5 transplanted cells). From this, the clinical dose 1 was set at 3.5 × 10^6 STEM-PD cells/putamen, distributed over 5 needle tracts in each putamen—two tracts targeting the anterior putamen and three tracts targeting the posterior putamen—with 4 × 2.5 μL deposits per tract and 7.1 × 10^4 cells/μL to achieve optimal coverage of the host tissue. Magnetic resonance imaging (MRI) is performed on each patient prior to surgery for individual calculation of target region, burr hole locations, and needle angles to achieve optimal targeting of the putamen while avoiding major blood vessels during the implantation surgery. To explore the optimal dose range, a second dose (dose 2) with twice the number of transplanted cells, targeting 2 × 10^5 TH^+ neurons/putamen = 7.1 × 10^6 STEM-PD cells administered per putamen at the same cell density but in the double number of deposits (i.e., 5 tracts with 8 deposits/tract in each putamen) has been included in the trial design. The human SN contains approximately 4 × 10^5 TH^+ neurons,[36] and based on the assumption that approximately half of these innervate the putamen, this corresponds to approximately 2 × 10^5 DA neurons. Thus, the selected doses (1 × 10^5 and 2 × 10^5 TH^+ neurons/hemisphere) are within the range of the number of DA neurons that normally innervate the putamen in healthy humans. At present, we plan for 4 patients to receive dose 1, and 4 patients to receive dose 2; however, a decision on dose escalation will be dependent on an interim safety assessment and positron emission tomography (PET) imaging of the first 4 patients transplanted with dose 1.

STEM-PD clinical trial design

The STEM-PD trial design is similar to that of the TRANSEURO trial (NCT01898390) and is based on lessons learned from previous fetal tissue transplantation trials, with rationales discussed in part previously.[14,37] Clinical trial participants will be transplanted in a pre-determined staggered schedule in which each patient will receive bilateral grafts of the STEM-PD product in a single neurosurgical session (Figure 6A). The primary objective of the trial is to assess the safety, tolerability, and feasibility of intraputamenal transplantation of the STEM-PD product in patients with moderate PD at 12 months post-transplantation. Secondary objectives related to efficacy are assessed at 36 months (Figure 6B). The secondary objectives are 3-fold, (1) to evaluate the course and efficacy on clinical features post-grafting using validated clinical measures such as the unified PD rating scale (UPDRS) in the defined OFF state (i.e., the patient is off DAergic medication during testing), (2) to assess the survival of DAergic cells following transplantation using F-DOPA and PE2I PET imaging to label functional DA terminals, (3) to determine the safety and clinical efficacy between doses (if dose escalation is undertaken) of the STEM-PD product, including assessment of whether there

(B) Viability and cell yield results from device delivery of STEM-PD cells *in vitro* at various angles relevant for the surgical setup. The top graph shows cell viability per simulated tract and the bottom graph shows the cell number delivered per simulated tract as a percentage of the number of cells loaded into the device. Predefined acceptance criteria for each test are noted in red. Data are represented as mean ± SD.

(C and D) A minipig transplanted with STEM-PD batch #3 using the R-L device was sacrificed at 11 months post-transplantation, and brain sections were immunostained for TH and hNCAM.

(C) TH^+ staining at the coronal level of the grafted neurons in the putamen of the 6-OHDA-lesioned minipig. Insert: high magnification images of the grafted neurons and fibers.

(D) hNCAM-positive graft staining of coronal minipig sections. Insert: images of positive fiber staining within the graft. The coronal sections are presented in anterior to posterior direction (corresponding to the minipig atlas S9, S11, and S13 https://auinstallation34.cs.au.dk/fileadmin/cense.au.dk/Atlas/atlas/index.html). Scale bars, 1,000 μm (coronal section); Scale bars, 100 μm (insets). Cd, caudate nucleus; ic, internal capsule; Pu, putamen; Acb, accumbens nucleus.

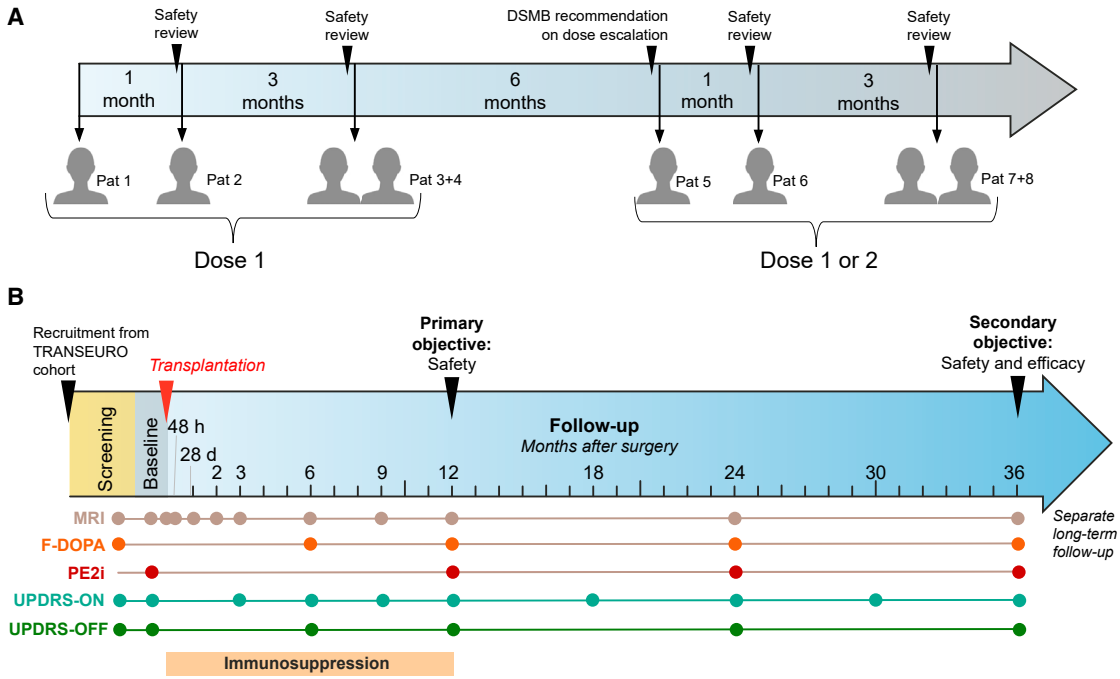


Figure 6. Overview of the STEM-PD trial

The STEM-PD trial is a 36-month trial including 8 patients.

(A) Schematic summarizing the staggered dosing scheme for the 8 patients. Patients 1–4 will be staggered with interim safety assessments, and all 4 patients will receive dose 1. At 6 months after patient 4 is dosed, once data are available for safety, imaging, and clinical measurements, the Data and Safety Monitoring Board (DSMB) for the trial will make a recommendation for the dosing of the remaining 4 patients. The DSMB can recommend to either (1) remain at dose 1, (2) proceed to dose 2, or, (3) wait longer to collect more data. The final decision will be made by the Trial Management Group, after receiving the DSMB's recommendation.

(B) Overview of trial design. Patients will be recruited from the ongoing TRANSEURO observational study[14] and at the end of the STEM-PD trial, the patients will be asked to take part in a long-term follow-up study. Patients will be assessed over the first 12 months for the primary objectives concerning safety, tolerability, and feasibility. Secondary objectives at 36 months include safety as well efficacy of the graft, based on clinical measures and DAergic PET imaging (F-DOPA and PE2i). Structural imaging with MRI is included to check for space-occupying lesions at the graft site as well as transplant placement. Clinical assessments include validated tests such as the unified Parkinson's disease rating scale (UPDRS) in ON and OFF states (i.e., the patient is on or off DAergic medication at the time of testing). Additional clinical assessments not shown here include a range of timed motor tasks as well as cognitive, affective, and quality-of-life assessments.

is a dose-response effect. All participants will be followed up for 3 years within the trial and also offered participation in a separate long-term follow-on study where they would be followed up indefinitely.

The trial will be open label involving 8 participants with moderately advanced PD who are entering the phase of the condition when an advanced interventional therapy such as deep brain stimulation (DBS), DuoDopa, or apomorphine infusions are being considered. Furthermore, and in line with TRANSEURO, patients with significant LIDs are excluded from the trial, as pre-existing LIDs may predispose participants to develop GIDs and may further require additional medical interventions, which could interfere with the trial. In addition, a relatively preserved DAergic innervation of the ventral striatum has been linked to better clinical outcome in patients grafted with human fetal VM tissue[10,38]; therefore, patients are required to show a good response to L-DOPA clinically and to be devoid of any major cognitive, neuropsychiatric, or autonomic problems. The patients selected have all been part of the TRANSEURO observational study for a minimum of 12 months, providing good long-term pre-intervention data on their disease progression.

All participants will receive immunotherapy similar to that used in renal transplant programs: basiliximab, tacrolimus (or cyclosporine if tacrolimus is not tolerated), azathioprine, and steroids for 12 months post-grafting.

DISCUSSION

There are currently at least two other ongoing clinical trials using induced pluripotent stem cell (iPSC)—and hESC—derived DAergic progenitor cells in patients with PD.[23,24] These trials, like the TRANSEURO trial and our STEM-PD trial, are allogeneic investigational therapies with immune suppression given to the recipients for 12 months post-grafting. The iPSC-based trial in Japan is using a fresh cell product that is manufactured separately for each patient starting from the same iPSC line, while the BlueRock trial in the US and Canada makes use of a cryopreserved cell product from the H9 hESC line, which was re-derived from a research-grade cell line. When designing STEM-PD, we opted for a cryopreserved cell product from the hESC line RC17, which was derived fully under GMP conditions.[39] The STEM-PD product has been manufactured in batches large enough to allow for quality assessment and *in vivo* safety and efficacy, as well as use in the entire clinical trial. While full characterization of all cells in the grafts after maturation *in vivo* was not requested and has not been performed, thorough testing of the

clinically administered batch for safety and efficacy in long-term animal studies gives high confidence in the functionality and safety of the exact product batch used in patients. This is particularly important for first-in-human studies using neural progenitor cells derived from pluripotent stem cells, since such cells undergo terminal differentiation, diversify, and become functional only after transplantation *in vivo*. In addition, the use of a single clinical batch ensures minimal product-related variation between participants, and the availability of a banked, cryopreserved product makes surgical and clinical planning logistically easier. However, in view of developing a global therapy at a scale that demands serial batch manufacturing, it is unlikely to be practically feasible or financially viable to perform extensive long-term *in vivo* testing of all batches intended for clinical use. It is therefore important to verify that the product is efficacious also across several GMP batches, as we have done here.

Whereas the other ongoing stem cell trials for PD involve patients with more advanced PD,[40,41] the STEM-PD trial targets patients with moderate disease, many of which would under normal circumstances be eligible for some form of invasive interventional therapy in the next few years, such as DBS, MRI-focused ultrasound lesioning, DuoDopa, and/or apomorphine infusions. The "moderate" inclusion criterion developed for the STEM-PD trial aims to strike a balance by avoiding the treatment of early-stage patients who may be managed well on their current standard of care for years, and late-stage patients who may be more vulnerable to lengthy surgery and less likely to respond to DA cell replacement therapy. If cell replacement therapies such as STEM-PD are shown to be safe and effective, we anticipate that earlier-stage patients, who would potentially benefit the most from cell replacement therapy, could be recruited for future trials. The STEM-PD trial has the major advantage that it is designed in a nested fashion, with patients recruited to the STEM-PD trial directly from the TRANSEURO observational study which began in 2011. This means that some of the patients in the trial have been followed for up to 10 years prior to the intervention with regular motor testing, ensuring that the clinical team has considerable data on the patients' pre-intervention disease trajectories. Although the approach of recruiting patients from a pre-existing cohort may potentially introduce a risk of selection bias toward patients with less aggressive disease progression, we will nevertheless be able to crucially monitor disease progression trajectories of each patient pre- and post-transplantation given that they will have been followed and assessed over many years. The TRANSEURO patients are further accustomed to the trial assessment tests over several years, thereby minimizing any initial training effects, which could confound the interpretation of clinical assessment scores.

Preclinical studies in rats have shown that transplanted hESC-derived VM progenitor cells can survive and mature into DA-producing neurons, innervate the host striatum, integrate into the host circuitry, mediate functional recovery in DA depletion models through DA secretion, and alleviate LIDs in those same models through DA reuptake.[16,17,19,42–45] Based on these data, the aim for the ongoing stem cell trials is that the grafted cells will survive, produce DA, and mediate clinical benefits in patients. As a first step, it will be crucial to demonstrate through PET imaging, that the stem-cell-derived neurons survive and give rise to local DAergic innervation of the putamen in humans, as has been shown in animal models. If safety and feasibility as well as cell survival and functionality can be shown, the ability to generate virtually unlimited numbers of standardized and quality-controlled DA cells from stem cells could transform cell replacement into a clinically useful therapy for PD. A standardized cell product further allows one to conduct properly controlled studies to determine the effect of the transplantation procedure, cell placement, and dose. It is estimated that approximately 100,000 surviving DA neurons per grafted putamen are needed to mediate clinical benefit, based on human fetal VM transplant data. A controlled dose-finding study with fetal tissue has previously not been possible, given severe issues with tissue availability and the inherent variation in tissue quality and preparation. Using manufactured stem cell-derived DA neurons, dose escalation studies can now be performed, such that the optimal number of cells can be deduced. Detailed analysis of the quantitative cellular composition of grafts from the clinical batches of STEM-PD was not requested by the authorities. However, as complementary data, experimental studies from our group have shown that stem-cell-derived VM grafts as well as in vitro organoids generated with same protocol as STEM-PD also contain other cell types including non-DA neurons, astrocytes, and vascular leptomeningeal cells (VLMCs).[46] The potential effect or function of these other cell types in the grafts is not known, although one study suggests that grafted astrocytes may be protective against alpha-synuclein pathology,[47] and data presented in this study suggest that mixing of non-DA neurons to the graft might be beneficial for maturation of the DA neurons. Future studies are required to elucidate this, and to determine if a controlled cellular composition may be harnessed for further improving graft function. Lastly, to enable upscaling and development of globally available therapies, refinements in manufacturing procedures, identification of validated potency markers, and exploration of safety assessments that rely less on long-term *in vivo* experiments are required. However, global regulatory consensus is needed, such that comparable data packages and standards can be applied for approval in the US, UK, EU, and Asia Pacific. With stem cell-based products now in clinical trials across three different continents, the prospects of such products going toward market authorization are getting closer.

Limitations of the study

The data presented in this study reflect the data package provided to the regulatory authorities in Sweden for the purpose of demonstrating preclinical safety and efficacy of the STEM-PD product to gain clinical trial authorization. Although additional experiments aimed at investigating biological or mechanistic details of the product may be of interest from a scientific point of view, further studies looking at this have not been requested by the authorities and have therefore not been performed. In general, preclinical assessment of clinical products does not involve exploratory studies, since each analysis performed on the product must have clear and predefined acceptance criteria that the product must adhere to. Furthermore, most of these analyses must be performed under highly controlled GMP/GLP conditions at outsourced facilities, which severely limits further scientific explorations. To investigate comparability of GMP manufacturing runs, ideally more than 2 production batches

should be tested *in vivo*; however, additional *in vivo* batch testing has not been possible due to the high costs incurred with manufacturing and testing. When setting the dose in STEM-PD, we made the assumption that the stem-cell-derived DA neurons have similar potency to fetal DA neurons based on previous experimental data[19] and the few individual animals in the efficacy study which contained a low number of TH^+ cells and yet demonstrated full recovery. The limited experimental data on potency, combined with potential host species differences in graft survival and differences in instrument used in rats and humans, bring uncertainty in the clinical dose calculation but are the best data available from the experimental models at hand. The clinical trial design presented here reflects the trial design at the initiation of the clinical study. Changes to the trial design may potentially occur during the study in agreement with the relevant authorities, if prompted by safety concerns, patient considerations, clinical data, or regulatory requirements.

CONSORTIA

The members of Novo Nordisk Cell Therapy R&D Consortium are Josefine Rågård Christiansen, Nicolaj Strøyer Christophersen, Simona Graziano, Kevin A. Keane, Ida Stenfeldt Mathiasen, Jonathan Christos Niclis, Lisbeth Palm, Sonja Pikkupeura, Lina A. Thorén, and Dorthe Bach Toft.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Cell lines
  - Animals and housing conditions, GLP safety study
  - Animals and housing conditions, non-GLP efficacy study
  - Minipig model for device testing
- METHOD DETAILS
  - STEM-PD manufacturing
  - Genetic testing on MCB and STEM-PD
  - Flow cytometry release testing
  - Quantitative real-time PCR
  - Immunocytochemistry for FOXA2 and LMX1A
  - STEM-PD stability testing
  - Trilineage test
  - Procedure for preparation of STEM-PD cells for transplantation and device testing
  - Surgical procedures, GLP safety study
  - GLP biodistribution testing
  - Surgical procedures, efficacy study
  - Amphetamine-induced rotations
  - Immunohistochemistry, efficacy study
  - Immunohistochemical quantification, efficacy study
  - Single nucleus RNA-sequencing of research-grade grafts, and bioinformatic analysis
  - Necropsy, histopathology and immunohistochemistry of animals from safety study
  - *In vitro* device testing
  - Transplantation to minipig
- QUANTIFICATION AND STATISTICAL ANALYSIS
- ADDITIONAL RESOURCES

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.stem.2023.08.014.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the New York Stem Cell Foundation (grant number #NYSCF-R-I37), EU (FP7 HEALTH-F4-2013-602278 Neurostemcellrepair, H2020 NSC-Reconstruct 874758 and Eurostars projekt E!10838, DOPALAM 2017-01484), the European Research Council (ERC grant agreement no. 771427), Novo Nordisk A/S, the Swedish Research Council (grant agreements 2016-00873, 2021-00661), The Swedish Parkinson's Association (Parkinsonfonden), the Swedish Brain Foundation (Hjärnfonden FO2019-0301), the Strategic Research Areas at Lund University MultiPark (Multidisciplinary research in Parkinson's disease) and StemTherapy/Lund Stem Cell Center, the Danish Innovation Fund, The Lundbeck Foundation (R303-2018-3058 and R211-2015-3784), Parkinsonforeningen, the Knut and Alice Wallenberg Foundation and the Novo Nordisk Foundation (NNF21CC0073729 and NNF18OC0030286). We would like to thank Ulla Jarl, Michael Sparennius, and Amalie Holm as well as the staff at the Aarhus University Hospital PET Centre (Aage Kristian Olsen Alstrup, David Brooks) and Center for Experimental Neuroscience (CENSE, Dariusz Orlowski) for assistance with the minipig study and Petra Johnels, Ulf Malmqvist, Lisbet Norlander, Sebastian Strunk Möller, Kristina Johansson for support on regulatory and ethical issues. We thank Josefina Dykes, Anette Malmberg, Sabine Apitzsch, and Anna Weddig for clinical and clean room support. We further thank Charlotte Bailey, Sharon Barkatullah, Regine Bergholdt, David Campard, Jeppe Dehli, Frederikke Bjergvang Flagstad, Constanze Hammerle, Kirsten Høstgaard-Jensen, Janne Jensen, Rikke Kaae Kirk, Kristian Kolind, Helle Frimer Larsen, Inge Holm Lauritzen, Karsten Nielsen, Anne Mette Nøhr, Gabriel Pemberton, Christian de Vos Petersen, Johannes Roubroeks, Jonas Lyng Thomsen, Alice Troy, Tanja Villumsen, Johanne Press Wegeberg, Tina Zinck, Anja Dahl Østergren, and Anya Østergård from Novo Nordisk A/S for guidance and support on CMC statistics, technical, and regulatory issues and Marie Jönsson for help with graphical abstract.

AUTHOR CONTRIBUTIONS

The authors confirm contribution to the paper as follows: trial conception and design, A.K., J.N., D.B.H., N.R., H.B., N.N.C.T., A.M.L., J.C.S., O.L., T.v.V., A.B., B.M., B.H., E.C., H.W., G.P., R.A.B., and M.P.; performance of experiments, data collection, analysis, and interpretation of results, A.K., J.N., D.B.H., N.R., H.B., N.N.C.T., A.F., A.F.A., T.C., S.V., J.M., P.S., Y.Z., A.M.L., T.P.L., M.L., C.C., O.B., B.M., and M.P.; draft manuscript preparation, A.K., M.P., and R.A.B. All authors reviewed the results, gave input to the manuscript, and approved the final version of the manuscript.

DECLARATION OF INTERESTS

M.P. is the owner of Parmar Cells AB. A.K. is the owner of Kirkeby Cell Therapy APS. M.P. and A.K. are co-inventors on patents WO2016162747A2/A3 and WO2019016113A1. M.P., A.K., R.A.B., H.W., H.B., A.B., E.C., D.B.H., G.P., and B.H. have performed paid consultancy for Novo Nordisk A/S, and members of NNCT R&D are current or previous employees of Novo Nordisk A/S. T.C., A.F.A., Y.Z., S.V., and D.B.H. performed the work as employees of Lund University but are currently employed by Novo Nordisk A/S (T.C., A.F.A., and S.V.), Takara Bio (Y.Z.), and D.B.H. at Eli Lilly and Company, where

she is also a minor share holder. Novo Nordisk A/S is developing the STEM-PD product for commercial use.

INCLUSION AND DIVERSITY

We support inclusive, diverse, and equitable conduct of research.

Received: March 17, 2023
Revised: July 6, 2023
Accepted: August 31, 2023
Published: October 5, 2023

REFERENCES

1. Kehagia, A.A., Barker, R.A., and Robbins, T.W. (2013). Cognitive impairment in Parkinson's disease: the dual syndrome hypothesis. Neurodegener. Dis. *11*, 79–92. https://doi.org/10.1159/000341998.
2. Bloem, B.R., Okun, M.S., and Klein, C. (2021). Parkinson's disease. Lancet *397*, 2284–2303. https://doi.org/10.1016/S0140-6736(21)00218-X.
3. Obeso, J.A., Rodriguez-Oroz, M.C., Chana, P., Lera, G., Rodriguez, M., and Olanow, C.W. (2000). The evolution and origin of motor complications in Parkinson's disease. discussion S21–S13. Neurology *55*, S13–S20.
4. Barker, R.A., and Björklund, A. (2023). Restorative cell and gene therapies for Parkinson's disease. Handb. Clin. Neurol. *193*, 211–226. https://doi.org/10.1016/B978-0-323-85555-6.00012-6.
5. Lindvall, O., Brundin, P., Widner, H., Rehncrona, S., Gustavii, B., Frackowiak, R., Leenders, K.L., Sawle, G., Rothwell, J.C., and Marsden, C.D. (1990). Grafts of fetal dopamine neurons survive and improve motor function in Parkinson's disease. Science *247*, 574–577.
6. Lindvall, O., Sawle, G., Widner, H., Rothwell, J.C., Björklund, A., Brooks, D., Brundin, P., Frackowiak, R., Marsden, C.D., and Odin, P. (1994). Evidence for long-term survival and function of dopaminergic grafts in progressive Parkinson's disease. Ann. Neurol. *35*, 172–180. https://doi.org/10.1002/ana.410350208.
7. Freed, C.R., Greene, P.E., Breeze, R.E., Tsai, W.Y., DuMouchel, W., Kao, R., Dillon, S., Winfield, H., Culver, S., Trojanowski, J.Q., et al. (2001). Transplantation of embryonic dopamine neurons for severe Parkinson's disease. N. Engl. J. Med. *344*, 710–719. https://doi.org/10.1056/NEJM200103083441002.
8. Kordower, J.H., Goetz, C.G., Freeman, T.B., and Olanow, C.W. (1997). Dopaminergic transplants in patients with Parkinson's disease: neuroanatomical correlates of clinical recovery. Exp. Neurol. *144*, 41–46. https://doi.org/10.1006/exnr.1996.6386.
9. Mendez, I., Sanchez-Pernaute, R., Cooper, O., Viñuela, A., Ferrari, D., Björklund, L., Dagher, A., and Isacson, O. (2005). Cell type analysis of functional fetal dopamine cell suspension transplants in the striatum and substantia nigra of patients with Parkinson's disease. Brain *128*, 1498–1510. https://doi.org/10.1093/brain/awh510.
10. Piccini, P., Pavese, N., Hagell, P., Reimer, J., Björklund, A., Oertel, W.H., Quinn, N.P., Brooks, D.J., and Lindvall, O. (2005). Factors affecting the clinical outcome after neural transplantation in Parkinson's disease. Brain *128*, 2977–2986. https://doi.org/10.1093/brain/awh649.
11. Kefalopoulou, Z., Politis, M., Piccini, P., Mencacci, N., Bhatia, K., Jahanshahi, M., Widner, H., Rehncrona, S., Brundin, P., Björklund, A., et al. (2014). Long-term clinical outcome of fetal cell transplantation for Parkinson disease: two case reports. JAMA Neurol. *71*, 83–87. https://doi.org/10.1001/jamaneurol.2013.4749.
12. Olanow, C.W., Goetz, C.G., Kordower, J.H., Stoessl, A.J., Sossi, V., Brin, M.F., Shannon, K.M., Nauert, G.M., Perl, D.P., Godbold, J., and Freeman, T.B. (2003). A double-blind controlled trial of bilateral fetal nigral transplantation in Parkinson's disease. Ann. Neurol. *54*, 403–414. https://doi.org/10.1002/ana.10720.
13. Barker, R.A., Barrett, J., Mason, S.L., and Björklund, A. (2013). Fetal dopaminergic transplantation trials and the future of neural grafting in Parkinson's disease. Lancet Neurol. *12*, 84–91. https://doi.org/10.1016/S1474-4422(12)70295-8.
14. Barker, R.A.; TRANSEURO consortium (2019). Designing stem-cell-based dopamine cell replacement trials for Parkinson's disease. Nat. Med. *25*, 1045–1053. https://doi.org/10.1038/s41591-019-0507-2.
15. Kikuchi, T., Morizane, A., Doi, D., Onoe, H., Hayashi, T., Kawasaki, T., Saiki, H., Miyamoto, S., and Takahashi, J. (2011). Survival of human induced pluripotent stem cell-derived midbrain dopaminergic neurons in the brain of a primate model of Parkinson's disease. J. Parkinsons Dis. *1*, 395–412. https://doi.org/10.3233/JPD-2011-11070.
16. Kirkeby, A., Grealish, S., Wolf, D.A., Nelander, J., Wood, J., Lundblad, M., Lindvall, O., and Parmar, M. (2012). Generation of regionally specified neural progenitors and functional neurons from human embryonic stem cells under defined conditions. Cell Rep. *1*, 703–714. https://doi.org/10.1016/j.celrep.2012.04.009.
17. Kirkeby, A., Nolbrant, S., Tiklova, K., Heuer, A., Kee, N., Cardoso, T., Ottosson, D.R., Lelos, M.J., Rifes, P., Dunnett, S.B., et al. (2017). Predictive markers guide differentiation to improve graft outcome in clinical translation of hESC-based therapy for Parkinson's disease. Cell Stem Cell *20*, 135–148. https://doi.org/10.1016/j.stem.2016.09.004.
18. Kriks, S., Shim, J.W., Piao, J., Ganat, Y.M., Wakeman, D.R., Xie, Z., Carrillo-Reid, L., Auyeung, G., Antonacci, C., Buch, A., et al. (2011). Dopamine neurons derived from human ES cells efficiently engraft in animal models of Parkinson's disease. Nature *480*, 547–551. https://doi.org/10.1038/nature10648.
19. Grealish, S., Diguet, E., Kirkeby, A., Mattsson, B., Heuer, A., Bramoulle, Y., Van Camp, N., Perrier, A.L., Hantraye, P., Björklund, A., and Parmar, M. (2014). Human ESC-derived dopamine neurons show similar preclinical efficacy and potency to fetal neurons when grafted in a rat model of Parkinson's disease. Cell Stem Cell *15*, 653–665. https://doi.org/10.1016/j.stem.2014.09.017.
20. Doi, D., Magotani, H., Kikuchi, T., Ikeda, M., Hiramatsu, S., Yoshida, K., Amano, N., Nomura, M., Umekage, M., Morizane, A., and Takahashi, J. (2020). Pre-clinical study of induced pluripotent stem cell-derived dopaminergic progenitor cells for Parkinson's disease. Nat. Commun. *11*, 3369. https://doi.org/10.1038/s41467-020-17165-w.
21. Kim, T.W., Piao, J., Koo, S.Y., Kriks, S., Chung, S.Y., Betel, D., Socci, N.D., Choi, S.J., Zabierowski, S., Dubose, B.N., et al. (2021). Biphasic activation of WNT signaling facilitates the derivation of midbrain dopamine neurons from hESCs for translational use. Cell Stem Cell *28*, 343–355.e5. https://doi.org/10.1016/j.stem.2021.01.005.
22. Piao, J., Zabierowski, S., Dubose, B.N., Hill, E.J., Navare, M., Claros, N., Rosen, S., Ramnarine, K., Horn, C., Fredrickson, C., et al. (2021). Preclinical efficacy and safety of a human embryonic stem cell-derived midbrain dopamine progenitor product, MSK-DA01. Cell Stem Cell *28*, 217–229.e7. https://doi.org/10.1016/j.stem.2021.01.004.
23. BlueRock_Therapeutics (2021). BlueRock therapeutics in collaboration with memorial sloan kettering cancer center receives IND clearance for DA01 in Parkinson's disease. https://bluerocktx.com/bluerock-therapeutics-in-collaboration-with-memorial-sloan-kettering-cancer-center-receives-ind-clearance-for-da01-in-parkinsons-disease/.
24. CiRA (2017). News and events. https://www.cira.kyoto-u.ac.jp/e/pressrelease/news/180730-170000.html.
25. Schweitzer, J.S., Song, B., Herrington, T.M., Park, T.Y., Lee, N., Ko, S., Jeon, J., Cha, Y., Kim, K., Li, Q., et al. (2020). Personalized iPSC-derived dopamine progenitor cells for Parkinson's disease. N. Engl. J. Med. *382*, 1926–1932. https://doi.org/10.1056/NEJMoa1915872.
26. Lund_University (2022). First patient receives milestone stem cell-based transplant for Parkinson's disease. https://www.lunduniversity.lu.se/article/swedish-medical-products-agency-grants-approval-clinical-study-new-stem-cell-based-parkinsons. https://www.lunduniversity.lu.se/article/first-patient-receives-milestone-stem-cell-based-transplant-parkinsons-disease.
27. Nolbrant, S., Heuer, A., Parmar, M., and Kirkeby, A. (2017). Generation of high-purity human ventral midbrain dopaminergic progenitors for in vitro maturation and intracerebral transplantation. Nat. Protoc. *12*, 1962–1979. https://doi.org/10.1038/nprot.2017.078.

28. Chambers, S.M., Fasano, C.A., Papapetrou, E.P., Tomishima, M., Sadelain, M., and Studer, L. (2009). Highly efficient neural conversion of human ES and iPS cells by dual inhibition of SMAD signaling. Nat. Biotechnol. *27*, 275–280. https://doi.org/10.1038/nbt.1529.

29. Bolon, B., Garman, R.H., Pardo, I.D., Jensen, K., Sills, R.C., Roulois, A., Radovsky, A., Bradley, A., Andrews-Jones, L., Butt, M., and Gumprecht, L. (2013). STP position paper: recommended practices for sampling and processing the nervous system (brain, spinal cord, nerve, and eye) during nonclinical general toxicity studies. Toxicol. Pathol. *41*, 1028–1048. https://doi.org/10.1177/0192623312474865.

30. Politis, M., Oertel, W.H., Wu, K., Quinn, N.P., Pogarell, O., Brooks, D.J., Bjorklund, A., Lindvall, O., and Piccini, P. (2011). Graft-induced dyskinesias in Parkinson's disease: high striatal serotonin/dopamine transporter ratio. Mov. Disord. *26*, 1997–2003. https://doi.org/10.1002/mds.23743.

31. Politis, M., Wu, K., Loane, C., Quinn, N.P., Brooks, D.J., Rehncrona, S., Bjorklund, A., Lindvall, O., and Piccini, P. (2010). Serotonergic neurons mediate dyskinesia side effects in Parkinson's patients with neural transplants. Sci. Transl. Med. *2*, 38ra46. https://doi.org/10.1126/scitranslmed.3000976.

32. Rath, A., Klein, A., Papazoglou, A., Pruszak, J., Garcia, J., Krause, M., Maciaczyk, J., Dunnett, S.B., and Nikkhah, G. (2013). Survival and functional restoration of human fetal ventral mesencephalon following transplantation in a rat model of Parkinson's disease. Cell Transplant. *22*, 1281–1293. https://doi.org/10.3727/096368912X654984.

33. Li, W., Englund, E., Widner, H., Mattsson, B., van Westen, D., Lätt, J., Rehncrona, S., Brundin, P., Björklund, A., Lindvall, O., and Li, J.Y. (2016). Extensive graft-derived dopaminergic innervation is maintained 24 years after transplantation in the degenerating parkinsonian brain. Proc. Natl. Acad. Sci. USA *113*, 6544–6549. https://doi.org/10.1073/pnas.1605245113.

34. Widner, H., Tetrud, J., Rehncrona, S., Snow, B., Brundin, P., Gustavii, B., Björklund, A., Lindvall, O., and Langston, J.W. (1992). Bilateral fetal mesencephalic grafting in two patients with parkinsonism induced by 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP). N. Engl. J. Med. *327*, 1556–1563. https://doi.org/10.1056/NEJM199211263272203.

35. Lindvall, O., Rehncrona, S., Brundin, P., Gustavii, B., Astedt, B., Widner, H., Lindholm, T., Björklund, A., Leenders, K.L., Rothwell, J.C., et al. (1989). Human fetal dopamine neurons grafted into the striatum in two patients with severe Parkinson's disease. A detailed account of methodology and a 6-month follow-up. Arch. Neurol. *46*, 615–631.

36. Hardman, C.D., Henderson, J.M., Finkelstein, D.I., Horne, M.K., Paxinos, G., and Halliday, G.M. (2002). Comparison of the basal ganglia in rats, marmosets, macaques, baboons, and humans: volume and neuronal number for the output, internal relay, and striatal modulating nuclei. J. Comp. Neurol. *445*, 238–255. https://doi.org/10.1002/cne.10165.

37. Kirkeby, A., Parmar, M., and Barker, R.A. (2017). Strategies for bringing stem cell-derived dopamine neurons to the clinic: A European approach (STEM-PD). Prog. Brain Res. *230*, 165–190. https://doi.org/10.1016/bs.pbr.2016.11.011.

38. Ma, Y., Tang, C., Chaly, T., Greene, P., Breeze, R., Fahn, S., Freed, C., Dhawan, V., and Eidelberg, D. (2010). Dopamine cell implantation in Parkinson's disease: long-term clinical and (18)F-FDOPA PET outcomes. J. Nucl. Med. *51*, 7–15. https://doi.org/10.2967/jnumed.109.066811.

39. hPSCReg (2011). RCe021-A. https://hpscreg.eu/cell-line/RCe021-A.

40. Takahashi, J. (2020). iPS cell-based therapy for Parkinson's disease: A Kyoto trial. Regen. Ther. *13*, 18–22. https://doi.org/10.1016/j.reth.2020.06.002.

41. NIH (2021). Phase 1 Safety and Tolerability Study of MSK-DA01 Cell Therapy for Advanced Parkinson's Disease. https://classic.clinicaltrials.gov/ct2/show/NCT04802733.

42. Cardoso, T., Adler, A.F., Mattsson, B., Hoban, D.B., Nolbrant, S., Wahlestedt, J.N., Kirkeby, A., Grealish, S., Björklund, A., and Parmar, M. (2018). Target-specific forebrain projections and appropriate synaptic inputs of hESC-derived dopamine neurons grafted to the midbrain of parkinsonian rats. J. Comp. Neurol. *526*, 2133–2146. https://doi.org/10.1002/cne.24500.

43. Elabi, O.F., Pass, R., Sormonta, I., Nolbrant, S., Drummond, N., Kirkeby, A., Kunath, T., Parmar, M., and Lane, E.L. (2022). Human embryonic stem cell-derived dopaminergic grafts alleviate L-dopa induced dyskinesia. J. Parkinsons Dis. *12*, 1881–1896. https://doi.org/10.3233/JPD-212920.

44. Grealish, S., Heuer, A., Cardoso, T., Kirkeby, A., Jönsson, M., Johansson, J., Björklund, A., Jakobsson, J., and Parmar, M. (2015). Monosynaptic tracing using modified rabies virus reveals early and extensive circuit integration of human embryonic stem cell-derived neurons. Stem Cell Rep. *4*, 975–983. https://doi.org/10.1016/j.stemcr.2015.04.011.

45. Steinbeck, J.A., Choi, S.J., Mrejeru, A., Ganat, Y., Deisseroth, K., Sulzer, D., Mosharov, E.V., and Studer, L. (2015). Optogenetics enables functional analysis of human embryonic stem cell-derived grafts in a Parkinson's disease model. Nat. Biotechnol. *33*, 204–209. https://doi.org/10.1038/nbt.3124.

46. Tiklová, K., Nolbrant, S., Fiorenzano, A., Björklund, Å.K., Sharma, Y., Heuer, A., Gillberg, L., Hoban, D.B., Cardoso, T., Adler, A.F., et al. (2020). Single cell transcriptomics identifies stem cell-derived graft composition in a model of Parkinson's disease. Nat. Commun. *11*, 2434. https://doi.org/10.1038/s41467-020-16225-5.

47. Yang, Y., Song, J.J., Choi, Y.R., Kim, S.H., Seok, M.J., Wulansari, N., Darsono, W.H.W., Kwon, O.C., Chang, M.Y., Park, S.M., and Lee, S.H. (2022). Therapeutic functions of astrocytes to treat alpha-synuclein pathology in Parkinson's disease. Proc. Natl. Acad. Sci. USA *119*, e2110746119. https://doi.org/10.1073/pnas.2110746119.

48. Christensen, A.B., Sørensen, J.C.H., Ettrup, K.S., Orlowski, D., and Bjarkam, C.R. (2018). Pirouetting pigs: A large non-primate animal model based on unilateral 6-hydroxydopamine lesioning of the nigrostriatal pathway. Brain Res. Bull. *139*, 167–173. https://doi.org/10.1016/j.brainresbull.2018.02.010.

49. Glud, A.N., Bech, J., Tvilling, L., Zaer, H., Orlowski, D., Fitting, L.M., Ziedler, D., Geneser, M., Sangill, R., Alstrup, A.K.O., et al. (2017). A fiducial skull marker for precise MRI-based stereotaxic surgery in large animal models. J. Neurosci. Methods *285*, 45–48. https://doi.org/10.1016/j.jneumeth.2017.04.017.

50. Lillethorup, T.P., Glud, A.N., Alstrup, A.K.O., Mikkelsen, T.W., Nielsen, E.H., Zaer, H., Doudet, D.J., Brooks, D.J., Sørensen, J.C.H., Orlowski, D., and Landau, A.M. (2018). Nigrostriatal proteasome inhibition impairs dopamine neurotransmission and motor function in minipigs. Exp. Neurol. *303*, 142–152. https://doi.org/10.1016/j.expneurol.2018.02.005.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
ALDH1A1: Rabbit polyclonal anti-ALDH1A1	Abcam	Cat#ab24343, RRID:AB_2224007
FOXA2: Recombinant human monoclonal anti-FOXA2 (clone REA506), APC	Miltenyi Biotec	Cat#130-123-850, RRID:AB_2819525
FOXA2: Polyclonal goat anti-FOXA2	R&D Systems	Cat#AF2400, RRID:AB_2294104
GIRK2: Rabbit polyclonal anti-GIRK2	Alomone Labs	Cat#APC-006, RRID:AB_2040115
Ki67: Rabbit polyclonal anti-Ki67	Novus	Cat#NB110-89717, RRID:AB_1217074
KU-80: Rabbit monoclonal anti-KU80 (clone C48E7)	Cell Signaling Technology	Cat#2180, RRID:AB_2218736
LMX1A: Rabbit polyclonal anti-LMX1A	Millipore	Cat#AB10533, RRID:AB_10805970
NANOG: Mouse monoclonal anti-Nanog (clone 23D2-3C6), Alexa Fluor® 647	BioLegend	Cat#674010, RRID:AB_2632605
NCAM: Mouse monoclonal anti-NCAM (clone ERIC 1)	Santa Cruz Biotechnology	Cat#sc-106, RRID:AB_627128
NCAM: Rabbit monoclonal anti-NCAM1 (clone EP2567Y)	Abcam	Cat#ab75813, RRID:AB_2632384
OTX2: Recombinant human monoclonal anti-OTX2 (clone REA1178), VioB515	Miltenyi Biotec	Cat#130-121-193, RRID:AB_2801807
PAX6: Recombinant human monoclonal anti-PAX6 (clone REA507), PE	Miltenyi Biotec	Cat#130-123-250, RRID:AB_2819456
PDGFRB: Recombinant human anti-CD140b (clone REA363), PE	Miltenyi Biotec	Cat#130-123-772, RRID:AB_2819521
SOX1: Recombinant human monoclonal anti-SOX1 (clone REA698), APC	Miltenyi Biotec	Cat#130-111-044, RRID:AB_2653491
SOX2: Mouse monoclonal anti-SOX2 (clone O30-678), V450	BD Biosciences	Cat#561610, RRID:AB_10712763
SOX17: Mouse monoclonal anti-SOX17 (clone P7-969), Alexa Fluor® 488	BD Biosciences	Cat#562205, RRID:AB_10893402
OCT3/4: Mouse monoclonal anti-OCT3/4 (clone 40/Oct-3), PE	BD Biosciences	Cat#560186, RRID:AB_1645331
TH: Rabbit polyclonal anti-Tyrosine Hydroxylase	Millipore	Cat#AB152, RRID:AB_390204
TH: Sheep polyclonal anti-Tyrosine Hydroxylase	Millipore	Cat#AB1542, RRID:AB_90755
TH: Rabbit polyclonal anti-Tyrosine Hydroxylase	Abcam	Cat#ab112, RRID:AB_297840
VE-Cadherin: Recombinant human anti-CD144 (clone REA199), FITC	Miltenyi Biotec	Cat# 130-123-688, RRID:AB_2819510
5-HT: Rabbit polycloncal anti-5-HT antibody	Immunostar	Cat#20080 RRID: AB_572263
Chemicals, peptides, and recombinant proteins		
iPS Brew GMP Basal Medium and Supplement	Miltenyi Biotec	Cat#170-076-317, Cat#170-076-318
MACS® GMP Recombinant Human TGF-β1	Miltenyi Biotec	Cat#170-076-166
Recombinant Human Laminin 521 CTG	Biolamina	Cat#CT521

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Recombinant Human Laminin 111	Biolamina	Cat#LN111
CryoStor® CS10	Stem Cell Technologies	Cat#07930
StemPro™ Accutase™	Thermo Fisher Scientific	Cat#A1110501
UltraPure™ 0.5M EDTA	Thermo Fisher Scientific	Cat#15575020
DMEM/F-12 medium	Thermo Fisher Scientific	Cat#21331-020
CTS™ Neurobasal™ Medium	Thermo Fisher Scientific	Cat#A13712-01
L-Glutamine	Thermo Fisher Scientific	Cat#25030-081
CTS™ N-2 Supplement	Thermo Fisher Scientific	Cat#A13707-01
CTS™ B-27 Supplement without vitamin A	Thermo Fisher Scientific	Cat#A3353501
StemMACS™ Y27632	Miltenyi Biotec	Cat#130-106-538, CAS:129830-38-2
SB431542	Miltenyi Biotec	Cat#130-106-543, CAS:301836-41-9
CHIR99021	Miltenyi Biotec	Cat#130-106-539, CAS:252917-06-9
Recombinant Human Noggin, GMP	R&D Systems	Cat#6057-GMP
Recombinant Human Sonic Hedgehog (C24II), GMP	R&D Systems	Cat#1845-GMP
Recombinant Human BDNF, GMP	R&D Systems	Cat#248-GMP
Recombinant Human FGF8b, Premium Grade	R&D Systems	Cat#130-095-740
L-Ascorbic Acid	Sigma	Cat#A4403-100MG, CAS:50-81-7
Zenalb® 4.5 (Human Serum Albumin 4.5%)	Bio Products Laboratory Ltd	N/A
Zenalb® 20 (Human Serum Albumin 20%)	Bio Products Laboratory Ltd	N/A
Dimethyl Sulfoxide (DMSO)	WAK-Chemie Medical GmbH	Cat#WAK-DMSO, CAS:67-68-5
LIVE/DEAD™ Fixable Violet Dead Cell Stain	Thermo Fisher Scientific	Cat#L34955
UltraComp eBeads™ Compensation Beads	Thermo Fisher Scientific	Cat#01-2222-41
MACS® Comp Bead Kit, anti-REA	Miltenyi Biotec	Cat#130-104-693
Transcription Factor Buffer Set	BD Biosciences	Cat#562574
LightCycler® 480 SYBR Green I Master	Roche Life Sciences	Cat#04707516001
Pulmozyme® (Dornase alpha, 2500U/2.5ml) ampoules	Roche	N/A
6-OHDA	Sigma	Cat# H4381
Amphetamine	Skåne University Hospital Pharmacy	Item no. 699258
Critical commercial assays		
HumanCytoSNP-12 v2.1 BeadChip Kit	Illumina	Cat#WG-320-2101
QIAamp DNA Mini Kit	QIAGEN	Cat#51306
AmpliSeq™ Cancer HotSpot Panel v2	Illumina	Cat#20019161
RNeasy Micro Kit	QIAGEN	Cat#74004
Maxima First Strand cDNA Synthesis Kit	Thermo Fisher Scientific	Cat#K1641
StemMACS™ Trilineage Differentiation Kit	Miltenyi Biotec	Cat#130-115-660
Deposited data		
CytoSNP data on MCB and STEM-PD	This paper, https://www.ncbi.nlm.nih.gov/geo	GSE229769
Illumina cancer Hot-spot panel data on MCB	This paper, https://www.ncbi.nlm.nih.gov/geo	GSE229769
Experimental models: Cell lines		
hPSCReg RCe021-A (RC17), passage 12, 46XX	Roslin Cells Ltd (now available through University of Edinburgh)	RRID:CVCL_L206
Experimental models: Organisms/strains		
Hsd:RH-*Foxn1*^{rnu} nude rats (female)	Envigo, France	Item no. 505F
Göttingen minipig (female)	Ellegaard Minipigs ApS	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Oligonucleotides		
Primers for qRT-PCR, see Table S9	This paper	N/A
Software and algorithms		
Beeline, version 2.0.3.3	Illumina	https://support.illumina.com/downloads/beeline-software-2-0.html
BlueFuse Multi, version 4.5	Illumina	https://support.illumina.com/downloads/bluefuse-multi-software-download.html
Illumina Variant Studio, version 3.0	Illumina	https://support.illumina.com/sequencing/sequencing_software/variantstudio/downloads.html
FACSDiva™, version 9.0	BD Biosciences	https://www.bdbiosciences.com/en-us/products/software/instrument-software/bd-facsdiva-software
FlowJo™, version 10.6.2	BD Biosciences	https://www.flowjo.com/solutions/flowjo/downloads/previous-versions
GraphPad Prism 9	GraphPad Software, Inc.	https://www.graphpad.com/updates/prism-900-release-notes
Other		
STEM-PD Clinical trial design, NCT05635409	STEM-PD clinical protocol, deposited at NIH clinical trial registry	https://clinicaltrials.gov/ct2/show/NCT05635409

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Malin Parmar (malin.parmar@med.lu.se).

Materials availability

The RC17 cell line is available under an MTA agreement with the University of Edinburgh and the UK Stem Cell Bank. The GMP-grade STEM-PD product is a clinical product which is currently not available for distribution, however, research-grade versions of the product may be supplied under a collaboration agreement. This study did not generate any other new unique reagents.

Data and code availability

- Raw data for the Illumina HumanCytoSNP-12 v2.1 array and the Illumina Cancer Hot-Spot Panel v2 sequencing are deposited and publicly available at https://www.ncbi.nlm.nih.gov/geo. Accession numbers are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Cell lines

GMP-grade RC17 hESCs (hPSCReg RCe021-A, 46XX, passage 12) were obtained from Roslin Cells Ltd., and a master cell bank (MCB) was generated at passage 14 at the Royal Free Hospital (London, United Kingdom). A vial from a parallel research-grade stock of the same cell line was also obtained from Roslin Cells at passage 20 and applied for research-grade studies (i.e. single nuclear sequencing of grafts in this study).

Animals and housing conditions, GLP safety study

The GLP safety study was performed at Covance Laboratories Ltd in Huntingdon, UK, study no. RC53QY for the main toxicity, tumourigenicity and biodistribution study, and study no. 8433861 for RC17 hESC positive tumour control group. A total of 134 athymic nude rats (Hsd:RH-Foxn1^{rnu} strain from Envigo RMS UK Ltd., 68 males and 66 females) were used for these studies, and the animals were housed in individually ventilated cages (IVC), with up to 4 animals of the same sex per cage, at 20-24^{o}C with a 12h:12h light/dark cycle and unrestricted access to sterilized water and food pellets. Interventions were performed after a minimum of 16 weeks of age. All experimental groups contained male and female animals to allow for detection of potential sex-specific responses, and all study animals were subject to weekly assessments of body weight, food consumption and palpation for peripheral tumour masses.

Animals and housing conditions, non-GLP efficacy study

The non-GLP efficacy study was performed in the lab of Malin Parmar, Lund University, Sweden under ethical permit no. M-8579/2017. A total of 88 female athymic nude rats (Hsd:RH-Foxn1^{rnu} strain from Envigo RMS UK Ltd.) were used for this study, and the animals were housed in individually ventilated cages (IVC) with up to 2-3 animals per cage at room temperature with a 12h:12h light/dark cycle and unrestricted access to sterilized water and food pellets. Interventions were performed after 12 weeks of age.

Minipig model for device testing

The Danish Animal Experiments Inspectorate (2016-15-0201-00878) approved the minipig study and all procedures were conducted in compliance with the 2010/63/EU directive for animal experiments and reported according to the ARRIVE guidelines. One female Göttingen minipig (6.2 months, 16.4 kg) from Ellegaard Minipigs ApS (Dalmose, Denmark) was housed with a cage-mate (4.6 m^2, humidity 50-55% and temperature 20°C) at the Aarhus University farm with free access to tap water and fed with hay and a restricted pellet diet.

METHOD DETAILS

STEM-PD manufacturing

The RC17 MCB and the STEM-PD product were both manufactured under GMP by the Centre for Cell, Gene and Tissue Therapeutics, Royal Free Hospital (London, United Kingdom). The cells were cultured in humidified temperature- and gas-controlled incubators at 37°C and 5% CO2. For MCB manufacturing, a vial of GMP-grade RC17 hESCs (hPSCReg RCe021-A, passage 12) from Roslin Cells Ltd. (Edinburgh, United Kingdom) was thawed and expanded in iPS Brew medium (Miltenyi Biotec, GMP-grade) on laminin-521 (1 μg/cm^2, Biolamina, cell therapy grade). The cells were expanded for two passages with daily media changes and cryopreserved in CryoStor® CS10 cryopreservation medium (Stem Cell Technologies). For STEM-PD manufacturing, RC17 MCB cells were thawed on day -4 of differentiation and seeded onto laminin-521-coated T25 flasks at 13,000 cells/cm^2 in iPS Brew medium with 10 μM Y27632 (Miltenyi Biotec). The differentiation protocol applied for STEM-PD manufacturing is a scaled-up version of a previously published protocol.[27] On day 0 of differentiation, the hESCs were dissociated with 0.5 mM EDTA (Thermo Fisher Scientific) and seeded onto T75 flasks coated with laminin-111 (2 μg/cm^2, Biolamina) at 8,000 cells/cm^2 in N2 medium consisting of a 1:1 mix of DMEM/F12 medium (Thermo Fisher Scientific) and CTS Neurobasal medium (Thermo Fisher Scientific) supplemented with 1% CTS N2 (Thermo Fisher Scientific) and 2 mM L-glutamine (Thermo Fisher Scientific), with 10 μM SB431542 (Miltenyi Biotec), 100 ng/mL Noggin (R&D Systems, GMP-grade), 0.7 μM CHIR99021 (Miltenyi Biotec), 200 ng/mL SHH-C24II (R&D Systems, GMP-grade), and 10 μM Y27632. The medium was changed on days 2, 4, and 7, without Y27632. On day 9, the medium was changed to N2 medium supplemented with 100 ng/mL FGF8b (Miltenyi Biotec, premium grade). On day 11, the cells were dissociated with Accutase (Thermo Fisher Scientific) and replated onto T175 flasks coated with laminin-111 at 800,000 cells/cm^2 in B27 medium consisting of CTS Neurobasal medium, 2% CTS B27 without vitamin A (Thermo Fisher Scientific), and 2 mM L-glutamine with 100 ng/mL FGF8b, 20 ng/mL BDNF (R&D Systems, GMP-grade), 200 μM L-ascorbic acid (Sigma), and 10 μM Y27632. The medium was changed on day 14, without Y27632. After 16 days of differentiation, the cells were dissociated with Accutase and cryopreserved in vials of 5×10^6 cells in a cryopreservation medium consisting of N2 medium supplemented with 2% CTS B27 without vitamin A and 10% DMSO (WAK-Chemie Medical). MCB and STEM-PD cryovials were stored in a temperature-monitored vapour-phase nitrogen tank at the Royal Free Hospital.

Genetic testing on MCB and STEM-PD

Karyotyping analysis of expanded MCB cells was carried out under GMP by BioReliance Ltd (UK) by fixing live cells in metaphase and a total of 50 cells in metaphase were analysed by standard G-banding by microscopical inspection of the chromosomes for aberrations. STR analysis of the MCB and STEM-PD was performed by determining the alleles of 20 autosomal short tandem repeat (STR) markers (D3S1358, D1S1656, D6S1043, D13S317, Penta E, D16S539, D18S51, D2S1338, CSF1PO, Penta D, TH01, vWA, D21S11, D7S820, D5S818, TPOX, D8S1179, D12S391, D19S433 and FGA) and a sex chromosome marker (AMEL). The STR profile of STEM-PD was compared to the STR profile of the MCB and the initial RC17 cell line from Roslin Cells for confirmation of identity. For assessment of potential novel copy number variations (CNVs) or copy neutral loss of heterozygosity (CN-LOH) events, both the MCB as well as one batch of STEM-PD was analysed by CytoSNP analysis. To assess the genetic stability of the MCB and STEM-PD cells over time, the MCB was analysed directly after thawing as well as after expansion for an additional 2 and 4 passages. Likewise, STEM-PD cells were analysed after thawing and additionally after 14 days of extended culturing. The CytoSNP analysis was done using the Illumina HumanCytoSNP-12 v2.1. gDNA was extracted using a QIAamp DNA Mini Kit (QIAGEN) and arrays were scanned with an Illumina iScan System. The raw data files were analysed using Beeline 2.0.3.3 and BlueFuse Multi v4.5 software. Copy number variations (CNVs) were identified using the log R ratio and B allele frequency. CNV regions >75Kb and copy-neutral loss of heterozygosity (CN-LOH) >5Mb were reported. The list of genes encompassing each CNV/CN-LOH region was cross-referenced to public databases to look for any association with human malignancies and/or diseases. CNVs were further cross-references to the Database of Genomic Variants for known constitutive genomic variants. The presence of somatic mutations in potential oncogenes was investigated in the RC17 MCB, in two different STEM-PD GMP batches, by targeted next generation sequencing (NGS) with the Illumina Cancer Hot-Spot Panel v2. This Panel covers approximately 2,800 COSMIC mutations in 50 cancer and tumour suppressor genes: *ABL1, JAK3, AKT1, KDR, ALK, KIT, APC, KRAS, ATM, MET, BRAF,MLH1, CDH1,MPL, CDKN2A, NOTCH1, CSF1R, NPM1,*

CTNNB1, NRAS, EGFR, PDGFRA, ERBB2, PIK3CA, ERBB4, PTEN, EZH2, PTPN11, FBXW7, RB1, FGFR1, RET, FGFR2, SMAD4, FGFR3, SMARCB1, FLT3, SMO, GNA11, SRC, GNAQ, STK11, GNAS, TP53, HNF1A, VHL, HRAS, IDH1, IDH2, JAK2, with a sensitivity of 5% mosaicism. This analysis identified only benign genomic variants across all analysed samples, allowing us to conclude that there were no detectable oncogenic variants in the cell line or the manufactured product. To ensure that the RC17 cell line was free from PD-associated mutations which might compromise DA neuron function, DNA from the RC17 MCB cells was sequenced for pathogenic variants in a panel of genes known to be associated with PD and complex parkinsonism. The panel involved whole-exome sequencing of 31 PD-associated genes at a depth of minimum 20X. The genes in the panel were: *ATP13A2, ATP1A3, CSF1R, DCTN1, DNAJC6, FBXO7, FTL, GCH1, GRN, LRRK2, LYST, MAPT, OPA3, PANK2, PARK7, PINK1, PLA2G6, PRKN, PRKRA, RAB39B, SLC30A10, SLC6A3, SNCA, SPG11, SPR, SYNJ1, TH, TUBB4A, VPS13A, VPS35, WDR45*. From this analysis, no clear pathogenic variants were detected.

Flow cytometry release testing

The STEM-PD product was characterised by flow cytometry analysis using three panels of fluorophore-conjugated antibodies against on-target ventral midbrain markers FOXA2 (clone REA506, Miltenyi Biotec) together with OTX2 (clone REA1178, Miltenyi Biotec), off-target neural markers PAX6 (clone REA507, Miltenyi Biotec) together with SOX1 (clone REA698, Miltenyi Biotec), and pluripotency markers OCT3/4 (clone 40/Oct-3, BD Biosciences) together with NANOG (clone 23D2-3C6, BioLegend). Cryovials of STEM-PD were thawed with the automated ThawSTAR Thawing System (Biocision), and after washing and counting with an automated cell counter (NC-200, Chemometec), the cells were stained with a violet LIVE/DEAD™ Fixable Dead Cell Stain (Thermo Fisher Scientific) for 15 minutes at room temperature and then fixed and permeabilised using the Transcription Factor Buffer Set (BD Biosciences) according to the manufacturer's instructions. For the pluripotency panel and on- and off-target panels, 1.0×10^6 and 0.5×10^6 cells, respectively, were stained by incubation with 100 μL of an antibody cocktail per 0.5×10^6 cells in 1X Perm/Wash buffer (BD Biosciences) for 30 minutes at 4°C, washed once with 1X Perm/Wash buffer, then washed once with FACS buffer (1% bovine serum albumin in PBS without Ca^{2+} and Mg^{2+}), and finally resuspended in 200 μL of FACS buffer per 0.5×10^6 cells. The stained cells were acquired on a BD Celesta flow cytometer (BD Biosciences) with BD FACSDiva™ software (BD Biosciences) and 50,000–100,000 single live cell events were recorded per sample. Compensation was performed using UltraComp eBeads (Thermo Fisher Scientific) and Anti-REA Compensation Beads (Miltenyi Biotec). FCS files were exported and analyzed using FACS Diva™ software. Debris, doublets, and dead cells were filtered out, and the fluorescent channels were gated based on fluorescence-minus-one (FMO) control samples. Cells from the MCB were analysed using the same method but applying only the pluripotency panel (OCT3/4 and NANOG).

Quantitative real-time PCR

RNA was isolated from STEM-PD cells using RNeasy Micro kit (both from Qiagen), running on a QIAcube instrument, according to the manufacturer's procedures. Reverse transcription was performed with random hexamer primers and a Maxima First Strand cDNA Synthesis Kit (Thermo Scientific) using up to 1 μg of RNA from each sample. The complementary DNA was pipetted onto a 384-well plate, together with SYBR green Mastermix (Roche Life Sciences) and primers using an automated liquid handler (I.DOT One, Dispendix). Samples were analyzed by real-time quantitative PCR on a LightCycler 480 instrument (Roche Life Sciences) using a two-step protocol with a 60°C annealing/elongation step, for 40 cycles (Ct calculations capped at 35). All qRT-PCR samples were run in duplicate wells, and the averaged Ct values were used for calculations. For each gene, the fold change was calculated as the average fold change relative to undifferentiated RC17 hESCs using the ΔΔCt method against two different housekeeping genes (ACTB and GAPDH). Each sample was analysed at least twice in separate qRT-PCR runs, representing separate dots in the graphs. Primer sequences are provided in Table S9.

Immunocytochemistry for FOXA2 and LMX1A

STEM-PD cells were thawed and cultured for 2 days in B27 medium supplemented with 20 ng/mL BDNF, 200 μM L-ascorbic acid and 10 μM Y27632 prior to fixation and immunofluorescent labelling for key vmDA markers LMX1A and FOXA2. The cells were fixed in 4% paraformaldehyde (PFA) for 15 min and washed three times in PBS (Gibco). The fixed cells were incubated with primary antibodies, goat anti-FOXA2 (1.1000, R&D AF2400) and rabbit anti-LMX1A (1:1000, Millipore 10533) diluted in blocking solution ($PBS^{-Ca2+/-Mg2++}$ + 0.1% Triton X-100 + 5% donkey serum) at 4°C overnight followed by 3x wash, 30 minutes pre-incubation in blocking solution and 2 h incubation with a fluorophore-conjugated secondary antibody (Molecular Probes or Jackson Laboratories, 1:200 in blocking solution) at room temperature.

STEM-PD stability testing

STEM-PD batches were placed in a stability programme prior to trial initiation. The batches will be monitored at long-term storage condition in vapour phase liquid nitrogen at < -135°C, for up to 60 months. The stability programme contains tests for appearance, purity, impurity, yield, viability, sterility and endotoxin, the first five being stability indicating parameters as they are most likely to change systematically. Sterility and endotoxin are tested at a minimum at the beginning and the end of the stability study.

Trilineage test

Trilineage differentiation of MCB cells was performed using the StemMACS Trilineage kit from Miltenyi following the manufacturer's instructions, with optimised coating and seeding conditions. After differentiation, the cells were harvested using Accutase (StemPro)

and fixed using 4% formaldehyde (Cell Signaling) prior to flow cytometry analysis for detection of ectoderm (PAX6, 1:50, Miltenyi and SOX2, 1:50, BD Biosciences), endoderm (SOX17, 1:100, BD Biosciences and FOXA2, 1:80, Miltenyi) and mesoderm (CD140b, 1:50, Miltenyi and CD144, 1:50, Miltenyi). The cells were stained by incubation for 30 minutes at room temperature and protected from light. The stained cells were acquired on a BD A5 Symphony flow cytometer (BD Biosciences) with BD FACSDiva™ software (BD Biosciences) and 20,000-single live cell events were recorded per sample. Compensation was performed using UltraComp eBeads (Thermo Fisher Scientific) and Anti-REA Compensation Beads (Miltenyi Biotec). FCS files were exported and analysed using FlowJo™ software. Debris and doublets cells were gated out, and the quadrant gate on fluorescent channels was based on a control sample fixed before the start of the trilineage differentiation, i.e., the undifferentiated MCB sample.

Procedure for preparation of STEM-PD cells for transplantation and device testing

To prepare STEM-PD cells for transplantation in both the GLP safety study and the non-GLP efficacy study for device testing, we applied the same washing and resuspension procedure as is applied in the clinical trial for patient cell preparation. STEM-PD cells were thawed using a ThawStar automated cell thawing system (BioCision), and then washed twice in wash buffer consisting of Hank's Balanced Salt Solution (HBSS, Thermo Fisher or Lonza) without Ca^{2+} and Mg^{2+}, supplemented with 0.5% human serum albumin (ZENALB, Bio Products Laboratory). During a second wash, prior to centrifugation, cells were counted using an automated NC-200 Nucleocounter (ChemoMetec), and the resulting cell pellet was resuspended in delivery vehicle (HBSS without Ca^{2+} and Mg^{2+}, supplemented with 20 U/ml Pulmozyme from Roche) at a concentration of 75,000-116,000 cells/μl. During the transplantation procedure, cells were kept at a cooled environment (2-8°C) for up to 10 hours.

Surgical procedures, GLP safety study

Surgery was conducted under isofluorane gaseous anesthesia. A bupivacaine/lidocaine solution was injected subcutaneous (s.c.) as local anesthesia. The animals were placed in a stereotactic frame and adjusted to a "flat head" position (dorsoventral vertical difference between bregma and lambda is ±0.2 mm or less). Cell implantations were performed with a glass capillary attached to a blunt needle of a Hamilton syringe. Intrastriatal cell transplantations of 700,000 cells were made via two tracts with two deposits per tract (1.5 μl/deposit, 117,000 cells/μL, 1 μL/min, 2 min diffusion time per deposit) using the following coordinates: A/P: +0.9, M/L: -3.0, D/V: -4.0/-5.0 and A/P: +1.4, M/L: -2.6, D/V: -4.0/-5.0

GLP biodistribution testing

Tissues were collected from the animals from the GLP safety study at necropsy (see Table S5 for full list of collected tissues) for assessment of the presence of human DNA by qPCR analysis. Each organ was cut in half upon collection, and one half was used for histopathology whereas the other half was used for qPCR. In case of bilateral organs, one organ on each side was allocated to either qPCR or histopathology. To avoid contamination of brain samples with fibers from the intrastriatal graft cells, the most lateral 2/3 of the untransplanted hemisphere was collected for qPCR analysis to assess the potential biodistribution of human cells to the contralateral side of the brain. DNA was extracted from each tissue and 1 μg of DNA (or 5 μl of eluate in case of liquid samples) was used for each qPCR reaction. Full validation of the DNA extraction and qPCR procedure was performed prior to analysis, and the method was determined to reliably detect human DNA in a rat DNA background with a sensitivity of LoD = 0.1 human cell per 1 μg of rat DNA and LoQ = 1 human cell per 1 μg of rat DNA. Technical triplicates were performed on all qPCR samples, and data was reported as the mean value of the triplicate values. A Taqman probe was applied for AluY detection, and the PCR reaction was performed with the following primers/probe: Alu Y, forward: 5'-CCGAGGCGGGCTGATC-3', Alu Y, reverse: 5'-TCACTGTGT TAGCCAGGATGGT-3', Alu Y, probe: 5'6FAM-CTA GGT CAG GAG ATC GA-MGB 3'

Surgical procedures, efficacy study

All surgical procedures on the nude rats for the efficacy study took place under general anesthesia by intraperitoneal (i.p.) injection of a mixture of ketaminol® (ketamine hydrochloride, 45 mg/kg) and domitor® (medetomidine, 0.3 mg/kg) according to weight. Marcain® (bupivacaine, 0.1 ml) was injected s.c. as local anesthesia. A post-operative s.c. injection of Antisedan® (atipamezole, 0.28 mg/kg) and Temgesic® (buprenorphine, 0.04 mg/kg) was administered for anesthesia-reversal and analgesia, respectively. The animals were placed in a stereotactic frame and adjusted to a "flat head" position (dorsoventral vertical difference between bregma and lambda is ±0.2 mm or less). Surgical procedures were performed with a glass capillary attached to a blunt needle of a Hamilton syringe. 6-OHDA (3.5 μg/μl free base, 3 μL, 0.3 μL/min, no diffusion time) lesions were carried out in the MFB with the following coordinates A/P: -3.9, M/L: -1.2, D/V: -7.3 Intrastriatal cell transplantations were made via two tracts with two deposits per tract (1 μl/deposit, 75.000 cells/μL, 1 μL/min, 2 min diffusion time per deposit) using the following coordinates: A/P: +0.9, M/L: -3.0, D/V: -4.0/-5.0 and A/P: +1.4, M/L: -2.6, D/V: -4.0/-5.0 and intranigral in one tract A/P: -4.9, M/L: -2.3, D/V: -6.4

Amphetamine-induced rotations

Amphetamine-induced rotation tests were used to evaluate the extent of the 6-OHDA lesions at a minimum of 3 weeks after lesioning, and to assess potential recovery after transplantation. The rats were injected with amphetamine (3.5 mg/kg) i.p., placed in a harness and put in a bowl which is connected to an automated counter that records their rotations over 90 minutes. Clockwise turns were recorded with a positive value and counter-clockwise turns were recorded with a negative value. The criterion for a complete lesion was set to $\geq$5 net clockwise turns per minute over a 90-minute period. Only animals with complete lesions were included in Group 1

for assessment of transplant efficacy. The lesion was confirmed with histology post-mortem. For full motor recovery there should be a statistically significant decrease in net turns per minute on a group level combined with the requirement that more than half the group should have a net turn value of 2 or less. Animals with surviving grafts (e.g. evidence of graft detection based on hNCAM staining) and appropriate graft placement (defined as majority of cells placed in the dorso-lateral striatum) were retained in study and used to determine graft efficacy. Out of a total group size of n=20 for Group 1 with STEM-PD batch #3, 2 animals were prematurely euthanized and were not included in the rotational analysis. An additional two animals were found at post-mortem to have graft placement outside of the target site in the striatum, and their data was excluded from the rotational analyses. For STEM-PD batch #1, 5 out 26 animals from Group 1 were prematurely euthanized and 1 animal was excluded due to poor graft survival.

Immunohistochemistry, efficacy study

At the end of the experiment, the animals were terminally anaesthetised with 1.2 mL pentobarbital injected i.p. and perfused transcardially with room temperature 0.9% saline solution for 3–5 min followed by ice-cold 4% PFA solution (pH 7.4 ± 0.2) for 5 minutes. Thereafter, the brains were taken out and put in 4% PFA at 4°C for 24 hours. Subsequently, the PFA was poured off and the brains were put in 25% sucrose solution at 4°C for 2-3 days until the brains had sunk to the bottom of the vial. The brains were sectioned coronally on a freezing microtome at 35 μm thickness in a 1:8 series. Sections that were analyzed immediately were put in 0.1M KPBS and sections analyzed later were put in antifreeze and stored at -20°C until analysis.

Sections were incubated with primary antibodies overnight in 0.1 M KPBS solution containing 0.25% Triton-X and 5% serum for the species specific to the secondary antibody. Sections were then incubated with fluorophore-conjugated (fluorescent detection) or biotin-coupled (DAB detection) secondary antibodies for 1 hr in the same solution. All stained sections were mounted on gelatin-coated microscope slides. Fluorescent sections were coverslipped using polyvinyl alcohol mounting medium with DABCO (Sigma-Aldrich). DAB-developed sections were dehydrated in an ascending series of alcohols, cleared with xylene, coverslipped using DPX mountant and left to dry overnight.

The primary antibodies for DAB staining were mouse anti-hNCAM (1:1000, Santa Cruz Biotechnology sc106), rabbit anti-TH (1:2000, Merck Millipore ab152) rabbit anti-5-HT (1:10,000, Immunostar 20080). The sections were incubated with secondary biotinylated antibodies (1:200, Vector Laboratories). The primary antibodies for fluorescent immunolabeling were: rabbit anti-TH (1:2000, Merck Millipore ab152), sheep anti-TH (1:1000, Merck Millipore ab1542), rabbit anti-GIRK2 (1:2000, Alamone APC 006), goat anti-FOXA2 (1:500 R%D systems AF2400) and rabbit anti-ALDH1A1 (1:1000 Abcam AB24343). The sections were incubated with fluorophore-conjugated secondary antibodies (1:200, Jackson ImmunoResearch Laboratories)

Immunohistochemical quantification, efficacy study

Images of hNCAM stained coronal sections were taken at the level of the striatum by an Epson Perfection V850 Pro scanner. To determine the DA neuron yield, the number of DAB-stained TH+ neurons in each section was counted manually using the Olympus AX70 microscope at 20x magnification in brightfield. Final counts were adjusted for the number of series (1:8) to get an estimate of the total number of TH+ cells within the graft. The serotonergic content was assessed by the analysis of one section from the graft core per animal in the group transplanted with 300,000 cells to striatum (group 1). All quantifications were performed by a single investigator who was blinded to the treatment. For graft volume quantification, images of HuNu stained coronal sections were taken at the level of the striatum by Epson Perfection V850 Pro scanner. To determine graft volume, the area of the cellular graft core in every eighth section through the graft was measured using ImageJ (version: 2.9.0) and calibrated by associating the number of pixels with a known measurement. The graft volume was calculated according to Cavalieri's principle, given the known distance between each section and the known section thickness.

Single nucleus RNA-sequencing of research-grade grafts, and bioinformatic analysis

Individual nuclei were isolated from tissue transplanted to the striatum (3 and 6 months) or the nigra (at 6, 9, and 12 months after transplantation) using a dounce homogenizer and nuclei lysis buffer. This suspension was sorted based on size using the BD FACSAria III Cell Sorter to remove cell debris and fractured nuclei. 7-10,000 nuclei per sample were loaded onto a 10x Chromium Next GEM Chip G following the manufacturer's instructions and processed in a Chromium controller (both 10x Genomics, Pleasanton, USA). Briefly, after encapsulating single nuclei with barcoded beads, cell lysis and reverse transcription were carried out in droplets containing polyT primers with cell-specific barcodes, Unique Molecular Identifiers (UMI), and sequencing adaptor sequences. Successful library preparation was confirmed using the Bioanalyzer (DNA HS kit, Agilent), and the libraries were sequenced on the Illumina NovaSeq 6000 100 cycles flow cell (Illumina) with run settings 28-10-10-90 cycles. Cell Ranger (version 3.0, 10x Genomics) was used to demultiplex base-call files to FastQ files and align the reads. Default alignment parameters were used, and a combined human/rat reference (Both version 93 from Ensembl) was utilized. Seurat (version 4) was applied for preprocessing and downstream analysis of the snRNA-seq data. Cell cycle analysis was performed using Seurat's CellCycleScoring with default parameters, with a score > 0.4 being taken as indicative of cycling cells. The percentage of Ki67+ cells was defined as cells with $\geq$1 transcript detected for *MKI67*. Individual cell values for cell cycle scores and MKI67 transcript numbers are shown in Table S10.

Necropsy, histopathology and immunohistochemistry of animals from safety study

The animals were killed by carbon dioxide asphyxiation with subsequent exsanguination and a detailed necropsy was performed. Tissues were routinely preserved in 10% Neutral Buffered Formalin (NBF) with the exception of testes, eye and CNS fluid that

was preserved initally in Davidsons fluid and then in NBF, in Davidsons's fluid or as smears, respectively. Tissue samples were dehydrated, embedded in paraffin wax, sectioned at a nominal four to five micron thickness and stained with hematoxylin and eosin.

The non-transplanted brain hemisphere was divided into two parts – 1/3 (sagittal) for qPCR and the remaining non-transplanted hemisphere remained intact along with the transplanted hemisphere for histology. The transplanted brain hemisphere was assessed by histology in all transplanted animals. Seven coronal samples were sectioned as per "Sampling and preparation according to Bolon et al.,[29] see also Figure 2B. For the section including the graft (level 2) the brain was embedded on the rostral surface and ~2.8 mm (2800 μm or 560 x 5 μm) was discarded. For animals transplanted with 700,000 STEM-PD or RC17 hESC cells (group 2 and 3) the first and then every 5th 5 μm section up to a total of 20 sections (representing 0.1 mm or 100 μm of tissue) were collected and mounted onto slides for histopathological assessment of the graft tissue. Five interspersed sections were studied with a routine hematoxylin and eosin (H&E) stain under GLP by a study pathologist with peer review from an external pathologist. The remaining unstained sections were used for IHC investigations (non-GLP).

The staining was performed in blocking buffer consisting of 0.1M Tris-HCl with 0.15M NaCl and 0.5 % TSA® Blocking reagent (Perkin Elmer). The hNCAM antibody was a rabbit monoclonal obtained from Abcam (Ab75813) and was diluted to a final concentration of 1:800. The KU80 stain was a rabbit monoclonal antibody obtained from Cell Signalling (2180) and was diluted to a final concentration of 1:200. Positive and negative control slides were included in each run. Detection was performed with an anti-rabbit antibody and horseradish peroxidase – diaminobenzidine (DAB) detection system in an automated immunostainer (Ventana Medical Systems) which produces a brown reaction product in a nucleus staining pattern. Tissues were counterstained with haematoxylin.

The Ki67 antibody was a rabbit polyclonal antibody obtained from Novus (NB110-89717) and was diluted to a final concentration of 1:2000. Positive and negative control slides were included in each run. Detection was performed using the anti-rabbit Bright Vision detection system in an automated immunostainer (Ventana Medical Systems) which produces a purple reaction product in a cell nucleus staining pattern. Tissues were counterstained with haematoxylin. Histopathological evaluation of the IHC slides was performed by the contributing scientist and the results were compared to the routine H&E sections from the study.

In vitro device testing

A preclinical batch of VM DA progenitor cells was prepared according to the procedure outlined for STEM-PD preparation for transplantation. The R-L device was filled with 20 μl cell suspension and the STEM-PD solution was delivered as 2.5 μl deposits into 8 separate vials containing HBSS buffer with 0.5% HSA. The 8 × 2.5 μl deposits correspond to the number of deposits delivered per tract during clinical transplantation of the high dose (Dose 2). Each deposit was delivered at the same speed and inter-deposit waiting times as is applied during clinical transplantation. The number of viable cells delivered per aliquot *in vitro* was quantified using the NC200 automated cell counting system (ChemoMetec). The cell delivery process was performed at 6 different angles of the R-L device within the stereotactic frame to simulate all possible injection angles during surgery (i.e. at 0°, 10°, 15°, 20°, 25° and 30° from horizontal plane). Delivery at each angle was performed in triplicate experiments.

Transplantation to minipig

The minipig was injected unilaterally with 6-hydroxydopamine (6-OHDA) and two months later with stem cells using MR stereotaxic surgery. Eleven months post-grafting, the brain was removed and studied using immunohistochemistry. Unilateral lesion with 6-OHDA was done in the anesthetised minipig according to,[48] targeting the medial forebrain bundle with 2 injections of 25 μL of 8 μg/μL 6-OHDA hydrobromide mixed in 0.9% NaCl with ascorbic acid (Sigma Aldrich, Denmark).[48] Infusion rates were 5 μL/min and retraction of needle 1 mm/min for the first 2 mm and then full retraction. Anaesthesia and stereotaxic procedures were done according to Glud et al.[49] and Lillethorup et al.[50] The lesion was verified on post-MRI and confirmed with amphetamine-induced rotations 3 weeks later (3.9 mg/kg amphetamine).

For transplantation of STEM-PD, the minipig was anesthetised, intubated and ventilated, similar to previous reports[49,50] and kept under anaesthesia using 2% sevoflurane during the whole procedure. Coordinates were determined through 3D MRI-based guidance, and the clinical R-L device was used for the implantation procedure. The minipig received a total dose of 1,250,000 STEM-PD Batch #3 cells into 5 tracts in the right putamen (2 mm apart in anterior-posterior direction) with 2 deposits/tract (2-3 mm apart ventral-dorsal) and 2.5 μL/deposit with a cell suspension of 50,000 cells/μL (total volume of 25 μL). The infusion rate was 2.5 μL/min and diffusion time was 1 min between injections and needle retractions. The minipig received immunosuppression with 10 mg Basiliximab (Simulect) IV 2 hours pre-grafting and 4 days post-grafting. Furthermore, Tacrolimus (Envarsus) was administered twice daily (approximately 0.7 mg/kg PO, Chiesi Farmaceutici). Post-operative tacrolimus was monitored for adjustment of dose every 2-3 weeks by measuring trough blood levels (target blood range 10-15 μg/L). The minipig was perfused at 11 months post-grafting and the brain processed for histological analysis. Immunohistochemistry was performed as previously described[50] on free-floating sections incubated with rabbit anti-TH (1:1000, ab112, abcam) and mouse anti-hNCAM (1:1000, SC-106, Santa Cruz Biotechnology).

QUANTIFICATION AND STATISTICAL ANALYSIS

Procedures for graft quantification are described in sections above. Procedures for statistical analysis of rotational data in group 1 of the efficacy study were as follows: Only animals with pre-transplantation scores ≥5 were included in Group 1 for assessment of transplant efficacy. For information on excluded animals, see section on *Amphetamine-induced rotations*. Criteria for successful recovery was defined as > 50% of the animals in the tested group having a reduction in rotational scores to <2 rotations/minute

post-transplantation with a group analysis showing a significant reduction from pre-transplantation rotation scores ($p < 0.05$ by student's paired t-test, parametric) The statistical analyses were performed by using the software GraphPad Prism 8/9 and Excel.

ADDITIONAL RESOURCES

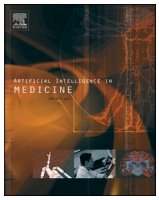
More information on the STEM-PD trial can be found here: https://stem-pd.org/. Press release from the first patient: https://www.lunduniversity.lu.se/article/first-patient-receives-milestone-stem-cell-based-transplant-parkinsons-disease. Details on the clinical trial design are available at https://clinicaltrials.gov/ct2/show/NCT05635409.

Artificial Intelligence In Medicine 147 (2024) 102698

Contents lists available at ScienceDirect

Artificial Intelligence In Medicine

journal homepage: www.elsevier.com/locate/artmed



Understanding the factors influencing acceptability of AI in medical imaging domains among healthcare professionals: A scoping review

David Hua [a,b], Neysa Petrina [a], Noel Young [c,e], Jin-Gun Cho [c,d,e], Simon K. Poon [a,d,*]

[a] School of Computer Science, The University of Sydney, Australia
[b] Sydney Law School, The University of Sydney, Australia
[c] Sydney Medical School, The University of Sydney, Australia
[d] Western Sydney Local Health District, Australia
[e] Lumus Imaging, Australia

AI ดี → ยอมรับ ใช้งานได้ ลดภาระ
→ ไม่ยอมรับ เสี่ยง work flow

ARTICLE INFO

Keywords:
Artificial intelligence
Diagnostic imaging
Acceptability
Healthcare professionals
Healthcare intervention

ABSTRACT

Background: Artificial intelligence (AI) technology has the potential to transform medical practice within the medical imaging industry and materially improve productivity and patient outcomes. However, low acceptability of AI as a digital healthcare intervention among medical professionals threatens to undermine user uptake levels, hinder meaningful and optimal value-added engagement, and ultimately prevent these promising benefits from being realised. Understanding the factors underpinning AI acceptability will be vital for medical institutions to pinpoint areas of deficiency and improvement within their AI implementation strategies. This scoping review aims to survey the literature to provide a comprehensive summary of the key factors influencing AI acceptability among healthcare professionals in medical imaging domains and the different approaches which have been taken to investigate them.

Methods: A systematic literature search was performed across five academic databases including Medline, Cochrane Library, Web of Science, Compendex, and Scopus from January 2013 to September 2023. This was done in adherence to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR) guidelines. Overall, 31 articles were deemed appropriate for inclusion in the scoping review.

Results: The literature has converged towards three overarching categories of factors underpinning AI acceptability including: user factors involving trust, system understanding, AI literacy, and technology receptiveness; system usage factors entailing value proposition, self-efficacy, burden, and workflow integration; and socio-organisational-cultural factors encompassing social influence, organisational readiness, ethicality, and perceived threat to professional identity. Yet, numerous studies have overlooked a meaningful subset of these factors that are integral to the use of medical AI systems such as the impact on clinical workflow practices, trust based on perceived risk and safety, and compatibility with the norms of medical professions. This is attributable to reliance on theoretical frameworks or ad-hoc approaches which do not explicitly account for healthcare-specific factors, the novelties of AI as software as a medical device (SaMD), and the nuances of human-AI interaction from the perspective of medical professionals rather than lay consumer or business end users.

Conclusion: This is the first scoping review to survey the health informatics literature around the key factors influencing the acceptability of AI as a digital healthcare intervention in medical imaging contexts. The factors identified in this review suggest that existing theoretical frameworks used to study AI acceptability need to be modified to better capture the nuances of AI deployment in healthcare contexts where the user is a healthcare professional influenced by expert knowledge and disciplinary norms. Increasing AI acceptability among medical professionals will critically require designing human-centred AI systems which go beyond high algorithmic performance to consider accessibility to users with varying degrees of AI literacy, clinical workflow practices, the institutional and deployment context, and the cultural, ethical, and safety norms of healthcare professions. As investment into AI for healthcare increases, it would be valuable to conduct a systematic review and meta-analysis of the causal contribution of these factors to achieving high levels of AI acceptability among medical professionals.

* Corresponding author at: School of Computer Science, The University of Sydney, Australia.
E-mail address: simon.poon@sydney.edu.au (S.K. Poon).

https://doi.org/10.1016/j.artmed.2023.102698
Received 1 March 2023; Received in revised form 29 September 2023; Accepted 29 October 2023
Available online 9 November 2023

1. Introduction and motivation

It is anticipated that the adoption of AI technology in the medical imaging industry will be a paradigm shifting trend which will radically increase the speed, quality, and value of work done by healthcare professionals [1]. AI refers to computer systems that are capable of performing tasks that ordinarily require human intelligence [2]. Machine learning is a subset of AI which involves algorithms autonomously extracting patterns and trends embedded in data to produce a mathematical model that can provide predictive outputs based on new, unseen inputs [2]. Medical AI systems of this nature have wide-ranging use cases to support clinical decision-making and optimise workflows to improve productivity and clinical outcomes [1,3,4]. Some examples of this include interpreting medical images to generate diagnostic recommendations and personalised treatment plans which act as a second medical opinion and form the basis of pre-populated preliminary medical reports, triaging patient cases based on severity, predicting patient admissions to inform resource and staffing allocation, and scheduling patient consultations. Hence, AI promises to be a valuable tool in addressing the systemic issues of increasing diagnostic imaging workloads and human error which threaten to compromise the quality of care provided by healthcare professionals in medical imaging contexts [5,6]. This is especially so for the radiology discipline which has been at the pioneering forefront of innovating and introducing AI in medical practice compared to other imaging fields [1]. Studies indicate that the average radiologist needs to interpret an image every three to four seconds in an eight-hour workday to satisfy workload requirements and that it is unlikely this onerous workload will stabilise or decline in the foreseeable future [7]. Meanwhile, misdiagnosis rates range between 3% and 5% in daily practice and is around 30% in retrospective radiologic studies [5]. The Institute of Medicine estimates that in the United States alone human error is responsible for more than 12 million misdiagnoses among adults and for 251,000 patient deaths annually [8].

Past studies and reviews concerning medical AI systems in healthcare imaging domains have largely focused on gathering evidence for the utility and safety of AI to demonstrate their feasibility (or lack thereof) for real-world deployment particularly where it concerns diagnostic use cases [1,3,4,9,10]. This often involves evaluating the technical performance of AI systems using diagnostic accuracy metrics (e.g. sensitivity, specificity, diagnostic odds ratio) and assessing their impact on task efficiency and patient health outcomes. However, most of this research has been conducted during the model development phase in controlled, laboratory settings with limited clinical studies set in real-world environments. While this research is important to establish the technical capabilities of AI as being trustworthy for use, greater emphasis needs to be placed on understanding the driving factors behind the acceptability of AI among healthcare professionals particularly as medical institutions begin to increase their investments in medical AI technologies. This is heightened by sociotechnical issues that are unique to AI technology such as poor explainability being a barrier to user understanding and trust, and the diagnostic capabilities of AI being viewed as a looming threat that will replace users rather than being an empowering tool to support them [1,3]. Accordingly, there is much research that recognises the importance of acceptability in the successful design, implementation, and value realisation of digital healthcare interventions [11–13]. Studies have shown that healthcare interventions with poor acceptability among healthcare professionals can lead to lower user uptake, a lack of meaningful and optimal value-added engagement, improper adherence with how they should be used as intended by intervention designers, and ultimately failure to realise their intended benefits even where the underlying technology was functional without error [11–13]. Indeed, many medical AI systems show promising results in theoretical lab settings but are often unsuccessful in yielding their desired benefits when deployed in medical practice because of low acceptability from healthcare professionals [14]. This has largely been caused by poor interaction design and lack of consideration for the clinical and user context among other factors [14]. Therefore, investigating the factors underpinning AI acceptability will be vital for medical institutions to pinpoint areas of deficiency and improvement within their AI implementation strategies to better ensure its intended benefits are realised.

Despite growing international interest and investment of AI in healthcare, evident in how the value of the global medical AI market is projected to increase from $13.82 billion in 2022 to $164.1 billion in 2029, research around AI acceptability in healthcare is limited [15]. The literature in this domain is sparse and fragmented which risks undermining efforts to make sense of AI acceptability. Recent reviews have been conducted that examine the literature for user perceptions and needs of AI alongside human-centred design approaches to developing AI systems to improve adoption in healthcare settings among patients and clinicians [16,17]. However, these reviews do not specifically target medical professionals in healthcare imaging fields nor do they thoroughly assess the research methodologies used to evaluate the factors underpinning the acceptability of AI. They are therefore necessarily restricted in the insights they can provide as to why different studies have reached certain results and conclusions and how researchers might improve upon these in the future. To address this gap, this scoping review aims to survey the health informatics literature to provide a comprehensive summary of the key factors influencing the acceptability of AI among medical professionals in healthcare imaging domains and to analyse the different approaches taken to investigate them. In the context of this review, the term “healthcare professionals” includes medical professionals (e.g. radiologists, radiation oncologists) who are clinicians that perform image interpretation and allied health professionals (e.g. radiographers) who typically do not perform interpretive work but may do so in some circumstances.

2. Methodology

A literature search was conducted to extract material to answer three research questions:

RQ1. What are the different ways in which acceptability has been defined or conceptualised by past studies?

RQ2. What theoretical frameworks and methodological approaches have been used to study the acceptability of AI in medical imaging domains for medical professionals?

RQ3. What are the key factors influencing the acceptability of AI in medical imaging domains for medical professionals which have been studied?

It should be noted that there is no consensus in the literature around how acceptability should be defined [12]. Four key formulations have emerged including user affective attitude towards the suitability of a system for medical usage, behavioural intention to use a system, actual system usage behaviour, and satisfaction following system usage [12]. Accordingly, acceptability in this context is not concerned with a technical perspective of the diagnostic accuracy of AI but rather the end user perspective of using AI. Restricting the ambit of this review to one specific interpretation may yield limited results given the scarcity of research in this domain. To ensure there is adequate material for analysis this scoping review will consider studies which use any of these conceptualisations of acceptability. Despite being distinctive, they all share some conceptual commonalities and therefore the insights derived from studies using different definitions of acceptability can still be informative [12].

2.1. Search strategy

An extensive search strategy was formulated in consultation with a research librarian for five academic databases including Medline, Cochrane Library, Web of Science, Compendex, and Scopus (see Supplementary File A for the search string used). The search period covered

publications from January 1, 2013 to September 15, 2023. The search strategy employed subject terms where they were available and free-text terms to address the following concepts of the scoping review:

(1) **Population:** Healthcare professionals specialising in medical imaging domains particularly radiology;
(2) **Intervention:** Medical AI technology used by healthcare professionals in a medical imaging field particularly radiology;
(3) **Outcome:** Healthcare professional acceptability of AI technology denoted by the term acceptability and closely related terms including user experience, user evaluation, implementation, integration, acceptance, satisfaction, and usability.

Given the paucity of research around the acceptability of AI among healthcare professionals, the search strategy encompassed any medical imaging domain as restricting the scope to one focus area would likely yield limited results. An emphasis was placed on the radiology discipline given its leading position in AI implementation which naturally has resulted in there being more research concerning AI in radiology compared to other diagnostic imaging fields [1]. Additional articles were selected by perusing the reference lists of relevant articles. Google Scholar was further used to identify grey literature which cited pertinent studies. The search strategy only concerns material published since 2013 as this captures the time period in which AI in the healthcare imaging industry has reached a reasonable level of technical maturity for real-world deployment and as AI technology prior to this time frame will likely be significantly different in nature (e.g. user interfaces, technical performance) [1]. The search strategy process and outcome is depicted in Fig. 1.

2.2. Criteria for inclusion and exclusion

Publications were limited to journal articles, conference proceedings, and dissertations in English. Studies were only included if they were a primary study reporting on the acceptability of AI among medical professionals in healthcare imaging domains and the factors influencing it. Those which investigated acceptability at the broader organisational level or using a multi-stakeholder approach were included as long as they explicitly accounted for the end user perspective. No restriction was placed on the temporality of the study design being prospective or retrospective (i.e. before or after healthcare professional use of AI). Articles were not considered for analysis in the scoping review if they met any of the following exclusion criteria: 1) lack of any quantitative or qualitative method of data collection such as survey instruments, interviews, and focus groups; 2) the study population did not include any healthcare professionals specialising in a medical imaging domain; 3) measurement of perceptions, attitudes, and experiences of AI were not explicitly linked to the issue of acceptability; 4) the article was about

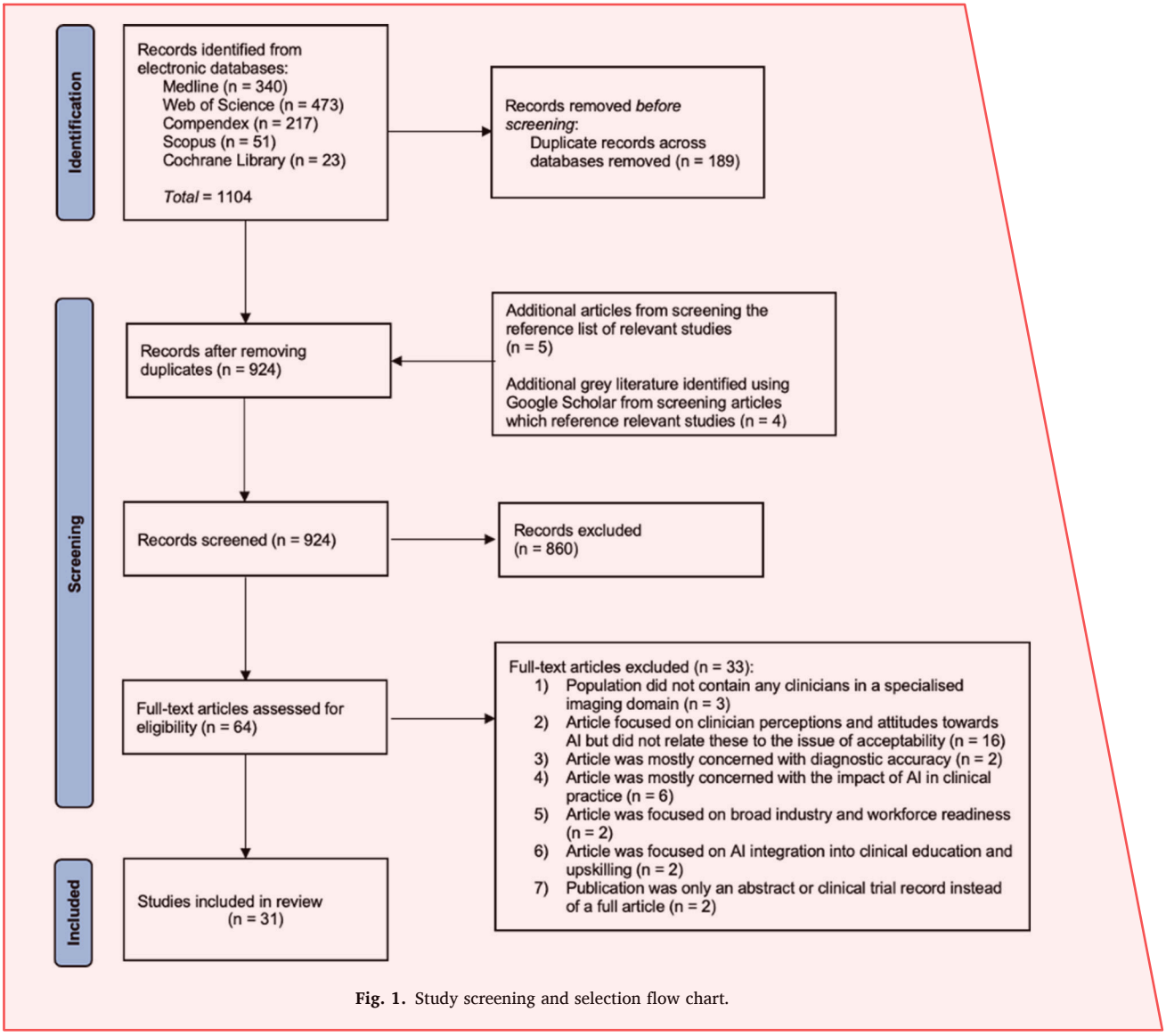


Fig. 1. Study screening and selection flow chart.

broad industry and workforce readiness; 5) the article was about AI integration into clinical education and upskilling rather than medical practice; 6) the article was predominantly or exclusively about acceptability from the perspective of the diagnostic accuracy and clinical impact of AI; and 7) the article was a review or protocol paper.

2.3. Data screening and extraction

The data screening and extraction process was completed in adherence to the PRISMA-ScR guidelines [18]. This was facilitated using the software programs of Microsoft Excel and EndNote 20. The criteria outlined above was used to screen studies firstly by title and abstract, and secondly by full-text assessment to determine whether they would be included in the scoping review. Qualitative data was collected to investigate the methodological approach adopted by each study to conceptualise and measure AI acceptance among healthcare professionals. Standardised data points were extracted from the final selection of articles including the article details (authors and their disciplinary affiliation, journal title and type, publication year) and research context (research aim and domain, type of AI system studied, setting, methodology, population, theoretical framework(s) used) which is summarised in Table 1. Thematic analysis was conducted on this collected data to answer the research questions, identify knowledge gaps in the literature, and extract emerging patterns and anomalies in the methodologies and results of the studies (e.g. how acceptability is measured by different constructs and indicators, how findings are affected by the underlying assumptions of different theoretical frameworks).

The search strategy yielded 924 articles after eliminating 189 duplicate results across the five databases. During the initial screening stage, a further five articles were included after searching the reference lists of relevant articles and an additional four grey literature publications were added after using Google Scholar to identify material citing pertinent articles. 64 out of 924 potentially relevant articles qualified for full-text assessment with 33 of these being removed under the exclusion criteria. A second researcher screened 15% of the articles by title and abstract, establishing an inter-rater reliability of 93% (125/135 agreement). Disagreements in the screening process were resolved based on discussion between the two researchers and the input of a third researcher where consensus could not be reached. Overall, the scoping review had 31 articles which were all analysed by one researcher because of resource and time constraints.

3. Results

3.1. Article characteristics

The literature search yielded 31 publications from 25 unique academic journals, two conference proceedings, and one university repository. All studies were published within the last four years (2020 to 2023 inclusive). The studies were conducted in 18 different countries, with two studies being international in scope by including multiple countries.

There is diversity in the type of journals and the backgrounds of the lead authors which can be indicative of the approach taken to investigate AI acceptability. 18 studies were published in medical journals with the majority of the lead authors having a clinical background [19–26,28–37]. This suggests a more clinically-oriented approach although Strohm et al. was the exception as Strohm's background is in innovation science which possibly reflects a more business-centric perspective [36]. Five studies were published in an information systems and management context (journal, conference, or university repository) with all first authors having this background, likely signifying a more engineering-centric approach [38–42]. Eight studies were published in interdisciplinary journals which may indicate a more holistic approach drawing from clinical and engineering perspectives [43–50]. The leader author in these papers either had an information systems or medical background but tended to collaborate with people from different disciplines.

3.2. Research aim and design

All the studies investigated AI acceptability in some capacity although their specific research aim and design varied. 24 studies sought to examine the willingness of healthcare professionals to accept the adoption of AI into medical practice and how this related to their broader perceptions, expectations, or understanding of AI; some were especially focused on exploring the nature of the facilitating and enabling factors underpinning AI acceptability while others adopted a broader institutional perspective to examine how acceptability begins at the individual level and diffuses throughout a medical organisation [19–22,24–26,28–34,36,37,39,41–43,45–48]. Four studies were empirically validating a measurement model for predicting the acceptability of AI among healthcare professionals based on established theoretical frameworks [23,38,40,50]. Three studies were a workflow-centric evaluation study seeking to investigate how healthcare professionals interact with, accept, and are affected by an AI system either in a simulated environment or real-world practice [35,44,49].

In terms of the temporality of the research design, 25 studies evaluated acceptability prospectively where participants had no opportunity to interact with a concrete AI system for the purposes of the study [19–25,28–34,37–43,45–48]. Five studies assessed acceptability retrospectively where participants had practical experience with the concrete AI system being studied [26,35,36,44,49]. Notably, only one study considered acceptability from both a prospective and retrospective perspective although this was based on whether healthcare professionals had experience with using the AI system under consideration rather than being a longitudinal study [50].

3.3. Types of AI system studied

Eight studies investigated AI acceptability in relation to an existing concrete AI system [26,35,36,39,40,44,49,50]. Six of these involved participants having practical interactions [26,35,36,44,49,50] with it while two were based on giving video demonstrations, verbal explanations, or simply informing participants of the particular system under study [39,50]. The different type of concrete AI systems studied included commercially available products (Lunit INSIGHT, BoneXpert, ATBM Master, and various diagnostic tools developed by an Israeli health technology company called Aidoc) [26,35,36,39,40], a software tool developed by a university and deployed for clinical use at a hospital [50], an application developed and validated for use at the emergency department of a health centre [49], and a prototype created for research purposes to test the usefulness of AI in medical practice [44]. The remaining 23 studies did not involve a concrete AI system which participants could interact with to form an assessment of its acceptability and instead considered AI in a general, hypothetical context. This means that these studies focused on a hypothetical AI system for use in a given healthcare context either by providing a description of the functions of a system and how it might affect medical practice, or leaving participants to their own conceptions of AI in their working environment in the absence of any such description.

3.4. Study population

The study participants encompassed healthcare professionals working in, or with, various medical imaging domains including radiology [19,20,24,29,30,34,36,40,41,45,48,49], radiography [19,21,22,33,48], dental radiography or radiology [25,31,35,43,46], radiation oncology [37,42,50], mammography [28,39,44], pathology [47], dermatology [26,32], and primary care [23]. Most studies involved clinicians and occurred in a radiological context which reaffirms how radiology is at

หมวดหมู่ของปัจจัย:

1. User Factors (ปัจจัยที่เกี่ยวข้องกับผู้ใช้แต่ละคน):
   - Trust, system understanding, AI literacy, technology receptiveness
2. System Usage Factors (ปัจจัยที่เกี่ยวข้องกับการใช้งานระบบ):
   - Value proposition, self-efficacy, burden, workflow integration
3. Socio-organizational-cultural Factors (ปัจจัยทางสังคม องค์กร และวัฒนธรรม):
   - Social influence, organizational readiness, ethicality, perceived threat

the pioneering forefront of AI research and implementation. All studies involved practicing healthcare professionals although some had prospective healthcare professionals that were either still in medical school or undergoing placement [40,48,50]. Four studies took a multi-stakeholder approach and further considered the acceptability of AI for other types of participants who might have some influence on the perspectives that healthcare professionals hold of AI such as patients, nurses, IT staff, data scientists, medical physicists, and executive management [26,31,36,39].

3.5. Data collection methodology

The distribution of the data collection methodologies of the 31 studies included 20 quantitative studies [19–25,29,32–35,37–39,43,45,46,48,50], eight qualitative studies [26,28,30,31,36,41,42,47], and three mixed-methods studies [40,44,49]. Survey instruments were the most used data collection method overall, appearing in 23 studies (74%). As shown in Table 2, eight of these studies either directly used or extended from widely validated survey instruments, while 15 either developed their own surveys in an ad-hoc manner or adapted from an existing survey that was created in an ad-hoc fashion which had limited or no validation. The remaining data collection methods were all qualitative in nature and included semi-structured interviews [30,31,36,40–42,44,47], direct observation of user interactions with AI [44], netnography [40], and focus groups [26]. Semi-structured interviews were the most used qualitative data collection method used as they appeared in eight studies while the other qualitative methods were only used once each.

3.6. How studies have defined or conceptualised acceptability

All studies converged towards measuring a conceptualisation of acceptability which concerns the willingness or behavioural intention to adopt AI for use in daily medical practice. This notion of intended user uptake at an individual, and in some cases at the broader organisational level, was not explicitly explained by most studies. It was typically conveyed or inferred by the use of cognate terms such as “adoption”, “acceptance”, “implementation”, “integration”, and “incorporation” but surprisingly no studies used the term “acceptability”. This could be because studies treat the term “acceptability” as a linguistic variation of the term “acceptance” (i.e. the base word “accept” is retained but the suffix “ance” is replaced with “ability”) and therefore view both terms as interchangeable and expressing the same underlying concept of behavioural usage intention.

3.7. Use of theoretical frameworks

Only 14 studies employed theoretical frameworks to structure or inform their investigation of AI acceptability. Technology acceptance frameworks, which seek to explain and predict the behavioural intention of individuals to use and in turn accept an innovation, were most frequently used. This covered five studies and included extensions of UTAUT [38,40,50], TAM [49], and modified TAM 3 [39]. The second most used framework was NASSS which is designed to evaluate the successful implementation of digital healthcare technologies in a way that accounts for the individual healthcare professional, organisational systems, and wider contextual factors (reflecting micro, meso, and macro level concerns) [51]. Two studies applied NASSS with one integrating it with the technology-organisation-environment (TOE) framework to focus more on institutional readiness to adopt new technologies [36,41]. Other frameworks which were each used once included: the Diffusion of Innovations (DOI) model which seeks to explain the process behind how individuals adopt an innovation and how it subsequently permeates throughout an organisation across time [30]; the Dimension of Trust (DOT) model which focuses on how acceptability is established based on the perceived capabilities of a technological system and the degree to which users understand and view it as beneficial [52]; the Barriers and Facilitators Assessment (BFA) framework which evaluates the relevance of key factors influencing successful implementation of technologies in preventive healthcare contexts based on the characteristics of the innovation, patients, healthcare professional, and use case context [53]; the integrated Theoretical Domains Framework (TDF) and Capabilities, Opportunities, and Motivations influencing Behaviours (COM-B) model which focuses on the psychological factors governing behavioural responses and change [31]; SHAPI which examines how medical professionals perceive AI based on their preparedness for AI and beliefs about its professional impact [23]; the Consolidated Framework for Implementation Research which assess contextual factors underpinning successful implementation of innovations [26]; and the Context-Mechanism-Outcome configurations framework which captures the different combinations of aspects of interventions that work and under what circumstances [47]. The remaining 17 studies did not use a theoretical framework and instead took an ad-hoc approach by relying upon their own hypotheses and domain knowledge acquired from reviewing the literature [19–22,24,25,28,29,32–35,42,43,45,46,48].

Table 2
Survey instruments used.

Survey instrument	Studies
Extensions of the Unified Theory of Acceptance and Use of Technology (UTAUT)	[38,40,50]
Modified form of the Technology Acceptance Model 3 (TAM 3)	[39]
Dimensions of Trust (DOT)	[44]
System Usability Scale (SUS)	[44,49]
NASA Task Load Index (NASA-TLX)	[44]
Barriers and Facilitators Assessment (BFA)	[37]
Shinners Artificial Intelligence Perception Tool (SHAIP)	[23]
Ad-hoc survey	[19–22,24,25,29,32–35,43,45,46,48]

3.8. Factors influencing healthcare professional acceptability of AI in medical imaging domains

There were a diverse range of factors concerning AI acceptability in medical imaging domains which were investigated. Many of these studies overlapped by substantively measuring the same underlying concept although their wording or framing of it tended to differ. For example, the idea that AI should be perceived as successfully providing meaningful value to healthcare professionals in their work, especially when compared to their current workflow, is formulated as “perceived usefulness” by TAM 3, “performance expectancy” by UTAUT, “value proposition” by NASSS, and “relative advantage” by DOI. While these are different terms, they represent a common theme around the value proposition of AI. Overall, 12 key factors for AI acceptability were identified based on emergent themes from the studies which were measured quantitatively or qualitatively by at least two studies. These factors and their proposed definitions are outlined in Table 3.

The identified factors can further be grouped into broad, overarching categories based on commonalities in what they are measuring. These include: user factors (individual end user characteristics) involving trust, system understanding, AI literacy, technology receptiveness; system usage factors (human-computer interaction and user experience concerns) entailing value proposition, self-efficacy, burden, and workflow integration; and socio-organisational-cultural factors (contextual and environmental matters) encompassing social influence, organisational readiness, ethicality, and perceived threat. The distribution of how these conditions were examined across 31 studies is shown in Fig. 2. Value proposition was the only universally explored factor which appeared across all 31 studies. Perceived threat, AI literacy, and trust were the next most considered factors as they were evaluated 24, 19, and 18 times, respectively. Workflow integration was examined 17 times,

Table 3
Key factors underpinning healthcare professional acceptability of AI inductively extracted from past studies and their proposed definitions.

Category	Factor	Definition	Studies
System usage	Value proposition	The extent to which healthcare professionals view the use of AI as providing meaningful value to their work (e.g. improving task productivity and/or diagnostic accuracy)	[19–26,28–50]
	Burden	The perceived or experienced amount of effort required to engage with AI in daily medical practice as a reflection of the complexity it adds to the existing workflow process	[26,35,38–40,43,44,47,49,50]
	Self-efficacy	The extent to which healthcare professionals are confident in their ability to perform the necessary actions to use AI in their daily workflow (e.g. interpreting and acting upon the diagnostic recommendations of AI)	[20,26,33,35,39,44,49]
	Workflow integration	The extent to which healthcare professionals perceive the integration of AI as preserving and being compatible with their existing workflow practices	[19,20,23,25,28,31,35–37,39,40,42,44,47–50]
Socio-organisational-cultural factors	Social influence	The extent to which healthcare professionals perceive that their use of AI is expected and encouraged by influencing agents important to them (e.g. colleagues, department heads, management)	[22,30,31,36,38–40,42,50]
	Organisational readiness	The extent to which healthcare professionals believe their organisation has the necessary infrastructure, resources, or processes to facilitate the use of AI	[19–21,23,26,29,30,34,36,37,41,42,45–47,50]
	Perceived threat	The extent to which healthcare professionals perceive AI as threatening their professional role at an individual and industry level compared to being an empowering tool that enables them to be better healthcare professionals	[19–26,29–46,48,50]
	Ethicality	The extent to which healthcare professionals perceive AI as fitting with their value system (e.g. whether patient data used for AI is safely managed to protect their privacy) and that of the medical profession	[20–25,29,32,41,42,47]
User factors	Technology receptiveness	The extent to which healthcare professionals are willing to try out new innovations as part of their professional role	[21,26,31,32,35,37,38,45,48]
	AI literacy	The level of understanding and knowledge which healthcare professionals have concerning AI in general as reflected by their education, training, and experience with it	[19,20,22,23,25,29–34,36–45,48–50]
	System understanding	The extent to which healthcare professionals understand the specific AI system they are using with particular focus on how it produces diagnostic outputs alongside the purpose and role of its features in relation to their workflow	[26,30,31,35,37,38,41,44,48,50]
	Trust	The extent to which healthcare professionals believe AI is safe and dependable for use based on their level of confidence in its functional characteristics (e.g. ability to consistently provide safe and accurate diagnostic recommendations) outweighing the perceived risks of its use (e.g. incorrect and misleading recommendations)	[22,25,26,28–34,36,38,40,42,44,47,48,50]

organisational readiness 16 times, and ethicality 11 times. System understanding and burden were measured 10 times while technology receptiveness and social influence were considered 9 times. Finally, self-efficacy was only examined 7 times. It should be stressed that the frequency of any given condition is not necessarily reflective of its causal significance to acceptability.

4. Discussion

4.1. Overview of results

This scoping review has outlined the key studied factors influencing healthcare professional acceptability of AI and the varying approaches that have been employed to investigate them. Despite significant interest in the healthcare applications of AI, research around AI acceptability has been somewhat limited as only 31 studies were included in the scoping review. The lack of relevant papers prior to 2020 may reflect the slow diffusion of AI in diagnostic imaging fields which historically has been caused by scepticism towards its clinical utility, limited evidence of its diagnostic accuracy, and the lack of robust regulatory regimes [54]. The sudden recent increase in papers however suggests interest in AI usage within the healthcare imaging industry is swiftly gaining momentum as the level of technical performance becomes increasingly suitable for medical use, as the market and regulatory landscape gradually consolidates, and as evidence for the applications and benefits of AI become more convincing [54]. The broad representation of countries from North and South America, Europe, the Middle-East, Asia, and Africa reflects global interest in the use of AI for healthcare imaging domains with a slim majority of studies being concentrated in the Global North (53%). Notably, these studies were all published within the last four years which strongly indicates a growing realisation of the importance of researching AI acceptability to properly realise the benefits of AI implementation in healthcare. Although conclusive judgements cannot be drawn given the highly diverse characteristics of these studies, there are clear patterns and trends which can be observed.

4.2. Commentary on studied factors underpinning AI acceptability

The identified factors, representing emergent themes from past research, provide an extensive image of what AI acceptability entails and can help researchers to avoid a limited perspective associated with ad-hoc approaches or exclusively using one theoretical framework to inform their study design. The overarching categories into which these factors are organised reflect how AI acceptability is contingent upon a dynamic interplay of factors associated with the end user, the experience of interacting with an AI system, and the broader context in which AI is deployed. These should be kept in mind by medical organisations to ensure a more comprehensive, systematic approach to examining the strengths and deficiencies of their AI implementation strategy in relation to acceptability among healthcare professionals. This will be useful to minimising the risk of important considerations being overlooked and ensuring that the key components of AI acceptability are adequately addressed.

The parsimoniousness of the identified factors could be improved by merging factors which have a sufficient level of conceptual overlap. For example, AI literacy and system understanding share considerable overlap as they are concerned with what AI is and how it works. The point of difference – that AI literacy is about knowledge of AI generally while system understanding is about concrete knowledge of the specific system in the deployed context and appreciating its role and function in the medical workflow – may be considered as insufficient to warrant keeping them as separate factors. Conversely, some factors could be partitioned into multiple factors if there was sufficient nuance distinguishing them despite having some overlap. For example,

Table 1
Characteristics of the studies included in the scoping review.

Year	Authors and disciplinary affiliation	Journal title and type	Research aim and domain	AI system studied	Setting	Methodology	Population	Theoretical framework
2023	Agrawal et al. [20] (Clinical background)	Emergency Radiology (Medical journal)	Investigate the perceptions and expectations of clinicians towards AI implementation (emergency radiology)	No reference made to a concrete AI system. About AI in a general, hypothetical context	United States	Cross-sectional quantitative study using a survey instrument (prospective)	Radiologists from members of the American Society Of Emergency Radiology (n = 113)	No theoretical framework used
2023	Hamd et al. [46] (Clinical background)	Heliyon (Interdisciplinary journal)	Assess the knowledge, perceptions, and readiness of clinicians for AI implementation (dentistry)	No reference made to a concrete AI system. About AI in a general, hypothetical context	United Arab Emirates	Cross-sectional quantitative study using a survey instrument (prospective)	Dentists practicing in hospitals and universities (n = 134)	No theoretical framework used
2023	King et al. [47] (Clinical background)	Journal of the Medical Informatics Association (Interdisciplinary journal)	Analyse the contextual factors that support or hinder uptake of AI among clinicians (pathology)	No reference made to a concrete AI system. About AI in a general, hypothetical context	United Kingdom	Qualitative study using semi-structured interviews (prospective)	Pathologists and pathology trainees from the National Health Service (n = 25)	The Context-Mechanism-Outcome framework
2023	Verma et al. [42] (Information systems and management background)	Conference on Human Factors in Computing Systems (Information systems and management conference)	Investigate the factors affecting AI adoption in clinical settings and the desired attributes of AI needed to improve uptake (oncology)	No reference made to a concrete AI system. About AI in a general, hypothetical context	Switzerland	Qualitative study using semi-structured interviews (prospective)	Oncology practitioners from the Lausanne University Hospital (n = 7)	No theoretical framework used
2023	Catalina et al. [23] (Clinical background)	Digital Health (Medical journal)	Assess the perception of healthcare professionals towards AI, the need for AI implementation, and the impact of AI on radiology (primary care with some overlap of a medical imaging context)	No reference made to a concrete AI system. About AI in a general, hypothetical context	Spain	Cross-sectional quantitative study using a survey instrument (prospective)	Primary care medical and nursing professionals from the Catalan Institute of Health which has numerous image diagnostic centres (n = 301)	Shinners AI Perception Tool
2023	Edzie et al. [45] (Clinical background)	Heliyon (Interdisciplinary journal)	Evaluate the perspectives of clinicians towards AI and their willingness to accept it (radiology)	No reference made to a concrete AI system. About AI in a general, hypothetical context	Ghana	Cross-sectional quantitative study using a survey instrument (prospective)	Radiologists (n = 77)	No theoretical framework used
2023	Haugsten et al. [26] (Clinical background)	JMIR Dermatology (Medical journal)	Evaluate the use and implementation of an AI system to understand barriers to their adoption (dermatology)	"ATBM Master" which can detect changes in moles and estimate their probability of malignancy	Denmark	Qualitative study using semi-structured interviews (retrospective)	Doctors and nurses with dermatological work experience (n = 14)	Consolidated Framework for Implementation Research
2022	Aldhafeeri [38] (Clinical background)	Insights into Imaging (Medical journal)	Examines clinician views on the medical application of AI to better understand successful system integration (radiography)	No reference made to a concrete AI system. About AI in a general, hypothetical context	Saudi Arabia	Cross-sectional quantitative study using a survey instrument (prospective)	Radiographers from public, private, and university medical hospitals (n = 562)	No theoretical framework used
2022	Rainey et al. [33] (Clinical background)	Radiography (Medical journal)	Investigate perceptions of clinicians about AI to better understand how trust and acceptance can be	No reference made to a concrete AI system. About AI in a general, hypothetical context	United Kingdom	Cross-sectional quantitative study using a survey instrument (prospective)	Reporting radiographers (n = 86)	No theoretical framework used

(*continued on next page*)

Table 1 (*continued*)

Year	Authors and disciplinary affiliation	Journal title and type	Research aim and domain	AI system studied	Setting	Methodology	Population	Theoretical framework
			improved (radiography)					
2022	Rabinovich et al. [49] (Information system and management background)	Studies in Health Technology and Informatics (Interdisciplinary journal)	Evaluate user acceptance and satisfaction of clinicians towards an AI system (radiology)	AI system developed at a health centre called "TRx" that performs automated detection of diseases in chest X-rays in an emergency department setting	Argentina	Mixed-methods study using survey instruments and semi-structured interviews (retrospective)	Radiology residents and emergency physicians at the radiology and emergency department of a university hospital (n = 13 for surveys, n = 6 for interviews)	TAM
2022	Hendrix et al. [28] (Clinical background)	Journal of the American College of Radiology (Medical journal)	To assess how different attributes of AI systems affect clinician intention to use them (mammography)	Hypothetical AI decision support tools for breast cancer detection and future breast cancer risk prediction	United States	Qualitative study using focus group interviews (prospective)	Radiologists specialising in mammography (n = 66)	No theoretical framework used
2022	Eschert et al. [25] (Clinical background)	Medicinia (Medical journal)	Assess clinician knowledge and perceptions of AI to better understand their acceptance of it (dental radiology)	No reference made to a concrete AI system. About AI in a general, hypothetical context	Germany	Cross-sectional quantitative study using a survey instrument (prospective)	Dentists with membership in the Dental Association of Westfalen-Lippe, Germany (n = 302)	No theoretical framework used
2022	Shiang et al. [35] (Clinical background)	Clinical Imaging (Medical journal)	Evaluate the real-time experiences and perceptions of clinicians towards using AI and how this impacts attitudes towards system use (radiology)	AI systems developed by Aidoc (an Israeli health technology company) for CT detection of pulmonary embolism, intracranial haemorrhage, and acute cervical spine fractures	United States	Cross-sectional quantitative study using a survey instrument (retrospective)	Radiology residents (n = 15)	No theoretical framework used
2022	Alamoudi [43] (Clinical background)	International Journal of Advanced and Applied Sciences (Interdisciplinary journal)	Assessment of clinician acceptance and willingness to use AI (radiology)	No reference made to a concrete AI system. About AI in a general, hypothetical context	Saudi Arabia	Cross-sectional quantitative study using a survey instrument (prospective)	Radiology residents and faculty radiologists (n = 55)	No theoretical framework used
2022	Pangti et al. [32] (Clinical background)	Indian Journal of Dermatology, Venerology, and Leprology (Medical journal)	Investigation into the acceptability of AI among clinicians and their attitudes and apprehensions towards it (dermatology)	No reference made to a concrete AI system. About hypothetical imaging-based AI	India	Cross-sectional quantitative study using a survey instrument (prospective)	Dermatologists and dermatology trainees (n = 166)	No theoretical framework used
2022	Calisto et al. [44] (Information systems and management background)	Artificial Intelligence in Medicine (Interdisciplinary journal)	Evaluation of AI acceptance and interaction in a simulated trial environment (mammography)	Prototype AI system called "BreastScreening-AI" that performs breast cancer diagnosis and segmentation	Portugal	Mixed-methods study using survey instruments, semi-structured interviews, and quantitative data for productivity and diagnostic accuracy analysis (retrospective)	Clinicians from public hospitals, private clinics, and cancer institutes (n = 45)	Dimensions of Trust adapted from Cai et al. [52]
2022	Abuzaid et al. [19] (Clinical background)	Academic Radiology (Medical journal)	Assessment of willingness to accept AI integration (radiology, radiography)	No reference made to a concrete AI system. About AI in a general, hypothetical context	United Arab Emirates	Cross-sectional quantitative study using a descriptive survey instrument (prospective)	Radiologists and radiographers (n = 153)	No theoretical framework used

(*continued on next page*)

Table 1 (*continued*)

Year	Authors and disciplinary affiliation	Journal title and type	Research aim and domain	AI system studied	Setting	Methodology	Population	Theoretical framework
2021	Botwe et al. [22] (Clinical background)	Radiography (Medical journal)	Assessment of perspectives towards AI integration and how future adoption can be supported (radiography)	No reference made to a concrete AI system. About AI in a general, hypothetical context	Africa	Cross-sectional quantitative study using an exploratory survey instrument (prospective)	Radiographers (n = 1020)	No theoretical framework used
2021	Zhai et al. [50] (Clinical background)	Journal of Medical Internet Research (Interdisciplinary journal)	Empirically test a measurement model to investigate the factors driving AI acceptance (radiation oncology)	AI tool which automates the contouring process for cancer patient scans. Developed at the Sun Yat-sen University Cancer Center and deployed for clinical use since 2019	China	Quantitative study using a survey instrument (prospective or retrospective depending on user exposure)	Medical students as prospective radiation oncologists, practicing radiation oncologists (n = 307)	UTAUT extended with additional constructs
2021	Mugabe [37] (Clinical background)	Technical Innovations and Patient Support in Radiation Oncology (Medical journal)	Investigation of the facilitators and barriers to adoption of AI (radiation oncology)	No reference made to a concrete AI system. About AI in a general, hypothetical context	New Zealand	Quantitative study using a survey instrument (prospective)	Radiation oncologists, medical physicists, and senior radiation therapists (n = 101)	BFA framework adapted from Harmsen et al. [53]
2021	Pumplun et al. [41] (Information systems and management background)	Hawaii International Conference on System Sciences 2021 (Information systems and management conference)	Exploration of the factors driving the AI adoption process and clinical readiness for AI (radiology)	No reference made to a concrete AI system. About AI in a general, hypothetical context	Germany	Qualitative study using semi-structured interviews (prospective)	Healthcare leadership staff from various imaging-based domains with most being radiologists (n = 15)	TOE framework integrated with NASSS framework
2021	Huisman et al. [29] (Clinical background)	European Radiology (Medical journal)	Investigation of clinician willingness for AI integration into medical practice and resident programs, and the hurdles to AI implementation (radiology)	No reference made to a concrete AI system. About AI in a general, hypothetical context	International, spanning 54 countries	Cross-sectional and multi-centre quantitative study using a survey instrument (prospective)	Radiologists and radiology residents (n = 1041)	No theoretical framework used
2021	Gerigoorian and Kloub [39] (Information systems and management background)	Combined Undergraduate Thesis in Computer and Information Sciences (Information systems and management thesis)	Examine the factors driving acceptance of a deployed AI system prior to use in a hospital environment (mammography)	Commercial AI computer-aided detection (CAD) system for supporting breast cancer detection called "Lunit INSIGHT MMG"	Sweden	Quantitative study using a survey instrument (prospective)	Radiologists, nurses, and IT staff working in a hospital breast screening unit in Stockholm, Capio St Göran Hospital (n = 28)	Modified version of TAM 3
2021	Qurashi et al. [48] (Clinical background)	Journal of Multidisciplinary Healthcare (Interdisciplinary journal)	Assessment of perceptions and acceptance levels towards medical AI integration (radiography)	No reference made to a concrete AI system. About AI in a general, hypothetical context	Saudi Arabia	Cross-sectional quantitative study using a survey instrument (prospective)	Radiographers, radiologists, clinical application specialists, internship radiography students (n = 224)	No theoretical framework used
2021	Morrison [30] (Clinical background)	Future Healthcare Journal (Medical journal)	Investigation of the barriers and enablers for the adoption of AI among staff in the National Health Service (radiology, pathology)	No reference made to a concrete AI system. About AI in a general, hypothetical context	United Kingdom	Qualitative study using semi-structured interviews (prospective)	Healthcare professionals from radiology, pathology, and general practice (n = 12)	DOI
2021	Müller et al. [31] (Clinical background)	Journal of Clinical Medicine (Medical journal)	Investigation of the barriers and enablers for the implementation and acceptability	No reference made to a concrete AI system. About AI in a general, hypothetical context	Germany	Qualitative study using semi-structured interviews (prospective)	Patients and dentists specialised in radiographic	TDF integrated with COM-B model

(*continued on next page*)

Table 1 (*continued*)

Year	Authors and disciplinary affiliation	Journal title and type	Research aim and domain	AI system studied	Setting	Methodology	Population	Theoretical framework
			of AI (dental radiographic diagnostics)				diagnosis (n = 13)	
2021	Prakash and Das [40] (Information systems and management background)	Information and Management (Information systems and management journal)	Empirically test a measurement model to investigate the factors affecting behavioural intention to use AI (radiology)	AI system called "Lunit INSIGHT" providing diagnostic recommendations for use with chest X-rays and mammography	India	Mixed-method study using netnography, semi-structured interviews, and a survey instrument (prospective)	In-training and practicing radiologists, imaging department heads (n = 15 for interviews in Study 1, n = 183 for survey for Study 2)	Modified UTAUT integrated with status quo bias and technology trust theories
2021	Coppola et al. [24] (Clinical background)	La radiologia medica (Medical journal)	Examine clinician views towards AI implementation in practice and how this relates to adoption rates (radiology)	No reference made to a concrete AI system. About AI in a general, hypothetical context	Italy	Cross-sectional quantitative study using a survey instrument (prospective)	Radiologist members of the Italian Society of Medical and Interventional Radiology (n = 1032)	No theoretical framework used
2021	Shelmerdine et al. [34] [Clinical background]	Paediatric Radiology (Medical journal)	Evaluate the attitudes and perceptions of healthcare professionals towards AI and how this relates to willingness of use (paediatric radiology)	No reference made to a concrete AI system. About AI in a general, hypothetical context	International (United Kingdom, Australia, New Zealand, Europe, United States)	Cross-sectional quantitative study using a survey instrument (prospective)	Members of various international paediatric radiology societies (n = 240)	No theoretical framework used
2020	Strohm et al. [36] (Innovation sciences background)	European Radiology (Medical journal)	Investigate the barriers and facilitators to the implementation and acceptance of AI (radiology)	Commercial X-Ray AI software called "BoneXpert" which performs automated bone maturity assessments based on paediatric hand scans of patients	Netherlands	Qualitative study using semi-structured interviews (retrospective)	Radiologists, clinical physicists, data scientists, management staff, innovation and implementation executives (n = 24)	NASSS framework
2020	Fan et al. [38] (Information systems and management background)	Annals of Operations Research (Information systems and management journal)	Investigate the factors explaining adoption of AI (medical imaging field - unspecified)	No reference made to a concrete AI system. About AI in a general, hypothetical context	China	Quantitative study using a survey instrument (prospective)	Healthcare professionals from medical imaging and medical clinical departments (n = 191)	Modified UTAUT integrated with trust theory and additional constructs

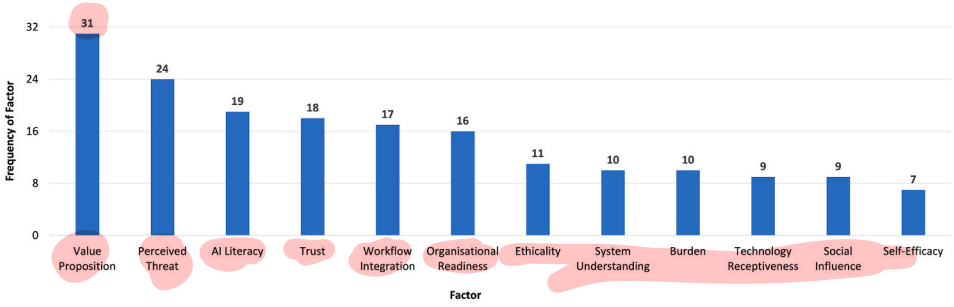


Fig. 2. Distribution of factors influencing healthcare professional acceptability of AI in the primary studies.

organisational readiness could be split into technical infrastructural readiness and organisation process readiness to address different aspects of institutional preparedness for AI. A possible approach which can address these concerns is to develop a conceptual framework designed specifically for AI acceptability which consists of unidimensional and multi-dimensional constructs. How the factors reported here are interpreted is a matter to be decided by researchers and medical organisations as the primary objective of this scoping review is to summarise what has been studied to provide a starting point for making sense of AI acceptability.

4.3. *Theoretical frameworks and ad-hoc approaches used*

Theoretical frameworks offer a structure of descriptive elements (e.g. concepts, constructs, variables) to guide research based on a formal theory or theories which provide a coherent interpretation of some

phenomenon [55]. They provide a theoretical underpinning and systematic approach to research, ensuring more comprehensive coverage of the important elements of the phenomenon of interest especially when compared to ad-hoc approaches [55]. However, they can also provide a narrow, constrained perspective because of their underlying design, assumptions, and intended use cases [55]. Many of the theoretical frameworks used by previous studies, in their original form, are arguably unsuitable for examining the intricacies and nuances of AI acceptability in healthcare. This limitation stems from the frameworks lacking consideration of healthcare-specific issues, treating AI no differently to past digital health technologies, and being consumer-centric or business-centric in their perspectives.

The majority of the frameworks used do not explicitly embed contextual healthcare and human factors which are crucial to AI being used as a healthcare intervention. Without further modification, they fail to account for critical concerns in a medical context such as the implications of AI on human life and patient safety, the qualitative and empathetic components of medical practice, organisational and professional culture, how integration into workflow processes and human-machine interactions might impact the medical decision-making process, the norms of healthcare professions especially around safety and risk, and trust and ethics which are foundational components to medical practice [56]. Only a minority of frameworks are specifically designed to consider technology in a healthcare context including NASSS, BFA, and the integrated TDF and COM-B framework, and the SHAIP.

All but one of the frameworks used are technology-agnostic and therefore treat AI the same as any other technology. This means they overlook the unique technical properties and challenges of AI (e.g. dynamic learning, extrapolating from the past to make predictions about the future, algorithmic bias) which distinguishes it from past technologies which are static and more simplistic in their behaviour. They further cannot capture nuanced sociotechnical issues that are specific to AI such as: issues of system transparency which can inhibit trust from being unable to critically interrogate the reasoning behind an output; the perceived threat of AI to professional autonomy based on fears about deskilling, replacement or redundancy, and overreliance grounded in the near-human performance of AI systems which can operate continuously at scale; and how AI literacy may affect the user interaction experience as effective AI usage arguably requires knowledge and skills that are qualitatively different from those associated with digital literacy to use general information technologies [3,17,40]. These frameworks also do not account for the status of AI as SaMD which is associated with the clinical ramifications it can have on patient health and management outcomes. This is because AI in medical imaging contexts primarily serves a medical function and naturally has more complex considerations such as medico-legal issues and requiring regulatory approval by governmental bodies. This is unlike other many digital healthcare technologies (e.g. electronic health records, telehealth) which are used to automate manual tasks or digitise workflow processes. Moreover, most of the theoretical frameworks were originally designed to study the behaviour of lay consumer and business end users where productivity and usability are the predominant concern. They do not account for healthcare professionals being the target audience who have specialist expertise and different priorities which inform their perspective on acceptability [57]. For example, studies have found that healthcare professionals tend to be more pragmatic by placing greater emphasis on factors which are important to improving outcomes and upholding patient safety (e.g. value proposition, trust, ethicality) while giving lesser weighting to factors which are not as crucial to achieving this objective (e.g. system burden, social influence) [57].

Studies using theoretical frameworks often made extensive modifications to account for these limitations, indicating that existing tools in their original state are ill-equipped to address the complexities of this domain. This does not necessarily mean that these frameworks have no value and should be avoided when investigating AI acceptability in healthcare. Although developing new frameworks is a possible option, these frameworks can still provide useful and relevant insights if they are extended to better reflect the realities of how healthcare professionals perceive and interact with AI in real-world medical settings as discussed above. Meanwhile, studies which did not use any theoretical frameworks tended to be less systematic and comprehensive in their coverage of the factors that might influence healthcare professional acceptability of AI. The use of an ad-hoc approach resulted in some studies overlooking factors that were embedded in the theoretical frameworks used (e.g. organisational readiness, social influence, burden). Yet, this approach often afforded them the flexibility to target more novel and specific aspects of AI in healthcare (e.g. workflow integration, AI literacy, ethicality) that are not embedded in the perspectives of some of these theoretical frameworks. Notably, the SHAIP was the only framework which is explicitly designed to address the use of AI in healthcare by medical professionals and captures a meaningful subset of the key factors observed in this review. However, it still overlooks some noteworthy factors (e.g. workflow integration, trust, AI literacy) which may be attributed to it only containing 10 questions [23].

4.4. The need for conceptual clarity and terminology consistency

There is a need for conceptual clarity concerning how acceptability is defined and terminology uniformity for how different terms considered synonymous or interchangeable with it are used. Rather than relying on the ordinary definition of words, studies should be explicitly clear in what they conceptually mean when they use the word “acceptability” or terminology associated with it (e.g. “adoption”, “acceptance”, “implementation”, “integration”, “incorporation”). This is important to distinguish acceptability from closely related but distinct concepts (e.g. usability, feasibility, enjoyment) which can often be conflated with acceptability [13,58]. This is particularly the case where researchers adopt a novel interpretation of acceptability based on the context of their study although all the studies in this review adopted the same conceptualisation of acceptability. Otherwise, this creates unnecessary confusion around how acceptability is being used which makes it more difficult to understand and compare the results of different studies. More uniform and clearer usage of terminology by the research community will be key to facilitating a more cohesive and consistent investigation into AI acceptability [12].

4.5. Accounting for study temporality when interpreting observations of AI acceptability

The temporality of each study (prospective or retrospective) is an important methodological factor that must be considered when interpreting the reported results on the acceptability of AI among medical professionals for each study. In prospective studies, where participants do not have the chance to practically engage with an AI system, it is possible that healthcare professionals could have a distorted perspective of AI (which could either be favourable or unfavourable) based on preconceived notions or speculative views around the value proposition of AI and how it would concretely operate in their specific clinical workflow context. In retrospective studies, where participants have interacted with a concrete AI system, the reported acceptability of AI will be grounded by the experiences of healthcare professionals using it and a contextualised understanding of how AI actually works in real-world or simulated medical practice. Therefore, the significance attributed to different factors and the outcome of AI acceptability among participants could be affected by the study temporality in a non-trivial manner. Differences in whether a study population was practically exposed to an AI system could indeed be a consequential or decisive factor in determining if AI is ultimately deemed acceptable. Future systematic reviews and meta-analyses on this topic should ensure to conduct a sub-group analysis based on temporality to better understand how AI acceptability outcomes might vary based on study temporality.

4.6. Gaps in existing research

There is some research that investigates the relationship between AI literacy and AI acceptability among healthcare professionals although there is limited work examining the interventional impact that different types of AI educational programs (e.g. self-learning, structured courses, work seminars and training) can have on improving acceptability. Studies have empirically shown that lower digital literacy is associated with more negative attitudes towards innovations in healthcare settings but that improving it can increase acceptability [59]. Validating whether this extends to AI in healthcare settings for AI literacy would be useful in informing the priority that medical organisations give to training and educating staff around AI and how precisely it should be delivered. Moreover, there is limited work examining the nature of human-AI interactions and the implications this can have on acceptability and desired outcomes. Calisto et al. was the only study to examine how system usability facilitated by human-centred design principles impacted diagnostic accuracy, productivity, and acceptability [44]. They found that the use of an AI system with high levels of acceptability on average contributed to reducing diagnosis time by three minutes, false positives by 27%, and false negatives by 4% [44]. This raises questions about the minimum level of expected improvement in outcomes needed to conclude with statistical significance the benefits of AI usage and the extent to which this can be attributed to acceptability.

In general, there is limited research which examines AI acceptability retrospectively with reference to a concrete AI system in a real-world rather than simulated setting. The scope of what can be examined is necessarily limited if considering AI prospectively in hypothetical terms (e.g. the user experience and how AI integrates into the medical workflow can only be examined meaningfully if a tangible AI system is involved) or if examining AI retrospectively in a simulated environment (e.g. organisational and cultural factors which impact how the workflow is approached is difficult to replicate in a controlled setting). It would also be worthwhile to examine AI acceptability through the lens of other formulations of acceptability. Other notable research gaps which warrant further investigation because of their implications on acceptability include: the use of explainable AI systems, perceptions of technical maturity compared to actual performance, the nature of the medical workflow and the specific deployment context, and user awareness of industry trends and the position of professional medical bodies and societies concerning AI.

4.7. Limitations

There are some limitations with this scoping review. The primary limitation concerns how the search strategy was formulated with respect to the intervention context. This review focused on the perspective of radiologists as they are most likely to be exposed to AI compared to other healthcare imaging domains given that significant AI research and development has been directed towards diagnostic radiological problems that require complex analysis of medical images. Although the search strategy was designed to include any medical imaging field, the search terms employed did not explicitly address other imaging contexts (e.g. radiography, radiation oncology, mammography) and attempted to capture all of them using the "imaging" free-text term and "Diagnostic Imaging" search term. In particular, radiography was excluded as a search term since the responsibilities of radiographers primarily involve capturing medical images rather than interpreting them and hence they were not a priority for this review despite their significant role in radiological systems. Nevertheless, this approach could have potentially caused some pertinent studies to be excluded if they were not associated with these terms in the electronic databases searched. Furthermore, only one researcher conducted the full screening process and qualitative analysis of the final set of 31 papers which could introduce some bias into the results. To help safeguard against this, another researcher screened a subset of all the papers including the final set of studies. Additionally, this review narrowly focuses on AI acceptability from the perspective of healthcare professionals and by design excludes the views of other important stakeholders (e.g. patients, nurses, hospital support staff) in healthcare imaging contexts. Finally, this research does not critically analyse the reported importance of different factors to AI acceptability for each study although this is an exercise in evidence synthesis that is more suited to a systematic review and meta-analysis.

4.8. Future work and direction

Further inquiry into AI in medical imaging domains should be pursued given the limited number of studies that exist. This will be useful from both a theoretical and practical perspective to develop a corpus of knowledge around how AI acceptability can best be achieved across varying imaging sectors to provide different medical organisations with more concrete and relevant insights to their circumstances. This will be valuable in facilitating the integration of AI into existing healthcare systems and workflows. This will require deeper consideration of questions which have been overlooked by the literature such as: What is the impact of AI on the medical workflow and operational practices? What are the different ways to approach human-AI collaboration to ensure effective co-existence and optimal augmentation of medical practice? What is the interventional impact of AI literacy on user acceptability?

A more robust study design for future research in this domain is warranted through utilising a mixed-methods approach rather than exclusively applying a qualitative or quantitative methodology which was what most studies used. The majority of studies exclusively used either semi-structured interviews (as part of a small-N qualitative study) or survey instruments (as part of a medium-N to large-N quantitative study) for data collection concerning AI acceptability. The subjective, self-reported nature of the data collected using these methods is highly susceptible to, and often skewed by, participant bias [60,61]. To improve result validity and reliability, it should be triangulated with observational data (e.g. participation observation, user experience and usability testing) alongside technical system data (e.g. task performance, keystroke and mouse click activity) to produce a more complete, objective image of acceptability. Where possible, researchers should endeavour to establish the user context and understand the nature of the concrete AI-assisted workflow for the participants being studied to better contextualise and inform their investigation of AI acceptability. This is absent from most studies and can lead to incorrect assumptions or interpretations concerning user perceptions of AI, and hinder a more nuanced analysis of results.

Studies which were empirically validating a measurement model tended to treat the factors underpinning AI acceptability as being causally independent and therefore focused on assessing the net impact of individual factors on acceptability in isolation from other factors [23,38,40,50]. This may limit the insights produced given that acceptability is likely a causally complex behavioural phenomenon driven by factors that are dynamically interacting with each other. Applying configurational approaches may be necessary to provide a combinatorial perspective of how different factors work together to holistically produce the outcome of AI acceptability. This will be beneficial to untangling the complexity underpinning different end user experiences of AI which each have different causally intertwined factors at play.

It can be expected that research in this domain will continue to expand as AI becomes increasingly used by healthcare professionals. However, future work must not neglect to investigate AI acceptability for patients in light of the growing uptake of AI-powered consumer digital health interventions (e.g. mental health chatbots, remote patient monitoring, smart health trackers). The nature of acceptability could plausibly vary between those that receive (patients) and administer (healthcare professionals) medical AI technologies because of differences in needs, preferences, and the context of the system use case [11,12]. The perspectives of medical practitioners will be significantly

shaped by their specialist expertise and professional norms, whereas the view of patients will tend to be affected more by social approval based on the recommendations of other consumers and healthcare professionals [11,12].

It would be worthwhile to perform an updated review in the coming years to see whether new factors influencing AI acceptability emerge and how methodological approaches for investigating acceptability have evolved as AI becomes increasingly integrated into medical systems. A systematic review accompanied by a meta-analysis would be particularly valuable to compile evidence for the causal impact, relative significance, and interrelationships of these factors to achieving high levels of healthcare professional acceptability of AI.

5. Conclusion

This is the first scoping review to survey the health informatics literature around the key factors influencing the acceptability of AI among healthcare professionals in medical imaging contexts and the methodological approaches adopted by past research to investigate it. It highlights the complex multiplicity of factors at the user, system usage, and socio-organisational-cultural level that need to be considered to properly address the nuances of AI acceptability in healthcare. This demonstrates how the acceptability of AI as a digital healthcare intervention is distinctive from the acceptability of other technologies which typically are targeted towards lay consumer or business end users and do not have significant implications on human welfare. To ensure a more comprehensive investigation of AI acceptability, future studies should ensure to address a meaningful combination of the key factors identified by this scoping review. As investment and research into AI for healthcare increases, it would be valuable to conduct a systematic review accompanied by a meta-analysis in the future to synthesise the empirical evidence around the contribution of these factors to achieving high levels of healthcare professional acceptability of AI.

Supplementary data to this article can be found online at https://doi.org/10.1016/j.artmed.2023.102698.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Pesapane F, Codari M, Sardanelli F. Artificial intelligence in medical imaging: threat or opportunity? Radiologists again at the forefront of innovation in medicine. Eur Radiol Exp 2018;2(1):35. 2018/10/24, https://doi.org/10.1186/s41747-018-0061-6.

[2] Do S, Song KD, Chung JW. Basics of deep learning: a radiologist's guide to understanding published radiology articles on deep learning. Korean J Radiol 2020;21(1):33–41.

[3] Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med 2019;25(1):44–56. 2019/01/01, https://doi.org/10.1038/s41591-018-0300-7.

[4] van Leeuwen KG, de Rooij M, Schalekamp S, van Ginneken B, Rutten MJCM. How does artificial intelligence in radiology improve efficiency and health outcomes? Pediatr Radiol 2021. https://doi.org/10.1007/s00247-021-05114-8. 2021/06/12.

[5] Berlin L. Radiologic errors and malpractice: a blurry distinction. Am J Roentgenol 2007;189(3):517–22. 2007/09/01, https://doi.org/10.2214/AJR.07.2209.

[6] Winder M, Owczarek AJ, Chudek J, Pilch-Kowalczyk J, Baron J. Are we overdoing it? Changes in diagnostic imaging workload during the years 2010–2020 including the impact of the SARS-CoV-2 pandemic. Healthcare 2021;9(11). https://doi.org/10.3390/healthcare9111557.

[7] McDonald RJ, et al. The effects of changes in utilization and technological advancements of cross-sectional imaging on radiologist workload (in eng) Acad Radiol Sep 2015;22(9):1191–8. https://doi.org/10.1016/j.acra.2015.05.007.

[8] I. o. Medicine, E. National Academies of Sciences, and Medicine. Improving diagnosis in health care. Washington, DC: The National Academies Press (in English); 2015. p. 472.

[9] Aggarwal R, et al. Diagnostic accuracy of deep learning in medical imaging: a systematic review and meta-analysis. npj Digit Med 2021;4(1):65. 2021/04/07, https://doi.org/10.1038/s41746-021-00438-z.

[10] Magrabi F, et al. Artificial intelligence in clinical decision support: challenges for evaluating AI and practical implications (in eng) Yearb Med Inform Aug 2019;28(1):128–34. https://doi.org/10.1055/s-0039-1677903.

[11] Sekhon M, Cartwright M, Francis JJ. Acceptability of healthcare interventions: an overview of reviews and development of a theoretical framework. BMC Health Serv Res 2017;17(1):88. 2017/01/26, https://doi.org/10.1186/s12913-017-2031-8.

[12] Perski O, Short CE. Acceptability of digital health interventions: embracing the complexity. Transl Behav Med 2021;11(7):1473–80. https://doi.org/10.1093/tbm/ibab048.

[13] Sekhon M, Cartwright M, Francis JJ. Acceptability of health care interventions: a theoretical framework and proposed research agenda (in eng) Br J Health Psychol Sep 2018;23(3):519–31. https://doi.org/10.1111/bjhp.12295.

[14] Yang Q, Zimmerman J, Steinfeld A. Review of medical decision support tools: emerging opportunity for interaction design. 2015.

[15] Artificial intelligence in healthcare market size, share & COVID-19 impact analysis, by platform (solutions and services), by application (robot-assisted surgery, virtual nursing assistant, administrative workflow assistance, clinical trials, preliminary diagnosis, and others), by end-user (hospitals & clinics, pharmaceutical & biotechnology companies, contract research organization (CRO), and Others), and regional forecasts, 2022–2029 [Online]. Available: https://www.fortunebusinessinsights.com/industry-reports/artificial-intelligence-in-healthcare-market-100534; 2022.

[16] Asan O, Choudhury A. Research trends in artificial intelligence applications in human factors health care: mapping review. JMIR Hum Factors 2021;8(2):e28236. 2021/6/18, https://doi.org/10.2196/28236.

[17] Chew HSJ, Achananuparp P. Perceptions and needs of artificial intelligence in health care to increase adoption: scoping review. J Med Internet Res 2022;24(1):e32939. 2022/1/14, https://doi.org/10.2196/32939.

[18] Tricco AC, et al. PRISMA extension for scoping reviews (PRISMA-ScR): checklist and explanation. Ann Intern Med 2018;169(7):467–73. 2018/10/02, https://doi.org/10.7326/M18-0850.

[19] Abuzaid MM, Elshami W, Tekin H, Issa B. Assessment of the willingness of radiologists and radiographers to accept the integration of artificial intelligence into radiology practice. Acad Radiol 2022;29(1):87–94. 2022/01/01/, https://doi.org/10.1016/j.acra.2020.09.014.

[20] Agrawal A, Khatri GD, Khurana B, Sodickson AD, Liang Y, Dreizin D. A survey of ASER members on artificial intelligence in emergency radiology: trends, perceptions, and expectations. Emerg Radiol 2023;30(3):267–77. https://doi.org/10.1007/s10140-023-02121-0. Article.

[21] Aldhafeeri FM. Perspectives of radiographers on the emergence of artificial intelligence in diagnostic imaging in Saudi Arabia. Insights Imaging Nov 2022;13(1):178. https://doi.org/10.1186/s13244-022-01319-z.

[22] Botwe BO, et al. The integration of artificial intelligence in medical imaging practice: perspectives of African radiographers. Radiography 2021;27(3):861–6. 2021/08/01/, https://doi.org/10.1016/j.radi.2021.01.008.

[23] Catalina QM, et al. Knowledge and perception of primary care healthcare professionals on the use of artificial intelligence as a healthcare tool. Digit Health Jun 2023;9:20552076231180511. https://doi.org/10.1177/20552076231180511.

[24] Coppola F, et al. Artificial intelligence: radiologists' expectations and opinions gleaned from a nationwide online survey (in English) Radiol Med 01 Jan 2021;126(1):63–71. https://doi.org/10.1007/s11547-020-01205-y.

[25] Eschert T, Schwendicke F, Krois J, Bohner L, Vinayahalingam S, Hanisch M. A survey on the use of artificial intelligence by clinicians in dentistry and oral and maxillofacial surgery. Medicina-Lithuania Aug 2022;58(8):1059. https://doi.org/10.3390/medicina58081059.

[26] Haugsten ER, Vestergaard T, Trettin B. Experiences regarding use and implementation of artificial intelligence-supported follow-up of atypical moles at a dermatological outpatient clinic: qualitative study (in English) JMIR Dermatol 2023;6:1. 2023 Jun 23 2023-08-31, https://doi.org/10.2196/44913.

[28] Hendrix N, et al. Radiologist preferences for artificial intelligence-based decision support during screening mammography interpretation. J Am Coll Radiol Oct 2022;19(10):1098–110. https://doi.org/10.1016/j.jacr.2022.06.019.

[29] Huisman M, et al. An international survey on AI in radiology in 1041 radiologists and radiology residents part 2: expectations, hurdles to implementation, and education. Eur Radiol 2021;31(11):8797–806. 2021/11/01, https://doi.org/10.1007/s00330-021-07782-4.

[30] Morrison K. Artificial intelligence and the NHS: a qualitative exploration of the factors influencing adoption. Future Healthc J 2021;8(3):e648. https://doi.org/10.7861/fhj.2020-0258.

[31] Müller A, Mertens SM, Göstemeyer G, Krois J, Schwendicke F. Barriers and enablers for artificial intelligence in dental diagnostics: a qualitative study. J Clin Med 2021;10(8):1612 [Online]. Available: https://www.mdpi.com/2077-0383/10/8/1612.

[32] Pangti R, Gupta S, Gupta P, Dixit A, Sati HC, Gupta S. Acceptability of artificial intelligence among Indian dermatologists. Indian J Dermatol Venereol Leprol 2022;88. https://doi.org/10.25259/IJDVL_210_2021.

[33] Rainey C, et al. UK reporting radiographers' perceptions of AI in radiographic image interpretation- current perspectives and future developments. Radiography Nov 2022;28(4):881–8. https://doi.org/10.1016/j.radi.2022.06.006.

[34] Shelmerdine SC, Rosendahl K, Arthurs OJ. Artificial intelligence in paediatric radiology: international survey of health care professionals' opinions. Pediatr Radiol 2022;52(1):30–41. https://doi.org/10.1007/s00247-021-05195-5.

[35] Shiang T, Garwood E, Debenedectis CM. Artificial intelligence-based decision support system (AI-DSS) implementation in radiology residency: introducing residents to AI in the clinical setting. Clin Imaging 2022;92:32–7. https://doi.org/10.1016/j.clinimag.2022.09.003. Article.

[36] Strohm L, Hehakaya C, Ranschaert ER, Boon WPC, Moors EHM. Implementation of artificial intelligence (AI) applications in radiology: hindering and facilitating factors. Eur Radiol 2020;30(10):5525–32. 2020/10/01, https://doi.org/10.1007/s00330-020-06946-y.

[37] Victor Mugabe K. Barriers and facilitators to the adoption of artificial intelligence in radiation oncology: a New Zealand study. Tech Innov Patient Support Radiat Oncol 2021;18:16–21. 2021/06/01/, https://doi.org/10.1016/j.tipsro.2021.03.004.

[38] Fan W, Liu J, Zhu S, Pardalos PM. Investigating the impacting factors for the healthcare professionals to adopt artificial intelligence-based medical diagnosis support system (AIMDSS). Ann Oper Res 2020;294(1):567–92. 2020/11/01, https://doi.org/10.1007/s10479-018-2818-y.

[39] Kloub M, Gerigoorian A. A cross-sectional technology acceptance study of an AI CAD system in a breast screening unit. Independent thesis Basic level (degree of Bachelor) Student thesis. TRITA-EECS-EX; 2021. p. 265. 2021. [Online]. Available: http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-299859.

[40] Prakash AV, Das S. Medical practitioner's adoption of intelligent clinical diagnostic decision support systems: a mixed-methods study. Inf Manage 2021;58(7):103524. 2021/11/01/, https://doi.org/10.1016/j.im.2021.103524.

[41] Pumplun L, Fecho M, Wahl-Islam N, Buxmann P. Machine learning systems in clinics – how mature is the adoption process in medical diagnostics?. presented at the Hawaii International Conference on System Sciences 2021. 2021.

[42] Verma H, et al. Rethinking the role of AI with physicians in oncology: revealing perspectives from clinical and research workflows. In: 2023 *CHI conference on human factors in computing systems, CHI 2023, April 23, 2023 - April 28, 2023*. Association for computing machinery, in conference on human factors in computing systems - proceedings; 2023. https://doi.org/10.1145/3544548.3581506. Hamburg, Germany. p. ACM SIGCHI; Apple; Bloomberg; Google; NSF; SIEMENS. [Online]. Available: https://doi.org/10.1145/3544548.3581506.

[43] Alamoudi A. Acceptance of artificial intelligence (AI) and machine learning (ML) among radiologists in Saudi Arabia. Int J Adv Appl Sci Jan 2022;9(1):154–7. https://doi.org/10.21833/ijaas.2022.01.018.

[44] Calisto FM, Santiago C, Nunes N, Nascimento JC. BreastScreening-AI: evaluating medical intelligent agents for human-AI interactions. Artif Intell Med 2022;127:102285. 2022/05/01/, https://doi.org/10.1016/j.artmed.2022.102285.

[45] Edzie EKM, et al. Perspectives of radiologists in Ghana about the emerging role of artificial intelligence in radiology. Heliyon 2023;9(5):e15558. https://doi.org/10.1016/j.heliyon.2023.e15558.

[46] Hamd ZY, Elshami W, Al Kawas S, Aljuaid H, Abuzaid MM. A closer look at the current knowledge and prospects of artificial intelligence integration in dentistry practice: a cross-sectional study. Heliyon Jun 2023;9(6):e17089. https://doi.org/10.1016/j.heliyon.2023.e17089.

[47] King H, Williams B, Treanor D, Randell R. How, for whom, and in what contexts will artificial intelligence be adopted in pathology? A realist interview study. J Am Med Inform Assoc Feb 2023;30(3):529–38. https://doi.org/10.1093/jamia/ocac254.

[48] Qurashi AA, Alanazi RK, Alhazmi YM, Almohammadi AS, Alsharif WM, Alshamrani KM. Saudi radiology personnel's perceptions of artificial intelligence implementation: a cross-sectional study (in eng) J Multidiscip Healthc 2021;14:3225–31. https://doi.org/10.2147/jmdh.S340786.

[49] Rabinovich D, Mosquera C, Torrens P, Aineseder M, Benitez S. User satisfaction with an AI system for chest X-ray analysis implemented in a hospital's emergency setting. Stud Health Technol Inform 2022;294:8–12. https://doi.org/10.3233/SHTI220386 [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85131107603&doi=10.3233%2fSHTI220386&partnerID=40&md5=ac73cd96bb99946966b4a33a84f9b61b.

[50] Zhai H, et al. Radiation oncologists' perceptions of adopting an artificial intelligence–assisted contouring technology: model development and questionnaire study. J Med Internet Res 2021;23(9):e27122. 2021/9/30, https://doi.org/10.2196/27122.

[51] Greenhalgh T, Abimbola S. The NASSS framework - a synthesis of multiple theories of technology implementation (in eng) Stud Health Technol Inform Jul 30 2019;263:193–204. https://doi.org/10.3233/shti190123.

[52] Cai CJ, Jongejan J, Holbrook J. The effects of example-based explanations in a machine learning interface. In: Presented at the proceedings of the 24th international conference on intelligent user interfaces, Marina del Ray, California; 2019. https://doi.org/10.1145/3301275.3302289 [Online]. Available.

[53] Harmsen M, Peters M, Wensing M. Barriers and facilitators assessment instrument: Introduction, instructions and instrument. Nijmegen: Scientific Institute for Quality of Healthcare; 2005.

[54] Alexander A, Jiang A, Ferreira C, Zurkiya D. An intelligent future for medical imaging: a market outlook on artificial intelligence for medical imaging (in eng) J Am Coll Radiol Jan 2020;17(1 Pt B):165–70. https://doi.org/10.1016/j.jacr.2019.07.019.

[55] Grant C, Osanloo A. Understanding, selecting, and integrating a theoretical framework in dissertation research: developing a 'blueprint' for your "house". Adm Issues J 01/01 2015;4. https://doi.org/10.5929/2014.4.2.9.

[56] Ward R. The application of technology acceptance and diffusion of innovation models in healthcare informatics. Health Policy Technol 2013;2(4):222–8. 2013/12/01/, https://doi.org/10.1016/j.hlpt.2013.07.002.

[57] Gücin NÖ, Berk ÖS. Technology acceptance in health care: an integrative review of predictive factors and intervention programs. Procedia Soc Behav Sci 2015;195:1698–704. 2015/07/03/, https://doi.org/10.1016/j.sbspro.2015.06.263.

[58] Bucyibaruta BJ, Peu D, Bamford L, van der Wath A. Closing the gaps in defining and conceptualising acceptability of healthcare: a synthesis review and thematic content analysis. 2020.

[59] Kuek A, Hakkennes S. Healthcare staff digital literacy levels and their attitudes towards information systems. Health Informatics J 2020;26(1):592–612. https://doi.org/10.1177/1460458219839613.

[60] Brod M, Tesler LE, Christensen TL. Qualitative research and content validity: developing best practices based on science and experience. Qual Life Res 2009;18(9):1263–78 [Online]. Available: http://www.jstor.org/stable/40539893.

[61] Kelley K, Clark B, Brown V, Sitzia J. Good practice in the conduct and reporting of survey research (in eng) International J Qual Health Care Jun 2003;15(3):261–6. https://doi.org/10.1093/intqhc/mzg031.

Radiography 30 (2024) 440–447

Contents lists available at ScienceDirect

Radiography

journal homepage: www.elsevier.com/locate/radi

Systematic Review

Workplace violence in medical radiation science: A systematic review

K.A. Shah [a], C.K.C. Ng [a, b, *]

[a] Curtin Medical School, Curtin University, GPO Box U1987, Perth, Western Australia 6845, Australia
[b] Curtin Health Innovation Research Institute (CHIRI), Faculty of Health Sciences, Curtin University, GPO Box U1987, Perth, Western Australia 6845, Australia

ARTICLE INFO

Article history:
Received 18 August 2023
Received in revised form
7 December 2023
Accepted 27 December 2023
Available online 9 January 2024

Keywords:
Bullying
Harassment
Medical sonography
Nuclear medicine
Radiation therapy
Radiography

ABSTRACT

Introduction: According to World Health Organization (WHO), workplace violence (WPV) is a significant issue in healthcare. However, no systematic review on WPV in medical radiation science (MRS) has been published yet. The purpose of this paper is to systematically review prevalence of WPV in MRS and its risk factors.
Methods: Electronic scholarly publication databases, namely EBSCOhost/Cumulative Index of Nursing and Allied Health Literature Ultimate, PubMed/Medline, ScienceDirect, Scopus, and Wiley Online Library were used for literature search to identify articles about WPV in MRS published over last 10 years as per preferred reporting items for systematic reviews and meta-analyses guidelines. To facilitate comparisons of the WPV prevalence and relative importance of individual risk factors across the included studies, their reported absolute figures of findings were used to synthesize respective percentages (if not stated).
Results: Twelve papers met the selection criteria and were included. This review shows that the WPV prevalence were 69.2–100 % (whole career) and 46.1–83.0 % (last 12 months) in diagnostic radiography, 63.0–84.0 % (whole career) in radiation therapy, 57.6 % in medical sonography (last 12 months), and 46.8 % (last 6 months) in nuclear medicine. The identified WPV risk factors included intoxicated patients, staff stress, feeling of inadequacy resulting in self-protection, more vulnerable practitioners (female, <40 years old and <5-year experience), working in radiation therapy treatment room, emergency department, examination room, general radiography, public hospital, and non-examination and waiting areas, long patient waiting time, night shift, overcrowding environment, unable to meet patients'/family members' expectations, miscommunication, patient handling, inadequate staff and security measures, interaction with colleagues, and lone working.
Conclusion: The WPV risk in diagnostic radiography and radiation therapy appears extremely high as a result of the aforementioned risk factors. Nevertheless, these study findings should be used with caution due to potential non-response bias.
Implications for practice: A WPV policy should be developed in every clinical workplace. Even if such policy is available, its enforcement including policy awareness boosting, and encouraging incident reporting and support seeking will be essential for reducing WPV. More survey studies based on WHO WPV questionnaire should be conducted for strengthening evidence base.

Introduction

According to World Health Organization (WHO), workplace violence (WPV) is a significant issue in healthcare. WPV can be defined as healthcare workers being abused, assaulted or threatened in work-related events which include but are not limited to travelling to and from workplace, and implicit or explicit conflict affecting their health, safety or wellness.[1] Common examples of WPV in healthcare include aggression, bullying, physical assault, sexual harassment, and verbal abuse and threats initiated by perpetrators such as patients, visitors and co-workers. Affected healthcare workers can experience from unnoticeable effects to fatality. Usually, these WPV incidents result in healthcare worker burnout, sleep disorder and stress, leading to reduction of their work performance as well as commitment and satisfaction. Hence, the WPV not only affects individual healthcare workers but also impacts on healthcare quality which in turn becomes academic,

* Corresponding author. Curtin Medical School, Curtin University, GPO Box U1987, Perth, Western Australia 6845, Australia.
E-mail addresses: curtise.ng@curtin.edu.au, curtise_ng@yahoo.com.hk (C.K.C. Ng).

https://doi.org/10.1016/j.radi.2023.12.018

clinical, professional, ethical, societal, political and government concerns.[1–5]

As a result of its significance, a number of systematic reviews about the WPV in healthcare have been published.[2–5] According to an umbrella review of meta-analyses covering 674,266 healthcare workers published in 2022, overall WPV prevalence was 58.7 % and commonest incidents were verbal abuse and threats (66.8 %), physical assault (20.8 %) and sexual harassment (10.5 %), respectively.[2] These findings are consistent with those of a previous systematic review and meta-analysis covering 331,544 healthcare workers published in 2019 (overall prevalence: 61.9 %, verbal abuse: 57.6 %, physical violence: 24.4 %, and sexual harassment: 12.4 %). However, it also highlighted that different countries and professions had variations of the WPV prevalence.[3] Hence, there are some literature reviews focussed on specific countries or regions, e.g. Africa,[6] Italy,[7] South-East Asian and Western Pacific Regions,[8] etc. and professions such as emergency medical services (EMS),[9] nursing[8] and pharmacy.[10]

To the best of our knowledge, no systematic review on WPV in medical radiation science (MRS) has been published yet. However, MRS is an indispensable part of modern healthcare.[11] Medical radiation practitioners (MRPs) including diagnostic radiographers, nuclear medicine technologists (NMTs), medical sonographers and radiation therapists work in various clinical areas such as radiology, nuclear medicine, emergency and radiation oncology departments, intensive care units and operating theatres.[12–23] As per the previous systematic reviews in healthcare, various clinical settings would have different risk levels. Common high risk areas that are related to MRS practice include emergency department (ED), evening shift work, and waiting room.[3–5] Although International Labour Office (ILO)/International Council of Nurses (ICN)/WHO/Public Services International (PSI) framework guidelines have provided suggested strategies to reduce the WPV in healthcare,[1] without understanding of the WPV prevalence and risk factors in MRS, effective strategies for reducing the incidents of WPV in this profession could not be determined.[2,3] Hence, it is timely to conduct a systematic review on the WPV in MRS. The purpose of this systematic review is to explore the published papers to answer the question “What was the prevalence of WPV in MRS and its risk factors?”

Methods

Preferred reporting items for systematic reviews and meta-analyses (PRISMA) guidelines were used for conducting this systematic review on the WPV in MRS.[24] Several major processes, namely literature search, article selection, and data extraction and synthesis, were involved.[11,25–27]

Literature search

The electronic scholarly publication databases, namely EBSCOhost/Cumulative Index of Nursing and Allied Health Literature (CINAHL) Ultimate, PubMed/Medline, ScienceDirect, Scopus, and Wiley Online Library, were used for literature search on 21st April 2023 to identify papers about the WPV in MRS published over the last 10 years. The search statement, (“Workplace Violence” OR “Bullying” OR “Harassment”) AND (“Radiographer” OR “Radiation Therapist” OR “Nuclear Medicine Technologist” OR “Sonographer”) was used. The publication year range was used for ensuring findings of this review relevant to current situation.[4,5] The search keywords were based on the review focus and the previous systematic reviews in healthcare.[2,3]

Article selection

Two reviewers (KAS and CKCN) independently selected the articles with any discrepancy resolved through discussion.[3,5] Only peer-reviewed original research papers written in English and focused on the WPV in any settings of MRS (against diagnostic radiographers, NMTs, medical sonographers and radiation therapists) were included. Grey literature, conference abstracts, editorials, review, perspective, opinion, commentary, and non-peer-reviewed articles were excluded. This is because well-developed methodological guidelines for appropriate selection of the grey literature were unavailable. Also, complete study information was not given in the conference abstracts. The editorials, review, perspective, opinion, and commentary papers only provided secondary information. Unsubstantiated information was presented in the non-peer-reviewed papers.[11,25–27]

The article selection process is shown in Fig. 1. After duplicate articles were removed from the results of the database search, article titles, abstracts and full texts were assessed against the selection criteria subsequently. Each non-duplicate paper within the search results was kept unless a decision on its removal could be made. Additional, relevant papers were identified by checking references lists of the included articles.[11,25–27]

Data extraction

Two systematic and one scoping reviews on the WPV in healthcare were used to devise a data extraction form (Table 1).[4,6,7] The data, namely author name and country, publication year, MRS division (diagnostic radiography, medical sonography, nuclear medicine or radiation therapy), study purpose and design, sample size and characteristics, and key findings, were extracted from the included papers.

Data synthesis

To facilitate comparisons of the WPV prevalence and relative importance of individual risk factors across the included studies, their reported absolute figures of findings were used to synthesize respective percentages (if not stated) as per the PRISMA guidelines.[24] Quality assessment tool for studies with diverse designs (QATSDD) was used to determine quality percentage of each article with <50 %, 50–70 % and >70 % representing low, moderate and high study quality, respectively.[11,25,27,28] The study quality percentages are presented in Table 1 as well.

Results

Twelve papers which met the selection criteria were included in this review. Table 1 shows the characteristics of these studies.[12–23] All but one study investigated the prevalence of WPV with the use of questionnaire,[12,13,15–23] and the only exception focused on the risk factors of WPV determined through a qualitative approach (interview).[14] Nonetheless, the WPV risk factors were also covered in all survey studies except the one by Trad and Johnson.[12,13,15–23] The included studies were mainly about the WPV situations in Africa ($n = 4$),[13–15,18] and North America ($n = 4$).[12,19,21,23] A quarter of them focused on the United States of America (USA) situation.[12,21,23] Two thirds of the included studies were published from 2019, indicating an increase of researchers' attention to the WPV issue recently.[12–16,18,20,23] Also, two thirds of them focussed on diagnostic radiography[12–15,17,18,20,22] with the others about radiation therapy,[19,21] medical sonography,[16] and nuclear medicine,[23] reflecting the typical characteristics of MRS.

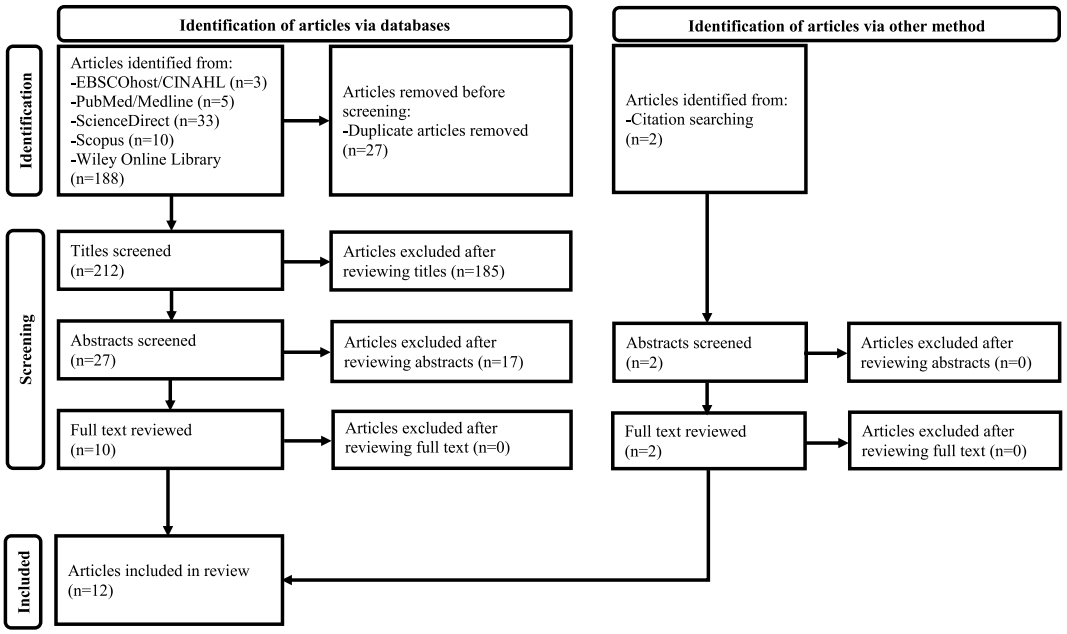


Figure 1. PRISMA flow diagram for systematic review of workplace violence in medical radiation science.

The overall WPV prevalence in diagnostic radiography was 69.2–100 % for radiographers' whole career[12,15,18,20] while it was 46.1–83.0 % for last 12 months.[13,17,22] For radiation therapy, 63.0–84.0 % of respondents experienced WPV in their career.[19,21] Similar prevalence percentages are noted in the medical sonography (57.6 % in last 12 months)[16] and nuclear medicine studies (46.8 % in last 6 months).[23] The prevalence of common WPV types were verbal abuse and threats (32.0–100 %),[12,13,15–20,22,23] sexual harassment (10.3–84.6 %),[12,15,19,22] and physical assault (3.0–51.0 %).[12,13,15–19,22] The WPV risk factors identified by the highest proportions of studies' participants included intoxicated patients (100 %),[15] staff stress (80.0 %), feeling of inadequacy resulting in self-protection (80.0 %),[17] more vulnerable MRPs (less than 40 years old [77.3 %][20] and female [68.9 %]),[23] working in radiation therapy treatment room (68.0 %),[19] long patient waiting time (61.5 %),[15] working in ED (56.5 %), examination room (54.1 %), and general radiography (46.8 %),[22] less than 5-year work experience (46.4 %), night shift (43.6 %),[20] working in public hospital (39.4 %),[16] overcrowding environment (30.8 %),[15] working in non-examination and waiting area (23.9 %),[22] unable to meet patients'/family members' expectations (23.1 %),[15] miscommunication (21.0 %),[22] patient handling (14.0 %),[18] inadequate staff (14.0 %),[22] inadequate security measures (12.6 %),[20] interaction with colleagues (12.0 %),[18] and lone working (11.8 %).[22] Although the major focus of all included studies was the WPV prevalence and/or risk factors, some reported the WPV impacts such as more irritable (65.0 %), difficult to concentrate (63.0 %),[21] affecting patient care (57.3 %),[23] more forgetful (46.1 %),[21] anxiety and stress (26.4 %), work motivation reduction (25.2 %), lower energy level (21.0 %),[22] and decrease of self-worth (24.5 %),[21] and coping strategies, e.g., nothing (72.7 %),[15] experience sharing with colleagues and family members (27.3 %),[15] calming down perpetrators (19.1 %), seeking help from colleagues (13.6 %), self-defence (12.7 %), being angry (12.7 %), and legal action (9.1 %) as well.[20]

For the quality of included studies, all but two were at least moderate.[12–23] Excluding the qualitative study by Chinene et al.,[14] their sample sizes ranged between 13 and 870 with the median size of 100.[12,13,15–23] Also, a wide range of response rate, 1.9–100 % is noted.[12,13,15,17–23]

Key risk factors identified include intoxicated patients, staff stress, feeling of inadequacy, and working in high-risk areas like emergency departments and radiation therapy rooms. Impacts of WPV include irritability, difficulty concentrating, and reduced patient care quality. Coping strategies range from doing nothing to seeking help and legal action.

Discussion

WPV prevalence

To the best of our knowledge, this is the first systematic review on the WPV in MRS. When compared with the overall prevalence figures from the two systematic reviews on the WPV in healthcare published in 2022 (58.7 %)[2] and 2019 (61.9 %),[3] the WPV prevalence in diagnostic radiography (69.2–100 %) and radiation therapy (63.0–84.0 %) reported by the included studies seems concerning.[12,15,18–21] Although the WPV prevalence range for diagnostic radiography over the last 12 months (46.1–83.0 %) was lower and more comparable to those figures for healthcare in general, the WPV prevalence for diagnostic radiographers' whole career should not be ignored because the aforementioned systematic reviews combined the prevalence figures for the 12-month and whole career periods from individual studies to determine the overall WPV prevalence in healthcare.[2,3,12,13,15,17,18,20,22] Usually, the literature including ILO/ICN/WHO/PSI framework guidelines for addressing WPV in the health sector indicates that nurses encounter WPV more frequently than other healthcare professionals.[1,3,5,8] As per the ILO/ICN/WHO/PSI framework guidelines, both nurses and EMS responders are classified as extremely high risk professions with regard to the WPV.[1] However, Varghese et al.'s systematic review on the WPV in nursing which covered 13 countries with 42,222 participants published in 2022 showed that the overall prevalence was only 58 %,[8] matching the figures for the whole healthcare sector reported by Sahebi et al.[2] (58.7 %) and Liu et al. (61.9 %).[3] It is well known that nursing is the largest healthcare workforce.[29–31] Hence, the figures of WPV prevalence in healthcare would be more representative for this profession.[2,3] Nonetheless, the overall WPV prevalence in diagnostic radiography (69.2–100 %) and radiation therapy (63.0–84.0 %) appears more similar to the one reported in a systematic review on the WPV in EMS (57.0–93.0 %) published in 2020 which covered 104 articles.[9] Based on the ILO/ICN/WHO/PSI framework guidelines, the WPV risk in diagnostic radiography and radiation therapy should be considered extremely high.[1]

Table 1
Characteristics of workplace violence (WPV) studies in medical radiation science (MRS).

Author, year and country	MRS division	Study purpose	Study design	Sample size and characteristics	Key findings	Quality
Beam et al. (2022)-United States of American (USA)[12]	Diagnostic radiography	Determination of prevalence of WPV	- Prospective - Cross-sectional - Questionnaire (unvalidated, not based on literature but with piloting)	- 193 out of 10,000 randomly selected radiographers in USA - Response rate: 1.9 %	- 69.2 % experienced WPV in their career - Prevalence of WPV types: verbal intimidation by patients (68 %) and physicians (60 %), insulted by visitors (56 %), sexual harassment by patients (55 %), and verbally intimated by visitors (54 %), physical assault by patients (51 %), and threat by patients (47 %) - Risk factor: work environment (increased chance of insult by visitors in CT, adult ED and MRI)	High (71.4 %)
Chinene et al. (2022)-Zimbabwe[13]	Mainly diagnostic radiography but also covering medical sonography, nuclear medicine, and radiation therapy	Determination of risk factors of WPV	- Prospective - Cross-sectional - Questionnaire (unvalidated but based on established scale and literature with piloting)	- 100 out of 110 randomly selected radiographers of 3 central hospitals in Zimbabwe - Response rate: 91 %	- 83 % experienced WPV in last 12 months - Prevalence of WPV types: verbal abuse (81 %), sexual abuse (21 %), and physical abuse (4 %) - Risk factors (scale 1–5): poor working conditions leading to frustration ($\bar{x}$: 3.93), long patient waiting time ($\bar{x}$: 3.91), power imbalance ($\bar{x}$: 3.87), burnout/fatigue ($\bar{x}$: 3.79), narcissism ($\bar{x}$: 3.79), communication style difference ($\bar{x}$: 3.68), opinion difference ($\bar{x}$: 3.59), personal/family issues ($\bar{x}$: 3.57), inadequacy feeling leading to self-protection ($\bar{x}$: 3.42); poor workplace culture ($\bar{x}$: 3.38) and bias caused by culture/generation/gender difference ($\bar{x}$: 3.32)	Moderate (64.3 %)
Chinene et al. (2022)-Zimbabwe[14]	Diagnostic radiography	Determination of risk factors of WPV	- Prospective - Exploratory qualitative - Individual semi-structured interview (guide based on literature with piloting)	11 radiography managers purposively selected from 3 hospitals in Zimbabwe	Risk factors: work environment (lack of leadership trust, radiographers' burnout/fatigue and low salary), power hierarchy (feeling of superiority, inappropriate professional boundary crossing and inadequate radiographer representation in healthcare), and lack of reporting framework (protocol and culture)	High (76.2 %)
Hattingh et al. (2019)-Namibia[15]	Diagnostic radiography	Determination of prevalence of WPV and coping strategies	- Prospective - Cross-sectional - Questionnaire (unvalidated, not based on literature but with piloting)	13 out of 15 (all) night shift radiographers of 1 state radiology department in Namibia - Response rate: 86.7 %	- 100 % experienced WPV in their career - Prevalence of WPV types: verbal abuse (100 %) and threats (84.6 %), sexual harassment (84.6 %), and physical assault (46.2 %) - Risk factors: patients with intoxication (100 %), long patient waiting time (61.5 %), overcrowding environment (30.8 %) and unable to meet expectations of patients/family members (23.1 %) - Coping strategies: Nothing (72.7 %), and experience sharing with colleagues and family members (27.3 %)	Low (42.9 %)
Lloyd–Jones et al. (2021)-Australia[16]	Medical sonography	Determination of prevalence of WPV	- Prospective - Cross-sectional - Questionnaire (unvalidated and without piloting but based on established questionnaire)	33 out of all Australasian Sonographers Association members	- 57.6 % experienced WPV in last 12 months - Prevalence of WPV types: verbal abuse (57.6 %), threatening behaviour (21.2 %), and physical violence (3 %) - Risk factors: work environment (public hospital) (39.4 %), >20-year experience (24.2 %), and 25–34 years old (15.2 %)	High (71.4 %)
Nyhsen et al. (2016)-United Kingdom[17]	Diagnostic radiography	Determination of prevalence of WPV	- Prospective - Cross-sectional	- 44 and 5 out of 51 general and 7 CT radiographers of 1	- 57 % experienced WPV initiated by junior doctors in last 12 months	Low (45.2 %)

(*continued on next page*)

Table 1 (*continued*)

Author, year and country	MRS division	Study purpose	Study design	Sample size and characteristics	Key findings	Quality
		initiated by junior doctors	- Questionnaire (unvalidated and without piloting but based on established questionnaire)	large acute hospital in United Kingdom, respectively - Response rate: 84.5 %	- Prevalence of WPV types: loud verbal abuse (57 %), verbal threat (51 %) and bullying (45 %), and physical threat (8 %) - Risk factors: stress (80 %) and inadequacy feeling leading to self-protection (80 %)	
Sethole et al. (2019)-South Africa[18]	Diagnostic radiography	Determination of prevalence of WPV	- Prospective - Cross-sectional - Questionnaire (unvalidated but based on established questionnaire with piloting)	- 37 of 65 (all) radiographers of 2 tertiary public hospitals in South Africa - Response rate: 57 %	- 73 % experienced WPV in the career - Prevalence of WPV types: verbal abuse (73 %), emotional abuse (46 %) and physical abuse (27 %) - Risk factors: patient handling for physical abuse (14 %), and interaction with co-workers for verbal (12 %) and emotional abuse (10 %)	Moderate (68.3 %)
Sperduti et al. (2018)-Canada[19]	Radiation therapy	Determination of prevalence of WPV, risk factors and impacts	- Prospective - Cross-sectional - Questionnaire (unvalidated but based on established questionnaire with piloting)	- 78 out of 165 (all) radiation therapists of 1 urban cancer centre in Canada - Response rate: 47 %	- 84 % experienced WPV in the career - Prevalence of WPV types: verbal abuse (76 %), harassment (59 %), verbal threat (32 %), and physical threatening (23 %) and assault (6 %) - Risk factors: work environment (treatment [68 %] and waiting areas [22 %]) - Impacts: anger, anxiety, depression, difficult to concentrate and sleep, fatigue, fear, flashback, frustration, headache, irritability, low self-esteem, nightmare, sadness, shame, and stress	High (76.2 %)
Tohidnia et al. (2019)-Iran[20]	Diagnostic radiography	Determination of prevalence of WPV and risk factors	- Prospective - Cross-sectional - Questionnaire (validated and based on established questionnaire with piloting)	- All (121) radiographers of 1 Iranian university's radiology department - Response rate: 100 %	- 72.7 % experienced WPV in the career - Common WPV type: verbal violence (77.3 %) - Risk factors: <40 years old (77.3 %), female (63.6 %), <5-year work experience (46.4 %), night shift (43.6 %), residents (42.7 %), over-crowding environment (21.0 %), inadequate security measures (12.6 %) and staff (11.7 %) - Coping strategies: calming down perpetrators (19.1 %), seeking help from colleagues (13.6 %), self-defence (12.7 %), being angry (12.7 %), legal action (9.1 %), and experience sharing with family (5.5 %)	Moderate (61.9 %)
Trad and Johnson (2014)-USA[21]	Radiation therapy	Determination of prevalence of WPV and impacts	- Prospective - Cross-sectional - Questionnaire (unvalidated but based on established questionnaire with piloting)	- 308 out of 665 radiation therapists of 48 radiation therapy centres in USA - Response rate: 46 %	- 63 % indicated WPV presented in current/previous centres - Impacts: more irritable (65.0 %), difficult to concentrate (63 %), more forgetful (46.1 %), and decrease of self-worth (24.5 %)	Moderate (66.7 %)
Tung et al. (2015)-Taiwan[22]	Mainly diagnostic radiography but also covering medical sonography, nuclear medicine, and radiation therapy	Determination of prevalence of WPV and risk factors	- Prospective - Cross-sectional - Questionnaire (unvalidated but based on established questionnaire and literature with piloting)	- 542 out of 4953 radiographers randomly selected in Taiwan - Response rate: 10.9 %	- 46.1 % experienced WPV in last 12 months - Prevalence of WPV types: verbal abuse (65.6 %), physical assault (21.8 %), and sexual harassment (10.3 %) - Risk factors: work environment (ED [56.5 %], examination room [54.1 %], general radiography [46.8 %], location other than examination room and waiting area [23.9 %], and waiting area [19.6 %]), long patient waiting time (21.6 %), miscommunication (21.0 %), influence of alcohol (19.7 %), lack of staff (14.0 %), and lone working (11.8 %) - Impacts: anxiety and stress (26.4 %), work motivation reduction (25.2 %), and lower energy level (21.0 %)	High (71.4 %)

Youngblood (2021)-USA[23]	Nuclear medicine	Determination of prevalence of WPV and risk factors, and awareness of WPV policy	- Prospective - Cross-sectional - Questionnaire (unvalidated but based on established questionnaire with piloting)	- 870 out of 20,385 nuclear medicine technologists (NMTs) in USA - Response rate: 4.3 %	- 46.8 % experienced WPV in last 6 months - Prevalence of WPV types: ignorance (61.2 %), information withholding for affecting performance (57.2 %), highlighting mistakes/errors (57.0 %), rumour/gossip spreading (53.1 %), constant criticism (50.1 %), hostile response (47.1 %), insult (46.6 %), and loud verbal abuse (42.3 %) - Risk factors: female (68.9 %), position (NMTs more likely to experience WPV than other roles such as students [59.9 %]), and 36–45 years old (55.7 %) - Impacts: affecting patient care (57.3 %) and attention (39.6 %) - WPV policy awareness: NMTs (38.0 %) and students (25.0 %)	High (76.2 %)

CT, computed tomography; ED, emergency department; MRI, magnetic resonance imaging.

Common WPV types

The most common WPV type noted in the included studies was verbal abuse and threats,[12,13,15–20,22,23] which is consistent with the findings from the systematics reviews in healthcare and nursing.[2,3,8] However, sexual harassment and physical assault were the second and third commonest WPV types in MRS.[12,13,15–19,22] In contrast, an opposite order of these two types is found in the whole healthcare and nursing sectors.[2,3,8] This could be attributed to lower awareness of sexual harassment of MRPs and perpetrators, resulting in more incidents in MRS.[15] Table 1 shows that four studies investigated the impacts of WPV on radiographers, NMTs and radiation therapists. All of them belong to psychological impacts,[19,21–23] and also match those stated in the systematic reviews on the WPV in healthcare, EMS, nursing and pharmacy because the verbal abuse and threats were the most common WPV type which could only cause the psychological impacts.[3–10]

According to a number of literature reviews on the WPV in healthcare and nursing, they highlighted that working in ED and with patients having mental health conditions were the major risk factors of WPV.[3–8] Although this review's findings reveal that working with intoxicated patients was the most important WPV risk factor in MRS,[15] ED is the major clinical area for managing these patients.[32–34] Also, mental health conditions are commonly associated with intoxication.[33–35] Hence, our findings of working with intoxicated patients as the most important WPV risk factor in MRS appear in line with the aforementioned literature reviews that working in ED and with patients having mental health conditions being the major risk factors.[3–8] Nonetheless, every WPV risk factor listed in Table 1 should not be ignored because the ILO/ICN/WHO/PSI indicated that the WPV has already spread from ED to all other areas of healthcare institutions. Also, all risk factors identified in this review are covered in their framework guidelines for addressing WPV in the health sector.[1]

WPV coping strategies

For the WPV coping strategies, only two included studies investigated these.[15,20] The most common coping strategy was doing nothing which is concerning.[15] However, several systematic reviews on the WPV in healthcare, nursing and EMS showed that not reporting WPV incidents was common.[6,8,9] This could be attributed to the general perception of healthcare professionals that the WPV is an inherent element of healthcare and such incident reporting can imply their lack of competence in delivering patient care and performing routine duties.[9] Such phenomenon is also consistent with the findings of Youngblood's study that only about one third of their NMTs aware of existence of WPV policy in the clinical workplace.[23] Similarly, less than 30 % of participants of Hattingh et al.'s[15] and Tohidnia et al.'s[20] studies were able to apply appropriate strategies such as experience sharing with colleagues and family members, calming down perpetrators, seeking help from colleagues, self-defence, and legal action for coping with the WPV in MRS.[1]

As per the ILO/ICN/WHO/PSI framework guidelines for addressing WPV in the health sector, a range of strategies could be applied for addressing the WPV risk factors identified in this review as follows[1]:

1. Intoxicated patients, working in radiation therapy treatment room, ED, examination room, general radiography, public hospital, non-examination and waiting areas, patient handling, and inadequate security measures: Provision of security services at departments' main entrances, multiple area accesses for staff but limiting public access and separate area for managing

mentally unstable patients, and installation of video surveillance and alarm systems.
2. Staff stress and feeling of inadequacy resulting in self-protection: Avoidance of staff overload, provision of support workers, sufficient rest period, time for problem solving, experience sharing and consultation, recreational area, quiet space, flexible work arrangement and regular work time schedule if feasible.
3. Long patient waiting time, overcrowding environment, unable to meet patients'/family members' expectations and miscommunication: Improvement of patient flow and appointment scheduling, and timely information and comfortable waiting area provided to patients and their families (e.g., television, newspapers, magazines, healthcare service brochures, toys, etc.).
4. More vulnerable MRPs (female, <40-year-old and <5-year experience), night shift, inadequate staff and lone working: Provision of training for coping with WPV and arrangement for team working.
5. Interaction with colleagues: Development of person-centred workplace culture focussing on cooperation, dignity, equal opportunity, non-discrimination, safety and tolerance.

Nevertheless, the most important measure for addressing the WPV issue in MRS should be development and implementation of the WPV policy in the workplace which includes increasing the MRPs awareness of such policy and conducting regular WPV survey with them.[1,36] Table 1 reveals that all but one study used the questionnaire as the data collection tool which is in line with the ILO/ICN/WHO/PSI recommendation because many WPV incidents might not be reported and recorded, making incident report review become less reliable data collection approach.[1,6,8,9,12,13,15–23] It is noted that some included studies had less representative sample sizes (e.g., 13, etc.) and response rates (such as 1.9 %), indicating potential non-response bias and study quality issue.[12,15,37] However, the study quality of all but two were at least moderate.[12–23] Besides, the included studies had a large variation of survey designs such as the WPV reporting periods (whole career, last 12 or 6 months).[12,13,15–23] Nonetheless, according to the ILO/ICN/WHO/PSI WPV in the health sector country case studies research instruments-survey questionnaire, the preferable reporting period should be 12 months,[38] consistent with their recommendation of conducting regular WPV survey.[1]

There are several limitations in this systematic review. Only English articles were included. This might affect its comprehensiveness. For example, no study from non-English European, South-East Asian and South American countries was covered in this review. Furthermore, two thirds of the included studies were about Africa[13–15,18] and North America.[12,19,21,23] Besides, only articles published over the last 10 years were selected but this could ensure our findings more relevant to current clinical practice.[39] It is also noted that no included study was about WPV in academic setting of MRS.[12–23] Hence, our findings should be used with caution although this is the first systematic review on the WPV in MRS.

Conclusion

As per the findings of the included studies, the WPV risk in diagnostic radiography and radiation therapy appears extremely high. Nevertheless, their findings should be used with caution due to the potential non-response bias. Hence, more studies based on the ILO/ICN/WHO/PSI WPV in the health sector country case studies research instruments-survey questionnaire should be conducted in all countries where there is limited WPV research for strengthening the evidence base in the future. Also, a WPV policy should be developed in every clinical workplace. Even if such policy is available, its enforcement including policy awareness boosting, and encouraging incident reporting and support seeking will be essential for reducing the WPV in MRS.

Conflict of interest statement

None.

Acknowledgement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

1. International Labour Office/International Council of Nurses/World Health Organization/Public Services International. *Framework guidelines for addressing workplace violence in the health sector*. Geneva: International Labour Office; 2002 [cited 2023 August 7], https://www.who.int/publications/i/item/9221134466.
2. Sahebi A, Golitaleb M, Moayedi S, Torres M, Sheikhbardsiri H. Prevalence of workplace violence against health care workers in hospital and pre-hospital settings: an umbrella review of meta-analyses. *Front Public Health* 2022;**10**:895818. https://doi.org/10.3389/fpubh.2022.895818.
3. Liu J, Gan Y, Jiang H, Li L, Dwyer R, Lu K, et al. Prevalence of workplace violence against healthcare workers: a systematic review and meta-analysis. *Occup Environ Med* 2019;**76**:927–37. https://doi.org/10.1136/oemed-2019-105849.
4. Mento C, Silvestri MC, Bruno A, Muscatello MRA, Cedro C, Pandolfo G, et al. Workplace violence against healthcare professionals: a systematic review. *Aggress Violent Behav* 2020;**51**:101381. https://doi.org/10.1016/j.avb.2020.101381.
5. Kumari A, Ranjan P, Sarkar S, Chopra S, Kaur T, Baitha U. Identifying predictors of workplace violence against healthcare professionals: a systematic review. *Indian J Occup Environ Med* 2022;**26**:207–24. https://doi.org/10.4103/ijoem.ijoem_164_21.
6. Njaka S, Edeogu OC, Oko CC, Goni MD, Nkadi N. Work place violence (WPV) against healthcare workers in Africa: a systematic review. *Heliyon* 2020;**6**:e04800. https://doi.org/10.1016/j.heliyon.2020.e04800.
7. Civilotti C, Berlanda S, Iozzino L. Hospital-based healthcare workers victims of workplace violence in Italy: a scoping review. *Int J Environ Res Publ Health* 2021;**18**:5860. https://doi.org/10.3390/ijerph18115860.
8. Varghese A, Joseph J, Vijay VR, Khakha DC, Dhandapani M, Gigini G, et al. Prevalence and determinants of workplace violence among nurses in the South-East Asian and Western Pacific Regions: a systematic review and meta-analysis. *J Clin Nurs* 2022;**31**:798–819. https://doi.org/10.1111/jocn.15987.
9. Murray RM, Davis AL, Shepler LJ, Moore-Merrell L, Troup WJ, Allen JA, et al. A systematic review of workplace violence against emergency medical services responders. *New Solut* 2020;**29**:487–503. https://doi.org/10.1177/1048291119893388.
10. Bhagavathula AS, Obamiro K, Hussain Z, Tesfaye W. Workplace violence against pharmacists: a systematic review and meta-analysis. *J Am Pharmaceut Assoc* 2023;**63**:23–31. https://doi.org/10.1016/j.japh.2022.07.012.
11. Ng CKC. Artificial intelligence for radiation dose optimization in pediatric radiology: a systematic review. *Children* 2022;**9**:1044. https://doi.org/10.3390/children9071044.
12. Beam AS, Brown L, Moore KG, Berry SD, Stovall H, Rainey M, et al. Workplace violence in imaging departments. *Radiol Technol* 2022;**94**:108–19.
13. Chinene B, Sibiya MN, Nkosi PB. Antecedents of disruptive behaviours involving radiographers at central hospitals in Harare Metropolitan Province, Zimbabwe. *Radiography* 2022;**28**:751–7. https://doi.org/10.1016/j.radi.2022.03.006.
14. Chinene B, Sibiya MN, Nkosi PB. Factors leading to disruptive behaviours at central hospitals in Harare Metropolitan Province: radiography managers perspectives. *J Med Imag Radiat Sci* 2022;**53**:580–90. https://doi.org/10.1016/j.jmir.2022.07.005.
15. Hattingh C, Nabasenja C, Daniels ER, Kalondo L, Karera A, Amkongo M. Workplace violence involving radiographers at a state radiology department in Windhoek Namibia. *South Afr Radiogr* 2019;**57**:19–23.
16. Lloyd-Jones C, Kennedy NJ, Quinton AE, Bligh L. The frequency and types of violence experienced by Australian sonographers from patients and visitors: a pilot study. *Sonography* 2021;**8**:155–61. https://doi.org/10.1002/sono.12286.
17. Nyhsen CM, Patel P, O'Connell JE. Bullying and harassment - are junior doctors always the victims? *Radiography* 2016;**22**:e264–8. https://doi.org/10.1016/j.radi.2016.05.005.
18. Sethole KM, van Deventer S, Chikontwe E. Workplace abuse: a survey of radiographers in public hospitals in Tshwane, South Africa. *J Radiol Nurs* 2019;**38**:272–6. https://doi.org/10.1016/j.jradnu.2019.09.005.
19. Sperduti A, Hindle D, Shessel A, Pidgeon B, Akmal H, Chaulk G, et al. Treating too lightly? Radiation therapists' experiences of workplace violence when providing care to cancer patients and their families. *J Med Imag Radiat Sci* 2018;**49**:56–61. https://doi.org/10.1016/j.jmir.2017.11.001.

20. Tohidnia MR, Rostami R, Moradi Ghomshei S, Moradi S, Abbasi Azizi S. Incidence rate of physical and verbal violence inflicted by patient and their companions on the radiology department staff of educational hospitals of medical university, Kermanshah, 2017. *Radiol Med* 2019;**124**:14–8. https://doi.org/10.1007/s11547-018-0933-0.
21. Trad M, Johnson J. Bullying among radiation therapists: effects on job performance and work environment. *Radiol Technol* 2014;**86**:122–31.
22. Tung Y, Lin C, Li S, Huang T. Workplace violence involving radiographers and its relevant factors in Taiwan. *J Radiol Sci* 2015;**40**:81–7.
23. Youngblood SN. Bullying in the nuclear medicine department and during clinical nuclear medicine education. *J Nucl Med Technol* 2021;**49**:156–63. https://doi.org/10.2967/jnmt.120.257204.
24. PRISMA. *Transparent reporting of systematic reviews and meta-analyses*. Ottawa: PRISMA; 2023 [cited 2023 August 8], http://www.prisma-statement.org/.
25. Ng CKC. A review of the impact of the COVID-19 pandemic on pre-registration medical radiation science education. *Radiography* 2022;**28**:222–31. https://doi.org/10.1016/j.radi.2021.07.026.
26. Ng CKC. Diagnostic performance of artificial intelligence-based computer-aided detection and diagnosis in pediatric radiology: a systematic review. *Children* 2023;**10**:525. https://doi.org/10.3390/children10030525.
27. Ng CKC. Generative adversarial network (generative artificial intelligence) in pediatric radiology: a systematic review. *Children* 2023;**10**:1372. https://doi.org/10.3390/children10081372.
28. Sirriyeh R, Lawton R, Gardner P, Armitage G. Reviewing studies with diverse designs: the development and evaluation of a new tool. *J Eval Clin Pract* 2012;**18**:746–52. https://doi.org/10.1111/j.1365-2753.2011.01662.x.
29. World Health Organization. *Nursing and midwifery*. Geneva: World Health Organization; 2023 [cited 2023 August 16], https://www.who.int/news-room/fact-sheets/detail/nursing-and-midwifery.
30. Australian Institute of Health and Welfare. *Health workforce*. Canberra: Australian Institute of Health and Welfare; 2023 [cited 2023 August 16], https://www.aihw.gov.au/reports/workforce/health-workforce.
31. American Association of Colleges of Nursing. *Nursing workforce fact sheet*. Washington: American Association of Colleges of Nursing; 2023 [cited 2023 August 16], https://www.aacnnursing.org/news-data/fact-sheets/nursing-workforce-fact-sheet.
32. Mäkelä KL, Kylmä J, Hakala T, Löyttyniemi E, Puolakka K, Koivunen M. Nursing staff's self-assessed skills concerning the care of acutely ill alcohol-intoxicated patients in emergency departments. *Nord J Nurs Res* 2020;**40**:25–32. https://doi.org/10.1177/2057158519864856.
33. Verheesen SMH, ten Doesschate F, van Schijndel MA, van der Gaag J, Cahn W, van Waarde JA. Intoxicated persons showing challenging behavior demand complexity interventions: a pilot study at the interface of the ER and the complexity intervention unit. *Eur Arch Psychiatr Clin Neurosci* 2021;**271**:903–13. https://doi.org/10.1007/s00406-020-01162-7.
34. Sarkar S, Bhatia G, Dhawan A. Clinical practice guidelines for assessment and management of patients with substance intoxication presenting to the emergency department. *Indian J Psychiatr* 2023;**65**:196–211. https://doi.org/10.4103/indianjpsychiatry.indianjpsychiatry_490_22.
35. Mustonen A, Alakokkare AE, Salom C, Hurtig T, Levola J, Scott JG, et al. Age of first alcohol intoxication and psychiatric disorders in young adulthood - a prospective birth cohort study. *Addict Behav* 2021;**118**:106910. https://doi.org/10.1016/j.addbeh.2021.106910.
36. Sarnese P. Safety precautions for radiology nurses. *J Radiol Nurs* 2023;**42**:271–4. https://doi.org/10.1016/j.jradnu.2023.05.004.
37. Mulvany JL, Hetherington VJ, VanGeest JB. Survey research in podiatric medicine: an analysis of the reporting of response rates and non-response bias. *Foot* 2019;**40**:92–7. https://doi.org/10.1016/j.foot.2019.05.005.
38. International Labour office/international Council of nurses/world health organization/public services international joint programme on workplace violence in the health sector. *Workplace violence in the health sector country case studies research instruments: survey questionnaire English*. Geneva: World Health Organization; 2003 [cited 2023 August 16], https://cdn.who.int/media/docs/default-source/documents/violence-against-health-workers/wvquestionnaire.pdf?sfvrsn=9f6810a5_2&download=true.
39. Al Mahrooqi KMS, Ng CKC, Sun Z. Pediatric computed tomography dose optimization strategies: a literature review. *J Med Imag Radiat Sci* 2015;**46**:241–9. https://doi.org/10.1016/j.jmir.2015.03.003.