



OPEN ACCESS

EDITED BY

Andrew Landstrom,
Duke University, United States

REVIEWED BY

Malgorzata Kasztan,
University of Alabama at Birmingham,
United States
Jane Hankins,
St. Jude Children's Research Hospital,
United States
Mohamed A Yassin,
Hamad Medical Corporation, Qatar

*CORRESPONDENCE

Obiageli Nnodu,
oennodu@gmail.com

SPECIALTY SECTION

This article was submitted to Genetics of
Common and Rare Diseases,
a section of the journal
Frontiers in Genetics

RECEIVED 10 September 2022

ACCEPTED 18 October 2022

PUBLISHED 04 November 2022

CITATION

Nnodu O, Madu A, Chianumba R, Isa HA,
Olanrewaju I, Osagie S, Oyekanmi N,
Sangeda RZ, Stewart A, Nembaware V,
Morrice J, Jonas M, Mazandu G,
Wonkam A and Owolabi O (2022),
Establishing a database for sickle cell
disease patient mapping and survival
tracking: The sickle pan-african
research consortium Nigeria example.
Front. Genet. 13:1041462.
doi: 10.3389/fgene.2022.1041462

COPYRIGHT

© 2022 Nnodu, Madu, Chianumba, Isa,
Olanrewaju, Osagie, Oyekanmi,
Sangeda, Stewart, Nembaware, Morrice,
Jonas, Mazandu, Wonkam and Owolabi.
This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s) are
credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Establishing a database for sickle cell disease patient mapping and survival tracking: The sickle pan-african research consortium Nigeria example

Obiageli Nnodu^{1,2*}, Anazoeze Madu³, Reuben Chianumba²,
Hezekiah Alkali Isa^{1,2}, Isaac Olanrewaju⁴, Samuel Osagie⁵,
Nash Oyekanmi⁶, Raphael Zozimus Sangeda⁷,
Annemie Stewart⁸, Victoria Nembaware⁹, Jack Morrice¹⁰,
Mario Jonas¹¹, Gaston Mazandu⁹, Ambroise Wonkam¹² and
Olumide Owolabi^{1,4}

¹Centre of Excellence for Sickle Cell Disease Research and Training (CESRTA), University of Abuja, Abuja, Nigeria, ²Department of Haematology and Blood Transfusion, University of Abuja, Abuja, Nigeria, ³Department of Haematology and Immunology, College of Medicine, University of Nigeria, Enugu, Nigeria, ⁴Department of Computer Science, University of Abuja, Abuja, Nigeria, ⁵ITMS Unit, University of Abuja, Abuja, Nigeria, ⁶H3Africa Bioinformatics Network (H3ABioNet) Node, Centre for Genomics Research and Innovation, Abuja, Nigeria, ⁷Department of Pharmaceutical Microbiology, Muhimbili University of Health and Allied Sciences, Dar Es Salaam, Tanzania, ⁸UCT Clinical Research Centre, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa, ⁹Computational Biology Division, IDM, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa, ¹⁰Division of Human Genetics, Department of Pathology, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa, ¹¹Faculty of Health Sciences, Department of Pathology, Division of Human Genetics, University of Cape Town, Cape Town, South Africa, ¹²Department of Medicine, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa

Background: The Sickle Pan-African Research Consortium (SPARCO) and Sickle Africa Data Coordinating Center (SADaCC) were set up with funding from the US National Institute of Health (NIH) for physicians, scientists, patients, support groups, and statisticians to collaborate to reduce the high disease burden and alleviate the impact of Sickle Cell Disease (SCD) in Africa. For 5 years, SPARCO and SADaCC have been collecting basic clinical and demographic data from Nigeria, Tanzania, and Ghana. The resulting database will support analyses to estimate significant clinical events and provide directions for targeting interventions and assessing their impacts.

Method: The Nigerian study sited at Centre of Excellence for Sickle Cell Disease Research and Training (CESRTA), University of Abuja, adopted REDCap for online database management. The case report form (CRF) was adapted from 1,400 data elements adopted by SPARCO sites. It captures 215 data elements of interest across sub-sites, i.e., demographic, social, diagnostic, clinical, laboratory, imaging, and others. These were harmonized using the SADaCC data dictionary. REDCap was installed on University of Abuja cloud server at <https://www.redcap.uniabuja.edu.ng>. Data collected at the sites are sent to CESRTA for collation, cleaning and uploading to the database.

Results: 7,767 people living with sickle cell disease were enrolled at 25 health institutions across the six zones in Nigeria with 5,295 having had at least one follow-up visit with their clinical data updated. They range from 44 to 1,180 from 3 centers from South East, 4 from South, 5 from South West, 8 from North Central, 4 in North West and 3 in the North East. North West has registered 1,383 patients, representing 17.8%; North East, 359 (4.6%); North Central, 2,947 (37.9%); South West, 1,609 (20.7%); South, 442 (5.7%) and South East, 1,027 patients (13.2%).

Conclusion: The database is being used to support studies including analysis of clinical phenotypes of SCD in Nigeria, and evaluation of Hydroxyurea use in SCD. Reports undergoing review in journals have relied on the ease of data access in REDCap. The database is regularly updated by batch and individual record uploads while we are utilizing REDCap's in-built functions to generate simple statistic.

KEYWORDS

database, REDCap, sickle cell disease, SPARCO, Nigeria, registry

Introduction

Sickle cell disease (SCD) is an inherited blood disorder arising from mutations in the gene coding for the β globin chain of haemoglobin. This leads to pathogenetic processes culminating in vessel occlusion and progressive ischemia and tissue necrosis. The disease is inherited in a simple Mendelian pattern (Piel et al., 2013) and affects more than 5 million people in Sub-Saharan Africa (SSA). About 150,000 children are born with this condition annually in this region. The incidence of SCD in Nigeria varies from 2% to 3% as reported by various investigators in the different regions of the Country (Grosse et al., 2011).

The disease is marked by acute exacerbations of pain and anemia warranting frequent hospital or clinic admissions and blood transfusion. Chronic complications such as osteonecrosis, leg ulcers, pulmonary hypertension and renal failure may occur. This requires the presence of appropriately trained medical personnel to provide care at various levels of healthcare. This has huge implications on the health economics and fiscal deployment in both adults and children (Baker et al., 2015). There are approximately 111.91 cases of homozygous S (SS) per 100,000 newborns globally, while Africa has a rate of 1,125.49 (Wastnedge et al., 2018). The disease burden of people living with SCD is currently estimated to be approximately 250 million people worldwide, with the majority living in the rural areas of Sub-Saharan Africa (SSA) and with limited access to health care.

The socio-economic impact of the disease on affected persons and families is massive with frequent morbidity associated with a high mortality rate (Olatunya et al., 2015; Nnodu et al., 2021). Health budgeting and intervention planning requires a good estimate of disease burden and patient mapping in order to effectively link patients to care and track other occurrences of significant disease-related events.

An extensive database is also necessary for effective and comprehensive documentation of the natural course of the disease and its complications in the different unique socio-cultural backgrounds found in Africa. The establishment and maintenance of a simple database would provide real time as well as retrospective information on the clinical state of patients across the country. This will also offer a platform for the estimation of significant clinical events including frequency of vaso occlusive crisis, transfusion needs, frequency of complications, pattern of drug prescription as well as survival. These will provide direction targeting interventions and assessing impact of such interventions.

The Sickle Pan-African Research Consortium (SPARCO) and Sickle Africa Data Coordinating Center (SADaCC) consists of physicians, scientists, patients support groups and statisticians who work together to reduce the disease burden and alleviate the impact of disease. These primary objectives are to be achieved by establishing a research database, carrying out research aimed at alleviating disease burden and reducing its impact on already affected persons. SPARCO and SADaCC have over the past 5 years accumulated and synchronized basic clinical and demographic data from Nigeria, Tanzania and Ghana. Challenges abound in a resource-limited setting but the benefits are enormous with regards to limiting waste and optimizing care.

Databases and other computer based systems

Databases were designed for persistent and integrated storage of data, allowing concurrent access to it by many users. They are collections of records related by referential integrity. Thus, a database is an organized collection of structured data meant to serve many applications with minimum redundancy (Elmasri

and Navathe, 2017). Database technology helps to alleviate many of the problems associated with conventional file organization methods, including data duplication, inflexibility and difficulties associated with online access by users.

Databases, in greatly aiding the availability and searchability of data, have become an indispensable tool for research in medical and other fields. For example, well-designed biological databases have become a powerful tool in biological research (Dong et al., 2015). Also, it is now well established that keeping electronic medical records leads to dramatic improvements in the level of care provided by health institutions. This was shown in a 2006 study on electronic medical records (Akor et al., 2018). The study showed, among others, that EMR facilitates easy access to medication administration records, sharing of consultation reports, and decreased transit time of test results by reducing the time taken to deliver paper versions.

Other developments in computing and computer technologies that have benefitted medical research include Data/text mining, Computer modeling, Medical imaging/anomaly detection and Machine learning. Applications of these techniques have been discussed in the literature (Hand et al., 2001; Lei and Karniadakis, 2013; Solanki, 2014; Khalaf et al., 2016). The use of Pattern Recognition techniques in medicine have also been studied (Benamrane et al., 2005; Bacardit et al., 2014). An interesting case of the application of text and data mining in SCD research is the development of an information exploration system, Dragon Exploration System for SCD (DESSCD), which aims to promote the easy exploration of SCD related data. The system processed 419,612 MEDLINE abstracts retrieved from a PubMed query using SCD-related keywords. The processed SCD-related data was then made available *via* the DESSCD web query interface that enables information retrieval using specified concepts, keywords and phrases, and the generation of inferred association networks and hypotheses (Essack, 2013).

REDCap and electronic data capture

One tool currently used by researchers in many countries to manage research data collections is the Research Electronic Data Capture (REDCap) application. REDCap is a mature, secure web application designed to support data capture and managing online databases. While REDCap can be used to collect virtually any type of data, it is specifically geared to support data capture for research studies. The application allows users to build and manage online surveys and databases quickly and securely, and is currently in production use or development build-status for more than 290,000 projects with over 398,000 users spanning numerous research focus areas across the consortium (Harris et al., 2019).

There are many other Electronic Data Capture software that are being used in medical research; some of these are Catalyst

Web Tools, OpenClinica, IBM Clinical Development and Videoc; REDCap has been found to compare well with some of these other systems (Franklin et al., 2011). It has even been suggested that the strength of REDCap as it pertains to its use in medical research is that it is primarily a clinical research database built by clinical researchers specifically for clinical research (Patridge and Bardyn, 2018). REDCap is a secondary survey tool that provides an easy-to-use interface and several desirable features for data collection, management, analysis, and reporting (Harris et al., 2009; Harris et al., 2019).

Projects are self-sufficient and workflow-based; the focus is on collecting data and exporting it to statistical programs and other data analysis software. REDCap is designed to provide a secure environment so that research teams can collect and store highly sensitive information. It is a flexible tool that can run on multiple operating systems such as Linux, UNIX, Windows, and Mac. We deployed REDCap at CESRTA, with an administrator and 4 data clerks who manage the recording of data into the database. The administrator for the system is a certified cloud infrastructure engineer from the University of Abuja's IT unit. The Principal Investigator (PI) and Data Manager were also assigned administrator roles. The users are verified members of the research team; the Site Lead determines the level of access and privileges granted to each user by the administrators.

Methodology

The study site was based in the Centre of Excellence for Sickle Cell Disease Research and Training (CESRTA), University of Abuja. The subjects are patients recruited into various research projects at CESRTA (<https://cesrta.uniabuja.edu.ng/>). Data was collected from consenting patients attending sickle cell clinics using case report forms (CRFs) designed for the project after ethical approval. The soft copy of the CRFs were produced on Excel sheets and shared with all the collaborating sites in Nigeria. This was preceded by a meeting of all the site leads who are doctors who run the adult and paediatric sickle cell clinics all over Nigeria and were available to be part of SPARCO. In these meetings data managers/collectors were trained on how to fill the CRFs and also to input the data electronically. The data collected with the Excel sheet are regularly sent for central collation at CESRTA where it is then batch uploaded into the REDCap database.

There are two studies currently ongoing on the REDCAP platform, 1) the SPARCO description of phenotype and 2) the survival study from the newborn screening. For the SPARCO phenotype study, we had a local care report form which was used to collect full demographic and clinical data from patients attending the SCD clinics at SPARCO centers across the country. This was then used to develop the CRF for enrollment of patients into the SPARCO registry. The hard copy CRF forms were deployed to the SPARCO Nigeria Sites for patient enrollment. Site leads were brought to

CESRTA for a data management workshop and training on the use of REDCap. Feedback was received and used to modify the first CRF, standardizing it and making it possible to separate features of past medical history, presenting symptoms and monitor follow-up data. The CRF was designed to achieve full clinical phenotyping and to obtain laboratory information as they become available. It included demographic, social, diagnostic, clinical, laboratory, imaging, and research data elements, all totaling 125 variables (“field name”), to be included in the database. These variables were harmonized using the SADaCC data dictionary; trainings were also conducted by the Big Data Analytics team in order to assist the team members harmonize the data elements in line with the SCD Ontology. Before implementation in REDCap, REDCap was installed on the University of Abuja dedicated cloud server by the University’s server Administrator/Webmaster at <https://www.redcap.uniabuja.edu.ng> alongside requisite applications such as MySQL database server and PHP. An Administrator, who can add users to the software and assign them permissions, was then designated. Users can create new projects or be assigned roles by project owners.

Creating a new project is done at the +New Project tab after logging in to REDCap. When a user creates a new project, the project is automatically set in development status, allowing the user to edit the project and test it before uploading real data. Once a project has moved to production status, it can no longer be edited.

The REDCap administrator created a project for our site with the record auto-numbering feature enabled. After thoroughly testing the project setup and being found satisfactory, it was moved to Production Status, which made it available online for duly authorized users.

Data tables were created in the Project to adequately represent all the data being gathered. Such Tables included: Basic demographics of the screened population, results of screening, contact numbers, exact location, names of schools and nearest primary health care centres. For those with sickle haemoglobin, SS or SC, only basic patient data (demographics), routine clinic visits, acute care visits, hospital admissions and laboratory studies are collected. Each table was developed with relevant data elements suggested by the requirements analysis and was translated into user interface forms and field-tested at the data collection sites. Changes were made in the design of the Tables based on the results of the field test. All these were created on the REDCap project with appropriate form interfaces for data capture. At CESRTA, the data collected from the screening points were initially prepared in MS Excel worksheets and uploaded in comma-separated (csv) format following the Data Dictionary template. The use of mobile devices were later introduced, allowing for secure on-site clinical use as well as the collection and transmission of research data to the online project. In addition, individual data records are sometimes entered directly as the need arises. All the submitted data are subjected to quality checks by the data management team.

Results

Data was obtained from 27 centers made up of 25 tertiary health institutions, 2 secondary and one private hospital currently engaged in SPARCO Nigeria patient enrollment. Table 1 shows the centres across the country that are currently enrolling patients on the SPARCO platform.

A total of 7,767 people living with SCD have been enrolled into the SPARCO Nigeria data platform and another 5,295 have had at least one follow up visit with their clinical data updated. This represents an estimated 0.001% of the entire population of sickle cell patients in Nigeria. The number of recruited patients range from 44 to 1,180 and included 3 centers from South East, 4 from South, 5 from South West, 8 from North Central, 4 in North West and 3 in the North East. As at the time of this report, the North West had registered 1,383 patients, representing 17.8%, North East, 359 patients (4.6%), North Central, 2,947 patients (37.9%), South West, 1,609 patients (20.7%), South, 442 patients (5.7%) and South East, 1,027 patients (13.2%). This is shown in Figure 1.

From the basic analysis done on the data with the REDCap analysis tools, the following were observed:

Age: the age of patients’ ranges from 0 to 64 years, with a mean value of 16.58 and median value of 16 years

Gender: 48.9% of the patients are Males, while 51.1% are Females. This displayed in Figure 2.

Genotype: the distribution of the patients by genotype is as follows, and is displayed in Figure 3.

HbSS—96.9%
HbSBThal—0.2%
HbSC—2.9%, and
HbAS—0%

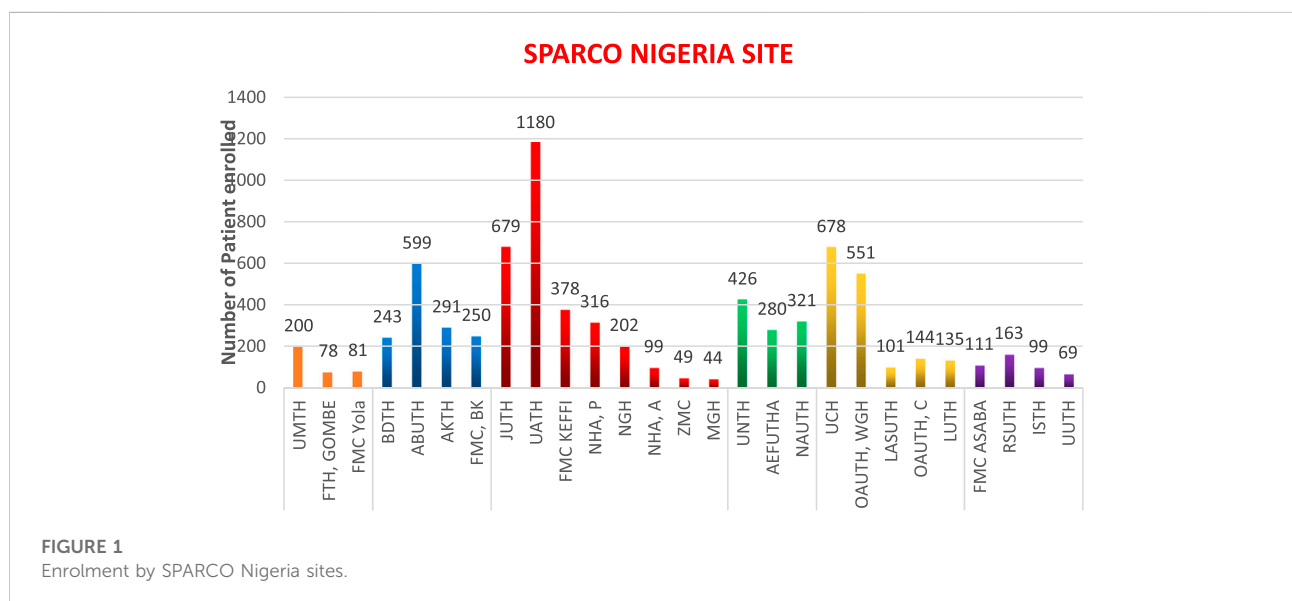
It could also be observed that, though the data was collected at 27 centres from across the country, all the States of the Nigerian Federation were represented in the data set. The database could thus be said to be generally representative of the Nigerian population.

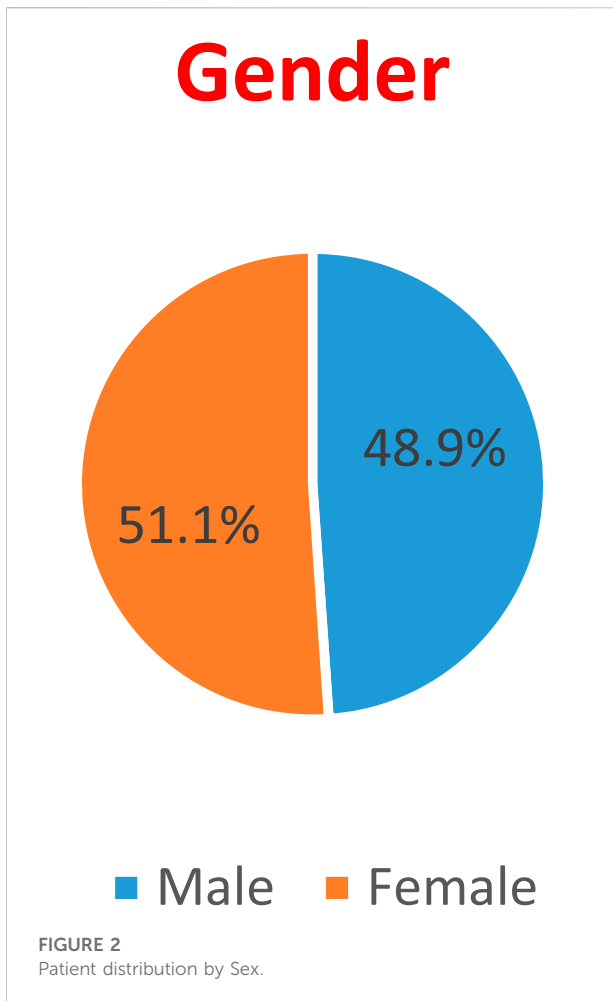
Discussion

Using REDCap features as described, we had designed data tables to store all the data elements defined in the Nigeria-specific Case Reporting Form (CRF) and SPARCO data dictionary. The online availability of the REDCap database makes it convenient for the Team members to access the database for data research, insertion, updates, viewing and reporting according to the access privileges of each member. This has been demonstrated by a preliminary analysis of the clinical phenotypes of SCD in Nigeria which was published (Isa et al., 2020) and a study on the evaluation of Hydroxyurea use in SCD (Chianumba et al., 2022). Other completed studies with submitted manuscripts undergoing review in different journals have relied heavily on the ease of access to data in the REDCap project. Data

TABLE 1 Enrolment by SPARCO centers across Nigeria.

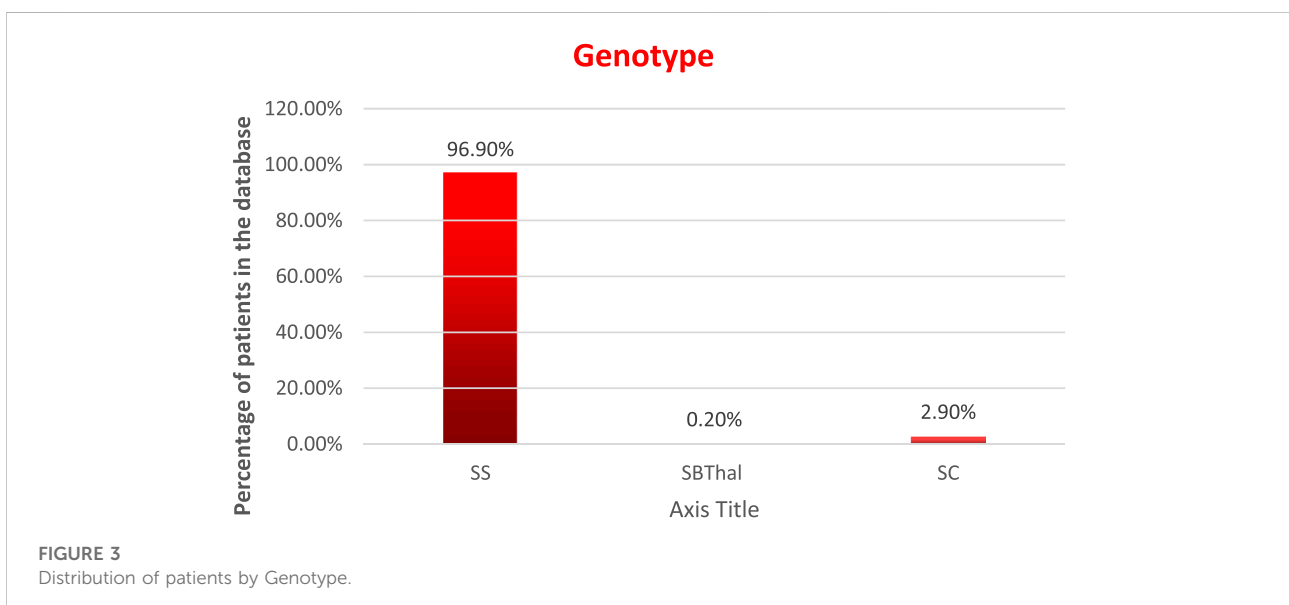
Hospital name abbreviation	Institution's full name	Geopolitical zone	Number
UMTH	University of Maidugiri Teaching Hospital	North East	200
FTH, GOMBE	Federal Teaching Hospital Gombe		78
FMC Yola	Federal Medical Centre, Yola		81
BDTH	BARAU DIKKO TEACHING HOSPITAL KADUNA	North West	243
ABUTH	Ahmadu Bello University Teaching Hospital, Zaria		599
AKTH	Aminu Kano Teaching Hospital		291
FMC, BK	Federal Medical Centre, Birnin Kebbi		250
JUTH	Jos University Teaching Hospital	North Central	679
UATH	University of Abuja Teaching Hospital, Gwagwalada		1,180
FMC KEFFI	Federal Medical Centre Keffi		378
NHA, P	National Hospital Abuja		316
NGH	General Hospital Nyanya		202
NHA, A	National Hospital Abuja, Adult		99
ZMC	ZANKLI MEDICAL CENTRE, Abuja		49
MGH	Maitama General Hospital		44
UNTH	University of Nigeria Teaching Hospital, Enugu	South East	426
AEFUTHA	Federal Teaching Hospital Abakaliki		280
NAUTH	Nnamdi Azikiwe University Teaching Hospital		321
UCH	University College Hospital, Ibadan	South West	678
OAUTH, WGH	Obafemi Awolowo University Teaching Hospital, WGH		551
LASUTH	Lagos State University Teaching Hospital		101
OAUTH, C	Obafemi Awolowo University Teaching Hospital College		144
LUTH	Lagos University Teaching Hospital		135
FMC ASABA	Federal Medical Asaba	South	111
RSUTH	River State University Teaching Hospital		163
ISTH	ISTH Irrua		99
UUTH	University of Uyo Teaching Hospital		69
	Total		7,767





upload is also regularly done both by batch upload and individual record insertion.

Several other studies on SCD have used REDCap for data storage and management and they report on the strengths of the software, which include rapid development of forms by technical staff with limited programming skills (Patridge and Bardyn, 2018), flexibility in data entry and analysis, efficient and secure methods for data collection, ease of use and quick turnaround, and very straightforward data exports, with exported data automatically coming in SPSS, Excel/CSV, SAS, R, and STATA formats (Franklin et al., 2011). Availability of copious amounts of training materials has also been cited as an advantage (Franklin et al., 2011). Using REDCap has aided the project’s data management function in a highly resourceful manner. From converting the CRF to representative data tables, and enforcing data quality and integrity, to executing data queries, REDCap provides the tools to effectively execute all desirable data management functions. We have also been able to utilize some of its in-built functions to generate some simple statistics on the data. To continually support security of the research data, and to maintain smooth data management operations, the University of Abuja REDCap instance has gone through a couple of upgrades resulting in the current version being used. This registry is an additional resource to the existing sickle cell databases in Africa and other parts of the world such as the Globin Research Network for Data and Discovery (GRNDaD), Sickle Cell Clinical Research and Intervention Program (SCCRIP), the Sickle Cell Data Collection (SCDC) program all in the United States. (Boye-Doe et al., 2020; Hankins, et al., 2018; CDC (Centers for Disease Control and Prevention), 2020). In the United Kingdom, some of the sickle cell registries include, the NHS Sickle Cell and Thalassaemia Screening Programme, National Haemoglobinopathy Register,



Screening Wales and Cardiff Sickle Cell and Thalassaemia Centre, and Paediatric and Adult Haematology Lead Republic of Ireland. (Dormandy et al., 2018). Most of these registries collect longitudinal data and are multicentred. The SCCRIP database is a single institution registry and in addition, has retrospective data. The main objective of these registries is the evaluation of outcomes of interventions and research with a view to improving the quality of life and survival of patients with SCD.

Conclusion

The initial challenge is that of getting like-minded researchers to agree with regards to the importance and value of available patient data on a particular disease. This required several correspondence and meetings and some funding. Another option is to rely on already existing networks, like SPARCO in this case.

Similar to the Tanzanian study reported in (Tluway and Makani, 2017), this work addresses the paucity of elaborate databases for many non-communicable diseases in Africa. The lack of cohesion and sustainability had hampered the availability of an online and broad-based data pool of sickle cell patients in Africa. Setting up such an online massive database is fraught with several challenges that were dealt with in this manuscript. This manuscript will therefore serve as a guide for future data collection of a large pool of patients by researchers across several countries and on a variety of diseases.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

ON and OO conceptualized the article, ON and OO drafted the manuscript, RC analysed the data and AM, RC, HA, SO, ON,

RS, AS, VN, JM, MJ, GM, AW, and OO all contributed. HI substantially contributed to the revision of the manuscript.

Funding

This research was funded by The Nigeria Tertiary Education Trust Fund 2020 IBR Award to ON and OO and the National Institutes of Health (NHLBI U24HL135881-01-04-Abuja), (SPARC-NEt, NIH NHLBI 1U01HL156942 2020-2022, SADACC NHLBI, 5U24HL135600-04). TETFUND and the NIH had no direct role in the study design, the collection, analysis and interpretation of data, in the writing of the report, or in the decision to submit the article for publication.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary Material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.1041462/full#supplementary-material>

References

- Akor, S., Uganneya Solomon, Ph.D., and Ogezi, I. C. (2018). "Evaluative study of digital record management system in the hospitals in minna metropolis. (A case study of general hospital minna, Niger state, Nigeria)," in *Library philosophy and practice (e-journal)*.
- Bacardit, J., Widera, P., Lazzarini, N., and Krasnogor, N. (2014). Hard data analytics problems make for better data analysis algorithms: Bioinformatics as an example. *Big Data* 2 (3), 164–176. doi:10.1089/big.2014.0023
- Baker, C., Grant, A. M., George, M. G., Grosse, S. D., and Adamkiewicz, T. V. (2015). Contribution of sickle cell disease to the pediatric stroke burden among hospital discharges of african-Americans-United States, 1997-2012. *Pediatr. Blood Cancer* 62 (12), 2076–2081. doi:10.1002/pbc.25655
- Benamrane, N., Fréville, A., and Nekkache, R. (2005). A hybrid fuzzy neural networks for the detection of tumors in medical images. *Am. J. Appl. Sci.* 2 (4), 892–896. doi:10.3844/ajassp.2005.892.896
- Boye-Doe, A., Brown, E., Puri-Sharma, C., Chawla, A., Field, J., Neumayr, L. D., et al. (2020). The grndad registry: Contemporary natural history data and an analysis of real-world patterns of use and limitations of disease modifying therapy in adults with SCD. *Blood* 136, 34–36. doi:10.1182/blood-2020-138895
- CDC (Centers for Disease Control and Prevention) (2020). Sickle cell data collection (SCDC) program. Available at <https://www.cdc.gov/ncbddd/hemoglobinopathies/scdc.html> (accessed October 15, 2022).

- Chianumba, R. I., Ofakunrin, A., Morrice, J., Olanrewaju, O., Oniyangi, O., Kuliya-Gwarzo, A., et al. (2022). Outcome of Hydroxyurea use in SCD and evaluation of patients' perception and experience in Nigeria. *Front. Genet.* 13, 826132. doi:10.3389/fgene.2022.826132
- Dong, Z., Ma, L., Yu, J., and Zhang, Z. (2015). Biological databases for human research, genomics. *Proteomics Bioinforma.* 13 (1), 55–63. doi:10.1016/j.gpb.2015.01.006
- Dormandy, E., James, J., Inusa, B., and Rees, D. (2018). How many people have sickle cell disease in the UK? *J. Public Health* 40 (3), e291–e295. doi:10.1093/pubmed/idx172
- Elmasri, R., and Navathe, S. B. (2017). *Fundamentals of database systems*. Boston: Pearson Education.
- Essack, M. (2013). Radovanovic A, Bajic VB. Information exploration system for sickle cell disease and repurposing of hydroxyfasudil. *PLoS One* 8 (6), e65190. doi:10.1371/journal.pone.0065190
- Franklin, J. D., Guidry, A., and Brinkley, J. F. (2011). A partnership approach for Electronic Data Capture in small-scale clinical trials. *J. Biomed. Inf.* 44, S103–S108. doi:10.1016/j.jbi.2011.05.008
- Grosse, S. D., Odame, I., Atrash, H. K., Amendah, D. D., Piel, F. B., and Williams, T. N. (2011). Sickle cell disease in Africa: A neglected cause of early childhood mortality. *Am. J. Prev. Med.* 41, S398–S405. doi:10.1016/j.amepre.2011.09.013
- Hand, D., Mannila, H., and Smyth, P. (2001). *Principles of data mining*. Cambridge: MIT Press.
- Hankins, J. S., Estep, J. H., Hodges, J. R., Villavicencio, M. A., Robison, L. L., Weiss, M. J., et al. (2018). Sickle cell clinical research and intervention program (SCCRIP): A lifespan cohort study for sickle cell disease progression from the pediatric stage into adulthood. *Pediatr. Blood Cancer* 65, 9. doi:10.1002/xbc.27228
- Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., et al. (2019). The REDCap consortium: Building an international community of software platform partners. *J. Biomed. Inf.* 95, 103208.
- Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., and Conde, J. G. (2009). Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inf.* 42, 377–381. doi:10.1016/j.jbi.2008.08.010
- Isa, H., Adegoke, S., Madu, A., Hassan, A. A., Ohiaeri, C., Chianumba, R., et al. (2020). Sickle Africa data coordinating center SADaCC, Nnodu OSickle cell disease clinical phenotypes in Nigeria: A preliminary analysis of the sickle Pan Africa research consortium Nigeria database. *Blood Cells Mol. Dis.* 84, 102438. doi:10.1016/j.bcmd.2020.102438
- Khalaf, M., Hussain, A. J., Keight, R., Tso, P., Fergus, P., Keenan, R., et al. (2016). Machine learning approaches to the application of disease modifying therapy for sickle cell using classification models. *Neurocomputing* 228, 154–164. doi:10.1016/j.neucom.2016.10.043
- Lei, H., and Karniadakis, G. E. (2013). Probing vasoocclusion phenomena in sickle cell anemia via mesoscopic simulations. *Proc. Natl. Acad. Sci. U. S. A.* 110, 11326–11330. doi:10.1073/pnas.1221297110
- Nnodu, O. E., Oron, A. P., Sopekan, A., Akaba, G. O., Piel, F. B., and Chao, D. L. (2021). Child mortality from sickle cell disease in Nigeria: A model-estimated, population-level analysis of data from the 2018 demographic and health survey. *Lancet. Haematol.* 8 (10), e723–e731. doi:10.1016/S2352-3026(21)00216-7
- Olatunya, O. S., Ogundare, E. O., Fadare, J. O., Oluwayemi, I. O., Agaja, O. T., Adeyefa, B. S., et al. (2015). The financial burden of sickle cell disease on households in Ekiti, Southwest Nigeria. *Clin. Outcomes Res.* 7, 545–553. doi:10.2147/CEOR.S86599
- Patridge, E. F., and Bardyn, T. P. (2018). Research electronic data capture (REDCap). *J. Med. Libr. Assoc.* 106 (1), 142–144. doi:10.5195/jmla.2018.319
- Piel, F. B., Hay, S. I., Gupta, S., Weatherall, D. J., and Williams, T. N. (2013). Global burden of sickle cell anaemia in children under five, 2010–2050: Modelling based on demographics, excess mortality, and interventions. *PLoS Med.* 10, e1001484. doi:10.1371/journal.pmed.1001484
- Solanki, A. V. (2014). Data mining techniques using WEKA classification for sickle cell disease. *Int. J. Comput. Sci. Inf. Technol.* 5 (4), 5857–5860.
- Tluway, F., and Makani, J. (2017). Sickle cell disease in Africa: An overview of the integrated approach to health, research, education and advocacy in Tanzania, 2004–2016. *Br. J. Haematol.* 177 (6), 919–929. doi:10.1111/bjh.14594
- Wastnedge, E., Waters, D., Patel, S., Morrison, K., Goh, M. Y., Adeloje, D., et al. (2018). The global burden of sickle cell disease in children under five years of age: A systematic review and meta-analysis. *J. Glob. Health* 8 (2), 021103. doi:10.7189/jogh.08.021103