CHARACTERIZATION OF THE GENETIC AND PHENOTYPIC LANDSCAPE OF PATIENTS WITH CNVS IN THE CHROMOSOME 22q11.2 REGION IN A COHORT OF SOUTH AFRICAN PATIENTS

Dylan McCarthy

Student number: 2115219

Supervisor: Mrs Bianca Rossouw

Co-Supervisor: Dr Bronwyn Dillon

A research report (in the format of a “submissible” paper) submitted to the Faculty of Health Sciences, University of the Witwatersrand, in partial fulfilment of the degree of Master of Science in Medicine (Genetic Counselling) by Research and Coursework.

Johannesburg, 2025

NATIONAL HEALTH LABORATORY SERVICE
School of Pathology, University of the Witwatersrand
DIVISION OF HUMAN GENETICS
Hospital Street, Johannesburg, 2001 | PO Box 1038, Johannesburg, 2000
[T] +27 11 489 9223 | [M] +27 78 080 8841 | [F] +27 11 489 9226 | [E] human.genetics@nhls.ac.za

Dear examiner, 14/01/2025

Re: Letter to the examiner: Research Report submitted to the Faculty of Health Sciences, University of the Witwatersrand, by Mr Dylan McCarthy in partial fulfilment for the degree of Master of Science in Medicine (Genetic Counselling)

Thank you for agreeing to examine the research report submitted by Mr Dylan McCarthy (student number 2115219) entitled, “Characterization of the genetic and phenotypic landscape of patients with CNVs in the 22q11.2 region in a cohort of South African patients”. This research report is being submitted by the candidate in the format of a “submissible” paper as per the style guide for Thesis, Dissertation, and Research Reports (Updated March 2016) by the University of Witwatersrand, Faculty of Health Sciences. This research paper contributes 30% towards the final mark of the MSc (Med): Genetic Counselling degree for which the candidate is enrolled.

The candidate and his supervisors have chosen to submit the paper to the Journal of Molecular Genetics and Genomic Medicine (https://onlinelibrary.wiley.com/page/journal/23249269/homepage/forauthors.html). This journal is a Wiley peer-reviewed, open access, international journal. The submissible paper is attached therefore written in accordance with the author guidelines of the stipulated journal. The author guidelines specify any referencing style but recommend APA 7th edition. We have therefore selected APA 7th edition as the referencing style for this research report. The only exceptions we have made to the author guidelines are for the benefit of presentation for this assessment and include exclusion of line numbers in the submission, as well as the placement of the author contributions, acknowledgements, data availability, consent, ethics approval, and conflict of interest statements as well as funding, keywords and correspondence sections. The placements of these sections mirror published papers in this journal as their placement according to the author guidelines was sometimes vague and contradictory.

The research protocol has been attached as an appendix for the purpose of providing an extended literature review and further contextualise the study. The protocol has already been

assessed and approved by the Faculty of Health Sciences Post Graduate Assessors Committee (See Appendix A) and is therefore not for examination.

Please refer to the Table of Contents for a complete list of the contents provided in this research report.

Sincerely,

Dylan McCarthy (Candidate)

Mrs Bianca Rossouw (Primary supervisor)

Dr Bronwyn Dillon (Co-supervisor)

Declaration

I, Dylan McCarthy, declare that this Research Report (in the format of a “submissible” paper) is my own, unaided work. It is being submitted for the Degree of Master of Science in Medicine (Genetic Counselling) by Research and Coursework at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at any other University.

Dylan McCarthy

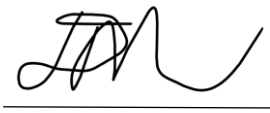
Student number: 2115219

14th Day of July 2025

Contribution of Candidate

Student's contribution to article(s) and agreement of co-author(s)

I, Dylan McCarthy, declare that this research report entitled, "Characterization of the genetic and phenotypic landscape of CNVs in the 22q11.2 region in a cohort of South African patients", is my own work and that I contributed significantly to the research findings presented in the paper intended for publication below.



Signature of student

14/07/2025

Date

Signature of primary supervisor

14/07/2025

Date

Agreement of co-author(s):

By signing this declaration, the co-authors listed below agree to the use of the research findings intended for a published article by the student as part of their Research Report.

Authors	Name	Signature	Date
1st	Mr Dylan McCarthy		14/07/2025
2nd	Dr Bronwyn Dillon		14/07/2025
3rd	Mrs Bianca Rossouw		14/07/2025

Presentations arising from this research

1. South African Society of Human Genetics 2024 Conference, 26th-29th October 2024 (poster presentation)

   Title: “Characterization of the genetic and phenotypic landscape of patients with copy number variants in the chromosome 22q11.2 region in a cohort of South African patients”,

   Authors: D. McCarthy, B Dillon, F Essop, B Rossouw

2. Genetic Counsellors South Africa Research Showcase 2024, 21st November 2024 (oral presentation)

   Title: “Characterization of the genetic and phenotypic landscape of patients with copy number variants in the chromosome 22q11.2 region in a cohort of South African patients”,

   Authors: D. McCarthy, B Dillon, B Rossouw

Abstract

Copy number variants (CNVs) in the chromosome 22q11.2 region are among the most frequently identified genomic rearrangements. However, their phenotypic variability in African populations remains poorly characterised due to limited access to diagnostics and underrepresentation in genomic research. We retrospectively reviewed 129 sub-Saharan African patients diagnosed with CNVs in the chromosome 22q11.2 region over ten years. Clinical, demographic, and genetic data were collected from fluorescence *in situ* hybridisation, multiplex ligation-dependent probe amplification, and chromosomal microarray analysis (CMA). CNVs were classified by low copy repeat involvement, and genotype-phenotype correlations were explored. A subset of 11 patients underwent analysis using Face2Gene, an AI-driven next-generation phenotyping tool. Deletions accounted for 71% and duplications 29% of cases. Classic CNVs were observed in 68% of patients and atypical CNVs in 32%. A statistically significant association was found between classic CNVs and congenital heart defects (CHDs). The higher prevalence of CHDs, immunological, endocrinological, and developmental abnormalities mirrored global trends; however, a lower frequency of palatal defects was observed. In total, 181 distinct phenotypic features were documented, highlighting the vast phenotypic spectrum associated with these CNVs. Face2Gene correctly identified all deletions using a gestalt-only approach, but performance decreased with a combined gestalt and feature score approach. This study addresses gaps in genomic data from African populations, while demonstrating the vast phenotypic spectrum of CNVs in the chromosome 22q11.2 region. The introduction of CMA markedly improved diagnostic yield annually. While Face2Gene performed well, further validation in larger,more diverse cohorts is needed for improved clinical implementation.

Acknowledgements

I would like to acknowledge both of my supervisors for all the help and guidance they offered me throughout the completion of this research report. Thank you for taking the time to carefully review each draft and assisting me at every step of the way throughout the research process. You have both helped me develop so much over the past year specifically in my writing and taught me so much about how this type of research is carried out. To Bianca Rossouw, thank you for constantly inspiring me to always reach higher and to tackle everything that I take on with 100% of my effort. The work you do as a genetic counsellor and everything you are involved in is so admirable. To Dr Bronwyn Dillon thank you for your help and clinical expertise, without you this project would not have been possible. The way that you approach all that you do with so much grace and positivity is so uplifting, and it was a joy to work with you on this project. Thank you both for managing to carve out time for me with your immensely busy schedules, I am forever grateful to you both.

I would like to thank everyone at the Division of Human Genetics at the NHLS, Braamfontein who has taught me so much over the past year and contributed to my professional development as a genetic counselling student.

Thank you to the Wits Postgraduate Merit award, which made the year financially attainable for me. The opportunities and connections that have been afforded to me through the completion of this degree and the skills that I have developed through being in this environment are priceless.

I would like to thank the other students in my MSc class, Asante, Reagan, Sli and Sam for all the support throughout the year. It was always comforting knowing there were others around experiencing the same thing as me. You are all fantastic people and superstar colleagues, and I have no doubt we will all go on to achieve great things.

Lastly, I would like to thank my family: Julie, John, Kayla, Dani, Hloni and Francina, Will, and all of my friends for always encouraging me to keep going and for reminding me that all of the work that I was doing was important would hopefully change lives.

Table of Contents

List of Appendices

List of Figures

List of Tables

List of Abbreviations

ACMG/AMP	American College of Medical Genetics and Genomics and the Association for Molecular Pathology
CMA	Chromosomal microarray analysis
CHDs	Congenital heart defects
CNS	Central nervous system
CNVs	Copy number variants
DD	Developmental delay
DECIPHER	Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources
FISH	Fluorescence *in situ* hybridization
HPO	Human phenotype ontology
ID	Intellectual disability
LCR	Low copy repeats
MLPA	Multiplex ligation-dependent probe amplification
NAHR	Non-allelic homologous recombination
NHLS	National Health Laboratory Service
NGP	Next-generation phenotyping
REDCap	Research Electronic Data Capture
VUS	Variant of unknown significance

Research Report in a “submissible” paper format

To be submitted to Molecular Genetics and Genomic Medicine ISSN:2324-9269

Characterization of the genetic and phenotypic landscape of patients with CNVs in the chromosome 22q11.2 region in a cohort of South African patients

Dylan McCarthy[1], Bronwyn Dillon[1], Bianca Rossouw[1]

Correspondence: Dylan McCarthy (dylan.james.mccarthy2711@gmail.com) | Bianca Rossouw (bianca.carzis@wits.ac.za) | Bronwyn Dillon (bronwyn.dillon@wits.ac.za)

Funding: This study was a retrospective analysis of freely available data and thus did not require funding from any entity.

Keywords: 22q11.2 deletion syndrome | 22q11.2 duplication syndrome | genotype-phenotype correlations | Face2Gene | South Africa

[1] Division of Human Genetics, National Health Laboratory Service and School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, 2001 South Africa

ABSTRACT

Background

Copy number variants (CNVs) in the chromosome 22q11.2 region are among the most frequently identified genomic rearrangements. However, their phenotypic variability in African populations remains poorly characterised due to limited access to diagnostics and underrepresentation in genomic research.

Methods

We retrospectively reviewed 129 sub-Saharan African patients diagnosed with CNVs in the chromosome 22q11.2 region over ten years. Clinical, demographic, and genetic data were collected from fluorescence *in situ* hybridisation, multiplex ligation-dependent probe amplification, and chromosomal microarray analysis (CMA). CNVs were classified by low copy repeat involvement, and genotype-phenotype correlations were explored. A subset of 11 patients underwent analysis using Face2Gene, an AI-driven next-generation phenotyping tool.

Results

Deletions accounted for 71% and duplications 29% of cases. Classic CNVs were observed in 68% patients and atypical CNVs in 32%. A statistically significant association was found between classic CNVs and congenital heart defects (CHDs). The higher frequency of CHDs, immunological, endocrinological, and developmental abnormalities mirrored global trends however, there was a lower frequency of palatal defects observed. In total, 181 distinct phenotypic features were documented, highlighting the vast phenotypic spectrum associated with these CNVs. Face2Gene correctly identified all deletions using a gestalt-only approach, but performance decreased with a combined gestalt and feature score approach.

Conclusion

This study addresses gaps in genomic data from African populations, while demonstrating the vast phenotypic spectrum of CNVs in the chromosome 22q11.2 region. The introduction of CMA markedly improved diagnostic yield annually. While Face2Gene performed well, further validation in larger,more diverse cohorts is needed for improved clinical implementation.

1 Introduction

Microdeletion and microduplication syndromes represent a group of genetic disorders characterised by submicroscopic deletions and duplications of contiguous genes located within specific chromosomal regions (Watson et al., 2014). Among these, 22q11.2 deletion syndrome and 22q11.2 duplication syndrome are two of the more commonly identified microdeletion and microduplication syndromes in human genetics clinics worldwide (Catusi et al. 2020), with 22q11.2 deletion syndrome being the most commonly identified microdeletion syndrome in previous literature (McDonald-McGinn et al., 2015). Global incidences for both 22q11.2 deletion syndrome and 22q11.2 duplication syndrome are yet to be reported but North American studies have estimated the incidence of 22q11.2 deletion syndrome to be between 1:2000 and 1:4000 (Bartik et al., 2022; Blagojevic et al., 2021). However, the true incidence is likely higher than this as many cases go undiagnosed due to the range of clinical severity observed in affected individuals (Bartik et al., 2022). A Danish study conducted by Olsen et al. (2018) reported the incidence of 22q11.2 duplication syndrome to be 1:1606. This is also likely to be an underestimation as the clinical features observed in microduplication carriers are often milder than the clinical features observed in microdeletion carriers, resulting in 22q11.2 microduplications being identified less frequently than 22q11.2 microdeletions (Portnoï, 2009; Van Campenhout et al., 2012).

1.1 Genomic architecture and classification of CNVs in the chromosome 22q11.2 region

The chromosome 22q11.2 region consists of eight highly homologous low copy repeat (LCR) blocks, designated LCR22A-LCR22H. These LCRs predispose individuals to chromosomal rearrangements via non- allelic homologous recombination (NAHR) and their presence facilitates recurrent copy number variants (CNVs) within the chromosome 22q11.2 region (Shaikh, 2000; Shaikh et al., 2001). These recurrent CNVs can range in size depending on their LCR involvement, the most common of which being an ~3Mb CNV occurring between the LCR22A-D blocks. The advent of newer molecular diagnostic technologies has since enabled the identification of smaller nested and atypical CNVs within the region, expanding our understanding of its genomic complexity. These atypical/nested CNVs are labelled according to their position within the chromosome 22q11.2 region, either referred to as proximal, central or distal CNVs (Gavril et al., 2022). Proximal CNVs span the region between LCR22A and LCR22B, central CNVs are those which span the LCR22B-LCR22D region and distal CNVs span the LCR22D to LCR22H region (Figure 1). Historically, CNVs in the chromosome 22q11.2 region were classified according to their associated phenotypic presentations,

implicating a variety of genetic disorders with the region, including DiGeorge syndrome, Velocardiofacial syndrome, Cayler syndrome, Conotruncal anomaly syndrome, Opitz G syndrome, CATCH 22 syndrome, and Sedlackova syndrome (Burnside, 2015). These previously recognized conditions were subsequently found to be primarily associated with the recurrent ~3Mb deletion spanning the LCR22A-D blocks, now referred to as the 'classic' deletion. Currently however, all deletions and duplications identified within this region are clinically classified as either 22q11.2 deletion syndrome or 22q11.2 duplication syndrome thus encompassing all of these previously recognized conditions (Burnside, 2015). Other molecular mechanisms such as translocations, mosaicism, and supernumerary marker chromosomes may result in triplications involving the chromosome 22q11.2 region, however these CNVs often result in distinct phenotypes such as Emmanuel syndrome and Cat eye syndrome (Carter et al., 2009).

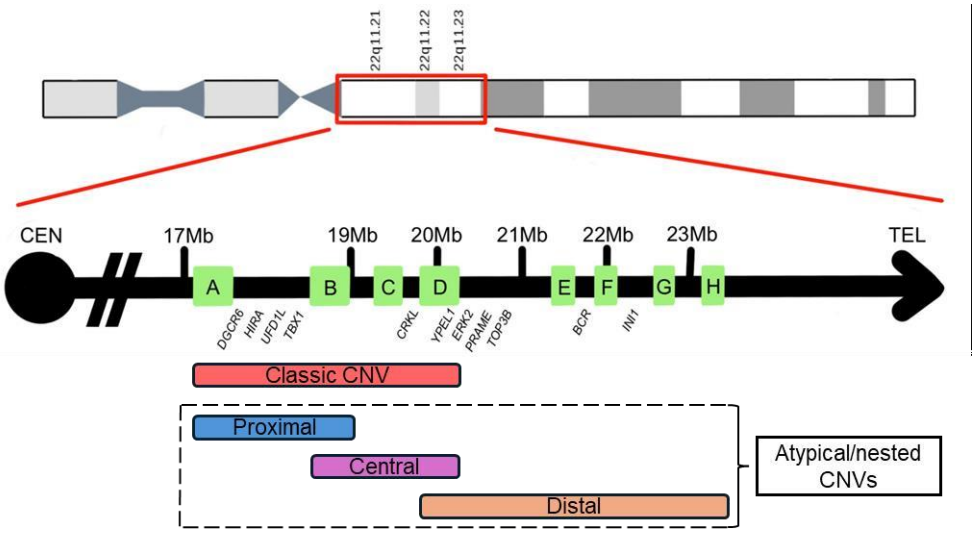


Figure 1 | Schematic representing the recurrent classic and atypical CNVs identified in the 22q11.2 region.

The classic 3Mb CNV is represented in red, while atypical proximal, central and distal CNVs are represented in blue, purple and orange respectively.

1.2 Clinical features associated with CNVs in the chromosome 22q11.2 region

Individuals affected by a microdeletion within the chromosome 22q11.2 region display a broad range of clinical manifestations, which most commonly include congenital heart defects (CHDs), craniofacial dysmorphism, immunodeficiency, endocrine abnormalities, developmental delay (DD) and neuropsychiatric disease (Table 1). Musculoskeletal, gastrointestinal, skin and genitourinary system involvement has also been reported to variable degrees (Bartik et al., 2022; Cirillo et al., 2022; Kruszka et al., 2017). In contrast, 22q11.2

duplication syndrome often presents with milder clinical presentations such as learning disabilities, DD, and other neuropsychiatric manifestations (Clements et al., 2017; Olsen et al., 2018; Van Campenhout et al., 2012) (Table 1).

Table 1 | System involvement observed in patients with chromosome 22q11.2 deletions and duplications respectively.

The data was adapted from from Bartik et al. (2022); Cirillo et al. (2022); Du et al., (2020); and Nehme et al., 2022

Clinical phenotype	% occurrence	
	Deletions	Duplications
Dysmorphic craniofacial features	50	26
Congenital heart defects	75	42
Gastrointestinal anomalies	30	84
Renal anomalies	14	52
Skeletal defects	60	56
Hypocalcaemia	35	6
Immune deficiency	50-70	54
Developmental delay	50	58
Learning disabilities	70	22
Psychiatric disorders	30	59

1.3 Diagnostic Challenges associated with CNVs in the chromosome 22q11.2 region, Evolving Technologies, and Access to Chromosomal Microarray in South Africa

The inter- and intrafamilial variability in the clinical presentation of chromosome 22q11.2 syndromes often delays early detection and diagnosis, as symptoms can range from subtle to severe and may involve multiple organ systems (Bartik et al., 2022; Cirillo et al., 2022). Some patients may present with atypical features that do not immediately suggest the underlying genetic condition, resulting in misdiagnosis or prolonged diagnostic odysseys (Kruszka et al., 2017; Sgardioli et al., 2019). Moreover, genotype- phenotype correlations for both 22q11.2 deletion syndrome and 22q11.2 duplication syndrome have not been well established in prior studies conducted in North American and European populations, leading to uncertainty regarding whether specific genetic variants in the 22q11.2 region correlate with any particular

clinical outcomes (Gavril et al., 2022). Furthermore, studies that have included a few individuals of African ancestry have mostly commented on CHDs but have also indicated that these individuals exhibit the least consistent phenotypic presentations associated with 22q11.2 deletion syndrome when compared to other populations (Kruszka et al., 2017; Wonkam et al., 2017). The evolution of molecular genetic testing has significantly enhanced our understanding of CNVs in the 22q11.2 region. Early diagnostic approaches primarily involved fluorescence in situ hybridization (FISH), which allowed for the detection of the classic 3Mb LCR22A–LCR22D CNV (Shaikh et al., 2001). However, FISH was limited in its ability to identify smaller nested and atypical microdeletions or microduplications (Jalali et al., 2008). The introduction of multiplex ligation-dependent probe amplification (MLPA) and chromosomal microarray analysis (CMA) has revolutionized the field by enabling the identification of atypical CNVs with greater accuracy (Jalali et al., 2008). CMA, in particular, has emerged as a first-tier diagnostic tool because of its ability to provide genome-wide copy number analysis and allow for the potential establishment of genotype-phenotype correlations (Sgardioli et al., 2019).

In the South African public healthcare system, CMA was officially implemented at the National Health Laboratory Service (NHLS) in 2021, with limited availability prior to this time (F. Baine-Savanhu, personal communication, June 19, 2025). CMA is currently only accessible to state patients who meet specific clinical criteria. These include individuals with developmental delay or intellectual disability of suspected genetic origin, multiple congenital abnormalities, autism spectrum disorder with developmental concerns, dysmorphic features suggestive of a genetic syndrome, or a suspected microdeletion/microduplication syndrome such as 22q11.2 deletion or Williams syndrome. Patients may also qualify if they are suspected, following clinical genetics assessment, to have a chromosomal abnormality not typically detected by other methods (e.g. aneuploidy). Testing is not available to private patients through the NHLS; instead, private referrals are directed to laboratories such as Lancet and Ampath that offer microarray services. Access to CMA testing typically requires referral by a clinician with appropriate suspicion of an underlying genetic condition. Referrals can be made by external doctors working within the public health system or following assessment by a clinical geneticist or genetic counsellor. Medical geneticists primarily consult within state hospital settings across various provinces, where they see patients through tertiary or quaternary referral centres. Testing is only processed when supported by detailed clinical information, including developmental history, dysmorphology, and family background, to ensure appropriate analysis and interpretation (F. Baine-Savanhu, personal communication, June 19, 2025).

1.4 Limitations of Next-Generation Phenotyping and the Need for Population Diversity in Genomic Research, particulary with respect to CNVs in the chromosome 22q11.2

region

As an augment to clinical diagnosis, next-generation phenotyping (NGP) technologies such as the Face2Gene application (FDNATM Inc.; https://www.face2gene.com) have begun to show promise in enhancing diagnostic accuracy and guiding clinical diagnoses through advanced computational methods and machine learning algorithms (Gurovich et al., 2019). However, these NGP tools have been designed using cohorts of mostly European ancestry and therefore require testing on independent datasets to validate the use of the technology in understudied and underrepresented populations before the software is to be trusted in a clinical setting (Hennocq et al., 2024). Despite these technological advancements, significant gaps remain in our understanding of the clinical variability associated with 22q11.2 syndromes, particularly in underrepresented populations (De Decker et al., 2016; Kruszka et al., 2017). The lack of comprehensive genomic and phenotypic data from African populations presents a significant barrier to the effective diagnosis and management of these conditions. Historically, African individuals have been underrepresented in genomic studies, which often focus on populations of European descent (Ramsay, 2022). This underrepresentation raises concerns regarding the generalisability of findings and the potential for missed diagnoses in diverse populations which may display more diverse presentations.

While much is known about the molecular genetics of 22q11.2 deletion syndrome and duplication syndrome, there is still much to be discovered with respect to understanding the causes behind the variable phenotypic presentation of the conditions. Newer technologies such as those incorporating NGP aim to improve clinical diagnoses of many genetically heterogeneous conditions, including 22q11.2 deletion and duplication syndrome, but are yet to be validated in South African patients. With the routine implementation of CMA technology, more precise genotypic information is now available for the establishment of genotype-phenotype correlations in patients of African ancestry. Therefore, the aim of this study is to characterize the CNVs that have been previously identified in the 22q11.2 region, and to determine any genotype-phenotype correlations in a South African patient cohort over a 10-year period.

2 Materials and Methods

2.1 Editorial Policies and Ethical Considerations

This research was approved by the higher degrees committee of the Faculty of Health Sciences, University of the Witwatersrand (Appendix A1) and the National Health Laboratory Service (NHLS) Academic Affairs and Research Management System (Appendix B). Ethics

approval was obtained from the University of the Witwatersrand Human Research Ethics Committee (Medical) (WITS-HREC Ref no. R14/49) with ethics clearance number: M240614. This ethics number approves the anonymous retrospective review of patient data and was approved under the proviso that parental informed consent that was obtained for the storage of patient photographs in their medical records and for their use in research. This consent was obtained by a clinician upon capturing the photograph and has been approved to be sufficient for this research by the University of the Witwatersrand Human Research Ethics Committee (Medical) (WITS-HREC Ref no. R14/49) as outlined in the research protocol (Appendix A).

2.2 Study Subjects

In this study, 138 patient records were retrospectively analysed. These patients underwent diagnostic testing at a single study site, the Division of Human Genetics, National Health Laboratory Service (NHLS) and School of Pathology, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa, between the years 2013 and 2023. The patient records included both physical patient files present at the Division of Human Genetics, NHLS Braamfontein and digitized files from an online Research Electronic Data Capture (REDCap) database (https://redcap.core.wits.ac.za) (Harris et al., 2009, 2019). The patients were classified as either internal patients or external patients. Internal patients were defined as patients who had been referred for a consultation with a medical geneticist at the Division of Human Genetics and therefore had a clinical file available which included comprehensive clinical information, phenotypic data and patient photographs. External patients were defined as patients who had only undergone diagnostic testing within the Division of Human Genetics and thus the only available clinical information included the information documented on the test request form. All of the patients included in this study were either referred for testing for possibly having a suspected deletion or duplication in the chromosome 22q11.2 region, or in other instances referred more generally for broad genetic testing on account of a characteristic clinical feature which warranted genetic testing, such as CHDs, DD or intellectual disability (ID) being indicative of CMA testing for example.

2.2.1 Inclusion Criteria

Records were only included in the study if they underwent genetic testing dated between January 2013 and December 2023 and if these patients tested positive for a CNV in the 22q11.2 region detected by FISH (Vysis LSI DiGeorge Region Probe Kit; Abbott Laboratories, Abbott Park, IL, USA), MLPA (SALSA® MLPA® P245 Microdeletion Kit and the SALSA® MLPA® P070 DiGeorge Kit; MRC- Holland, Amsterdam, The Netherlands), or CMA (Agilent SurePrint G3 CGH Microarray, 8x60K; Agilent Technologies, Santa Clara, CA, USA). with data

interpretation performed using Agilent CytoGenomics software, version 5.3 (Agilent Technologies, Santa Clara, CA, USA).

A total of 138 patients in whom CNVs in the chromosome 22q11.2 region were identified in this study. The CNVs identified in these patients were classified as pathogenic, likely pathogenic, variants of uncertain significance (VUS), likely benign, or benign based on the American College of Medical Genetics and Genomics and Association for Molecular Pathology (ACMG/AMP) criteria (Riggs et al, 2020). Notably, patients with CNVs classified as VUS, likely benign, or benign were included in the study despite their non-pathogenic classifications as these classifications were influenced by factors such as phenotypic variability, incomplete penetrance, low dosage, or insufficient sensitivity and curation of evidence, leaving them with inconclusive data for a pathogenic classification at the time of testing. However, their phenotypes were still considered relevant for the purposes of this research, and therefore all of these variants were analysed as if they were pathogenic.

Of the 138 patients initially identified, nine were excluded due to the presence of additional genetic factors that could independently influence/confound their phenotype. These cases included four patients with Emanuel syndrome, resulting from a translocation of chromosomes 11 and 22 that caused a triplication of the 22q11.2 region, one patient with both a 6p27 deletion and a 22q11.2 deletion, one patient who was mosaic for trisomy 22, one patient with a marker chromosome 22 causing a triplication, one patient with a classic 22q11.2 deletion as well as a pathogenic *NF1* variant, and one patient with a 22q11.2 triplication detected by MLPA, for which the underlying mechanism was unknown. After excluding these nine patients, the final study cohort consisted of 129 patients.

2.3 Retrospective File Review

All 129 records were retrospectively analysed. Information extracted from the files included the type of diagnostic test requested (FISH, MLPA, CMA), the genetic testing result, demographic information including age of referral, sex, ethnicity, and referring hospital of the patient. This information was captured using a data collection sheet designed using the online REDCap database (Appendix A.1) (Harris et al., 2009, 2019). All available clinical phenotypic data were collected for the patients included in the study. For internal patients, phenotypic data were retrieved from each patient's clinical file as well as from their test request form, because internal patients may have been assessed on multiple occasions during follow-up appointments, updated clinical information from each entry was collected. For external

patients, clinical phenotypic data were limited only to information documented on the patient's test request form. The phenotypic data was also captured using a clinical tick-sheet designed using REDCap (Appendix A.2). Microsoft® Excel® (Microsoft Corporation, Redmond, VA) for Microsoft 365 (Version 2410 Build 16.0.18129.20100) as well as the REDCap database were used for data handling and statistical analysis. The data were analysed using descriptive statistical methods, where categorical variables were summarised as frequencies and percentages, while continuous variables were summarised as medians and interquartile ranges for non-normally distributed data and means and standard deviations for normally distributed data. Statistical analyses included using the Shapiro-Wilk test for normality, followed by the Wilcoxon rank-sum test to identify statistically significant differences between groups. Statistical analyses were conducted using Microsoft® Excel® for Microsoft 365 MSO (Version 2410, Build 16.0.18129.20100) with the XRealStats add-in software (Release 9.2.2).

2.4 Genotype-Phenotype Correlations

All 129 patient files were included in this aspect of the study, comprising 45 internal patients and 84 external patients. As the phenotypic data available for external patients were limited to information documented on their test request forms, each phenotypic feature was categorised as either assessed or unassessed, with assessed values being denoted as either present or absent for the particular phenotype and reported using Human Phenotype Ontology (HPO) terms. As a result, the number of internal and external patients considered for each system and phenotypic feature varied, with only those patients for whom the phenotype was assessed being included in the analysis. The genotypes of the patients were categorised depending on their LCR involvement. For CNVs identified using MLPA this was determined by assessing the probe involvement of either the standard p245 MLPA kit (MRC, Holland) or the chromosome 22q11.2 specific p070 MLPA kit (MRC, Holland). For CNVs identified using CMA the Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources (DECIPHER) (Firth et al., 2009) was used to clarify the breakpoints observed in the patients. Phenotypic counts were recorded and summarised as frequencies and percentages. As the amount of clinical information available in each patient's clinical file or referral form varied, the number of clinical features recorded for each patient also varied. Genotype-phenotype correlations were assessed using Chi-square or Fisher's exact tests where appropriate. Bonferroni correction was applied to adjust for multiple comparisons, with a significance level set at p-value< 0.05/n. Statistical analyses were conducted using Microsoft® Excel® for Microsoft 365 MSO (Version 2410, Build 16.0.18129.20100) with the XRealStats add-in software (Release 9.2.2).

2.5 Assessment of Face2Gene for the Identification of chromosome 22q11.2 Deletion/Duplication Syndrome

Eleven patients were included in the Face2Gene aspect of the study, as they were the only patients in the cohort with facial photographs recorded in their files. The patient's facial photographs present in their file were uploaded to the Face2Gene software, for each Face2Gene analysis, a single patient photograph of the patient's frontal facial profile can be uploaded (FDNATM Inc.; (https://www.face2gene.com). Face2Gene is an AI-powered NGP tool that analyses facial features from patient photographs to assist in the clinical diagnosis of genetic syndromes. It compares the input image to a large database of known syndromes to generate a ranked list of 30 possible diagnoses based on facial similarity and clinical features. Face2Gene maintains patient anonymity and privacy as it converts patient photographs into deidentified facial descriptors, while the original images are considered personal health information. Face2Gene does not store any personal health information, and this information is only accessible through a password protected Face2Gene user account. The data generated by Face2Gene included the facial dysmorphism-score (D-score), gestalt score, feature score, and a ranked list of the 30 most likely genetic syndromes suggested for the patient. This data was recorded using Microsoft® Excel® (Microsoft Corporation, Redmond, VA) for Microsoft 365 MSO (Version 2410 Build 16.0.18129.20100). The Face2Gene analysis was conducted in multiple stages: (1) assignment of D- scores to classify dysmorphism levels as low, unknown significance, or high; (2) analysis using facial photographs alone to generate a gestalt score assigned as low, medium or high based on the patient's facial features using the computational software designed by Face2Gene (DeepGestalt™) followed by the generation of a ranked list of the top 30 recommended genetic diagnoses; and (3) analysis combining facial photographs and HPO terms to calculate feature scores assigned as low, medium or high using the FeatureMatcher algorithm and refine syndrome suggestions and rankings. Gestalt scores were recorded as a quantitative measure of how closely the patient's facial features matched reference gestalt profiles, while feature scores provided a ranking based on uploaded HPO terms. For each patient, the HPO term analysis was limited to seven terms, which were carefully selected by a medical geneticist to best represent the phenotypic features described in the patient's file. When examining the ranked lists, to account for possible lower rankings influenced by the ethnicity of the patients, if 22q11.2 deletion syndrome or 22q11.2 duplication syndrome was listed anywhere within the list of 30 suggested genetic diagnoses, they were classified as consistent for CNVs in the chromosome 22q11.2 region. If neither syndrome appeared in the recommended genetic diagnoses, they were classified as inconsistent for CNVs in the 22q11.2 region. Consistency rates were then reported for 22q11.2 deletion syndrome and 22q11.2 duplication syndrome to assess Face2Gene's ability to

accurately identify CNVs in the chromosome 22q11.2 region.

3 Results

3.1 Retrospective File Audit of Patient Records in whom CNVs in the 22q11.2 Region were Identified

The demographics of this cohort are described in Table 2. All patients were of self-reported ancestry and the cohort was comprised of patients identified in South Africa from various ancestral groups. Most patients however were of African ancestry. There were more male (61%) patients than female patients (39%) and in total deletions (88%) were more common than duplications (12%). The age distributions identified in this cohort were not found to be normally distributed following Shapiro-Wilk statistical analyses (Supplementary Table 1). Most patients were referred for testing before 3 years of age. Additionally, when comparing the ages of referral between classic and atypical CNVs, classic CNVs were associated with a statistically significant earlier median age of referral when compared to atypical CNVs (p-value= 0.004264) (Supplementary Table 1). Among the identified CNVs, classic deletions (n=84) were the most commonly identified, followed by atypical deletions (n=29). Atypical duplications (n=12) were more common than classic duplications (n=4) (Table 1)

Table 2 | Demographic data and comparison between patients with classic and atypical CNVs in the chromosome 22q11.2 region in a cohort of sub-Saharan African patients

Demographic data observed in n=129 patients			
Race	**Number (%)**	**Ethnolinguistic Group**[§]	**Number (%)**
African	111 (86)	Zulu	14 (10)
European	9 (7)	Tsonga	6 (4)
Mixed race	4 (3)	Sotho	5 (4)
Indian	2 (2)	Afrikaans	5 (4)
Unknown	3 (2)	English	4 (3)
Internally/externally referred	**Number (%)**	Tswana	3 (2)
Internally referred patients	44 (34)	Other	10 (7)
Externally referred patients	85 (66)	Unknown	92 (66)
Patients with atypical CNVs (n=41)		Patients with classic LCR22A-D CNVs (n=88)	
Sex	**Number (%)**	**Sex**	**Number (%)**
Male	25 (61)	Male	54 (61)
Female	16 (29)	Female	34 (29)
Type of CNV	**Number (%)**	**Type of CNV**	**Number (%)**
Deletions	29 (71)	Deletions	84 (96)
Duplications	12 (29)	Duplications	4 (4)
Age at referral for genetic testing[*]		**Age at referral for genetic testing**[*]	
Median	1 year 5 months	Median	6 months
IQR[†]	4 years 8 months	IQR[†]	1 year 9 months
Minimum	1 day	Minimum	1 day
Maximum	40 years 6 months	Maximum	29 years
Age when test requested[‡]	**Number (%)**	**Age when test requested**[‡]	**Number (%)**
Neonate (0–27 days)	3 (7)	Neonate (0–27 days)	25 (28)
Infant (30 days–1 year)	14 (34)	Infant (30 days–1 year)	36 (40)

Toddler (1–2 years)	12 (29)	Toddler (1–2 years)	10 (11)
Early childhood (3–5 years)	4 (10)	Early childhood (3–5 years)	7 (8)
Middle childhood (6–11 years)	4 (10)	Middle childhood (6–11 years)	8 (9)
Adolescence (12–17 years)	1 (2)	Adolescence (12–17 years)	1 (1)
Adulthood (> 17 years)	3 (7)	Adulthood (> 17 years)	1 (1)

* = P-value<0,05; †= Interquartile range; ‡= Ages stratified according to the Munich system (Althammer et al., 2023); §= Where individuals have parents from differing ethnolinguistic groups, both were recorded. Other ethnolinguistic groups include Sepedi, Venda, Hindi, Muslim, Xhosa, Shona, Swati, Mixed ancestry, Congolese, Ethiopian.

In general, the number of CNVs identified per year increased over time with the introduction of newer testing technologies such as MLPA and CMA (Figure 2). The number of atypical CNVs identified per year also increased over time, with the highest number of atypical CNVs identified occurring in 2021 and 2022 (Figure 2). Trends were also observed when separately comparing the number of atypical and classic deletions and duplications per year (Supplementary Figure 1 and Supplementary Figure 2). The number of classic and atypical deletions identified per year increased steadily with the introduction of newer technologies such as MLPA and CMA. In contrast, the number of classic and atypical duplications showed only a marginal increase following the introduction of MLPA and fluctuated inconsistently across the years thereafter. Thus, it is evident that the increase in diagnostic yield mirrors the introduction of the new diagnostic technologies, with increases being observed with the introduction of MLPA and then further increases observed with the introduction of CMA.

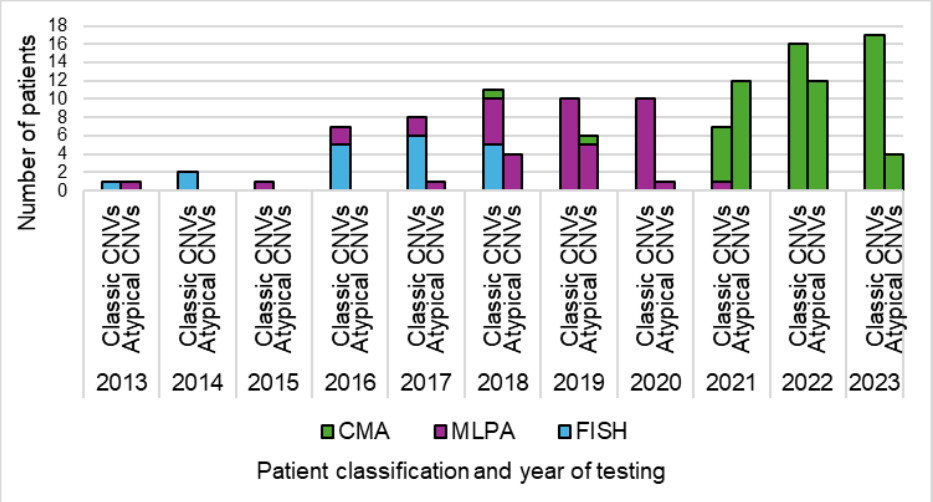


Figure 2 | The annual number of classic LCR22A-D CNVs and atypical CNVs identified in the chromosome 22q11.2 region using various testing methodologies over 10 years in a cohort of South African patients.

The total number of CNVs identified annually includes both deletions and duplications.

When comparing the LCR region involvement observed in patients with atypical deletions, proximal CNVs were the most commonly identified CNVs (Figure 3). Central CNVs accounted for the next highest proportion of atypical CNVs and distal CNVs were the least frequently identified group of CNVs (Figure 3).

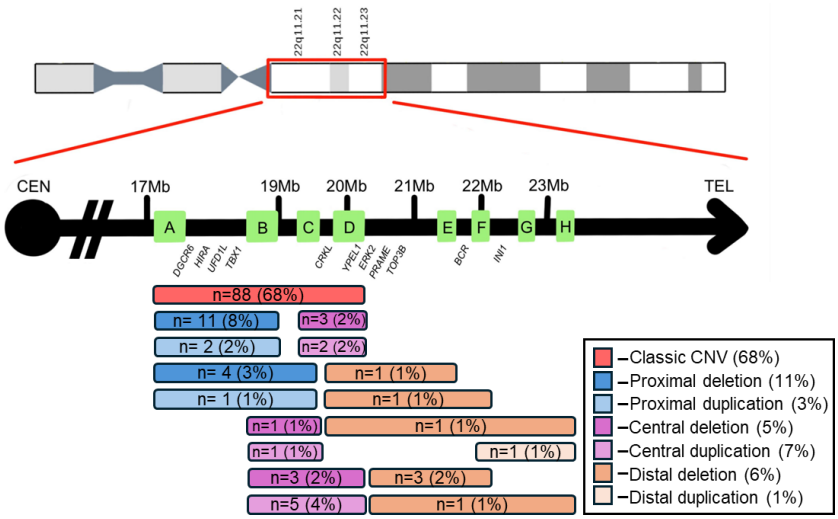


Figure 3 | Schematic representation of the distribution CNVs identified in the chromosome 22q11.2 region in a cohort of sub-Saharan African patients, n=129.

In terms of the referral indication for patients with classic CNVs (Figure 4), the majority of the patients presented with either dysmorphic features and/or CHDs/pulmonary/chest abnormalities with all the other reasons for referral being noted at low frequencies. The CHDs/pulmonary/chest category included structural heart defects (such as septal defects, outflow tract anomalies, and valve abnormalities), pulmonary findings (including pulmonary hypoplasia and respiratory difficulties), and selected thoracic features (such as widely spaced nipples). This grouping was used to reflect anomalies affecting the heart, lungs, and chest wall region as documented in clinical records. Patients were counted multiple times if multiple reasons for referral were indicated. The maximum number of reasons for referral observed in a single patient was seven and the minimum number of reasons being one. When considering patients with atypical CNVs, a wider array of reasons for referral were observed (Figure 4). When conducting similar comparisons between internally referred and externally referred patients, internally referred patients predominantly presented with dysmorphic features and central nervous system (CNS)/developmental abnormalities. Externally referred patients were most commonly referred for testing because of CHDs/pulmonary/chest abnormalities and dysmorphic features (Figure 5). Additionally, immunodeficiency and endocrine abnormalities were notably more commonly reported as reasons for referral among external patients when

compared to internal patients (Figure 5).

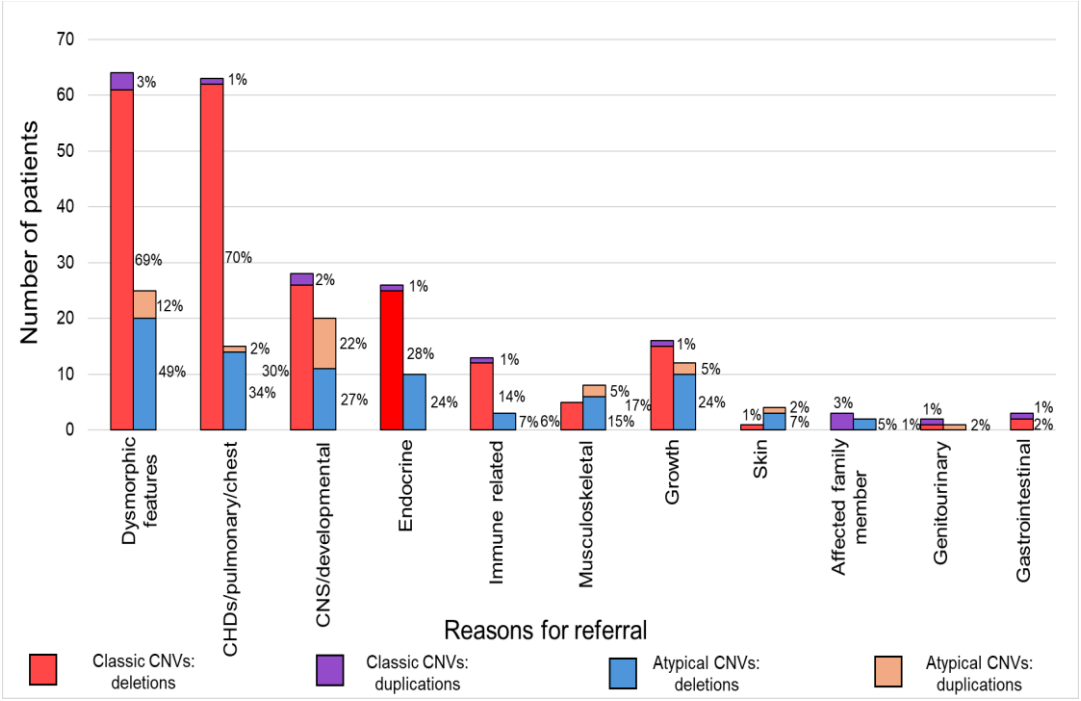


Figure 4 | The reasons for referral for patients with classic LCR22A-D CNVs, and patients with atypical CNVs in the chromosome 22q11.2 region.

Patients with classic CNVs are represented by the bars on the left and patients with atypical CNVs are represented by the bars on the right, patients could be counted multiple times if multiple reasons for referral were indicated on the referral form or in their clinical file (n=88 patients with classic CNVs and n=41 patients with atypical CNVs)

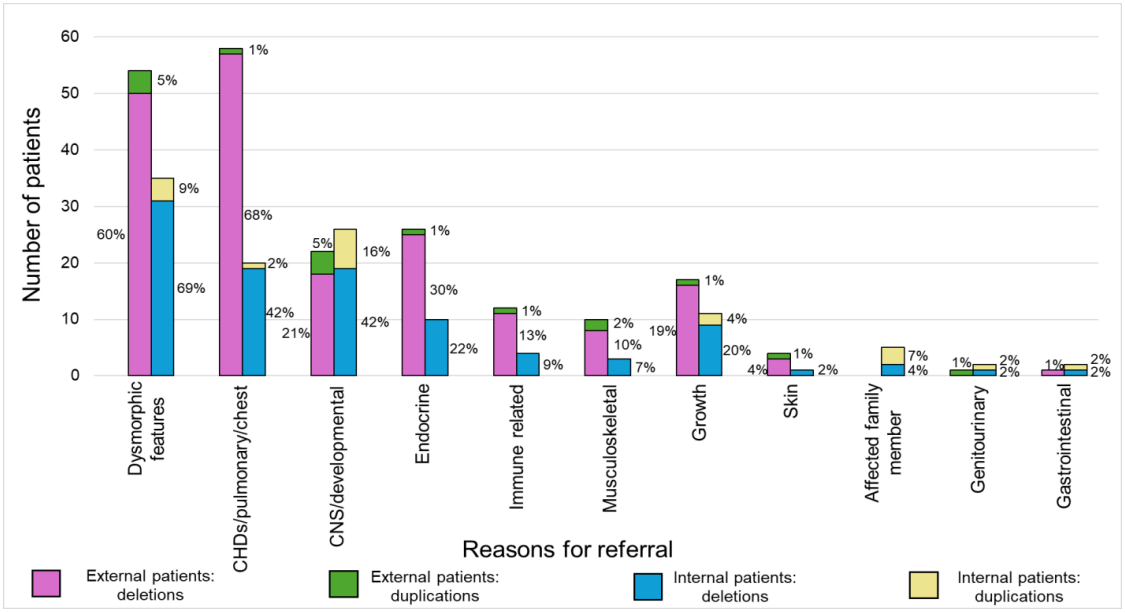


Figure 5 | The reasons for referral for external patients and internal patients with CNVs in the chromosome 22q11.2 region.

External patients are represented by the bars on the left and internal patients are represented by the bars on the right, patients could be counted multiple times if multiple reasons for referral were indicated on the referral form or in their clinical file (n=84 external patients, and n=45 internal patients).

3.2 Genotype-Phenotype Correlations

3.2.1 Genotype-phenotype correlations in patients with classic LCR22A-D CNVs

Patients with classic LCR22A-D CNVs in the 22q11.2 region exhibited a wide range of abnormalities across multiple systems, detailed in figure 6 and supplementary table 2. Among those with classic CNVs, deletions were identified far more frequently than duplications. The body systems observed at the highest frequencies in patients with classic deletions included dysmorphic features, CHDs/pulmonary/chest abnormalities and CNS/developmental abnormalities. Some of the notably prevalent clinical features observed in these systems included various abnormalities of the ears (52%) and nose (62%), conotruncal heart defects (>50%), DD (55%) and learning difficulties (26%) (Supplementary table 2). Growth abnormalities were common, with features associated with growth restriction (short stature, low body mass, and microcephaly) being prevalent patients with classic deletions. Immune related and endocrinological abnormalities were also prevalent patients with classic deletions, with patients mostly presenting with recurrent infections (45%) and hypocalcaemia (34%). The musculoskeletal abnormalities observed in patients with classic deletions mostly involved various abnormalities of the limb, the most common of which being unilateral clubfoot being observed in 11% of patients in which musculoskeletal features were assessed. Skin abnormalities, genitourinary abnormalities gastrointestinal abnormalities were rarely observed in patients with classic deletions.

When considering classic duplications, the most commonly affected body systems were dysmorphic features, CNS/developmental abnormalities and genitourinary abnormalities, with CHDs/pulmonary/chest abnormalities, musculoskeletal abnormalities, immune deficiency and endocrine abnormalities only observed in one third of the patients with classic duplications.

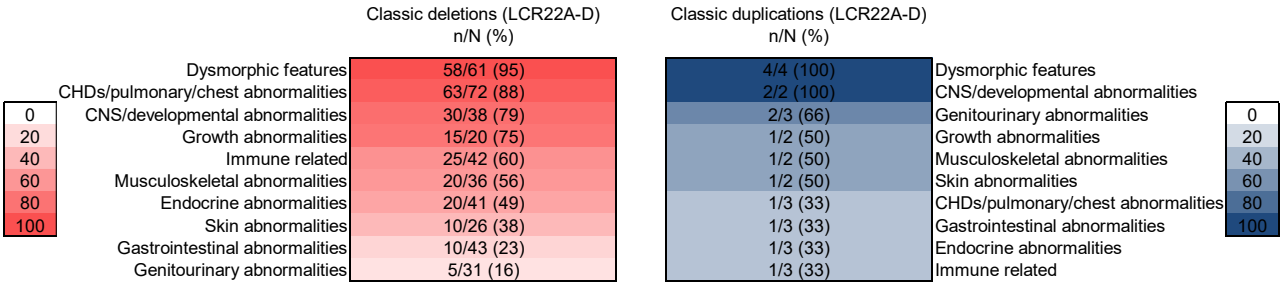


Figure 6 | Heatmaps representing the frequencies of affected body systems observed in sub-Saharan African patients with classic LCR22A-D CNVs in the chromosome 22q11.2 region (n=88).

There were at total of 84 patients with classic deletions and four patients with classic duplications. Frequencies of each affected body system are represented as the number of patients observed whereby the body system

was affected out of the number of patients assessed for clinical features belonging to the particular body system (n/N (%)).

3.2.2 Genotype-phenotype correlations in patients with atypical CNVs in the chromosome 22q11.2 region

Patients with atypical deletions and duplications in the 22q11.2 region also exhibited a range of abnormalities across multiple systems (Supplementary table 3). For atypical deletions, most of the clinical features were present in over half of the individuals assessed, however none of the patients presented with features in the gastrointestinal or genitourinary systems. Interestingly, 82% of the CHDs/pulmonary/chest abnormalities and 75% of musculoskeletal abnormalities observed in patients with atypical CNVs, were observed in patients with atypical deletions in the proximal and central LCR regions (Figure 7). Additionally, patients with atypical deletions in the proximal or central reasons also accounted for >50% of the CNS/developmental abnormalities as well as growth abnormalities observed in patients with atypical deletions. Most notably proximal deletions also accounted for all of the patients affected with immunodeficiency and all of the patients affected with endocrine abnormalities (Figure 7). In contrast, the only systems which affected more than 25% of patients with atypical duplications included musculoskeletal abnormalities, skin abnormalities, CNS/developmental abnormalities and dysmorphic features. CNS/developmental abnormalities and skin abnormalities were observed regularly in patients with central duplications, while musculoskeletal abnormalities and dysmorphic features were observed in patients with proximal, central and distal deletions (Figure 7).

Following statistical analyses using Fisher's exact test, statistically significant relationships were observed between CHDs/pulmonary/chest abnormalities and various genotypes within the cohort (Supplementary Table 2). Most interestingly, a statistically significant correlation was observed between CHDs/pulmonary/chest abnormalities and classic CNVs (deletion or duplication) when compared to atypical CNVs (Supplementary Table 2).

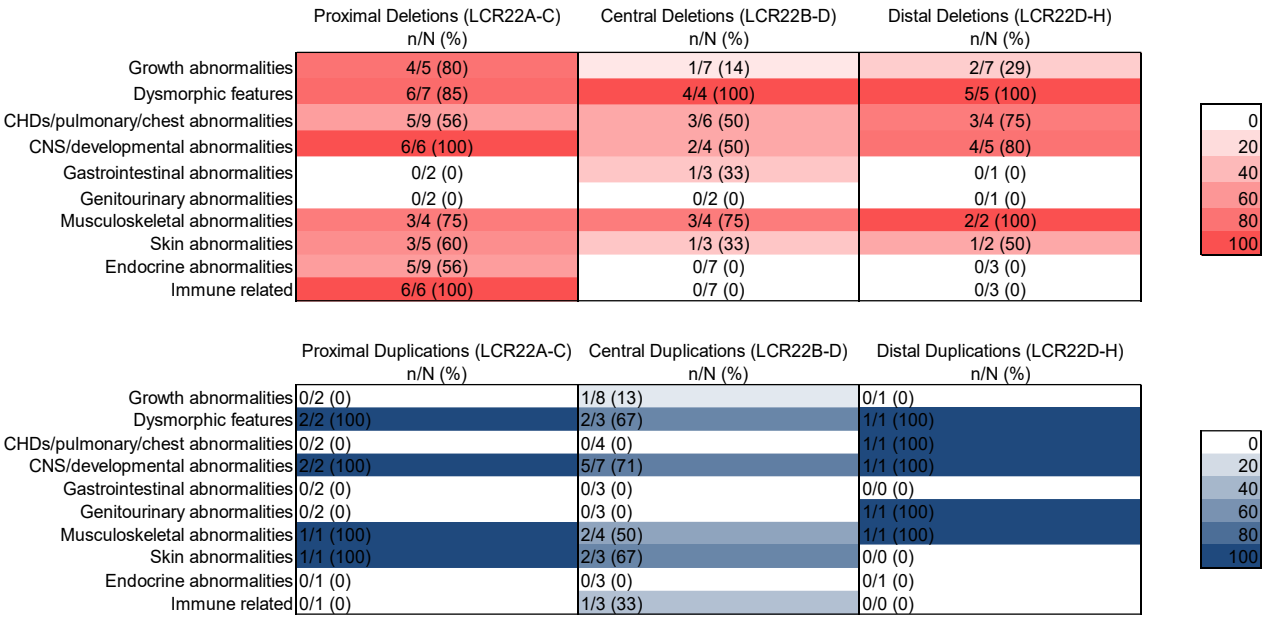


Figure 7 | The clinical features observed in sub-Saharan African patients with atypical CNVs in the chromosome 22q11.2 region (n=41). The clinical features were reported using HPO terms. There were at total of 29 patients with atypical deletions and 12 patients with atypical duplications. Frequencies of each affected body system are represented as the number of patients observed whereby the body system was affected out of the number of patients assessed for clinical features belonging to the particular body system (n/N (%))

3.3 Assessment of Face2Gene for the Identification of chromosome 22q11.2 Deletion/Duplication Syndrome

Altogether 11 patients were assessed using Face2Gene, six with classic deletions, three with atypical deletions, one with a classic duplication and one with an atypical duplication. Patient 74 was the only patient with two analyses conducted, due to having both high- and low-quality photos available. The seven most prominent HPO terms uploaded for each patient varied depending on the assessment by a medical geneticist. HPO terms where selected to include features that would not be identifiable on the photo (e.g., developmental delay, seizures, growth abnormalities), prominent dysmorphic features, as well as any significant dysmorphology that would not have been noted on a frontal photo (e.g., an ear abnormality). Using only facial photographs, 22q11.2 deletion syndrome was suggested as a possible diagnosis in all nine patients with deletions in the chromosome 22q11.2 region and thus demonstrated a 100% consistency rate. However, when using both facial photographs and feature scores 22q11.2 number decreased to seven out of nine patients, resulting in a consistency rate of 78%. In the two instances whereby Face2Gene did not suggest 22q11.2 deletion syndrome as a possible diagnosis, one patient had a classic deletion, and one patient had an atypical deletion in the 22q11.2 region (Table 4). For the two patients with 22q11.2 duplications, Face2Gene did not rank 22q11.2 duplication syndrome or 22q11.2 deletion syndrome among the 30 suggested diagnoses, regardless of whether the analysis used facial photographs alone or included both facial photographs and feature scores (0% consistency rate).

Additional factors influencing Face2Gene's classifications were also evaluated. The inclusion of feature scores consistently lowered the ranking of 22q11.2 deletion syndrome or excluded it entirely. Photograph quality also affected performance as high-quality photos (subjectively assessed) often ranked 22q11.2 deletion syndrome within the top five diagnoses, whereas medium and low-quality photos frequently ranked it below 10th or omitted it altogether (Table 4). For patients with classic deletions, photograph quality perfectly correlated with gestalt scores (Supplementary Table 3) and subsequent high/low rankings, but this pattern was not observed for atypical deletions. In patients with duplications, photograph quality appeared irrelevant, as Face2Gene was unable to rank 22q11.2 deletion syndrome or duplication syndrome for one patient (P74) with both high- and low-quality photos. Additionally, D-scores (Supplementary Table 3) had no clear impact, as some patients with high facial D-scores were misclassified, while others with low or uncertain D- scores correctly ranked 22q11.2 deletion syndrome. All but two patients were of sub-Suharan African ancestry with patients P85 and P87 being of European descent. Majority of the African patients were correctly classified using Face2Gene, demonstrating a 100% (7/7) consistency rate rate for African patients with deletions in the 22q11.2 region when using the photograph-only approach and an 86% consistencey rate (6/7) when using the gestalt and feature score approach.

Table 3 | Results of Face2Gene analysis conducted on sub-Saharan African patients with CNVs in the chromosome 22q11.2 region (n=11)

Patient code	Type of CNV	Ethnicity	Picture quality	22q11.2 CNV rank† using facial photographs only	22q11.2 CNV rank‡ using facial photographs and feature scores
P2	Classic deletion	African	High§	1	2
P6	Classic deletion	African	High¶	1	2
P87	Classic deletion	European	High¶	1	1
P91	Classic deletion	African	High¶	1	2
P88	Classic deletion	African	Medium§	3	5
P85	Classic deletion	European	Low§	10	N/A
P77	Atypical deletion	African	High§	1	1
P9	Atypical deletion	African	High¶	5	11
P59	Atypical deletion	African	Medium§	17	N/A
P5	Classic duplication	African	Medium§	N/A	N/A
P74	Atypical duplication	African	Low§	N/A	N/A

Notes: †- The rank at which 22q11.2 deletion syndrome or 22q11.2 duplication syndrome appeared in Face2Gene's list of 30 suggested genetic syndromes when only using facial photographs (gestalt scores); ‡- The rank at which 22q11.2 deletion syndrome or 22q11.2 duplication syndrome appeared in Face2Gene's list of the top 30 suggested genetic syndromes when using both facial photographs (gestalt scores) and feature scores; §- A scanned photograph was uploaded, subjectively assessed as low/medium/high quality; ¶- An original photograph was uploaded. Shading represents different types of CNVs. Picture quality was subjectively assessed depending on resolution, shadowing, obscuring factors and alignment. Abbreviations: N/A= 22q11.2 deletion syndrome was not identified in the ranked list of syndromes.

4 Discussion

4.1 Key Findings

This study provides insight into the genetic and phenotypic characteristics of sub-Saharan African individuals, predominantly of African ancestry, with CNVs in the chromosome 22q11.2 region. The demographic distribution of this cohort addresses a critical gap in global genomic research, where African populations remain underrepresented (Ramsay, 2022). Studies such as De Decker et al. (2016), Kruszka et al. (2017), and Wonkam et al. (2017) highlight how ancestral diversity can influence the phenotypic expression of genetic syndromes, with African populations often displaying less stereotypical features. Kruszka et al. (2017) found that while certain core features of 22q11.2 deletion syndrome are conserved across populations, the expression of other traits, particularly craniofacial features, varies by ancestry, underscoring the need for population-aware diagnostic approaches. Wonkam et al. (2017) demonstrated that the prevalence of 22q11.2 deletions among Cameroonian patients with CHDs was comparable to international data, providing initial evidence that hallmark clinical features such as CHDs are concordant in individuals of African ancestry. However, they emphasised the need for further validation in larger African cohorts. In the South African context, De Decker et al. (2016) reported that routine clinical suspicion failed to identify most cases of 22q11.2 deletions in children with CHDs, with prospective FISH screening revealing a prevalence (4.8%) nearly three times higher than the clinician-recognised rate. They also observed that patients of non-European ancestry often presented with subtle or absent dysmorphic features, contributing to underdiagnosis. Together, these findings reinforce the importance of molecular testing and diverse reference data in improving recognition and diagnosis of CNVs in the chromosome 22q11.2 region in African populations. Through the analysis of 129 patients, this study details the distribution of CNVs in the chromosome 22q11.2 region in a predominantly South African cohort, while also identifying trends in referral pathways. It emphasises the importance of newer diagnostic technologies for improved CNV detection and corroborates previous findings on genotype-phenotype correlations, guiding clinical management of chromosome 22q11.2 CNV syndromes in South Africa. Additionally, the potential utility of Face2Gene for diagnosing South African patients is highlighted, however further validation of Face2Gene with larger sample sizes is recommended.

4.2 Retrospective File Review

This study highlights key insights into the demographics, diagnostic advancements and referral patterns related to 22q11.2 deletion syndrome and 22q11.2 duplication syndrome in South Africa, emphasising the molecular and clinical landscape of CNVs in the chromosome 22q11.2 region. A male predominance was observed in the cohort in addition to the majority of patients receiving a confirmed diagnosis of a chromosome 22q11.2 CNV prior to the age of

3 years. This may be a result of sampling bias, however these findings mirror those of Napolitano et al. (2022) and Loomes et al. (2017), who suggest that male patients with neurodevelopmental disorders tend to be diagnosed at an earlier age, attributed to more noticeable atypical behaviours in males, and potential neurological differences between sexes. Deletions were far more prevalent than duplications, in keeping with global trends that associate 22q11.2 deletions with more severe clinical phenotypes and earlier recognition compared to 22q11.2 duplications (Cirillo et al., 2022; De Decker et al., 2016). Classic CNVs, predominantly involving the LCR22A-D region, were more common than atypical CNVs, corroborating their higher global incidence and historical research focus (Burnside, 2015; Clements et al., 2017; Gavril et al., 2022). Additionally, previous research has shown that the transition from FISH to MLPA and CMA has significantly improved CNV detection, particularly for atypical CNVs (Sgardioli et al. ,2019 and Jalali et al., 2008). The present study's findings support these results, as the introduction of newer technologies resulted in increased detection rates for both classic and atypical CNVs in the chromosome 22q11.2 region, further reinforcing that CMA should be a first-line diagnostic tool for phenotypically variable conditions such as 22q11.2 CNVs, consistent with recommendations by Catusi et al. (2020). Patients with classic CNVs were referred significantly earlier than those with atypical CNVs, reflecting the more distinct and classic phenotypic features of classic 22q11.2 deletion syndrome, such as CHDs and dysmorphic traits (Bartik et al., 2022). In contrast, atypical CNVs were marked by subtler or variable presentations (Burnside, 2015; Gavril et al., 2022) and were associated with delayed diagnoses. This trend was also observed when examining externally referred patients with the majority of external patients presenting with the classic features associated with 22q11.2 deletion syndrome. This suggests that patients originating from centres without access to clinical genetics services are less likely to be referred for genetic testing unless they present with these classic features, highlighting the potential underdiagnosis of CNVs in the chromosome 22q11.2 region on a national level. These findings emphasise the need to expand awareness of genetic conditions and improve national referral guidelines, addressing both diagnostic delays and regional inequities in patient care. Enhanced access to clinical genetic services and provider training are critical for early detection and equitable management of CNVs in the chromosome 22q11.2 region across South Africa

4.3 Genotype-Phenotype Correlations

Although many phenotypic presentations associated with CNVs in the chromosome 22q11.2 region in the present study align with existing literature, the broader phenotypic variability in CNVs in the 22q11.2 region highlights the challenges in delineating clear genotype-phenotype correlations, with statistically significant results only being achieved when investigating CHDs/pulmonary/chest abnormalities. This finding is unsurprising as it reflects previous

literature, whereby classic CNVs were strongly associated with CHDs, with 85% of patients presenting with cardiac anomalies, comparable to the 64% prevalence reported by Campbell et al. (2018). Moreover, the high prevalence of CHDs in this cohort may also reflect ascertainment bias as it is likely that patients with CHDs would have been referred to specialist centres with access to and knowledge of clinical genetic services. Majority of the CHDs identified in the present study were identified in individuals with proximal CNVs, supporting previous findings which emphasise the roles of genes found in the proximal 22q11.2 region such as *TBX1* (MIM: 602054)*, DGCR6* (MIM: 601279), and *CRKL* (MIM: 602007) in cardiac development (Carli et al., 2021). Similarly, DD and other CNS abnormalities affected 80% of patients with classic CNVs in the present study, aligning with Van Campenhout et al. (2012) and Catusi et al. (2020). Neuropsychiatric features were most common in central duplications, reinforcing the neurodevelopmental role of LCR22B-D involvement (Cancrini et al., 2014) and the potential neurodevelopmental effects of dosage changes in the *CRKL*, *SCARF2* (MIM: 613619), and *PI4KA* (MIM: 600286) genes previously suggested by Motahari et al. (2019). The frequency of dysmorphic features, particularly nasal and ear anomalies, were consistent with the findings of Campbell et al. (2018), while immunodeficiency and endocrine abnormalities were observed mostly in individuals with proximal deletions, reflected in the findings of Campbell et al. (2018) and Cheung et al. (2014). The growth abnormalities observed in this study suggest that CNVs in the chromosome 22q11.2 region are associated with disrupted growth as most patients presented with short stature, microcephaly and low weight, also aligning with previous reports by Campbell et al. (2018) and McDonald-McGinn et al. (2015), while notably there were only two patients observed with macrocephaly. The frequency of musculoskeletal abnormalities (60%) is in keeping with previous frequencies identified by Homans et al. (2018) who reported musculoskeletal abnormalities affecting up to 60% of individuals with CNVs in the chromosome 22q11.2 region. Gastrointestinal and genitourinary abnormalities were less frequently observed when compared to other body systems, consistent with global trends and previous findings (Campbell et al., 2018; Gavril et al., 2022). However, it is important to note that the low number of classic duplications (four patients) and the limited assessment of certain features in both classic and atypical CNVs, may limit the generalisability of these findings. The retrospective and cross-sectional design of this study also is emphasised, as the data relied solely on information available in patient files, without the opportunity to reexamine or further investigate patient features to improve the accuracy of the results.

There was a higher incidence of skin abnormalities observed in a small subset of patients in this study with atypical CNVs (five with atypical deletions and three with atypical duplications). While dermatological findings (particularly those secondary to immune dysfunction) such as

eczema, have been documented, skin abnormalities are otherwise infrequently reported in association with CNVs in the chromosome 22q11.2 region. In this study the numbers are too limited to draw firm conclusions. The most frequently observed skin features involved skin pigmentary abnormalities such as hyperpigmentation and hypopigmentation. Given the small sample size, these findings are limited in their significance but warrant further investigation in future larger studies. Interestingly palatal abnormalities were present in only ~17% of patients in whom dysmorphic features were assessed. This contrasts with prior reports, such as Jackson et al. (2019), who reported frequencies as high as 67%. Ascertainment bias may partially explain this difference, as it is notable that our cohort included both patients specifically referred for suspected 22q11.2 deletion syndrome or 22q11.2 duplication syndrome as well as those with incidental diagnoses following broader genetic testing, therefore patients may not have been referred for palatal defects, as perhaps referring clinicians were not aware of its association with CNVs in the 22q11.2 region. Another plausible explanation was that the phenotype may have been present in our cohort but perhaps not documented on any clinical records or request forms, resulting in these patients being under-identified. It is however also possible that our findings may reflect a broader phenotypic spectrum for 22q11.2 deletion syndrome and 22q11.2 duplication syndrome in African populations, as proposed by Kruzka et al. (2017) and Wonkam et al. (2017), however this it to be explored further in future studies using larger cohorts.

4.4 Assessment of Face2Gene for the Identification of 22q11.2 Deletion/Duplication Syndrome

Face2Gene demonstrated strong performance in identifying 22q11.2 deletion syndrome within the top suggested syndromes among the patients with deletions in the region, including both classic and atypical deletions using both the gestalt only and the combined gestalt and feature score approach. This result is consistent with previous studies (McDonald-McGinn et al., 2023; Roalf et al., 2024), highlighting Face2Gene's potential utility as a NGP tool for diagnosing this condition in African patients. Face2Gene was unable to identify 22q11.2 duplication syndrome in the two patients with 22q11.2 duplications indicating a potential limitation of the application. It is worth acknowledging that Face2Gene has not yet been trained to identify 22q11.2 duplication syndrome and therefore was not expected to correctly classify these two patients, however, 22q11.2 deletion syndrome was also not suggested in the list of potential diagnoses for them. This underscores the phenotypic variability and distinctiveness between deletions and duplications, despite some overlapping features (Bartik et al., 2022). Among factors potentially influencing Face2Gene's performance, genotype (deletion versus duplication) appears to be the primary contributor. Photograph quality was observed to possibly be a factor which influenced Face2Gene's classification as high-quality photographs repeatedly resulted in high gestalt scores for patients with classic deletions. However, this was not always the

case as for some patients' low-quality images did not universally inhibit accurate diagnosis, and conversely, high-quality images did not guarantee correct identification. Additionally, and interestingly the introduction of feature scores decreased the current reliability of Face2Gene, often lowering the ranking of 22q11.2 deletion syndrome or removing it entirely from the recommended syndromes. This reduction in accuracy may be attributed to the variable phenotypic presentations of 22q11.2 deletion syndrome (Gavril et al., 2022), particularly in African patients (Kruszka et al., 2017), where broader phenotypic diversity could affect feature-based classification. These findings suggest that in African populations, using both the gestalt-only approach and the gestalt and feature score approach in tandem may yield better diagnostic results, as each method may complement the other in addressing the phenotypic variability observed in this population. Additionally, although all 30 suggestions provided by Face2Gene were included in this study and Face2Gene was considered consistent if 22q11.2 deletion syndrome appeared anywhere within the list, it must be acknowledged that it is unlikely that the diagnosis would be seriously considered unless it appeared within the top 10 syndromes higher rankings are more suggestive of an accurate match. While most of the rankings in this did appear in the top 10 diagnoses, further validating the potential utility of Face2Gene for 22q11.2 deletion syndrome, future studies may consider using a higher cutoff threshold to simulate a more practical and realistic clinical approach. Despite the promising consistency exhibited by Face2Gene in this study, the small sample size limits the robustness of these findings. Larger studies with more diverse and representative cohorts are necessary to validate the utility of Face2Gene for the diagnosis of CNVs in the 22q11.2 region in African populations. Nevertheless, this study provides encouraging preliminary evidence that Face2Gene may hold promise as a supplemental diagnostic tool in this population.

4.5 Study Limitations

The cross-sectional design inherently restricts the scope of the data to a single point in time, capturing only a snapshot of each patient's clinical and phenotypic profile. Longitudinal studies would be better suited to provide a more comprehensive understanding of the evolving manifestations of 22q11.2 deletion syndrome. Furthermore, the retrospective nature of the analysis introduced the potential for incomplete or inaccurate records, which could influence the robustness of the findings. This further limit the robustness of some of the associations identified in this study as there were varying amounts of data collected depending on the information available in the patient's clinical files or on the patient's referral forms – this means that incidentally some of the genotypes may have had more data available in comparison to others allowing for skewing of the data. This would be particularly impactful when comparing the clinical features of internal patients who by virtue have more clinical information available, to external patients who were limited to the information available on the test request forms.

Ascertainment bias may also have been present in the study, as some patients were referred for testing with a suspected diagnosis of a CNV in the chromosome 22q11.2 region, while others had CNVs in the 22q11.2 region incidentally identified following broader genetic testing for more generalised referral indications. This has the potential to skew some of the findings of the study as some clinical characteristics may be overrepresented due to their historic association with CNVs in the 22q11.2 region, with other less frequently associated clinical characteristics potentially being underrepresented. For internal patients specifically, there may have been bias introduced into the data as they were clinically assessed post-diagnosis via genetic testing, allowing for increased identification of clinical features that are characteristic of CNVs in the 22q11.2 region. A future recommendation is therefore to delineate the number of patients in which a CNV in the chromosome 22q11.2 region was discovered incidentally, and how many patients were referred and assessed with a suspected or confirmed diagnosis of a CNV in the chromosome 22q11.2 region. Additionally, a small subset of 6 patients possessed CNVs without a pathogenic classification but were analysed in this study as if the CNVs were pathogenic. This has the potential to skew the data as the clinical features recorded for these patients may be a consequence of other undiscovered pathogenic variants. Finally, the small cohort size and variable photograph quality used for the Face2Gene assessment, along with the limited number of assessed individuals for particular clinical characteristics during the exploration of genotype-phenotype correlations further constrains the generalisability of these results. Lastly, Larger, more diverse studies are essential to expand these findings and to establish the broader applicability of Face2Gene as a diagnostic tool in varied populations.

5 Conclusion

In conclusion, this study provides valuable insights into the genetic and phenotypic characteristics of CNVs in the chromosome 22q11.2 region in a cohort of predominantly African ancestry, addressing critical gaps in global genomic research. By highlighting the phenotypic variability and referral patterns observed for patients with CNVs in the chromosome 22q11.2 region, it underscores the need for improved access to clinical and laboratory genetic services, and equitable healthcare services across South Africa. The findings affirm the utility of CMA as a first-line diagnostic tool for identifying both classic and atypical CNVs evidenced by the increased detection rates observed with the implementation of newer technologies. Moreover, the findings of this study corroborate the findings of previous studies investigating genotype-phenotype correlations in patients with CNVs in the 22q11.2 region but also suggest some divergence from international findings specifically with respect to skin and palatal abnormalities. Overall, the small samples sizes of individuals assessed for some of the clinical characteristics identified in this cohort limits the generalisability of some of these findings, particularly for the skin abnormalities and palatal defects, however this study provides

preliminary evidence of population specific differences that should be further explored in future studies. The extensive phenotypic variability associated with CNVs in the chromosome 22q11.2 region is well demonstrated in this study, which contributes to the growing body of literature characterising the broad clinical spectrum of 22q11.2 deletion and duplication syndromes. By documenting the range of observed features, the study may aid clinicians in recognising and considering these syndromes in patients who present with atypical or subtle manifestations. The present study also demonstrates the potential of NGP tools, such as Face2Gene, in diagnosing 22q11.2 deletion syndrome, while emphasising the need for further validation in African populations. Limitations such as a small sample size and retrospective design constrain generalisability, however this work lays the foundation for future research on the genomic and clinical diversity of CNVs in the chromosome 22q11.2 region in underrepresented populations. Efforts to expand genomic studies and enhance diagnostic tools tailored to African patients are critical to improving outcomes and reducing diagnostic delays for these complex conditions.

Author contributions

DM, BR and BD designed the study. BR and BD contributed to the acquisition of the data. DM analysed the data. DM drafted the manuscript. DM, BR and BD revised the manuscript for important intellectual content. BR and BD supervised the research. All authors approved the final version of the manuscript.

Acknowledgements

The authors would like to acknowledge the contribution of the NHLS Braamfontein as the source of the data. The authors would also like to acknowledge the contributions of Fahmida Essop who assisted with the procurement of MLPA data for this study, Fiona Baine-Savanhu who assisted with the procurement of CMA data for this study, and Anneline Lochan who assisted with the procurement of patient photographs for this study.

Data availability statement

The data and materials used to support the findings of this study are available from the corresponding author upon request

Conflict of interest

All authors declare no competing interests.

Consent

Departmental consent for the use of patient facial photographs for research purposes was obtained during routine clinic visits.

Ethics approval

Ethics approval was obtained from the University of the Witwatersrand Human Research Ethics Committee (Medical) (WITS-HREC Ref no. R14/49) with ethics clearance number: M240614

6 References

Althammer, A., Prückner, S., Gehring, G. C., Lieftüchter, V., Trentzsch, H., & Hoffmann, F. (2023). Systemic review of age brackets in pediatric emergency medicine literature and the development of a universal age classification for pediatric emergency patients - the Munich Age Classification System (MACS). *BMC Emergency Medicine*, *23*, 77. https://doi.org/10.1186/s12873-023-00851-5

Bartik, L. E., Hughes, S. S., Tracy, M., Feldt, M. M., Zhang, L., Arganbright, J., & Kaye, A. (2022). 22q11.2 duplications: Expanding the clinical presentation. *American Journal of Medical Genetics Part A*, *188*(3), 779–787. https://doi.org/10.1002/ajmg.a.62577

Blagojevic, C., Heung, T., Theriault, M., Tomita-Mitchell, A., Chakraborty, P., Kernohan, K., Bulman, D. E., & Bassett, A. S. (2021). Estimate of the contemporary live-birth prevalence of recurrent 22q11.2 deletions: a cross-sectional analysis from population-based newborn screening. *CMAJ Open*, *9*(3), 802–809. https://doi.org/10.9778/cmajo.20200294

Burnside, R. D. (2015). 22q11.21 Deletion Syndromes: A Review of Proximal, Central, and Distal Deletions and Their Associated Features. *Cytogenetic and Genome Research*, *146*(2), 89–99. https://doi.org/10.1159/000438708

Campbell, I. M., Sheppard, S. E., Crowley, T. B., McGinn, D. E., Bailey, A., McGinn, M. J., Unolt, M., Homans, J. F., Chen, E. Y., Salmons, H. I., Gaynor, J. W., Goldmuntz, E., Jackson, O. A., Katz, L. E., Mascarenhas, M. R., Deeney, V. F. X., Castelein, R. M., Zur, K. B., Elden, L., … McDonald-McGinn, D. M. (2018). What is new with 22q? An update from the 22q and You Center at the Children's Hospital of Philadelphia. *American Journal of Medical Genetics Part A*, *176*(10), 2058–2069. https://doi.org/10.1002/ajmg.a.40637

Cancrini, C., Puliafito, P., Digilio, M. C., Soresina, A., Martino, S., Rondelli, R., Consolini, R., Ruga, E. M., Cardinale, F., Finocchi, A., Romiti, M. L., Martire, B., Bacchetta, R., Albano, V., Carotti, A., Specchia, F., Montin, D., Cirillo, E., Cocchi, G., … Rossi, P. (2014). Clinical Features and Follow-Up in Patients with 22q11.2 Deletion Syndrome. *The Journal of Pediatrics*, *164*(6), 1475-1480. https://doi.org/10.1016/j.jpeds.2014.01.056

Carli, D., Moroni, A., Eleonora, D. G., Zonta, A., Montin, D., Licciardi, F., Aidala, E., Bordese, R., Carlo, P. N., Brusco, A., Giovanni Battista, F., & Mussa, A. (2021). Atypical microdeletion 22q11.2 in a patient with tetralogy of Fallot. *Journal of Genetics*, *100*, 5. https://doi.org/10.1007/s12041-020-01257-z

Carter, M. T., St Pierre, S. A., Zackai, E. H., Emanuel, B. S., & Boycott, K. M. (2009). Phenotypic delineation of Emanuel syndrome (supernumerary derivative 22 syndrome): Clinical features of 63 individuals. *American journal of medical genetics. Part A*, *149* (8), 1712–1721. https://doi.org/10.1002/ajmg.a.32957Catusi, I., Recalcati, M. P., Bestetti, I., Garzo, M., Valtorta, C., Alfonsi, M., Alghisi, A., Cappellani, S., Casalone, R., Caselli, R., Ceccarini, C., Ceglia, C., Ciaschini, A. M., Coviello, D., Crosti, F., D'Aprile, A., Fabretto, A., Genesio, R., Giagnacovo, M., … Giardino, D. (2020). Testing single/combined clinical categories on 5110 Italian patients with developmental phenotypes to improve array-based detection rate. *Molecular Genetics & Genomic Medicine*, *8*, e1056. https://doi.org/10.1002/mgg3.1056

Cheung, E. N. M., George, S. R., Costain, G. A., Andrade, D. M., Chow, E. W. C., Silversides, C. K., & Bassett, A. S. (2014). Prevalence of hypocalcaemia and its associated features in 22q11·2 deletion syndrome. *Clinical Endocrinology*, *81*(2), 190–196. https://doi.org/10.1111/cen.12466

Cirillo, A., Lioncino, M., Maratea, A., Passariello, A., Fusco, A., Fratta, F., Monda, E., Caiazza, M., Signore, G., Esposito, A., Baban, A., Versacci, P., Putotto, C., Marino, B., Pignata, C., Cirillo, E., Giardino, G., Sarubbi, B., Limongelli, G., & Russo, M. G. (2022). Clinical Manifestations of 22q11.2 Deletion Syndrome. *Heart Failure Clinics*, *18*(1), 155–164. https://doi.org/10.1016/j.hfc.2021.07.009

Clements, C. C., Wenger, T. L., Zoltowski, A. R., Bertollo, J. R., Miller, J. S., de Marchena, A. B., Mitteer, L. M., Carey, J. C., Yerys, B. E., Zackai, E. H., Emanuel, B. S., McDonald-McGinn, D. M., & Schultz, R. T. (2017). Critical region within 22q11.2 linked to higher rate of autism spectrum disorder. *Molecular Autism*, *8*, 58. https://doi.org/10.1186/s13229-017-0171-7

De Decker, R., Bruwer, Z., Hendricks, L., Schoeman, M., Schutte, G., & Lawrenson, J. (2016). predicted v. real prevalence of the 22q11.2 deletion syndrome in children with congenital heart disease presenting to Red Cross War Memorial Children's Hospital, South Africa: A prospective study. *South African Medical Journal*, *106*, S82-S86. https://doi.org/10.7196/SAMJ.2016.v106i6.11003

Firth, H. V., Richards, S. M., Bevan, A. P., Clayton, S., Corpas, M., Rajan, D., Vooren, S. Van, Moreau, Y., Pettett, R. M., & Carter, N. P. (2009). DECIPHER: Database of

Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *The American Journal of Human Genetics*, *84*(4), 524–533. https://doi.org/10.1016/j.ajhg.2009.03.010

Gavril, E.-C., Popescu, R., Nucă, I., Ciobanu, C.-G., Butnariu, L. I., Rusu, C., & Pânzaru, M.-C. (2022). Different Types of Deletions Created by Low-Copy Repeats Sequences Location in 22q11.2 Deletion Syndrome: Genotype–Phenotype Correlation. *Genes*, *13*(11), 2083. https://doi.org/10.3390/genes13112083

Gurovich, Y., Hanani, Y., Bar, O., Nadav, G., Fleischer, N., Gelbman, D., Basel-Salmon, L., Krawitz, P. M., Kamphausen, S. B., Zenker, M., Bird, L. M., & Gripp, K. W. (2019). Identifying facial phenotypes of genetic disorders using deep learning. *Nature Medicine*, *25*(1), 60–64. https://doi.org/10.1038/s41591-018-0279-0

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L., Delacqua, G., Delacqua, F., Kirby, J., & Duda, S. N. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*, *95*, 103208. https://doi.org/10.1016/j.jbi.2019.103208

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, *42*(2), 377–381. https://doi.org/10.1016/j.jbi.2008.08.010

Hennocq, Q., Willems, M., Amiel, J., Arpin, S., Attie-Bitach, T., Bongibault, T., Bouygues, T., Cormier-Daire, V., Corre, P., Dieterich, K., Douillet, M., Feydy, J., Galliani, E., Giuliano, F., Lyonnet, S., Picard, A., Porntaveetus, T., Rio, M., Rouxel, F., … Garcelon, N. (2024). Next generation phenotyping for diagnosis and phenotype–genotype correlations in Kabuki syndrome. *Scientific Reports*, *14*(1), 2330. https://doi.org/10.1038/s41598-024-52691-3

Homans, J. F., Tromp, I. N., Colo, D., Schlösser, T. P. C., Kruyt, M. C., Deeney, V. F. X., Crowley, T. B., McDonald-McGinn, D. M., & Castelein, R. M. (2018). Orthopaedic manifestations within the 22q11.2 Deletion syndrome: A systematic review. *American Journal of Medical Genetics Part A*, *176*(10), 2104–2120. https://doi.org/10.1002/ajmg.a.38545

Jackson, O., Crowley, T. B., Sharkus, R., Smith, R., Jeong, S., Solot, C., & McDonald-Mcginn, D. (2019). Palatal evaluation and treatment in 22q11.2 deletion syndrome. *American Journal of Medical Genetics Part A*, *179*(7), 1184–1195. https://doi.org/10.1002/ajmg.a.61152

Jalali, G. R., Vorstman, J. A., Errami, A., Vijzelaar, R., Biegel, J., Shaikh, T., & Emanuel, B. S. (2008). Detailed analysis of 22q11.2 with a high density MLPA probe set. *Human mutation*, *29*(3), 433–440. https://doi.org/10.1002/humu.20640

Kruszka, P., Addissie, Y. A., McGinn, D. E., Porras, A. R., Biggs, E., Share, M., Crowley, T. B., Chung, B. H. Y., Mok, G. T. K., Mak, C. C. Y., Muthukumarasamy, P., Thong, M., Sirisena, N. D., Dissanayake, V. H. W., Paththinige, C. S., Prabodha, L. B. L., Mishra, R., Shotelersuk, V., Ekure, E. N., … Muenke, M. (2017). 22q11.2 deletion syndrome in diverse populations. *American Journal of Medical Genetics Part A*, *173*(4), 879–888. https://doi.org/10.1002/ajmg.a.38199

Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. *Journal of the American Academy of Child & Adolescent Psychiatry*, *56*(6), 466–474. https://doi.org/10.1016/j.jaac.2017.03.013

McDonald-McGinn, D., Crowley, T., McGinn, D., Gaiser, K., Giunta, V., Lairson, L., Zackai, E., Tran, O., Share, M., Valverde, K., Pastor, S., & Emanuel, B. (2023). P248: Is 22q11.2 deletion syndrome truly less common in Black patients? *Genetics in Medicine Open*, *1*(1), 100276. https://doi.org/10.1016/j.gimo.2023.100276

McDonald-McGinn, D. M., Sullivan, K. E., Marino, B., Philip, N., Swillen, A., Vorstman, J. A. S., Zackai, E. H., Emanuel, B. S., Vermeesch, J. R., Morrow, B. E., Scambler, P. J., & Bassett, A. S. (2015). 22q11.2 deletion syndrome. *Nature Reviews Disease Primers*, *1*(1), 15071. https://doi.org/10.1038/nrdp.2015.71Motahari, Z., Moody, S. A., Maynard, T. M., & LaMantia, A.-S. (2019). In the line-up: Deleted genes associated with DiGeorge/22q11.2 deletion syndrome: Are they all suspects? *Journal of Neurodevelopmental Disorders, 11*(1), 7. https://doi.org/10.1186/s11689-019-9267-z

Napolitano, A., Schiavi, S., La Rosa, P., Rossi-Espagnet, M. C., Petrillo, S., Bottino, F., Tagliente, E., Longo, D., Lupi, E., Casula, L., Valeri, G., Piemonte, F., Trezza, V., & Vicari, S. (2022). Sex Differences in Autism Spectrum Disorder: Diagnostic, Neurobiological, and Behavioral Features. *Frontiers in Psychiatry*, *13,* 889636. https://doi.org/10.3389/fpsyt.2022.889636

Olsen, L., Sparsø, T., Weinsheimer, S. M., Dos Santos, M. B. Q., Mazin, W., Rosengren, A., Sanchez, X. C., Hoeffding, L. K., Schmock, H., Baekvad-Hansen, M., Bybjerg-Grauholm, J., Daly, M. J., Neale, B. M., Pedersen, M. G., Agerbo, E., Mors, O., Børglum, A., Nordentoft, M., Hougaard, D. M., … Werge, T. (2018). Prevalence of rearrangements in the 22q11.2 region and population-based risk of neuropsychiatric and developmental

disorders in a Danish population: a case-cohort study. *The Lancet Psychiatry*, *5*(7), 573–580. https://doi.org/10.1016/S2215-0366(18)30168-8

Portnoï, M.-F. (2009). Microduplication 22q11.2: A new chromosomal syndrome. *European Journal of Medical Genetics*, *52*(2–3), 88–93. https://doi.org/10.1016/j.ejmg.2009.02.008

Ramsay, M. (2022). African genomic data sharing and the struggle for equitable benefit. *Patterns*, *3*(1), 100412. https://doi.org/10.1016/j.patter.2021.100412

Roalf, D. R., McDonald-McGinn, D. M., Jee, J., Krall, M., Crowley, T. B., Moberg, P. J., Kohler, C., Calkins, M. E., Crow, A. J. D., Fleischer, N., Gallagher, R. S., Gonzenbach, V., Clark, K., Gur, R. C., McClellan, E., McGinn, D. E., Mordy, A., Ruparel, K., Turetsky, B. I., … Gur, R. E. (2024). Computer-vision analysis of craniofacial dysmorphology in 22q11.2 deletion syndrome and psychosis spectrum disorders. *Journal of Neurodevelopmental Disorders*, *16*(1), 35. https://doi.org/10.1186/s11689-024-09547-8

Riggs, E. R., Andersen, E. F., Cherry, A. M., Kantarci, S., Kearney, H., Patel, A., Raca, G., Ritter, D. I., South, S. T., Thorland, E. C., Pineda-Alvarez, D., Aradhya, S., & Martin, C. L. (2020). Technical standards for the interpretation and reporting of constitutional copy-number variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics (ACMG) and the Clinical Genome Resource (ClinGen). *Genetics in medicine : official journal of the American College of Medical Genetics*, *22*(2), 245–257. https://doi.org/10.1038/s41436-019-0686-8

Sgardioli, I. C., Paoli Monteiro, F., Fanti, P., Paiva Vieira, T., & Gil-da-Silva-Lopes, V. L. (2019). Testing criteria for 22q11.2 deletion syndrome: preliminary results of a low cost strategy for public health. *Orphanet Journal of Rare Diseases*, *14*(1), 123. https://doi.org/10.1186/s13023-019-1098-1

Shaikh, T. H. (2000). Chromosome 22-specific low copy repeats and the 22q11.2 deletion syndrome: genomic organization and deletion endpoint analysis. *Human Molecular Genetics*, *9*(4), 489–501. https://doi.org/10.1093/hmg/9.4.489

Shaikh, T. H., Kurahashi, H., & Emanuel, B. S. (2001). Evolutionarily conserved low copy repeats (LCRs) in 22q11 mediate deletions, duplications, translocations, and genomic instability: An update and literature review. *Genetics in Medicine*, *3*(1), 6–13. https://doi.org/10.1097/00125817-200101000-00003

Van Campenhout, S., Devriendt, K., Breckpot, J., Frijns, J.-P., Peeters, H., Van Buggenhout, G., Van Esch, H., Maes, B., & Swillen, A. (2012). Microduplication 22q11.2: a description of the clinical, developmental and behavioral characteristics during childhood. *Genetic*

Counseling (Geneva, Switzerland), *23*(2), 135–148.

Watson, C. T., Marques-Bonet, T., Sharp, A. J., & Mefford, H. C. (2014). The genetics of microdeletion and microduplication syndromes: an update. *Annual review of genomics and human genetics*, *15*, 215–244. https://doi.org/10.1146/annurev-genom-091212-153408

Wonkam, A., Toko, R., Chelo, D., Tekendo-Ngongang, C., Kingue, S., & Dahoun, S. (2017). The 22q11.2 Deletion Syndrome in Congenital Heart Defects: Prevalence of Microdeletion Syndrome in Cameroon. *Global Heart*, *12*(2), 115. https://doi.org/10.1016/j.gheart.2017.01.003

Supplementary Material

Supplementary table 1 | Significant results following Shapiro-Wilk test for normality (W statistic) and the Wilcoxon Rank Sum test (U statistic) for the comparison of the mean age of referral between groups

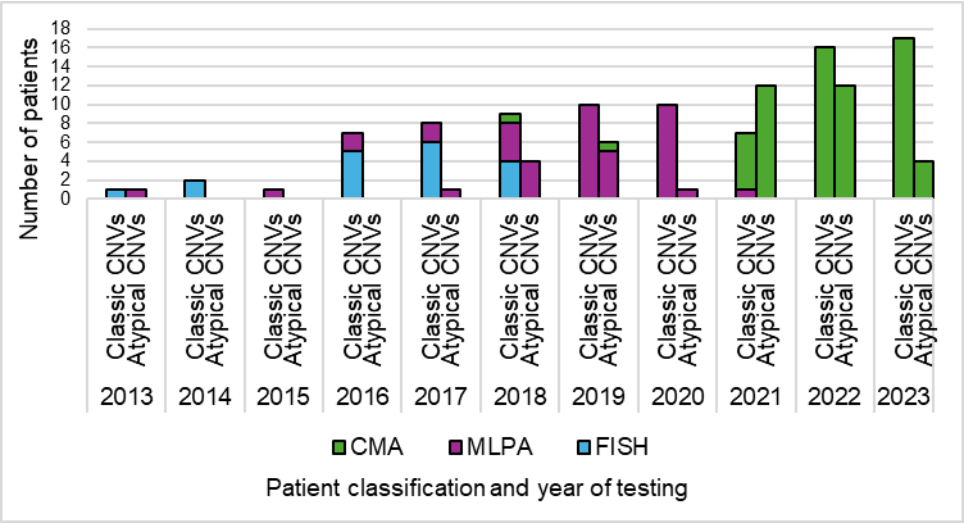
Test group	W statistic	P-value	Outcome
Atypical CNVs	0,51938355	2,112689E-10***	Non-Normal
Classic CNVs	0,511050820	1,264829E-15***	Non-Normal
External Patients	0,463337004	4,541732E-16***	Non-Normal
Internal Patients	0.562919974	3.070213E-10***	Non-Normal
Comparison of median age of referral	**U statistic**	**P-value**	**Outcome**
Atypical CNVs vs Classic CNVs	1241.5	0.004264*	Significant
External Patients vs Internal Patients	2572.5	0.00045***	Significant

CNVs= Copy Number Variants; *= p-value<0,05; ***=p-value<0,0005

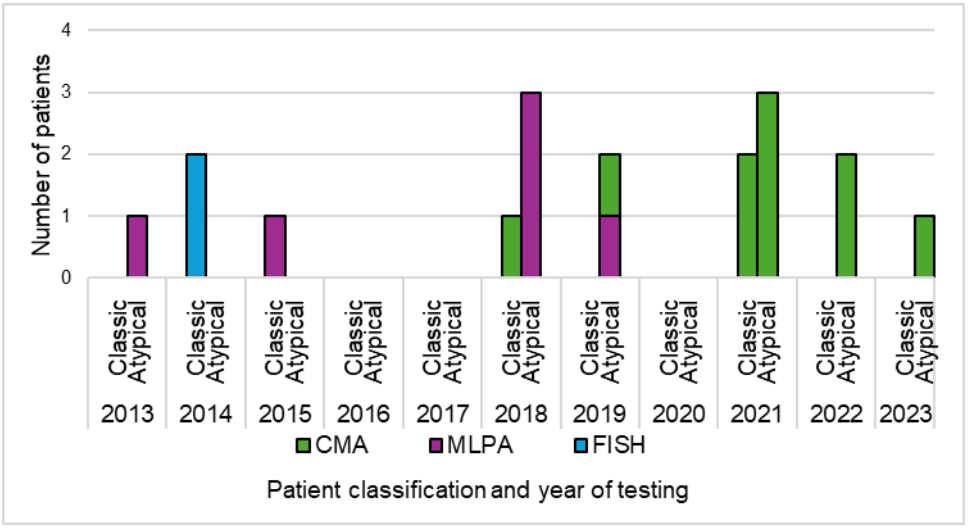
Supplementary table 2 | Significant results of Fisher's exact test assessing associations between various groups of patients with CNVs in the 22q11.2 region and Congenital heart defects (CHDs)/respiratory abnormalities

Total deletions vs total duplications	
Phenotype	CHDs/pulmonary/chest abnormalities
P-value	0,000168***
Adjusted P-value	0,001518**
Odds Ratio	17,411764
Interpretation	Significantly associated with deletions
Total classic CNVs vs total atypical CNVs	
Phenotype	CHDs/pulmonary/chest abnormalities
P-value	0,000168***
Adjusted P-value	0,001520**
Odds Ratio	6,78787
Interpretation	Significantly associated with classic CNVs
Classic deletions vs atypical deletions	
Phenotype	CHDs/pulmonary/chest abnormalities
P-value	0,006689*
Adjusted P-value	0,060208
Odds Ratio	5,090909
Interpretation	Evidence for potential association with classic deletions

CNVs= Copy Number Variants; *= p-value<0,05;**=p-value<0,005 ***=p-value<0,0005



Supplementary Figure 1 | The annual number of classic LCR22A-D deletions and atypical deletions identified in the 22q11.2 region using various testing methodologies over 10 years in a sub-Saharan African cohort. The total number of CNVs identified annually includes both deletions and duplications.



Supplementary Figure 2 | The annual number of classic LCR22A-D duplications and atypical duplications identified in the 22q11.2 region using various testing methodologies over 10 years in a South African cohort. The total number of CNVs identified annually includes both deletions and duplications.

Supplementary table 3 | The clinical features observed in sub-Suharan African patients with classic LCR22A-D CNVs in the chromosome 22q11.2 region (n=88). Frequencies of deletions and duplications are represented as the number of patients observed with the clinical feature out of the number of patients assessed for the phenotype (n/N (%)). The clinical features were reported using HPO terms.

Growth abnormalities n=22†/88‡ (25%)	Deletions n= 15/20† (75)	Duplications n=1/2† (50)
Microcephaly	9/20 (45)	0/2 (0)
Macrocephaly	2/20 (10)	0/2 (0)
Short stature	8/20 (40)	1/2 (50)
Underweight	7/20 (35)	0/2 (0)
Dysmorphic features n=65†/88‡ (74%)	**Deletions n= 58/61† (95)**	**Duplications n=4/4† (100)**
Abnormalities of the head/face	**9 (15)**	**1 (25)**
Skull asymmetry	3 (5)	1 (25)
Asymmetric crying facies	3 (5)	0 (0)
Malar hypoplasia	1 (2)	0 (0)
Low anterior hairline	1 (2)	0 (0)
Curly eyelashes	1 (2)	0 (0)
Synophrys	1 (2)	0 (0)
Sparse eyebrows	3 (5)	0 (0)
Broad eyebrow	1 (2)	0 (0)
Abnormalities of the eye	**28 (46)**	**3 (75)**
Hypertelorism	11 (18)	0 (0)
Upslanting palpebral fissures	7 (11)	0 (0)
Epicanthic folds	6 (10)	1
Hooded eyelids	6 (10)	0 (0)
Downslanting palpebral fissures	5 (8)	0 (0)
Abnormalities of the ear	**35 (57)**	**2 (50)**
Low-set ears	22 (36)	1 (25)
Cupped/cup-shaped	6 (10)	0 (0)
Small ears	4 (7)	0 (0)
Pre-auricular pit/s	1 (2)	1 (25)
Pre-auricular tag/s	1 (2)	0 (0)
Posteriorly rotated ears	1 (2)	1 (25)
Upturned earlobes	0 (0)	2 (50)
Prominant ears	1 (2)	0 (0)
Recurrent ear infections	1 (2)	0 (0)
Abnormalities of the nose	**38 (62)**	**3 (75)**
Wide nasal bridge	8 (13)	0 (0)
Long nose	7 (11)	0 (0)
Broad nose	7 (11)	0 (0)
Depressed nasal tip	6 (10)	0 (0)
Depressed nasal bridge	5 (8)	1 (25)
Hypoplastic alae nasi	4 (7)	0 (0)
Short nose	4 (7)	0 (0)
Short columella	2 (3)	1 (25)
Prominent nasal bridge	1 (2)	0 (0)
Nasal dimple/bifid nasal tip	1 (2)	0 (0)

Prominent nasal tip	1 (2)	0 (0)
Small nose	1 (2)	0 (0)
Beaked nose	1 (2)	0 (0)
Abnormalities of the mouth and pharynx	**35 (57)**	**2 (50)**
Feeding difficulties in infancy	3 (5)	0 (0)
Downturned mouth	3 (5)	0 (0)
Tented upper lip	3 (5)	0 (0)
Thick upper lip vermillion	1 (2)	0 (0)
Flat philtrum	3 (5)	0 (0)
Dental caries	7 (11)	0 (0)
Wide mouth	1 (2)	0 (0)
Ankyloglossia	1 (2)	0 (0)
Hypernasal speech	3 (5)	0 (0)
High arched palate	4 (7)	0 (0)
Cleft palate	3 (5)	0 (0)
Velopharyngeal incompetence	2 (3)	0 (0)
Short palate	1 (2)	0 (0)
Submucous cleft palate	1 (2)	0 (0)
Micrognathia	19 (31)	0 (0)
Retrognathia	3 (5)	0 (0)
Prognathism	0 (0)	1 (25)
Abnormalities of the larynx and trachea	**2 (3)**	**0 (0)**
Laryngotracheomalacia	1 (2)	0 (0)
Stridor	1 (2)	0 (0)
CNS/developmental abnormalities n=40†/88‡ (45%)	Deletions n= 30/38† (79)	Duplications n=2/2† (100)
Developmental delay	21 (55)	2 (100)
Learning difficulties	10 (26)	0 (0)
Intellectual disability	6 (16)	0 (0)
Hypotonia	4 (11)	1 (50)
Hypertonia	5 (13)	0 (0)
Hyperreflexia	1 (3)	0 (0)
Seizures	4 (11)	0 (0)
Neuropsychiatric differences	**5 (13)**	**1 (50)**
Temper/aggression	2 (5)	1 (50)
Abnormal brain imaging¶	2 (5)	0 (0)
Attention deficit hyperactivity disorder	2 (5)	0 (0)
Disinhibition/impulsiveness	1 (3)	1 (50)
Autism spectrum disorder	1 (3)	0 (0)
Schizophrenia	1 (3)	0 (0)
Anxiety	1 (3)	0 (0)
Social difficulties	1 (3)	0 (0)
Depression	1 (3)	0 (0)
CHDs/pulmonary/chest abnormalities n=75†/88‡ (85%)	Deletions n= 63/72† (88)	Duplications n=1/3† (33)
Ventricular septal defect	35 (48)	1 (33)

Overriding aorta	16 (22)	0 (0)
Pulmonary stenosis	16 (22)	0 (0)
Tetralogy of Fallot	15 (21)	0 (0)
Right ventricular hypertrophy	10 (14)	0 (0)
Atrial septal defect	8 (11)	0 (0)
Patent ductus arteriosus	6 (8)	1 (33)
Aortic arch anomaly	7 (10)	0 (0)
Interrupted aortic arch	7 (10)	0 (0)
Truncus arteriosus	6 (8)	0 (0)
Pulmonary atresia	6 (8)	0 (0)
Unspecified cardiac abnormality	5 (7)	1 (33)
Patent foramen ovale	1 (1)	1 (33)
Major aortopulmonary collateral arteries	2 (3)	0 (0)
Double outlet right ventricle	2 (3)	0 (0)
Bicuspid aortic valve	1 (1)	0 (0)
Mitral valve prolapse	1 (1)	0 (0)
Aortic regurgitation	1 (1)	0 (0)
Right aortic arch	1 (1)	0 (0)
Dyspnoea	1 (1)	0 (0)
Pulmonary hypoplasia	3 (4)	0 (0)
Widely spaced nipples	2 (3)	0 (0)
Gastrointestinal abnormalities n=46†/88‡ (52%)	Deletions n= 10/43† (23)	Duplications n=1/3† (33)
Oesophageal atresia	1 (2)	1 (33)
Umbilical hernia	1 (2)	0 (0)
Imperforate anus	2 (5)	0 (0)
Anteriorly placed anus	6 (14)	0 (0)
Genitourinary abnormalities n=31†/88‡ (35%)	Deletions n= 5/31† (16)	Duplications n=2/3† (66)
Renal agenesis (unilateral)	3 (10)	2 (66)
Cryptorchidism (bilateral/unilateral)	4 (13)	2 (66)
Cliteromegaly	2 (6)	0 (0)
Hypoplastic external genitalia§	1 (3)	0 (0)
Prominent midline raphe	0 (0)	1 (33)
Musculoskeletal abnormalities n=38†/88‡ (43%)	Deletions n= 20/36† (56)	Duplications n=1/2† (50)
Abnormalities of the spine	**2 (6)**	**0 (0)**
Rib anomaly	1 (3)	0 (0)
Scoliosis	1 (3)	0 (0)
Lumbar hyperlordosis	1 (3)	0 (0)
Abnormalities of the limb	**18 (50)**	**1 (50)**
Joint laxity	1 (3)	1 (50)
Syndactyly	1 (3)	0 (0)
Clinodactyly	1 (3)	0 (0)
Brachydactyly	3 (8)	0 (0)
Postaxial polydactyly	3 (8)	0 (0)
Tapered fingers	2 (6)	0 (0)

Long fingers	2 (6)	0 (0)
Genu valgum	1 (3)	0 (0)
Club foot (unilateral)	4 (11)	0 (0)
Rocker bottom feet	1 (3)	0 (0)
Overfolded toes	3 (8)	0 (0)
Sandal gap	1 (3)	0 (0)
Skin abnormalities n=28†/88‡ (36%)	Deletions n= 10/26† (38)	Duplications n=1/2† (50)
Cafe-au-lait macule/s	3 (12)	0 (0)
Hyperpigmentation of the skin	3 (12)	1 (50)
Hypopigmentation of the skin	3 (12)	0 (0)
Melanocytic macule/s	1 (4)	0 (0)
Thickened finger creases	1 (4)	0 (0)
Thickened nuchal skin fold	1 (4)	0 (0)
Thickened palmar creases	1 (4)	0 (0)
Decreased palmar creases	1 (4)	0 (0)
Hypoplastic fingernails	1 (4)	0 (0)
Hypoplastic toenails	1 (4)	0 (0)
Endocrine abnormalities n=44†/88‡ (50%)	Deletions n= 20/41† (49)	Duplications n=1/3† (33)
Hypocalcaemia	14 (34)	1 (33)
Hypoparathyroidism	3 (7)	0 (0)
Hypothyroidism	1 (2)	0 (0)
Hyperthyroidism	1 (2)	0 (0)
Immunodeficiency n=45†/88‡ (51%)	Deletions n= 25/42† (60)	Duplications n=1/3† (33)
Recurrent infections	19 (45)	1 (33)
Thymic hypoplasia	7 (17)	0 (0)

†- The number of patients for which the clinical feature was assessed;‡- The total number of patients with classic LCR22A-D CNVs in the chromosome 22q11.2 region; ¶- Patient A: Dilated ventricles, corpus callosum visible in sagittal view but not coronal view, Patient B: Hyperintensities of the centrum semi ovale bilaterally and bi-occipital deep white matter on T2 weighted image and fluid attenuated inversion recovery, sparing of the Rolandic cortices and basal ganglia, indicating a peripheral pattern of hypoxic ischaemic injury (chronic).§- Male

Supplementary table 4 | The clinical features observed in sub-Saharan African patients with atypical CNVs in the chromosome 22q11.2 region (n=41). The clinical features were reported using HPO terms.

Growth abnormalities n=10†/41‡ (24%)	LCR span involved in deletions n=7/9† (70%)										LCR span involved in duplications n=1/1† (100%)					
	A-B	A-C	B-C	B-D	C-D	C-E	C-F	C-H	D-F	D-H	A-B	A-C	B-C	B-D	C-D	F-H
Short stature	2	1	-	-	-	-	-	-	1	-	-	-	-	-	-	-
Tall stature	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-
Microcephaly	2	-	-	-	-	1	-	-	1	-	-	-	-	-	-	-
Underweight	1	-	-	1	-	-	-	-	1	-	-	-	-	-	-	-
Dysmorphic features n=24†/41‡ (56%)	**LCR span involved in deletions n=18/19† (95%)**										**LCR span involved in duplications n=4/5† (80%)**					
	A-B	A-C	B-C	B-D	C-D	C-E	C-F	C-H	D-F	D-H	A-B	A-C	B-C	B-D	C-D	F-H
Abnormalities of the head/face	**3**	**-**	**-**	**1**	**-**	**1**	**-**	**1**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**1**
Skull asymmetry	1	-	-	-	-	-	-	1	-	-	-	-	-	-	-	1
Malar hypoplasia	2	-	-	1	-	1	-	-	-	-	-	-	-	-	-	-
Sparse eyebrow	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Synophrys	-	-	-	-	-	1	-	-	-	-	-	-	-	1	-	-
Broad eyebrow	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-
Bushy eyebrows	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-
Curly eyelashes	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-
Long eyelashes	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-
Abnormalities of the eye	**2**	**1**	**-**	**-**	**1**	**1**	**-**	**1**	**1**	**1**	**-**	**1**	**-**	**2**	**-**	**-**
Epicanthic folds	2	-	-	-	-	1	-	1	-	-	-	1	-	-	-	-
Upslanting palpebral fissures	1	-	-	-	-	-	-	-	-	1	-	1	-	-	-	-
Hypertelorism	-	1	-	-	-	-	-	-	-	-	-	-	-	1	-	-
Sclerocornea	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-
Reduced visual acuity	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-
Abnormalities of the ear	**5**	**-**	**-**	**3**	**-**	**1**	**1**	**1**	**2**	**1**	**-**	**1**	**-**	**2**	**-**	**1**
Pre-auricular tag/s	-	-	-	-	-	1	2	-	-	-	-	1	-	-	-	-
Pre-auricular pit/s	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-
Cupped/cup-shaped	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-
Protuberant ears	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Low set ears	3	-	-	2	-	-	-	1	2	1	-	-	-	1	-	1

Posteriorly rotated	1	-	-	-	-	-	-	-	-	1	-	1	-	-	-	1
Abnormalities of the nose	**4**	**1**	**-**	**2**	**-**	**1**	**-**	**-**	**-**	**1**	**-**	**1**	**-**	**2**	**-**	**-**
Depressed nasal tip	1	-	-	1	-	-	-	-	-	-	-	-	-	2	-	-
Depressed nasal bridge	-	1	-	1	-	-	-	-	-	1	-	-	-	-	-	-
Anteverted nares	1	-	-	1	-	-	-	-	-	-	-	-	-	1	-	-
Short columella	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-
Abnormalities of the mouth and pharynx	**3**	**1**	**-**	**2**	**-**	**1**	**-**	**-**	**1**	**-**	**-**	**-**	**-**	**-**	**-**	**1**
Macroglossia	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Cleft lip and palate	-	-	-	1	-	-	-	-	1	-	-	-	-	-	-	-
High arched palate	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	1
Micrognathia	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
Retrognathia	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Abnormalities of the larynx and trachea	**-**	**-**	**-**	**-**	**1**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**
High pitched cry	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-
CNS/Developmental Abnormalities n=23†/41‡ (56%)	LCR span involved in deletions n=12/15† (80%)										LCR span involved in duplications n=6/8† (75%)					
	A-B	A-C	B-C	B-D	C-D	C-E	C-F	C-H	D-F	D-H	A-B	A-C	B-C	B-D	C-D	F-H
Developmental delay	3	1	-	-	1	1	1	1	1	4	-	1	-	-	-	-
Learning difficulties	1	-	-	-	-	-	-	-	-	-	1	1	1	1	1	-
Intellectual disability	1	-	-	-	-	-	-	-	-	1	-	-	-	1	-	-
Hypotonia	2	-	-	-	-	-	-	-	1	2	-	-	-	-	-	1
Hypertonia	-	-	-	-	-	-	-	-	-	1	-	-	-	-	-	1
Hyperreflexia	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Seizures	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Neuropsychiatric differences	**1**	**-**	**-**	**1**	**-**	**1**	**-**	**-**	**-**	**-**	**-**	**1**	**1**	**-**	**-**	**-**
Attention deficit hyperactivity disorder	1	-	-	-	-	1	-	-	-	-	-	-	-	1	-	-
Autism spectrum disorder	1	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-
Schizophrenia	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Temper/aggression	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-
Hyperactivity	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-
CHDs/pulmonary/chest abnormalities	LCR span involved in deletions										LCR span involved in duplications					

n=27†/41‡ (63%)	n=12/20† (60%)										n=1/7† (14%)					
	A-B	A-C	B-C	B-D	C-D	C-E	C-F	C-H	D-F	D-H	A-B	A-C	B-C	B-D	C-D	F-H
Ventricular septal defect	2	-	1	-	1	-	-	-	-	-	-	-	-	-	-	-
Tetralogy of Fallot	1	-	1	-	1	-	-	-	-	-	-	-	-	-	-	-
Overriding aorta	1	-	1	-	1	-	-	-	-	-	-	-	-	-	-	-
Pulmonary stenosis	1	-	1	-	1	-	-	-	-	-	-	-	-	-	-	-
Right ventricular hypertrophy	-	-	1	-	1	-	-	-	-	-	-	-	-	-	-	-
Atrioventricular septal defect	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-
Truncus arteriosus	-	-	-	-	-	-	-	-	1	-	-	-	-	-	-	-
Widely spaced nipples	3	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Unspecified cardiac abnormality	1	1	-	-	1	-	-	-	1	-	-	-	-	-	-	1
Gastrointestinal abnormalities n=10†/41‡ (24%)	LCR span involved in deletions n=0/5† (0%)										LCR span involved in duplications n=0/5† (0%)					
	A-B	A-C	B-C	B-D	C-D	C-E	C-F	C-H	D-F	D-H	A-B	A-C	B-C	B-D	C-D	F-H
Genitourinary abnormalities n=10†/41‡ (24%)	LCR span involved in deletions n=0/5† (0%)										LCR span involved in duplications n=1/5† (20%)					
	A-B	A-C	B-C	B-D	C-D	C-E	C-F	C-H	D-F	D-H	A-B	A-C	B-C	B-D	C-D	F-H
Cryptorchidism (bilateral)	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
Musculoskeletal abnormalities n=17†/41‡ (41%)	LCR span involved in deletions n=6/10† (60%)										LCR span involved in duplications n=4/7† (57%)					
	A-B	A-C	B-C	B-D	C-D	C-E	C-F	C-H	D-F	D-H	A-B	A-C	B-C	B-D	C-D	F-H
Abnormalities of the spine	**-**	**1**	**-**	**-**	**1**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**	**-**
Short neck	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Short trunk	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-
Abnormalities of the limb	**1**	**-**	**-**	**1**	**1**	**1**	**-**	**-**	**-**	**-**	**-**	**1**	**-**	**2**	**-**	**1**
Joint laxity	-	-	-	-	-	1	-	-	-	-	-	-	-	2	-	-
Tapered fingers	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Brachydactyly	1	-	-	-	1	1	-	-	-	-	-	1	-	-	-	-
Clinodactyly	-	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-
Syndactyly	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-	-
Clubfoot	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1
Skin abnormalities	LCR span involved in deletions										LCR span involved in duplications					

n=12†/41‡ (34%)	n=5/8† (63%)										n=3/4† (75%)					
	A-B	A-C	B-C	B-D	C-D	C-E	C-F	C-H	D-F	D-H	A-B	A-C	B-C	B-D	C-D	F-H
Cafe-au-lait macule/s	-	-	-	-	-	1	-	-	-	-	-	-	-	2	-	-
Hyperpigmentation of the skin	1	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-
Hypopigmentation of the skin	-	-	-	-	1	-	-	-	-	-	-	-	-	-	-	-
Hypopigmented macules	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-
Melanocytic macule/s	-	-	-	-	-	-	-	-	-	-	-	-	-	2	-	-
Follicular lesions	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-
Hypertrichosis	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Deep-set toe nails	-	-	-	-	-	-	-	-	-	-	-	-	-	1	-	-
Hypoplastic fingernails	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-
Hypoplastic toenails	-	-	-	-	-	-	-	-	-	-	-	1	-	-	-	-
Severe infantile eczema	-	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Endocrine abnormalities n=17†/41‡ (41%)	LCR span involved in deletions n=8/13† (62%)										LCR span involved in duplications n=0/4† (0%)					
	A-B	A-C	B-C	B-D	C-D	C-E	C-F	C-H	D-F	D-H	A-B	A-C	B-C	B-D	C-D	F-H
Hypocalcaemia	3	3	-	1	-	-	-	-	-	-	-	-	-	-	-	-
Hypoparathyroidism	1	1	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Immune related n=16†/41‡ (39%)	LCR span involved in deletions n=7/12† (58%)										LCR span involved in duplications n=1/4† (25%)					
	A-B	A-C	B-C	B-D	C-D	C-E	C-F	C-H	D-F	D-H	A-B	A-C	B-C	B-D	C-D	F-H
Thymic hypoplasia	2	2	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Recurrent infections	4	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Allergies	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

†- The number of patients for which the clinical feature was assessed; ‡- The total number of patients with atypical CNVs in the 22q11.2 region; '-' = there were 0 patients identified with this clinical feature. Shading in the LCR regions separates proximal, central and distal 22q11.2 regions.

Supplementary table 5 | Data generated for each patient after Face2Gene analysis and HPO terms uploaded for each patient (n=11).

Patient code	Facial D-score	Gestalt score	Feature score	HPO terms inputted
P2	High	High	Low	Conductive hearing impairment [HP:0008513]; overfolded helix [HP:0000396]; seizure [HP:0001250]; tricuspid regurgitation [HP:0001658]; macrocephaly [HP:0000256]; neurodevelopmental delay [HP:0012758]; hooded eyelid [HP:0000492]
P6	High	High	Medium	Hypoparathyroidism [HP:0000829]; short stature [HP:0004322]; ADHD [HP:0007018]; atrial septal defect [HP:0001631]; ID (moderate) [HP:0010864]; sparse lateral
P87	High	High	Medium	Eyebrow [HP:0005338]; vocal cord paralysis [HP:0001605] Smooth philtrum [HP:0000319]; epicanthus [HP:0000286]; atrial septal defect [HP:0001631]; rectovaginal fistula [HP:0000059]; ventricular septal defect [HP:0001629];
P91	High	High	Medium	unilateral facial palsy [HP:0012799]; hooded eyelid [HP:0000492] Hypertelorism [HP:0000316]; short stature [HP:0004322]; epicanthus [HP:0000286]; cupped ear [HP:0000378]; pointed chin [HP:0000307]; seizure [HP:0001250]; recurrent infections in infancy and early childhood [HP:0005423]
P88	High	Medium	Low	Hypertelorism [HP:0000316]; hyperactivity [HP:0000752]; high palate [HP:0000218]; depressed nasal tip [HP:0000420]; autistic behaviour [HP:0000729];
P85	High	Low	N/A	Tetralogy of Fallot [HP:0001636]; moderate global DD [HP:0011343] bulbous nose [HP:0000414]; ankyloglossia [HP:0010296]; preauricular pit [HP:0004467]; preauricular skin tag [HP:0100879]; short femur [HP:0002988]; common atrium [HP:0001670]; inlet ventricular septal defect [HP:0011670]
P77	US	Medium	Medium	Depressed nasal bridge [HP:0005280]; epicanthus [HP:0000286]; cleft palate [HP:0000175]; inguinal hernia [HP:0000023]; low-set ears [HP:0000369]; phimosis [HP:0000053]; unilateral cleft lip [HP:0011351]
P9	High	Low	Medium	Epicanthus [HP:0000286]; bulbous nose [HP:0000414]; Atrioventricular canal defect [HP:0001709]; ID (moderate) [HP:0010864]; preauricular skin tag [HP:0100879];
P59	Low	Low	N/A	Microcephaly [HP:0000252]; synophrys [HP:0000664] Postnatal growth retardation [HP:0008897]; malar flattening [HP:0000272]; clinodactyly of the 5th finger [HP:0004209]; medial flaring

Notes: Abbreviations: N/A = no value provided. ADHD = attention deficit hyperactivity disorder. DD = Developmental delay. ID =Intellectual disability.

Appendices

Appendix A

Approved research protocol

UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

Private Bag 3 Wits, 2050
Fax: 027117172119
Tel: 02711 7172076

Reference: Ms Sandra Munesar
E-mail: Sandra.munesar@wits.ac.za

13 August 2024
Person No: 2115219
PAG

Mr DJ Mccarthy
38 Paul Street
Northwold
Randburg
2188
South Africa

Dear Mr Dylan Mccarthy

Master of Science in Medicine: Approval of Title

We have pleasure in advising that your proposal entitled *Characterization of the genetic and phenotypic landscape of patients with CNVs in the chromosome 22q11.2 region in a cohort of South African patients* has been approved. Please note that any amendments to this title have to be endorsed by the Faculty's higher degrees committee and formally approved.

Yours sincerely

Ms Sandra Munesar
Faculty Registrar
Faculty of Health Sciences

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

FACULTY OF HEALTH SCIENCES

CANDIDATE'S SURNAME:	McCarthy	FIRST NAME/S:	Dylan James	STUDENT NUMBER:	2115219
CURRENT QUALIFICATIONS:					
TEL: N/A	CELL: 0714828565	E-MAIL: 2115219@students.wits.ac.za		FAX:N/A	
DEGREE FOR WHICH PROTOCOL IS BEING SUBMITTED: MSc (Med): Genetic Counselling					
PART-TIME OR FULL-TIME: Full-Time					
FIRST REGISTERED FOR THIS DEGREE:	TERM : January			YEAR: 2024	
DEPARTMENT: Division of Human Genetics, School of Pathology, Faculty of Health Science					
TITLE OF PROPOSED RESEARCH: Characterisation of the genotypic and phenotypic landscape of patients with CNVs in the 22q11.2 region in a cohort of South African patients					
CANDIDATE'S SIGNATURE:				DATE: 23/05/2024	
SUPERVISOR 1 (NAME & SURNAME): Mrs. Bianca Rossouw				% Supervision 60%	
SUPERVISOR'S QUALIFICATIONS: Master of Science in Medicine Genetic Counselling					
SUPERVISOR'S DEPARTMENT: Division of Human Genetics, School of Pathology, Faculty of Health Science					
SUPERVISOR'S ADDRESS / TEL / E-MAIL: 0114899236/9223 | bianca.rossouw@nhls.ac.za					
SUPERVISOR 2 (NAME & SURNAME): Dr. Bronwyn Dillon				% Supervision 40%	
SUPERVISOR'S QUALIFICATIONS Bachelor of Medicine and Bachelor of Surgery, Master of Medicine in Medical Genetics, Fellowship of the College of Medical Geneticists of South Africa					
SUPERVISOR'S ADDRESS / TEL / E-MAIL: 0114899338/9223 | Bronwyn.dillon@nhls.ac.za					
SUPERVISOR 3 (NAME & SURNAME): N/A				% Supervision N/A	
SUPERVISOR'S QUALIFICATIONS N/A					
SUPERVISOR'S ADDRESS / TEL / E-MAIL: N/A					

SYNOPSIS OF RESEARCH:*(Brief summary of proposed research project; between 200-300 words only; with sub-headings: an introduction and justification for study, aim/s, proposed methodology and expected outcome/s)*
[Use reverse side of this page if
more space is required]

Introduction

The chromosome 22q11.2 region is a hotspot for non-homologous recombination, which results in the appearance of recurrent copy number variants (CNVs) in this region. Deletions and duplications spanning the 22q11.2 region result in 22q11.2 deletion syndrome and 22q11.2 duplication syndrome respectively. These two syndromes are the most common microduplication and microdeletion syndromes globally, and present with a variable clinical presentation. This clinical presentation varies both in the affected organ systems and severity of affectedness. Over the years, diagnostic testing for CNVs in this region has progressed from Fluorescence *in* situ hybridisation (FISH), to multiplex probe-dependent amplification (MLPA) and finally chromosomal microarray (CMA). While much is known about the genetic aetiology of 22q11.2 deletion and 22q11.2 duplication syndrome, the relationship between patient's genotypic and phenotypic presentation is poorly understood. Additionally, the phenotypic presentation of 22q11.2 deletion syndrome and 22q11.2 duplication syndrome in patients of African ancestry is poorly described.

WITS ETHICS NOT REQUIRED: Yes No WITS ETHICS PENDING: Yes No WITS ETHICS APPROVED: Yes No (circle appropriate symbol)* ***Please note the final human ethics clearance certificate or animal ethics certificate must be available prior to starting research**		IF Y SUPPLY ETHICS CLEARANCE CERTIFICATE AS ATTACHMENT AND INCLUDE ETHICS NUMBER HERE:
As supervisor/s, I/we confirm that I have read the protocol which has been submitted for assessment.		
SIGNATURE OF SUPERVISOR/S:	Bassam. … ……..	……
SIGNATURE PG OFFICE STAFF ……..	REGISTERED YES….. NO…..	STAMP

SYNOPSIS OF RESEARCH CONTINUED

Machine learning, artificial intelligence tools such as Face2Gene have provided a revolutionary means to clinically diagnose patients according to their phenotypic presentation, however these technologies are yet to be verified in African populations. Therefore, the aim of the study is to characterize the CNVs that have been previously identified using CMA in the 22q11.2 region, and to determine any genotype-phenotype correlations in a South African patient cohort over a 10-year period.

Aim

To characterize the CNVs that have been previously identified using CMA in the 22q11.2 region, and to determine any genotype-phenotype correlations in a South African patient cohort over a 10-year period.

Methodology

22q11.2 CNV positive patient files between the years 2013 and 2023 available within the division of Human Genetics, at the NHLS Braamfontein will be retrospectively analyzed to extract relevant information. This information will include: the type of genetic test performed, genetic test results, patient ethnicity, age of presentation, patient sex, patient photographs, and any clinical phenotype data on record. This data will be analyzed using descriptive statistics to identify trends. The phenotypes of the patients who were tested using CMA will be captured and compared to their genotypes. Genotype-phenotype correlations will be assessed using a contingency table and Chi-Square analysis at a significance level of a p-value <0.05. Finally patient photographs will be uploaded to Face2Gene, whereby the sensitivity of Face2Gene will be assessed for the identification of 22q11.2 duplication syndrome or 22q11.2 deletion syndrome in patients of African ancestry.

Expected outcomes

Trends in testing patterns over the 10-year period as well as the frequencies of deletion positive and duplication positive patients are expected to be identified. Genotype-phenotype correlations in patients of African ancestry are expected to be established and the sensitivity of Face2Gene for the identification of the 22q11.2 region associated microduplication and microdeletion syndromes will be determined, therefore the utility of the Face2Gene software in an African setting will be assessed.

Characterization of the genetic and phenotypic landscape of patients with CNVs in the chromosome 22q11.2 region in a cohort of South African patients

Dylan McCarthy

2115219

MSc (Med): Genetic Counselling

Supervisor 1:

Mrs. Bianca Rossouw

(MSc (Med): Genetic counselling, Genetic Counsellor)

Supervisor 2:

Dr Bronwyn Dillon

(MBBCh, MMed: Medical Genetics, FCMG(SA), Medical Geneticist)

Table of Contents

1 Background literature analysis and critique

1.1 Introduction

Microdeletion and microduplication syndromes are a group of disorders caused by submicroscopic deletions and duplications of contiguous genes on parts of chromosomes, typically ranging from one to three megabases (Mb) in length (Sgardioli et al., 2017). Among the well-known microdeletion and microduplication syndromes, 22q11.2 deletion and 22q11.2 duplication syndrome are the most prevalent conditions observed in genetics clinics globally (Blagowidow et al., 2023). North American studies have estimated the live birth incidence of 22q11.2 deletion syndrome to be 1:2000 with a notably higher prenatal prevalence, affecting approximately 1:1000 pregnancies, however global prevalence figures are yet to be defined (Blagojevic et al., 2021; Maisenbacher et al., 2017). The prevalence of 22q11.2 duplication syndrome is less defined with most studies yet to identify a consistent prevalence estimate (Bartik et al., 2022). Molecular genetic testing for 22q11.2 deletion or duplication syndrome has progressed from fluorescence *in situ* hybridisation (FISH), to multiplex ligation-dependent probe amplification (MLPA), and most recently chromosomal microarray (CMA) (Sooknanan et al., 2023). Each technology has enabled a deeper understanding of each condition, however diagnostic challenges remain. One of these challenges is the variable clinical presentation observed in individuals harboring copy number variants (CNVs) within the 22q11.2 region (Cirillo et al., 2022). This variable clinical presentation can make clinical diagnoses challenging to attain, consequently complicating molecular genetic testing, and leaving many patients undiagnosed (Cirillo et al., 2022). Adding to diagnostic challenges is a lack of both phenotypic and genomic data on the manifestation of copy number variation in the 22q11.2 region in individuals of African descent (Kruszka et al., 2017).

Next generation phenotyping (NGP) technologies are a new and exciting development in the field of clinical diagnostics and may have the potential to improve the diagnosis and management of both 22q11.2 deletion and 22q11.2 duplication syndrome. However, these technologies are yet to be verified in African populations and require further validation before they are to be implemented in South African clinical settings (Elmas & Göğüş, 2022; Hennocq et al., 2024). Thus, while much has been done to advance our understanding of copy number variation with the 22q11.2 region, there is still a

large gap in the literature with respect to the presentations of these conditions in Southern African populations which remains to be explored.

1.2 Molecular genetics of the copy number variation observed in the 22q11.2 region

Copy number variation within the 22q11.2 region occurs as a result of non-allelic homologous recombination (NAHR) between clusters of low copy repeat (LCR) blocks, designated LCR22A-H depending on the size and location of the CNV within the region (Figure 1). The high incidence of NAHR in the 22q11.2 region can be attributed to the increased sequence homology observed between the LCR22A-H blocks (Rojnueangit et al., 2022). The most common CNV observed in this region is an approximate 3Mb CNV spanning the LCR22A-D blocks (Figure 1). Although relatively less common, smaller nested or atypical deletions and duplications can also be observed between the other LCR blocks (Burnside, 2015).

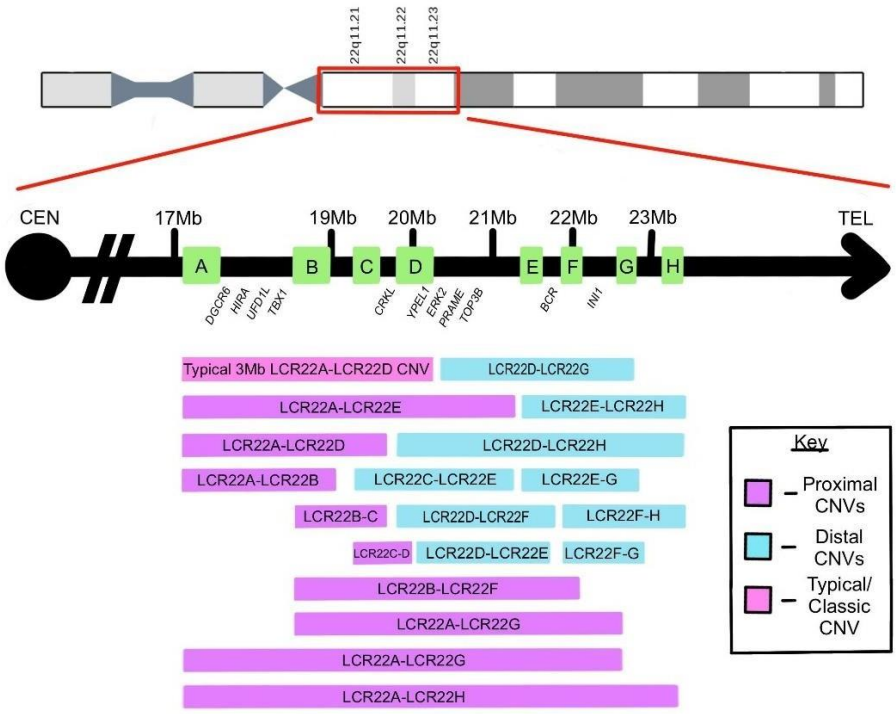


Figure 1: The chromosome 22q11.2 region as well as the genes of interest and LCR blocks observed within the region. The classic LCR22A-D CNV is represented in pink, with atypical proximal and distal CNVs represented in purple and blue.

1.3 Achieving a molecular diagnosis of 22q11.2 deletion/duplication syndrome

A confirmed 22q11.2 CNV molecular diagnosis allows for more targeted medical management and surveillance for the patient and provides valuable information on the inheritance and risk of recurrence in future generations (Blagowidow et al., 2023; McDonald-McGinn et al., 2020). A wide variety of testing options are available for the diagnosis of 22q11.2 deletion or duplication syndrome, namely FISH, MLPA and CMA. Each technology possesses both advantages and limitations (Table 1), with FISH being the most limited technology (Blagowidow et al., 2023). The use of MLPA kits has since improved upon the limitations associated with FISH and allowed for the use of a single assay for the detection of multiple microdeletion and microduplication syndromes, including 22q11.2 deletion syndrome and 22q11.2 duplication syndrome (Sooknanan et al., 2023). In addition to the standard p245 MLPA kit (MRC, Holland), condition specific MLPA kits are also available. The p070 MLPA kit (MRC, Holland) designed specifically to identify CNVs in the 22q11.2 region further improved the resolution of both FISH and the standard MLPA kit, however, even this kit has limitations with respect to its probe coverage and is unable to identify every 22q11.2 CNV or their precise breakpoints (Sgardioli et al., 2017). More recently, CMA has provided a means for genome-wide copy number analysis and allows for the visualization of precise breakpoints approximately 10 kilobases (Kb) or higher (Sooknanan et al., 2023). This enables accurate sizing and allows for the precise identification of any genes which may be affected by the CNV, thereby allowing for genotype-phenotype correlations to be made in individuals presenting with CNVS in the 22q11.2 region (Du et al., 2020).

Table 1: Advantages and limitations associated with FISH, the p070 MLPA kit and CMA respectively.

	FISH	MLPA	CMA	Reference
Types of CNVs identified	• LCR22A-D	• LCR22A-D • Atypical CNVs (limited accuracy)	• LCR22A-D • Atypical CNVs (precise accuracy)	(Blagowidow et al., 2023; Sooknanan et al., 2023)
Cost	~R2000	~R3000	~R8000	(Mudau et al., 2023)
Main limitation	Cannot identify CNVs outside of LCR22A-D region	Cannot accurately identify CNV breakpoints	More costly than FISH and MLPA	(Blagowidow et al., 2023; Sooknanan et al., 2023)

1.4 Phenotypic variability observed in patients with CNVs in the 22q11.2 region

Individuals presenting with CNVs in the 22q11.2 region can have variable phenotypic presentations, in terms of degree of symptom severity, and the nature and extent of the affected body systems (McDonald-McGinn et al., 2020). Some patients who harbor CNVs within the 22q11.2 region demonstrate fewer or milder phenotypic features and may go undiagnosed for most of their lives (Cirillo et al., 2022). Conversely, some individuals may have a more severe clinical presentation with multiple body systems affected (Cirillo et al., 2022). The major and more common clinical manifestations that have been frequently observed in patients with 22q11.2 deletions and duplications are outlined in Table 2.

Table 2: The major phenotypic features in patients with CNVs in the 22q11.2 region.

Clinical phenotype	% occurrence		Reference
	Deletions	Duplications	
Dysmorphic craniofacial features	50%	26%	(Bartik et al., 2022; Du et al., 2020)
Congenital heart defects	75%	42%	(Bartik et al., 2022; Cirillo et al., 2022; Du et al., 2020)
Gastrointestinal anomalies	30%	84%	(Bartik et al., 2022; Du et al., 2020)
Renal anomalies	14%	52%	(Bartik et al., 2022; Du et al., 2020)
Skeletal defects	60%	56%	(Bartik et al., 2022; Du et al., 2020)
Hypocalcaemia	35%	6%	(Bartik et al., 2022; Du et al., 2020)
Immune deficiency	50-70%	54%	(Bartik et al., 2022; Du et al., 2020)
Developmental delay	50%	58%	(Bartik et al., 2022; Du et al., 2020)
Learning disabilities	70%	22%	(Bartik et al., 2022; Du et al., 2020)
Psychiatric Disorders	30%	59%	((Bartik et al., 2022; Cirillo et al., 2022; Du et al., 2020; Nehme et al., 2022)

1.5 The impact of next generation phenotyping

Recently, the field of dysmorphology has been revolutionized by the introduction of artificial intelligence and NGP technologies for the characterization of patients with various genetic disorders (Hennocq et al., 2024). This new advancement serves as an adjunct screening tool to assist with

providing differential clinical diagnoses for patients with dysmorphic features. One such tool is the Face2Gene application (FDNA^{TM} Inc.) (https://www.face2gene.com). The Face2Gene classification algorithm has been trained to identify and characterize various genetic disorders using training cohorts comprised of affected individuals of various ethnicities (Gurovich et al., 2019). The software makes use of a patient's facial photograph, in combination with any additional phenotypic features described in human phenotype ontology (HPO) terms inputted by a clinician, to compile a matched list of suggested differential clinical diagnoses (Gurovich et al., 2019). The photograph itself is converted into a de-identified mathematical facial descriptor to protect the privacy and personal data of the patient involved (Gurovich et al., 2019). NGP is an exciting development in the clinical space, as it may be an effective tool for the clinical diagnosis of patients with relatively non-specific, or genetically heterogeneous, phenotypic features and improve clinical diagnoses, and subsequent molecular diagnoses (Elmas & Göğüş, 2022). NGP technologies could be especially valuable in low-middle income settings, allowing for the prioritization of patients who are most likely to test positive for a particular condition, thus maximizing the use of costly resources such as CMA or whole exome sequencing in resource constrained settings (Mudau et al., 2023). Although Face2Gene has been trained to identify over 100 genetic syndromes, ongoing research on independent datasets to validate the use of the technology is necessary for the software to be trusted in a clinical setting (Hennocq et al., 2024). One of the main limitations of such NGP tools is the underrepresentation of African patients during their algorithm training (Elmas & Göğüş, 2022; Hennocq et al., 2024). Thus, while NGP presents an exciting development in clinical diagnostics, more research is required to validate the use of the Face2Gene technology to clinically diagnose patients of South African ancestry.

1.6 22q11.2 deletion syndrome and 22q11.2 duplication syndrome in Africa

There is a general lack of genomic data in Africa, with respect to both genotypic and phenotypic data for various genetic conditions due to both the underrepresentation of African populations in genomic studies (Ramsay, 2022). The situation is no different for 22q11.2 deletion or duplication syndrome, as African populations have been historically underrepresented in most studies (De Decker et al., 2016). A study conducted by Kruszka et al. (2017) which included 22q11.2 patients of both African and African American patients, noted that patients of African descent display the least consistent phenotypes when compared to the literature. In addition, while there have been a vast number of studies investigating genotype-phenotype correlations in patients with 22q11.2 deletion or duplication

syndrome in North American, European, South American, Asian, and some African populations, (Burnside, 2015; Cirillo et al., 2022; Kruszka et al., 2017; Méndez-Rosado et al., 2022; Rojnueangit et al., 2022; Verbesselt et al., 2022) there have been no studies of this nature conducted in South African populations, representing a clear gap in the literature with respect to what is known about 22q11.2 duplication and deletion syndrome in these populations.

1.7 Study rationale

While much is known about the molecular genetics of 22q11.2 deletion syndrome and duplication syndrome, there is still much to be discovered with respect to understanding the causes behind the variable phenotypic presentation of the conditions. NGP technology aims to improve clinical diagnoses of many genetically heterogeneous conditions, including 22q11.2 deletion and duplication syndrome, but is yet to be validated in South African patients. With the routine implementation of CMA technology, more precise genotypic information is now available for the establishment of genotype-phenotype correlations in patients of African ancestry. Therefore, the aim of this study is to characterize the CNVs that have been previously identified using in the 22q11.2 region, and to determine any genotype-phenotype correlations in a South African patient cohort over a 10-year period.

2 Aim and objectives

2.1 Aim

To characterize the CNVs that have been identified in the 22q11.2 region and determine if there are any genotype-phenotype correlations in a South African cohort.

Objectives:

1. To perform a retrospective file review of all patients who have tested positive for a CNV in the chromosome 22q11.2 region at the Division of Human Genetics, NHLS Braamfontein between 2013 and 2023, by means of FISH, MLPA or CMA.

2. To identify genotype-phenotype correlations in patients with CNVs present in the 22q11.2 region.

3. To assess the utility of using Face2Gene, for the clinical diagnosis of South African patients with CNVs in the 22q11.2 region.

3 Methods

3.1 Study participants

Approximately 200-300 diagnostic test records from the diagnostic laboratory service in the Division of Human genetics at the NHLS Braamfontein will be accessed for this study. In some cases, patients may have also been consulted by the clinical genetics unit in the Division of Human Genetics, in these cases the patient's clinical file will also be accessed. These patient records may include both physical patient files present at the NHLS Braamfontein and digitized files from the online Research Electronic Data Capture (REDCap) database (Harris et al., 2009; Harris et al., 2019).

3.1.1 Inclusion criteria

Only records dated between 2013 and 2023 will be included in this study. Patients must have a positive molecular diagnosis of a CNV in the 22q11.2 region, detected using FISH, MLPA or CMA. For the Face2Gene objective of the study, only patients who possess a high-quality frontal facial photograph within their file and were consulted by the clinical genetics unit within the Division of Human Genetics at the NHLS Braamfontein will be included in this aspect of the study (Figure 2).

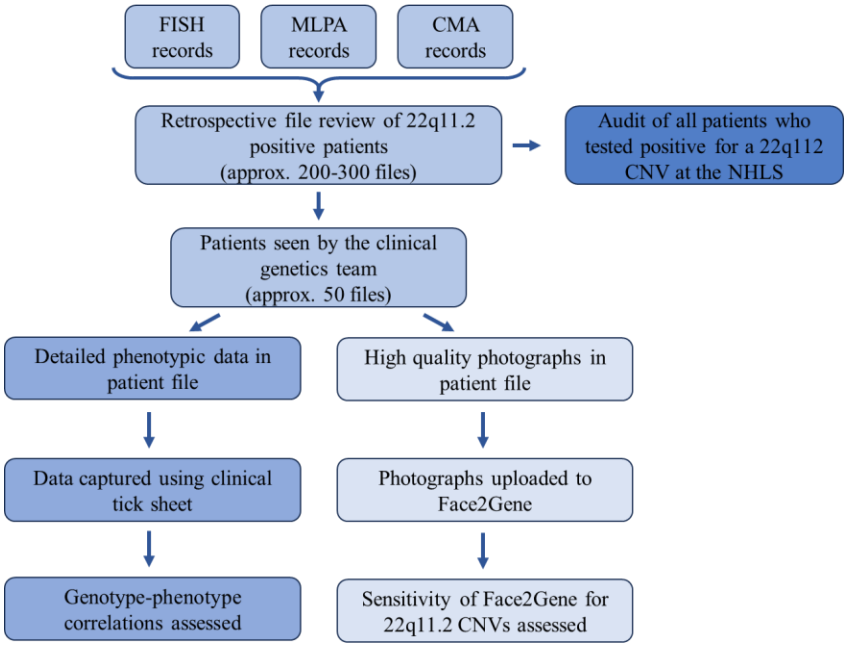


Figure 2: Proposed study design representing the process of collecting patient data, and the inclusion criteria for the various aspects of the study.

3.2 Audit of all patients with CNVs in the 22q11.2 region seen at the NHLS Braamfontein between 2013 and 2023

All patient records concerning patients with CNVs in the 22q11.2 region will be reviewed, and relevant information will be extracted from the records. All diagnostic testing records regarding the use of FISH, MLPA and CMA will be reviewed. The information extracted from the diagnostic testing records will be captured using a data collection sheet (Appendix 1), designed on REDCap which will include the type of genetic test used, the patient age, sex and demographics, the genetic test result, and the year of the genetic testing. The clinical files of any patient with diagnostic records who was also seen by the clinical genetics unit within the Division of Human Genetics at the NHLS Braamfontein, will also be accessed. Information extracted from the patient's clinical file (Appendix 2) will include any clinical and laboratory data on record, as well as any good quality photographic images present in the file. The data will be analyzed using descriptive statistical methods, where categorical variables will be summarized as frequencies, percentages, and figures while continuous variables will be summarized as medians for non-normally distributed data and means for normally distributed data. The data will be tested for normality using the Shapiro-Wilk test. The data will be analyzed using Microsoft Excel.

3.3 Identification of genotype-phenotype correlations

3.3.1 Establishment of patient genotypes and phenotypes

Genotype-phenotype correlations will be investigated only in patients who have been consulted by the clinical genetics team from the Division of Human Genetics, at the NHLS Braamfontein. Information such as the breakpoints of the CNVs present in the patients, as well as the patients' phenotypic descriptors will be recorded using a clinical tick sheet (Appendix 2) designed on REDCap. Cytogenomic tools such as the Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources (DECIPHER) (Firth et al., 2009) will then be used to characterize the CNV regions based upon the affected LCR blocks (Figure 1) and identify which genes have been deleted or duplicated in each patient.

3.3.2 Statistical analysis for the identification of genotype-phenotype correlations

Microsoft Excel will be used to record the frequencies of the various phenotypic features and their corresponding deleted or duplicated regions. Patients will be categorized according to the type of CNVs

they possess (Figure 1). Statistical analyses will include Chi-square tests or Fisher's exact tests where appropriate to determine if there are any statistically significant differences in the phenotypic presentations between the two groups at a significance level of a p-value <0.05.

3.4 Assessment of the utility Face2Gene for the identification 22q11.2 deletion/duplication syndrome

Only South African patients with clinical files which contain high-quality frontal facial photographs will be included in the subset of patients who will undergo Face2Gene analysis. The results generated by Face2Gene include the dysmorphism score (D-Score), feature score, gestalt score, and a list of possible diagnoses. The D-score is a measure of dysmorphism in the patient based upon a normalized probability that dysmorphism is present within a patient's photograph (Elmas & Göğüş, 2022). The gestalt score is an indication of how closely the patient's facial profile matches the profiles of individuals with the various disorders that the software has been trained upon, while the feature score is an indication of how closely the uploaded clinical features (HPO terms) match the typical clinical features observed in patients with the various conditions (Elmas & Göğüş, 2022). The list of possible diagnoses is generated based upon the patient's feature and gestalt scores and indicates the most likely diagnoses recommended by Face2Gene. Only the patient's facial photographs present in their file will be uploaded to the Face2Gene software and the data generated by the software including the D-score, gestalt score and list of possible diagnoses will be recorded. The feature score will not be captured, as this aspect of the study is not blinded, therefore bias will be introduced if the patient's phenotypic features are reported on Face2Gene. Microsoft Excel will be used to capture the Face2Gene data. The D-score will be compared to a clinician's assessment of dysmorphism to assess whether Face2Gene is an accurate predictor of dysmorphology in South African patients with CNVs in the 22q11.2 region. The gestalt score will be used as a metric to assess the extent of Face2Gene's ability to classify the facial profiles of South African individuals as being associated with 22q11.2 deletions or duplications. The list of recommended diagnoses will then be assessed to determine the accuracy of Face2Gene with respect to diagnosing South African patients with 22q11.2 deletions or duplications. This will be achieved by assessing the number of cases whereby 22q11.2 deletion or duplication syndrome falls within the top ten conditions suggested by Face2Gene for atypical and classic CNVs. The sensitivity of Face2Gene for the different CNVs, grouped according to their affected LCR regions, will be determined and comparisons will be made between the different CNVs observed in the patients.

4 Ethics

A submission to the Human Research Ethics Committee (Medical) of the University of the Witwatersrand was made on 18/04/2024, the submission is currently under review. Permission for the use of patient data related to any patient seen and/or tested at the NHLS Braamfontein with CNVs in the 22q11.2 region has been granted by the gatekeeper of the data (Appendix 3). All data accessed in this study will be anonymized and coded and only the researcher and his supervisors will have access to any identifiable information. This anonymous, coded data will be uploaded to Face2Gene along with patient photographs obtained from the patient files. It is emphasized that Face2Gene converts patient photos into deidentified facial descriptors, while the original images are considered personal health information. Face2Gene does not store any personal health information, instead it is saved on a separate disk volume, which will only be accessible to the researcher and his supervisors through a password protected Face2Gene user account on a password protected work computer. Consent for the storage of patient photos within their patient file and for their use in future diagnostic and other purposes has been obtained by a clinician. This consent form should be present within the patient's file. Any patient without a consent form which specifies their consent for their image to be taken and kept on record in their file for diagnostic and other purposes will be excluded from the Face2Gene aspect of this study.

5 Timing

The proposed timeline of the project has been outlined in the Gantt Chart below, with the final write up of the project intended to be submitted to Faculty by December 2024.

	Feb '24	Mar '24	Apr '24	May '24	June '24	July '24	Aug '24	Sept '24	Oct '24	Nov '24	Dec '24
Literature review											
Ethics application											
Protocol preparation											
Protocol assessment											
Data collection											
Record F2G data											
Data analysis											
Writing research report											

6 Funding

This project is retrospectively analyzing freely available data using freely available software, therefore there are no additional costs to the project. For this reason, funding is not required.

7 Anticipated problems/limitations

Due to the variable expressivity observed in patients with CNVs in the 22q11.2 region, previous studies have struggled to identify many genotype-phenotype correlations for the disorders (Gavril et al., 2022). This may also be the case in our cohort and may limit the findings of the study. Additionally, the sample size of the cohort is relatively small as it is limited to patients that have been seen through the Division of Human Genetics at the NHLS Braamfontein, this may limit the ability to draw statistically significant conclusions from the data. As this study is retrospectively analyzing the information captured in the patient files, there may be inconsistencies with respect to how the phenotypes of the patients are described as well as the amount of information available in each patient's file. Some files may be incomplete and therefore would have information omitted. Also, the patient images found within the file may be missing or be poor quality, preventing the patient from being included in the Face2Gene aspect of the study. Consent forms for the use of patient images for diagnostic and other purposes may also be absent from patient files, meaning that these patients would also need to be excluded. These exclusions would decrease the sample size of the Face2Gene aspect of this study, further limiting the findings. Lastly, because the Face2Gene aspect of the study is not blinded, this may introduce bias into the results.

8 References

Bartik, L. E., Hughes, S. S., Tracy, M., Feldt, M. M., Zhang, L., Arganbright, J., & Kaye, A. (2022). 22q11.2 duplications: Expanding the clinical presentation. *American Journal of Medical Genetics Part A*, *188*(3), 779–787. https://doi.org/10.1002/ajmg.a.62577

Blagojevic, C., Heung, T., Theriault, M., Tomita-Mitchell, A., Chakraborty, P., Kernohan, K., Bulman, D. E., & Bassett, A. S. (2021). Estimate of the contemporary live-birth prevalence of recurrent 22q11.2 deletions: a cross-sectional analysis from population-based newborn screening. *CMAJ Open*, *9*(3), E802–E809. https://doi.org/10.9778/cmajo.20200294

Blagowidow, N., Nowakowska, B., Schindewolf, E., Grati, F. R., Putotto, C., Breckpot, J., Swillen, A., Crowley, T. B., Loo, J. C. Y., Lairson, L. A., Óskarsdóttir, S., Boot, E., Garcia-Minaur, S., Cristina Digilio, M., Marino, B., Coleman, B., Moldenhauer, J. S., Bassett, A. S., & McDonald- McGinn, D. M. (2023). Prenatal Screening and Diagnostic Considerations for 22q11.2 Microdeletions. *Genes*, *14*(1), 160. https://doi.org/10.3390/genes14010160

Burnside, R. D. (2015). 22q11.21 Deletion Syndromes: A Review of Proximal, Central, and Distal Deletions and Their Associated Features. *Cytogenetic and Genome Research*, *146*(2), 89–99. https://doi.org/10.1159/000438708

Cirillo, A., Lioncino, M., Maratea, A., Passariello, A., Fusco, A., Fratta, F., Monda, E., Caiazza, M., Signore, G., Esposito, A., Baban, A., Versacci, P., Putotto, C., Marino, B., Pignata, C., Cirillo, E., Giardino, G., Sarubbi, B., Limongelli, G., & Russo, M. G. (2022). Clinical Manifestations of 22q11.2 Deletion Syndrome. *Heart Failure Clinics*, *18*(1), 155–164. https://doi.org/10.1016/j.hfc.2021.07.009

De Decker, R., Bruwer, Z., Hendricks, L., Schoeman, M., Schutte, G., & Lawrenson, J. (2016). predicted v. real prevalence of the 22q11.2 deletion syndrome in children with congenital heart disease presenting to Red Cross War Memorial Children's Hospital, South Africa: A prospective study. *South African Medical Journal*, *106*(6), 82. https://doi.org/10.7196/SAMJ.2016.v106i6.11003

Du, Q., de la Morena, M. T., & van Oers, N. S. C. (2020). The Genetics and Epigenetics of 22q11.2 Deletion Syndrome. *Frontiers in Genetics*, *10*. https://doi.org/10.3389/fgene.2019.01365

Elmas, M., & Göğüş, B. (2022). The road from mutation to next generation phenotyping: contribution of deep learning technology (Face2Gene) to diagnosis neurofibromatosis type 1. *The European Research Journal*, *8*(2), 145–154. https://doi.org/10.18621/eurj.894631

Gavril, E.-C., Popescu, R., Nucă, I., Ciobanu, C.-G., Butnariu, L. I., Rusu, C., & Pânzaru, M.-C. (2022). Different Types of Deletions Created by Low-Copy Repeats Sequences Location in 22q11.2 Deletion Syndrome: Genotype–Phenotype Correlation. *Genes*, *13*(11), 2083. https://doi.org/10.3390/genes13112083

Gurovich, Y., Hanani, Y., Bar, O., Nadav, G., Fleischer, N., Gelbman, D., Basel-Salmon, L., Krawitz, P. M., Kamphausen, S. B., Zenker, M., Bird, L. M., & Gripp, K. W. (2019). Identifying facial phenotypes of genetic disorders using deep learning. *Nature Medicine*, *25*(1), 60–64. https://doi.org/10.1038/s41591-018-0279-0

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L., Delacqua, G., Delacqua, F., Kirby, J., & Duda, S. N. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*, *95*, 103208. https://doi.org/10.1016/j.jbi.2019.103208

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, *42*(2), 377–381. https://doi.org/10.1016/j.jbi.2008.08.010

Hennocq, Q., Willems, M., Amiel, J., Arpin, S., Attie-Bitach, T., Bongibault, T., Bouygues, T., Cormier-Daire, V., Corre, P., Dieterich, K., Douillet, M., Feydy, J., Galliani, E., Giuliano, F., Lyonnet, S., Picard, A., Porntaveetus, T., Rio, M., Rouxel, F., … Garcelon, N. (2024). Next generation phenotyping for diagnosis and phenotype–genotype correlations in Kabuki syndrome. *Scientific Reports*, *14*(1), 2330. https://doi.org/10.1038/s41598-024-52691-3

Kruszka, P., Addissie, Y. A., McGinn, D. E., Porras, A. R., Biggs, E., Share, M., Crowley, T. B., Chung, B. H. Y., Mok, G. T. K., Mak, C. C. Y., Muthukumarasamy, P., Thong, M., Sirisena, N. D., Dissanayake, V. H. W., Paththinige, C. S., Prabodha, L. B. L., Mishra, R., Shotelersuk, V., Ekure, E. N., … Muenke, M. (2017). 22q11.2 deletion syndrome in diverse populations. *American Journal of Medical Genetics Part A*, *173*(4), 879–888. https://doi.org/10.1002/ajmg.a.38199

Maisenbacher, M. K., Merrion, K., Pettersen, B., Young, M., Paik, K., Iyengar, S., Kareht, S., Sigurjonsson, S., Demko, Z. P., & Martin, K. A. (2017). Incidence of the 22q11.2 deletion in a large cohort of miscarriage samples. *Molecular Cytogenetics*, *10*(1), 6. https://doi.org/10.1186/s13039-017-0308-6

McDonald-McGinn, D. M., Hain, H. S., Emanuel, B. S., & Zackai, E. H. (2020). *22q11.2 Deletion Syndrome*. https://www.ncbi.nlm.nih.gov/books/NBK1523

Méndez-Rosado, L. A., de León-Ojeda, N., García, A., Sheth, F., Gaadi, A., Bousfiha, A. A., Lehlimi, M., Natiq, A., Kurinnaia, O. S., Vorsanova, S. G., Iourov, I., Huhle, D., & Liehr, T. (2022). Clinical characterization of 72 patients with del(22)(q11.2q11.2) from different ethnic backgrounds. *Egyptian Journal of Medical Human Genetics*, *23*(1), 158. https://doi.org/10.1186/s43042-022-00374-x

Mudau, M. M., Seymour, H., Nevondwe, P., Kerr, R., Spencer, C., Feben, C., Lombard, Z., Honey, E., Krause, A., & Carstens, N. (2023). A feasible molecular diagnostic strategy for rare genetic disorders within resource-constrained environments. *Journal of Community Genetics*, *15*(1), 39–48. https://doi.org/10.1007/s12687-023-00674-8

Nehme, R., Pietiläinen, O., Artomov, M., Tegtmeyer, M., Valakh, V., Lehtonen, L., Bell, C., Singh, T., Trehan, A., Sherwood, J., Manning, D., Peirent, E., Malik, R., Guss, E. J., Hawes, D., Beccard, A., Bara, A. M., Hazelbaker, D. Z., Zuccaro, E., … Eggan, K. (2022). The 22q11.2 region regulates presynaptic gene-products linked to schizophrenia. *Nature Communications*, *13*(1), 3690. https://doi.org/10.1038/s41467-022-31436-8

Ramsay, M. (2022). African genomic data sharing and the struggle for equitable benefit. *Patterns*, *3*(1). https://doi.org/10.1016/j.patter.2021.100412

Rojnueangit, K., Khetkham, T., Onsod, P., & Chareonsirisuthigul, T. (2022). Clinical Features to Predict 22q11.2 Deletion Syndrome Proven by Molecular Genetic Testing. *Journal of Pediatric Genetics*, *11*(01), 022–027. https://doi.org/10.1055/s-0040-1718386

Sgardioli, I. C., de Mello Copelli, M., Monteiro, F. P., dos Santos, A. P., Lustosa Mendes, E., Paiva Vieira, T., & Gil-da-Silva-Lopes, V. L. (2017). Diagnostic Approach to Microdeletion Syndromes Based on 22q11.2 Investigation: Challenges in Four Cases. *Molecular Syndromology*, *8*(5), 244–252. https://doi.org/10.1159/000477598

Sooknanan, R., Baine-savanhu, F., & Ayuk, S. (2023). A review of 22q11.2 microdeletion syndrome: clinical and diagnostic perspective. *The Journal of Medical Laboratory Science and Technology of South Africa*, *5*(1), 31–38. https://doi.org/10.36303/JMLSTSA.92

Verbesselt, J., Zink, I., Breckpot, J., & Swillen, A. (2022). Cross-sectional and longitudinal findings in patients with proximal 22q11.2 duplication: A retrospective chart study. *American Journal of Medical Genetics Part A*, *188*(1), 46–57. https://doi.org/10.1002/ajmg.a.62487

Protocol appendix A.1: Data collection sheet

Data Collection Sheet

Record ID _____

Patient Demographics

Sex
- ◯ Female
- ◯ Male

Age at Referral _____

Referring Hospital
- ◯ CHBAH
- ◯ CMJAH
- ◯ RMMCH
- ◯ NMCH
- ◯ SBAH
- ◯ Other

If other, specify _____

Population Group
- ◯ Asian
- ◯ Black
- ◯ Indian
- ◯ Mixed Ancestry
- ◯ White
- ◯ Other

Ethnolinguistic group
- ☐ Afrikaner
- ☐ Ashkenazi Jewish
- ☐ European
- ☐ Hindi
- ☐ Muslim
- ☐ Ndebele
- ☐ Sepedi
- ☐ Sotho
- ☐ Tsonga
- ☐ Tswana
- ☐ Venda
- ☐ Zulu
- ☐ Xhosa
- ☐ Other
- ☐ Unknown

Genetic Testing History

Year of Testing _____

Deletion or Duplication?
- ◯ Deletion
- ◯ Duplication

Classic or Atypical CNV?
- ◯ Classic CNV (LCR22A-D)
- ◯ Atypical CNV

Which LCR regions are involved?	☐ LCR22A-B ☐ LCR22B-C ☐ LCR22C-D ☐ LCR22D-E ☐ LCR22E-F ☐ LCR22F-G ☐ LCR22G-H
Therefore, overall span of CNV?	_____
FISH	◯ Yes ◯ No
Genetic Test Result	_____
Upload test result (if possible)	
MLPA	◯ Yes ◯ No
MLPA kits used for testing	☐ P245 ☐ P036 ☐ P070 ☐ P250
Genetic Test Result	_____
Upload test result (if possible)	
CMA	◯ Yes ◯ No
Genetic Test Result	_____
Upload test result (if possible)	

Clinical History

Seen by the clinical genetics unit?	◯ Yes ◯ No
High quality image in file?	◯ Yes ◯ No
Phenotype well described in file?	◯ Yes ◯ No

Overall inclusion in the study

Which aspects of the study will this participant be included?

- ☐ Clinical Audit
- ☐ Face2Gene
- ☐ Genotype-Phenotype correlations

Protocol appendix A.2: Clinical ticksheet

Clinical Ticksheet

Record ID _____

Age at last assessment _____

Family History

Family History of CNVs in the 22q11.2 region
◯ Yes
◯ No

Specify _____

Upload Family pedigree

Prenatal Findings

Prenatal History on record?
◯ Yes
◯ No

Abnormalities noted
◯ Yes
◯ No
◯ Unknown

Specify _____

Decreased fetal movement?
◯ Yes
◯ No
◯ Unknown

Abnormal presentation/lie (eg. breech)
◯ Yes
◯ No
◯ Unknown

Specify _____

Polyhydramnios
◯ Yes
◯ No
◯ Unknown

Mode of delivery
◯ Vaginal Birth
◯ Caesarian section
◯ Unknown

APGAR scores known?
◯ Yes
◯ No

AGAR score @1min

_____

AGAR score @5min

_____

AGAR scores @10min

_____

Postnatal complications

◯ Yes
◯ No
◯ Unknown

Specify

_____

Any other finding/s of note

_____

Growth

Growth history at birth on record?

◯ Yes
◯ No

Weight (centile/Z-score):

_____

Interpretation:

_____

Length (centile/Z-score):

_____

Interpretation:

_____

OFC (centile/Z-score):

_____

Interpretation:

_____

Growth history at last consultation on record?

◯ Yes
◯ No

Weight (centile/Z-score):

_____

Interpretation:

_____

Length (centile/Z-score): _____

Interpretation: _____

OFC (centile/Z-score): _____

Interpretation: _____

Craniofacial anomalies and Dysmorphism

Craniofacial Anomalies assessed?
- ◯ Yes
- ◯ No

Skull/face

Asymmetric crying facies
- ◯ Present
- ◯ Absent
- ◯ Unknown/Not Assessed

Skull asymmetry or craniosynostosis (specify)
- ◯ Present
- ◯ Absent
- ◯ Unknown/Not Assessed

Other (specify) _____

Eyes

Hooded eyelids
- ◯ Present
- ◯ Absent
- ◯ Unknown/Not Assessed

Ptosis (specify uni/bilateral)
- ◯ Present (bilateral)
- ◯ Present (unilateral)
- ◯ Absent
- ◯ Unknown/Not Assessed

Epicanthic folds
- ◯ Present
- ◯ Absent
- ◯ Unknown/Not Assessed

USPD
- ◯ Present
- ◯ Absent
- ◯ Unknown/Not Assessed

DSPF
- ◯ Present
- ◯ Absent
- ◯ Unknown/Not Assessed

Hypertelorism (or impression thereof)	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Any other craniofacial anomalies? (specify)	_____
Formal ophthalmologic exam done	◯ Yes ◯ No ◯ Unknown
Abnormality (specify)	_____
Tortuous retinal vessels	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Sclerocornea	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Coloboma (specify uni/bilateral; location))	◯ Present (bilateral) ◯ Present (unilateral) ◯ Absent ◯ Unknown/Not Assessed
Location of Coloboma	_____
Cataract (specify uni/bilateral)	◯ Present (bilateral) ◯ Present (unilateral) ◯ Absent ◯ Unknown/Not Assessed
Location of Cateract	_____
Anophthalmia (specify uni/bilateral)	◯ Present (bilateral) ◯ Present (unilateral) ◯ Absent ◯ Unknown/Not Assessed
Strabismus (specify uni/bilateral)	◯ Present (bilateral) ◯ Present (unilateral) ◯ Absent ◯ Unknown/Not Assessed
Visual acuity problem	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify visual acuity problem	_____

Other (specify) _____

Nose

Abnormalities of the nose	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Prominent nasal bridge	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Bulbous nose	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Tubular nose	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Hypoplastic alae nasi	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Nasal dimple/bifid nasal tip	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Broad nose	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Flat nose	◯ Present ◯ Absent ◯ Unknown/Not Assessed

Other (specify) _____

Oropharngeal abnormalities

Oropharyngeal abnormalities present	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Cleft lip only (specify uni/bilateral)	◯ Present (Bilateral) ◯ Present (Unilateral) ◯ Absent ◯ Unknown/Not Assessed
Cleft lip and palate	◯ Present ◯ Absent ◯ Unknown/Not Assessed

Specify	_____
Cleft palate only	◯ Present (hard) ◯ Present (soft) ◯ Present (v-shaped) ◯ Present (u-shaped) ◯ Absent ◯ Unknown/Not Assessed
Submucous cleft palate	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Velopharyngeal incompetence (VPI)	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Bifid uvula	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Hypernasal speech	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Feeding difficulty	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify type and intervention	_____
Food allergy	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify Allergy	_____
Dysphagia	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Micrognathia	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Dental caries	◯ Present ◯ Absent ◯ Unknown/Not Assessed

Other (specify)

_____

Laryngotracheoesophageal

Laryngotracheoesophageal anomalies present	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Vascular ring	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Laryngeal web	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Laryngotracheomalacia	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Subglottic stenosis	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Stridor	◯ Present ◯ Absent ◯ Unknown/Not Assessed

Other (specify)

_____

Ear

Ear anomalies present	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Over-folded helices	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Cupped/cup-shaped	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Microtia	◯ Present (Bilateral) ◯ Present (Unilateral) ◯ Absent ◯ Unknown/Not Assessed

Grade of microtia

_____

Protuberant ears	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Pre-auricular pit/s	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Pre-auricular tag/s	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Narrow external auditory meatus/meati	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Any other ear anomalies of note	_____

Cardiac/chest

Formal cardiac assessment performed	◯ Yes ◯ No ◯ Unknown
Congenital heart defect/s	◯ Present ◯ Absent ◯ Unknown/Not Assessed
VSD	◯ Present ◯ Absent ◯ Unknown/Not Assessed
ASD	◯ Present ◯ Absent ◯ Unknown/Not Assessed
AVSD	◯ Present ◯ Absent ◯ Unknown/Not Assessed
PDA	◯ Present ◯ Absent ◯ Unknown/Not Assessed
PFO	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Tetralogy of Fallot	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Aortic arch anomaly	◯ Present ◯ Absent ◯ Unknown/Not Assessed

Coarctation of the aorta	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Interrupted aortic arch	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Truncus arteriosus	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Pulmonary atresia/stenosis	◯ Present (Atresia) ◯ Present (Stenosis) ◯ Absent ◯ Unknown/Not Assessed
Dilated aortic root	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Bicuspid aortic valve	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Mitral valve prolapse	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Abnormal heart rate	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify HR	_____
Abnormal heart rhythm	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify abnormal heart rhythm	_____
Abnormality of chest wall (e.g. pectus)	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify chest wall abnormality	_____
Abnormal lung lobation	◯ Present ◯ Absent ◯ Unknown/Not Assessed

Breathing difficulties	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify age and severity of breathing difficulties	_____
Any other chest/heart anomalies (specify)	_____

Gastrointestinal

Gastrointestinal anomalies	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Anteriorly placed anus	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Imperforate anus	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Oesophageal atresia	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Jejunal atresia	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Intestinal malrotation	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Hirschsprungs disease	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Accessory spleen	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Absent spleen	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Diaphragmatic hernia	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Umbilical hernia	◯ Present ◯ Absent ◯ Unknown/Not Assessed

Inguinal hernia	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Other gastrointestinal abnormalities (specify)	_____
Functional problems	☐ Constipation only ☐ Gastrointestinal reflux only ☐ Constipation and gastrointestinal reflux ☐ Absent ☐ Not assessed

Genitourinary

Renal ultrasound/imaging performed	◯ Yes ◯ No
Renal Abnormalities	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify	_____
Hydronephrosis	◯ Present (unilateral) ◯ Present (bilateral) ◯ Absent ◯ Unknown/Not Assessed
Renal agenesis	◯ Present (unilateral) ◯ Present (bilateral) ◯ Absent ◯ Unknown/Not Assessed
Multicystic kidneys	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Dysplastic kidneys	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Cryptorchidism	◯ Present (Bilateral) ◯ Present (Unilateral) ◯ Absent ◯ Unknown/Not Assessed
Hypospadias	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify location	_____

Vaginal agenesis	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Uterine agenesis	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Bladder exstrophy	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Any other renal anomalies (specify)	_____

CNS/Development

CNS/Developmental abnormality present	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Abnormal tone present?	◯ Present ◯ Absent ◯ Unassessed
Abnormal tone	◯ hypotonia ◯ hypertonia
Specify abnormal tone	_____
Abnormal reflexes present?	◯ Present ◯ Absent ◯ Unassessed
Abnormal reflexes	◯ hyporeflexia ◯ hyperreflexia
Specify abnormal reflexes	_____
Seizures present	◯ Present (unprovoked) ◯ Present (Associated with low Ca2+) ◯ Absent ◯ Unknown/Not Assessed
Type of seizures experienced	_____
Brain imaging performed	◯ Yes ◯ No ◯ Unknown/Not Assessed

Normal or abnormal	◯ Normal ◯ Abnormal
Specify any brain imaging abnormalities	_____
Developmental delay observed	◯ Yes ◯ No ◯ Unknown/Not Assessed
Specify domain and degree	_____
Learning difficulties	◯ Yes ◯ No ◯ Unknown/Not Assessed
Intellectual disability	◯ Yes ◯ No ◯ Unknown/Not Assessed
Specify degree	_____
Neuropsychiatric differences	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Any other CNS/Developmental Abnormalities noted? (Specify)	_____
ASD	◯ Yes ◯ No ◯ Unknown/Not Assessed
ADHD	◯ Yes ◯ No ◯ Unknown/Not Assessed
Schizophrenia	◯ Yes ◯ No ◯ Unknown/Not Assessed
Anxiety	◯ Yes ◯ No ◯ Unknown/Not Assessed
Social difficulties	◯ Yes ◯ No ◯ Unknown/Not Assessed

Disinhibition and impulsiveness	◯ Yes ◯ No ◯ Unknown/Not Assessed
Shyness and withdrawn	◯ Yes ◯ No ◯ Unknown/Not Assessed
Temper problem and aggression	◯ Yes ◯ No ◯ Unknown/Not Assessed
Parkinson disease	◯ Yes ◯ No ◯ Unknown/Not Assessed
Highly sociable	◯ Yes ◯ No ◯ Unknown/Not Assessed
Any other neuropsychiatric conditions (Specify)	_____

Musculoskeletal

Spinal imaging performed	◯ Yes ◯ No
Musculoskeletal abnormalities present	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Occipital cervical anomaly	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify occipital cervical anomaly	_____
Scoliosis	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify	_____
Kyphosis	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify	_____

Kyphoscoliosis	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify	_____
Rib anomaly	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify	_____
Vertebral anomaly	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify	_____
Club foot	◯ Present (Bilateral) ◯ Present (Unilateral) ◯ Absent ◯ Unknown/Not Assessed
Polydactyly	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify polydactyly present	_____
Overfolded toes	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Congenital hip dysplasia	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Contractures	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify contractures	_____
Other (specify)	_____

Hearing

Formal hearing assessment done	◯ Yes ◯ No
Hearing Loss	◯ Present (bilateral) ◯ Present (unilateral) ◯ Present (CHL) ◯ Present (SNHL) ◯ Present (mixed) ◯ Absent ◯ Unassessed
Specify type and extent of hearing loss	_____

Immune

Immune deficiency present	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify Immune deficiency	_____
Frequent infections	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Frequent infections (specify type, pathogen, age, duration, intervention)	_____
Thymic hypoplasia	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Abnormal response to vaccine	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify abnormal response	_____
Documented laboratory abnormality	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify lab abnormality	_____

Autoimmune

Autoimmune disorder	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify Autoimmune disorder	_____
Atopy (e.g. eczema, asthma)	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Grave's disease	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Hypothyroidism	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Juvenile RA	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Vitiligo	◯ Present ◯ Absent ◯ Unknown/Not Assessed
ITP	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Haemolytic anaemia	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Celiac disease	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Other autoimmune findings? (specify)	_____

Malignancy

Malignancy Present?	◯ Present ◯ Absent ◯ Unassessed
Specify Malignancy	_____

Laboratory findings

Laboratory findings on file?	◯ Present ◯ Absent ◯ Unassessed
Thyroid Function Test (TFT) performed	◯ Yes ◯ No
Thyroid abnormality present	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify type and age at diagnosis	_____
Parathyroid hormone (PTH) test performed	◯ Yes ◯ No
PTH abnormality present	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify type and age at diagnosis	_____
Blood calcium test performed	◯ Yes ◯ No
Calcium abnormality	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify type and age at diagnosis	_____
Growth hormone (GH) (or IGF1/IGFBP3) testing performed	◯ Yes ◯ No
GH Abnormality	◯ Present ◯ Absent ◯ Unknown/Not Assessed
Specify type and age at diagnosis	_____
FBC performed	◯ Yes ◯ No
FBC abnormality	◯ Present ◯ Absent ◯ Unknown/Not Assessed

Specify abnormality	_____
FBC with differential done	◯ Yes ◯ No
Abnormal FBC noted?	☐ Anaemia ☐ Thrombocytopaenia ☐ Lymphopaenia ☐ CD3/CD4 Lymphopaenia ☐ Neutropaenia ☐ Other
Specify cell number, range and age at diagnosis	_____
Immunoglobulin testing performed?	◯ Present ◯ Absent ◯ Unassessed
Immunoglobulin abnormality noted	☐ Hypogammaglobulinaemia ☐ Hypergammaglobulinaemia ☐ IgA deficiency ☐ Other
Specify lab value, range and age at diagnosis	_____
Other laboratory findings (sepecify)	_____

Protocol appendix A.3: Gatekeeper approval letter

NATIONAL HEALTH LABORATORY SERVICE
School of Pathology, University of the Witwatersrand

DIVISION OF HUMAN GENETICS

Hospital Street, Johannesburg, 2001 | PO Box 1038, Johannesburg, 2000
[T] +27 11 489 9223 | [M] +27 78 080 8841 | [F] +27 11 489 9226 | [E] human.genetics@nhls.ac.za

16 April 2024

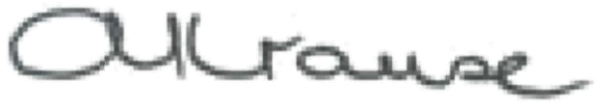
Wits Human Research Ethics Committee (Medical)
Research Office
Faculty of Health Sciences
Phillip V Tobias Building
Cnr York Road & Princess Wales Terrace
Parktown, 2193

PERMISSION TO ACCESS PATIENT DATA FILES IN THE DIVISION OF HUMAN GENETICS

Dylan McCarthy (ID 211521) has applied for Postgraduate admission as an MSc (Med) Genetic Counselling candidate in the Division of Human Genetics. His research project is attempting to characterise the genetic and phenotypic landscape of patients seen within the Division of Human genetics with CNVs in the 22q11.2 region.

I as gatekeeper, confirm Dylan may access clinical records of patients who presented to the Human Genetics Clinic between 2013 and 2023, residual diagnostic samples of patients who presented to the Human Genetics Clinic between 2013 and 2023, and residual diagnostic test records of patients who presented to the Human Genetics Clinic between 2013 and 2023.

Yours Sincerely,



Prof Amanda Krause | MBBCh, PhD
Medical Geneticist, Associate Professor
Head: Division of Human Genetics
Email: amanda.krause@nhls.ac.za
Tel: 011 489 9211/9

Appendix B

Ethics Clearance Certificate and NHLS AARMS Permission

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG

R14/49 Mr Dylan McCarthy M240614

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

CLEARANCE CERTIFICATE NO. M240614

NAME: (Principal Investigator)	Mr Dylan McCarthy
DEGREE:	MSc Genetic Counselling
STUDENT/STAFF NO:	2115219
DEPARTMENT:	School of Pathology Human Genetics NHLS
PROJECT TITLE:	Characterization of the genetic and phenotypic landscape of patients with CNVs in the chromosome 22q11.2 region in a cohort of South African patients
DATE CONSIDERED:	21/06/2024
DECISION:	Approved unconditionally
CONDITIONS:	
NOTE:	Ethics Online System ref no MED24-04-576
SUPERVISOR:	Mrs Bianca Rossouw and Dr Bronwyn Dillon
APPROVED BY:	Paul Ruff Paul Ruff (Aug 27, 2024 14:10 GMT+2) Prof P Ruff, Chair, HREC (Medical)
DATE OF APPROVAL:	29/07/2024 **EXPIRY DATE:** 29/07/2029

This clearance certificate is valid for 5 years from date of approval. Extension may be applied for.

DECLARATION OF INVESTIGATORS

I/we fully understand the conditions under which I am/we are authorized to carry out the above-mentioned research and I/we undertake to ensure compliance with these conditions. Should any departure be contemplated, from the research protocol as approved, I/we undertake to resubmit the application to Committee. **I agree to submit a yearly progress report** in July each year until study is closed. Failure to submit annual report will result in ethics clearance certificate suspension. The date for annual re-certification will be one year after the date of convened meeting where the study was initially reviewed. In this case, the study was initially reviewed in June and will therefore be due in the month of June each year. Unreported changes to the application invalidate the clearance given by the HREC (Medical). Email signed copy of this ethics clearance certificate prior to commencing with the study hrec-medical.researchoffice@wits.ac.za

DMcCarthy

Principal Investigator Signature

29-08-2024

Date

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

Academic Affairs and Research
1 Modderfontein Road, Sandringham, 2031
Tel: +27 (0)11 555 0367/0406
Email: babatyi.kgokong@nhls.ac.za
academic.research@nhls.ac.za
Web: www.nhls.ac.za

14 November 2024

Applicant: Dylan McCarthy
Institution: University of the Witwatersrand
E-mail Address: 2115219@students.wits.ac.za
Cell: 071 482 8565

Project Title: Characterisation of the genotypic and phenotypic landscape of patients with CNVs in the 22q11.2 region in a cohort of South African patients.

Reference Number: PR2455081

Research Application Type(s):
1. Request for Data

RE: APPROVAL LETTER: REQUEST TO ACCESS NHLS RESOURCES FOR RESEARCH PURPOSES

This letter serves to advise that the application requesting permission to conduct the above-mentioned research using the listed NHLS resources has been reviewed and **"Approved"**. Please note that the approval is granted **without undergoing the full internal peer review process** on the condition of the **urgency of the request and its time-sensitive nature**, **therefore further clarity may be required by the processing unit.** You are required to comply with the NHLS Research Material and Data Access Policy and requirements stated below.

1. All material and data requested shall be used as per the research protocol submitted to the NHLS and as approved by the relevant Health Research Ethics Committee (HREC) in South Africa.
2. Access to the NHLS material and/or data shall be limited to the minimum required for successful completion of the approved study and shall be made available ***without patient names and other patient identifiers (including, but not limited to, national identity numbers, hospital/clinic file numbers, addresses and telephone numbers)***.
3. Confidentiality shall be maintained at the participant and institutional level and there shall be no disclosure of personal information or confidential information.
4. Data and/or material shall not be shared with other parties unless approved by the NHLS
5. The material and/or data obtained from the NHLS shall be anonymised and not, for any reason, be used to track or recruit patients as no pre-approval/consent is obtained from patients.
6. Processes shall be discussed with the relevant NHLS departments (i.e. Corporate Data Warehouse (CDW), NHLS Laboratory Management, Operations Office, etc.) and agreed upon.
7. Any amendments to the study requirements, including the use of the material and/or data for purposes not initially disclosed to the NHLS) shall be cleared by an approved HREC and submitted to the NHLS for approval via the AARMS system – https://aarms.nhls.ac.za.
8. The NHLS shall be acknowledged as a source of material and/or data in any output, such as abstracts and journal articles, emanating from the project.
9. A final report of the research study and any published output resulting from this study shall be submitted to the NHLS via AARMS

Please note that this letter constitutes approval by the NHLS Academic Affairs and Research Office. The NHLS entities tasked with providing the material and/data may have additional requirements for access. Data related queries may be directed to NHLS CDW, email: zarina.sabat@nhls.ac.za; contact number: 011 386 6074 and sample related queries (if applicable) shall be directed to the relevant business manager.

Dr Babatyi Malope-Kgokong

National Manager: Academic Affairs and Research

Chairperson: Prof Jeffrey Mphahlele CEO: Prof Koleka Mlisana
Physical Address: 1 Modderfontein Road, Sandringham, Johannesburg, South Africa Postal Address: Private Bag X8, Sandringham, 2131, South Africa
Tel: +27 (0) 11 386 6000/ 0860 00 NHLS(6457) www.nhls.ac.za
Practice number: 5200296

Appendix C

Plagiarism Report

PLAGIARISM DECLARATION TO BE SIGNED BY ALL HIGHER DEGREE STUDENTS

SENATE PLAGIARISM POLICY: APPENDIX ONE

I Dylan McCarthy (Student number: 2115219) am a student registered for the degree of MSc(Med): Genetic Counselling in the academic year 2024.

I hereby declare the following:

- I am aware that plagiarism (the use of someone else's work without their permission and/or without acknowledging the original source) is wrong.
- I confirm that the work submitted for assessment for the above degree is my own unaided work except where I have explicitly indicated otherwise.
- I have followed the required conventions in referencing the thoughts and ideas of others.
- I understand that the University of the Witwatersrand may take disciplinary action against me if there is a belief that this is not my own unaided work or that I have failed to acknowledge the source of the ideas or words in my writing.
- I have included as an appendix a report from "Turnitin" (or other approved plagiarism detection) software indicating the level of plagiarism in my research document.

Signature: DMcCarthy Date: 14-01-2025

DM research report for plaigerism checker.pdf

ORIGINALITY REPORT

15%	12%	11%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.coursehero.com	Internet Source
2	Submitted to University of Witwatersrand	Student Paper
3	www.mdpi.com	Internet Source
4	onlinelibrary.wiley.com	Internet Source
5	wiredspace.wits.ac.za	Internet Source
6	www.omicsdi.org	Internet Source
7	Submitted to University of New Mexico	Student Paper
8	Sooknanan, Rachna. "Assessing the Diagnostic Utility of Fluorescent in Situ Hybridisation (Fish) as Follow-up Testing for	

Postnatal Microarray Re

sults.", University of Johannesburg (South Africa), 2024 Publication	2%
	1%
	1%
	1%
	<1%
	<1%
	<1%
	<1%

9	www.researchsquare.com Internet Source	< 1 %
10	biosciencetrends.com Internet Source	< 1 %
11	Zodanu, Gloria Kafui Esi. "Genetic Examination of Congenital Heart Defects and Non-Syndromic Intellectual Disability", Szegedi Tudomanyegyetem (Hungary), 2023 Publication	< 1 %
12	www.ashg.org Internet Source	< 1 %
	www.repository.up.ac.za Internet Source	< 1 %
14	pure.rug.nl Internet Source	< 1 %
15	www.researchgate.net Internet Source	< 1 %
16	Submitted to Turun yliopisto Student Paper	< 1 %
17	mdpi-res.com Internet Source	< 1 %
18	www.nature.com Internet Source	< 1 %
19	repository.escholarship.umassmed.edu Internet Source	

		< 1 %
20	spiral.imperial.ac.uk Internet Source	< 1 %
21	www.ncbi.nlm.nih.gov Internet Source	< 1 %
22	Lena Wallin, Christopher Gillberg, Elisabeth Fernell, Carina Gillberg, Eva Billstedt. "Neurodevelopmental and other psychiatric disorders in 22q11.2 deletion syndrome from childhood to adult age: Prospective longitudinal study of 100 individuals", American Journal of Medical Genetics Part C: Seminars in Medical Genetics, 2023 Publication	< 1 %
23	academic.oup.com Internet Source	< 1 %
24	archive-ouverte.unige.ch Internet Source	< 1 %
	www.orthopeden.org Internet Source	< 1 %
26	Kevin Antshel. "22q11.2 Deletion Syndrome: Genetics, Neuroanatomy and Cognitive/Behavioral Features Keywords", Child Neuropsychology, 2/1/2005 Publication	< 1 %

27	Sun-Chong Wang, Meng Hua Chen. "Herbal Medicine for Rare Diseases – Alleviating Symptoms by GMP Herbal Formulations", CRC Press, 2025 Publication	< 1 %
28	globalheartjournal.com Internet Source	< 1 %
29	Submitted to American Sentinel University Student Paper	< 1 %
30	Anne S. Bassett, Chelsea Lowther, Daniele Merico, Gregory Costain et al. "Rare Genome-Wide Copy Number Variation and Expression of Schizophrenia in 22q11.2 Deletion Syndrome", American Journal of Psychiatry, 2017 Publication	< 1 %
31	Anne S. Bassett, Eva W.C. Chow, Janice Husted, Rosanna Weksberg, Oana Caluseriu, Gary D. Webb, Michael A. Gatzoulis. "Clinical features of 78 adults with 22q11 deletion syndrome", American Journal of Medical Genetics Part A, 2005 Publication	< 1 %
32	James E. Maddux, Barbara A. Winstead. "Psychopathology – Foundations for a Contemporary Understanding", Routledge, 2024 Publication	< 1 %

33	Nieuwenhuis, M.H.. "Correlations between mutation site in APC and phenotype of familial adenomatous polyposis (FAP): A review of the literature", Critical Reviews in Oncology and Hematology, 200702 Publication	<1%
34	backend.orbit.dtu.dk Internet Source	<1%
35	orca.cf.ac.uk Internet Source	<1%
36	Elizabeth Goldmuntz. "The 22q11.2 Deletion Syndrome", S. Karger AG, 2015 Publication	<1%
	Ryan J. Huang, Sherri L. Smith, Libor Brezina, Kristal M. Riska. "A Comparison of Falls and Dizziness Handicap by Vestibular Diagnosis", American Journal of Audiology, 2021 Publication	<1%
38	Zhang, Xiaoqing, Yuejuan Xu, Deyuan Liu, Juan Geng, Sun Chen, Zhengwen Jiang, Qihua Fu, and Kun Sun. "A modified multiplex ligation-dependent probe amplification method for the detection of 22q11.2 copy number variations in patients with congenital heart disease", BMC Genomics, 2015. Publication	<1%
	www.frontiersin.org	

39	Internet Source	< 1 %
40	Irene E. Drmic, Bonnie MacKinnon Modi, Beth McConnell, Sanne Jilderda et al. "Neurodevelopmental functioning in probands and non-proband carriers of 22q11.2 microduplication", American Journal of Medical Genetics Part A, 2022 Publication	< 1 %
41	storage.googleapis.com Internet Source	< 1 %
42	www.legco.gov.hk Internet Source	< 1 %
43	habib.edu.pk Internet Source	< 1 %
44	healthdocbox.com Internet Source	< 1 %
45	pdffox.com Internet Source	< 1 %
46	unsworks.unsw.edu.au Internet Source	< 1 %
47	www.haematologica.org Internet Source	< 1 %
48	www.thieme-connect.de Internet Source	< 1 %

	Judy Bothamley, Maureen Boyle. "Medical Conditions Affecting Pregnancy and Childbirth", Routledge, 2020 Publication	< 1 %
50	Paul C. Helm, Marc-André Koerten, Hashim Abdul-Khaliq, Helmut Baumgartner, Deniz Kececioglu, Ulrike M. M. Bauer. "Representativeness of the German National Register for Congenital Heart Defects: a clinically oriented analysis", Cardiology in the Young, 2015 Publication	< 1 %
51	Reymundo Lozano, Catherine Gbekie, Paige M. Siper, Shubhika Srivastava et al. "FOXP1 syndrome: a review of the literature and practice parameters for medical assessment and monitoring", Journal of Neurodevelopmental Disorders, 2021 Publication	< 1 %
52	assets-eu.researchsquare.com Internet Source	< 1 %
53	dergipark.org.tr Internet Source	< 1 %
54	etd.uwc.ac.za Internet Source	< 1 %
55	orca.cardiff.ac.uk Internet Source	< 1 %

56	www.hindawi.com Internet Source	<1%
57	www.oncotarget.com Internet Source	<1%
58	"Cassidy and Allanson's Management of Genetic Syndromes", Wiley, 2021 Publication	<1%
59	Eve S. Kortanek, Nicole M. McDonald, Erin E. Nosco, Gabrielle A. MacNaughton, Amy Lin, Shafali S. Jeste, Carrie E. Bearden. "Early developmental concerns in 22q11.2 deletion and duplication carriers", Research in Autism Spectrum Disorders, 2022 Publication	<1%
60	M. Zamariolli, A. G. Dantas, N. Nunes, M. Moysés-Oliveira et al. " Phenotypic heterogeneity in 22q11.2 deletion syndrome: as genetic modifiers for congenital heart disease in a Brazilian cohort ", American Journal of Medical Genetics Part A, 2023 Publication	<1%
	Margaret C. Souders, Donna M. McDonald-McGinn, Kosha Ruparel, Tyler M. Moore et al. "Sleep difficulties related to psychopathology and neurocognition in people with 22q11.2 Deletion Syndrome", Psychiatry Research, 2024 Publication	<1%

62	Mary Ella Pierpont, Martina Brueckner, Wendy K. Chung, Vidu Garg et al. "Genetic Basis for Congenital Heart Disease: Revisited: A Scientific Statement From the American Heart Association", Circulation, 2018 Publication	< 1 %
63	Raissa Teteli, Annette Uwineza, Yvan Butera, Janvier Hitayezu et al. "Pattern of congenital heart diseases in Rwandan children with genetic defects", Pan African Medical Journal, 2014 Publication	< 1 %
64	core-cms.prod.aop.cambridge.org Internet Source	< 1 %
65	iris.uniroma1.it Internet Source	< 1 %
66	kclpure.kcl.ac.uk Internet Source	< 1 %
67	molecularautism.biomedcentral.com Internet Source	< 1 %
68	qmro.qmul.ac.uk Internet Source	< 1 %
69	repositorio.unesp.br Internet Source	< 1 %
70	repositorio.unifesp.br Internet Source	

<
1
%

71	v-des-dev-lnx1.nwu.ac.za Internet Source	<1%
72	3tonderzoek.sites.uu.nl Internet Source	<1%
	Aurelie Merlo, Audrey Khoury, Mohsin Shah, Tabitha Linville, John Ikonomidis, Matthew Cavender, John Vavalle, Thomas Caranasos. "Temporal Trends in Internal vs. External Referrals for TAVR in a Large Academic Center: Patients Characteristics and Outcomes", Journal of Interventional Cardiology, 2022 Publication	<1%
74	Daniel Bachiller. "Mutations in TBX1 genocopy the 22q11.2 deletion and duplication syndromes: a new susceptibility factor for mental retardation", European Journal of Human Genetics, 06/2007 Publication	<1%
75	Diego A. Forero. "Bioinformatics and Human Genomics Research", Routledge, 2021 Publication	<1%
76	Erik Boot, Sólveig Óskarsdóttir, Joanne C.Y. Loo, Terrence Blaine Crowley et al. "Updated clinical practice recommendations for managing adults with 22q11.2 deletion syndrome", Genetics in Medicine, 2023	<1%

Publication

77 Felix-Julian Campos-Garcia, Addy-Manuela Castillo-Espinola, Carolina-Elizabeth Medina-Escobedo, Juan Zenteno et al. "Screening Method for 22q11 Deletion Syndrome Involving the Use of TaqMan qPCR for TBX1 in Patients with Conotruncal Congenital Heart Disease", Cardiogenetics, 2022 <1%
Publication

78 H. Heinrich, A. G. H. Pijpers, I. H. Linskens, E. van Leeuwen, L. D. Eeftinck Schattenkerk, J. P. M. Derikx, E. Pajkrt. "Congenital small bowel obstruction: Prenatal detection and outcome", Prenatal Diagnosis, 2023 <1%
Publication

79 Harold Chen. "Atlas of Genetic Diagnosis and Counseling", Springer Science and Business Media LLC, 2017 <1%
Publication

80 Kaitlin J. Stanley, Kelsey J. Kalbfleisch, Olivia M. Moran, Rajiv R. Chaturvedi et al. "Expanding the phenotypic spectrum of NOTCH1 variants: clinical manifestations in families with congenital heart disease", European Journal of Human Genetics, 2024 <1%
Publication

81 Katarzyna Wiater, Tomasz Świtaj, Jacek Mackiewicz, Ewa Kalinka-Warzocha et al. <1%

"Efficacy and safety of ipilimumab therapy in patients with metastatic melanoma: a retrospective multicenter analysis", Współczesna Onkologia, 2013
Publication

---

82 Kathleen P. O'Hora, Carolyn M. Amir, Emily Chiem, Charles H. Schleifer et al. "Differential Inflammatory Profiles in Carriers of Reciprocal 22q11.2 Copy Number Variants", Psychoneuroendocrinology, 2024 <1%
Publication

---

83 Marta Smyk, Maciej Geremek, Kamila Ziemkiewicz, Tomasz Gambin et al. "Coexisting Conditions Modifying Phenotypes of Patients with 22q11.2 Deletion Syndrome", Genes, 2023 <1%
Publication

---

84 Samira Spineli-Silva, Luciana M. Bispo, Vera L. Gil-da-Silva-Lopes, Társis P. Vieira. "Distal deletion at 22q11.2 as differential diagnosis in Craniofacial Microsomia: Case report and literature review", European Journal of Medical Genetics, 2018 <1%
Publication

Somaya Salah, Hiba Jaber, Ayala Frumkin, Tamar Harel. " Homozygous 22q11.2 distal type microdeletion is associated with syndromic neurodevelopmental delay ", <1%

	American Journal of Medical Genetics Part A, 2023 Publication	
86	Yan Wang, Min Zhang, Lingji Chen, Hailong Huang, Liangpu Xu. "Prenatal diagnosis of BACs-on-Beads assay in 1520 cases from Fujian Province, China", Molecular Genetics & Genomic Medicine, 2020 Publication	< 1 %
87	cds.ismrm.org Internet Source	< 1 %
88	escholarship.org Internet Source	< 1 %
89	mohcsr.gov.om Internet Source	< 1 %
90	ojrd.biomedcentral.com Internet Source	< 1 %
91	repository.uel.ac.uk Internet Source	< 1 %
92	researchonline.lshtm.ac.uk Internet Source	< 1 %
93	warm.dovepress.com Internet Source	< 1 %
94	www.journalofhospitalmedicine.com Internet Source	< 1 %

95	www.medrxiv.org Internet Source	<1%
96	www.science.gov Internet Source	<1%
	Henrique Garcia Silveira, Carlos Eduardo Steiner, Giovana Toccoli, Luise Longo Angeloni et al. "Variants in KMT2A in Three Individuals with Previous Suspicion of 22q11.2 Deletion Syndrome", Genes, 2024 Publication	<1%
98	Malú Zamariolli, Chiara Auwerx, Marie C. Sadler, Adriaan van der Graaf et al. "The impact of 22q11.2 copy-number variants on human traits in the general population", The American Journal of Human Genetics, 2023 Publication	<1%
99	Rakonjac, Marijana S.. "Analiza Govorno-jezičkog, Socio-Emocionalnog i Kognitivnog Razvoja Bolesnika sa Heterozigotnom Mikrodelecijom Regiona Q11.2 na Hromozomu 22", University of Belgrade (Serbia), 2024 Publication	<1%
100	Ward, Holly. "An Exploration of the Experiences of Children and Young People With 22q11.2 Deletion Syndrome.",	<1%

Canterbury Christ Church University (United Kingdom), 2024
Publication

101 Enea Rezzonico, Sofiane Lariani, Caroline Barretto, Gabriella Cuanoud et al. "Global transcriptome analysis of the heat shock response of Bifidobacterium longum", FEMS Microbiology Letters, 2007 <1%
Publication

Exclude quotes	On	Exclude matches	Off
Exclude bibliography	On		

Appendix D

Molecular Genetics and Genomic Medicine Author Guidelines

1. SUBMISSION

Authors should kindly note that submission implies that the content has not been published *or* submitted for publication elsewhere except as a brief abstract in the proceedings of a scientific meeting or symposium.

Once the submission materials have been prepared in accordance with the Author Guidelines, new submissions should be made via the Research Exchange submission portal. ou may check the status of your submission at any time by logging on to submission.wiley.com and clicking the “My Submissions” button. For technical help with the submission system, please review our FAQs or contact submissionhelp@wiley.com.

The submission system will prompt the author to use an ORCiD ID (a unique author identifier) to help distinguish their work from that of other researchers. Click here to find out more.

Data Protection

By submitting a manuscript to or reviewing for this publication, your name, email address, and affiliation, and other contact details the publication might require, will be used for the regular operations of the publication, including, when necessary, sharing with the publisher (Wiley) and partners for production and publication. The publication and the publisher recognize the importance of protecting the personal information collected from users in the operation of these services, and have practices in place to ensure that steps are taken to maintain the security, integrity, and privacy of the personal data collected and processed. You can learn more at authorservices.wiley.com/statements/data-protection-policy.

For help with submissions, please contact the Editorial Office: mggm@wiley.com. When necessary, the Editorial Office staff may refer questions to the Editor-in-Chief.

Return to Guideline Sections

2. AIMS AND SCOPE

Molecular Genetics & Genomic Medicine is a peer-reviewed journal for rapid dissemination of quality research related to the dynamically developing areas of human, molecular and medical genetics. The journal publishes original research articles covering findings in phenotypic, molecular, biological, and genomic aspects of genomic variation, inherited disorders and birth defects. The broad publishing spectrum of *Molecular Genetics & Genomic Medicine* includes rare and common disorders from diagnosis to treatment. Examples of appropriate articles include reports of novel disease genes, functional studies of genetic variants, in-depth genotype-phenotype studies, genomic analysis of inherited disorders and cancer, molecular diagnostic methods, medical bioinformatics, ethical, legal, and social implications (ELSI), epidemiological studies, and approaches to clinical diagnosis. *Molecular Genetics & Genomic Medicine* provides a scientific home for next generation sequencing studies of rare and common disorders, which will make research in this fascinating area easily and rapidly accessible to the scientific community. This will serve as the basis for translating next generation sequencing studies into individualized diagnostics and therapeutics, for day-to-day medical care.

Molecular Genetics & Genomic Medicine publishes original research articles, reviews, and research methods papers, along with invited editorials and commentaries. Original research papers must report well-conducted research with conclusions supported by the data presented. Simple, purely descriptive, single-patient Case Reports are out of scope for this journal. For such Case Reports we recommend considering submission to *Case Reports in Genetics*.

Molecular Genetics & Genomic Medicine is a Wiley Open Access journal, one of a series of peer reviewed titles publishing quality research with speed and efficiency. Authors of accepted papers pay an Article Publication Charge and their papers are published under a Creative Commons license. With Creative Commons licenses, the author retains copyright and the public is allowed to reuse the content. The author grants Wiley a license to publish the article and identify as the original publisher. To find out which Created Commons Licenses are available for the journal, click here. For further information visit the Wiley Open Access website. For more information on this journal's APCs, please see the Open Access page.

Return to Guideline Sections

3. MANUSCRIPT CATEGORIES AND REQUIREMENTS

MANUSCRIPT TYPES

- Original Research Articles
- Clinical Reports
- Reviews
- Methods
- Invited Commentaries

GENERAL INSTRUCTIONS

Manuscripts must be submitted in grammatically correct English. Manuscripts that do not meet this standard cannot be reviewed. Authors for whom English is a second language may wish to consult an English-speaking colleague or consider having their manuscript professionally edited before submission to improve the English. A list of independent suppliers of editing services can be found at http://authorservices.wiley.com/bauthor/english_language.asp. All services are paid for and arranged by the author, and use of one of these services does not guarantee acceptance or preference for publication. A manuscript is considered for review and possible publication on the condition that it is submitted solely to *Molecular Genetics and Genomic Medicine*, and that the manuscript or a substantial portion of it is not under consideration elsewhere.

Novel nucleotide sequence data including genetic mutations must be submitted to a public database prior to publication. See the "Sequence Data" section, below.

MANUSCRIPT TERMINOLOGY

Manuscript wording and terminology should reflect the Journal's preferred criteria in describing and referring to human beings. Individuals described within a manuscript should be regarded with sensitivity. They should be referred to as patients rather than cases or as having a condition rather than being simply labeled by a specific terminology. Consider and

avoid any stigmatizing terms, such as simian. If it is necessary to identify an individual, use a numerical designation (e.g. Patient 1) rather than using any other identifying notations such as initials. Phenotypic descriptors should be standardized as much as possible and follow these guidelines from the Elements of Morphology
project: http://onlinelibrary.wiley.com/doi/10.1002/ajmg.a.v149a:1/issuetoc

ETHICAL COMPLIANCE

Please include a statement confirming that your study was approved by an ethics committee as the first sentence of your Methods section, under the subheading, “Editorial Policies and Ethical Considerations”.

INFORMED CONSENT

Molecular Genetics & Genomic Medicine requires that all appropriate steps be taken in obtaining informed consent of any and all human and/or experimental animal subjects participating in the research comprising the manuscript submitted for review and possible publication, and a statement to this effect must be included in the Methods section of the manuscript, under the subheading, “Editorial Policies and Ethical Considerations”. Identifying information should not be included in the manuscript unless the information is essential for scientific purposes and the study participants or patients (or parents or guardians) give written informed consent for publication. *See also in section 5 under "Human Studies and Subjects*". In addition, if your submission includes research on a specific ethnic population(s) or potentially vulnerable populations such as an ethnic or religious minorities, we require additional information on the steps taken by the authors to ensure that free and informed consent was given by the subjects/patients of the research study. Information on these additional steps can be found here.

PATIENT PHOTOGRAPHS

Molecular Genetics & Genomic Medicine strongly prefers to publish unmasked patient photos. We encourage all prospective authors to work with families prior to submission to address the issue of permission for review and possible publication of patient images. If your submission contains ANY identifiable patient images or other protected health information, you MUST provide documented permission from the patient (or the patient’s parent, guardian, or legal representative) before the specific material will be circulated among *Molecular Genetics & Genomic Medicine* editors, reviewers and staff for the purpose of possible publication in MGGM. The documented permission may be supplied as supplemental material uploaded with the submission. NOTE: If the consent documentation is in a language other than English, please also provide the best possible English translation of the document as a separate file, clearly marked as a translation.

While the manuscript will be processed upon submission, anything considered protected health information will be restricted from access prior to the receipt of documented permission. We caution you that the absence of material or cited figures may adversely impact the manuscript in the review process. The submission of masked photos without sufficient deidentification is strongly discouraged (i.e., facial photographs with only small dark geometric shapes over the eyes are insufficient). *See also in section 5 under "Human Studies and Subjects*".

DISCLOSURE STATEMENT

Molecular Genetics & Genomic Medicine requires that all authors disclose any potential sources of conflict of interest. Any interest or relationship, financial or otherwise, that might

be perceived as influencing an author's objectivity is considered a potential source of conflict of interest. These must be disclosed when directly relevant or directly related to the work that the authors describe in their manuscript. Potential sources of conflict of interest include, but are not limited to, patent or stock ownership, membership of a company board of directors, membership of an advisory board or committee for a company, and consultancy for or receipt of speaker's fees from a company. The existence of a conflict of interest does not preclude publication in this journal.

If the authors have no conflict of interest to declare, they must also state this at submission. It is the responsibility of the corresponding author to review this policy with all authors and collectively to list on the cover letter to the Editor-in-Chief, in the manuscript (under the Acknowledgements section), and in the online submission system ALL pertinent commercial and other relationships.

The above policies are in accordance with the Uniform Requirements for Manuscripts Submitted to Biomedical Journals produced by the International Committee of Medical Journal Editors (http://www.icmje.org/). See also the Editorial Policies and Ethical Considerations section, below.

Return to Guideline Sections

4. PREPARING THE SUBMISSION

Free Format Submission

Molecular Genetics & Genomic Medicine now offers Free Format submission for a simplified and streamlined submission process.

Before you submit, you will need:

- Your manuscript: this should be an editable file including text, figures, and tables, or separate files – whichever you prefer. All required sections should be contained in your manuscript, including abstract (which does need to be correctly styled), introduction, methods, results, and conclusions. Figures and tables should have legends. Figures should be uploaded in the highest resolution possible. If the figures are not of sufficiently high quality your manuscript may be delayed. We also encourage you to include your figures within the main document to make it easier for editors and reviewers to read your manuscript. References may be submitted in any style or format, as long as it is consistent throughout the manuscript. Supporting information should be submitted in separate files. If the manuscript, figures or tables are difficult for you to read, they will also be difficult for the editors and reviewers, and the editorial office will send it back to you for revision. Your manuscript may also be sent back to you for revision if the quality of English language is poor.
- The title page of the manuscript, including:
    - Your co-author details, including affiliation and email address. *(Why is this important? We need to keep all co-authors informed of the outcome of the peer review process.)*
    - Statements relating to our ethics and integrity policies, which may include any of the following *(Why are these important? We need to uphold rigorous ethical standards for the research we consider for publication):*
        - data availability statement

- funding statement
- conflict of interest disclosure
- ethics approval statement
- patient consent statement
- permission to reproduce material from other sources
- clinical trial registration

To submit, login at http://mc.manuscriptcentral.com/MGGM and create a new submission. Follow the submission steps as required and submit the manuscript.

Parts of the Manuscript

The manuscript should be submitted in separate files: main text file; figures.

Main Text File

Manuscripts can be uploaded either as a single document (containing the main text, tables and figures), or with figures and tables provided as separate files. Should your manuscript reach revision stage, figures and tables must be provided as separate files. The main manuscript file can be submitted in Microsoft Word (.doc or .docx) format.

Your main document file should include:

- A short informative title containing the major key words. The title should not contain abbreviations;
- The full names of the authors with institutional affiliations where the work was conducted, with a footnote for the author's present address if different from where the work was conducted;
- Acknowledgments;
- Abstract structured (intro/methods/results/conclusion);
- Up to seven keywords;
- Main body: formatted as introduction, materials & methods, results, discussion, conclusion;
- References;
- Tables (each table complete with title and footnotes);
- Figure legends: Legends should be supplied as a complete list in the text. Figures should be uploaded as separate files.

For the benefit of the reviewers and editors, please include line numbers in your main text files. To add line numbers using Microsoft Word, click on Page Layout and from there, go to Page Setup. In Page Setup, click on Line Numbers and then select "Restart Each Page".

Figures and supplementary/supporting information should be supplied as separate files (see below under "Additional Files"). Figures must be clearly labeled.

Authorship

Please refer to the journal's Authorship policy in the Editorial Policies and Ethical Considerations section for details on author listing eligibility.

Acknowledgements

Contributions from anyone who does not meet the criteria for authorship should be listed, with permission from the contributor, in an Acknowledgments section. Financial and material support should also be mentioned. Thanks to anonymous reviewers are not appropriate.

Conflict of Interest Statement

Authors will be asked to provide a conflict of interest statement during the submission process. For details on what to include in this section, see the Conflict of Interest section in the Editorial Policies and Ethical Considerations section below. Submitting authors should ensure they liaise with all co-authors to confirm agreement with the final statement.

Abstract

Please provide a structured abstract of 200 words containing the major keywords summarizing the article. Use the subheadings Background, Methods, Results, Conclusion.

Keywords

Please provide three to five keywords.

Main Text

The journal uses US spelling; however, authors may submit using either option, as spelling of accepted papers is converted during the production process.

References

The accuracy of references is the responsibility of the authors. Only published papers and those in press may be included in the reference list. Unpublished data and submitted manuscripts must be cited parenthetically within the text. Personal communications should also be cited within the text; permission in writing from the communicator is required.

This journal recommends that references be prepared according to the *Publication Manual of the American Psychological Association* (7th edition). The APA website includes a range of resources for authors learning to write in APA style, including an overview of the manual, free tutorials on APA Style basics, and an APA Style Blog. For more information about APA referencing style, please also refer to the APA FAQ. Please note that the editorial office will update the formatting into journal style when the manuscript is accepted for publication.

EndNote users can download the style here.

According to APA style, in text citations should follow the author-date method whereby the author's last name and the year of publication for the source should appear in the text, for example, (Jones, 1998). The complete reference list should appear alphabetically by name at the end of the paper.

Authors should note that the APA referencing style requires that a Digital Object Identifier (DOI) be provided for all references where available. Also, for journal articles, issue numbers are not included unless each issue in the volume begins with page one.

Reference examples follow:

Journal article

Beers, S. R., & De Bellis, M. D. (2002). Neuropsychological function in children with maltreatment-related posttraumatic stress disorder. *The American Journal of Psychiatry, 159,* 483–486. *doi:*10.1176/appi.ajp.159.3.483

Book

Bradley-Johnson, S. (1994). *Psychoeducational assessment of students who are visually impaired or blind: Infancy through high school* (2nd ed.). Austin, TX: Pro-ed.

Internet Document

Norton, R. (2006, November 4). How to train a cat to operate a light switch [Video file]. Retrieved from http://www.youtube.com/watch?v=Vja83KLQXZs

Footnotes

Footnotes should be placed as a list at the end of the paper only, not at the foot of each page. They should be kept to a minimum. Keep footnotes brief; they should contain only short comments tangential to the main argument of the paper and should not include references. They should be numbered in the list and referred to in the text with consecutive, superscript Arabic numerals.

Tables

Tables should be self-contained and complement, not duplicate, information contained in the text. They should be supplied as editable files, not pasted as images. Legends should be concise but comprehensive – the table, legend, and footnotes must be understandable without reference to the text. All abbreviations must be defined in footnotes. Footnote symbols: †, ‡, §, ¶, should be used (in that order) and *, **, *** should be reserved for P-values. Statistical measures such as SD or SEM should be identified in the headings.

Figure Legends

Legends should be concise but comprehensive – the figure and its legend must be understandable without reference to the text. Include definitions of any symbols used and define/explain all abbreviations and units of measurement.

Figures

Although authors are encouraged to send the highest quality figures possible, for peer-review purposes, a wide variety of formats, sizes, and resolutions are accepted. Click here for the basic figure requirements for figures submitted with manuscripts for initial peer review, as well as the more detailed post- acceptance figure requirements.

Figures submitted in color will be published in color free of charge. Please note, however, that it is preferable that line figures (e.g. graphs and charts) are supplied in black and white so that they are legible if printed by a reader in black and white.

Cover Submission GuidelinesIf you would like to submit an image from or related to your accepted manuscript for consideration of the cover of a journal issue, please follow these guidelines.

Additional Files

Appendices

Appendices will be published after the references. For submission they should be supplied as separate files but referred to in the text.

Supporting Information

Supporting information is information that is not essential to the article, but provides greater depth and background. It is hosted online and appears without editing or typesetting. It may include tables, figures, videos, datasets, etc.

Click here for Wiley's FAQs on Supporting Information.

Note: if data, scripts, or other artefacts used to generate the analyses presented in the paper are available via a publicly available data repository, authors should include a reference to the location of the material within their paper.

General Style Points

The following points provide general advice on formatting and style.

- Abbreviations: In general, terms should not be abbreviated unless they are used repeatedly and the abbreviation is helpful to the reader. Initially, use the word in full, followed by the abbreviation in parentheses. Thereafter use the abbreviation only.
- Units of measurement: Measurements should be given in SI or SI-derived units. Visit the Bureau International des Poids et Mesures (BIPM) website for more information about SI units.
- Numbers: numbers under 10 should be spelt out, except for: measurements with a unit (8 mmol/L); age (6 weeks old), or lists with other numbers (11 dogs, 9 cats, 4 gerbils).
- Trade Names: Chemical substances should be referred to by the generic name only. Trade names should not be used. Drugs should be referred to by their generic names. If proprietary drugs have been used in the study, refer to these by their generic name, mentioning the proprietary name and the name and location of the manufacturer in parentheses.

Wiley Author Resources

Manuscript Preparation Tips: Wiley has a range of resources for authors preparing manuscripts for submission available here. In particular, authors may benefit from referring to Wiley's best practice tips on Writing for Search Engine Optimization.

Editing, Translation, and Formatting Support: Wiley Editing Services offers expert help with English Language Editing, as well as translation, manuscript formatting, figure illustration, figure formatting, and graphical abstract design – so you can submit your manuscript with confidence.
Also, check out our resources for Preparing Your Article for general guidance about writing and preparing your manuscript.

Return to Guideline Sections

5. EDITORIAL POLICIES AND ETHICAL CONSIDERATIONS

Peer Review and Acceptance

The acceptance criteria for all papers are the quality and originality of the research and its significance to journal readership. Papers will only be sent to review if the Editors determine that the paper meets the appropriate quality and relevance requirements.

Except where otherwise stated, manuscripts are single-blind peer reviewed. Wiley's policy on the confidentiality of the review process is available here.

Refer and Transfer Program

Wiley believes that no valuable research should go unshared. This journal participates in Wiley’s Refer & Transfer program. If your manuscript is not accepted, you may receive a recommendation to transfer your manuscript to another suitable Wiley journal, either through a referral from the journal’s editor or through our Transfer Desk Assistant.

Data Sharing and Data Accessibility

Molecular Genetics & Genomic Medicine encourages data sharing wherever possible, unless this is prevented by ethical, privacy, or confidentiality matters. Authors publishing in the journal are therefore encouraged to make their data, scripts, and other artefacts used to generate the analyses presented in the paper available via a publicly available data repository; however, this is not mandatory. If the study includes original data, at least one author must confirm that he or she had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Human Studies and Subjects

For manuscripts reporting medical studies that involve human participants, a statement identifying the ethics committee that approved the study and confirmation that the study conforms to recognized standards is required, for example: Declaration of Helsinki; US Federal Policy for the Protection of Human Subjects; or European Medicines Agency Guidelines for Good Clinical Practice US Federal Policy for the Protection of Human Subjects;. It should also state clearly in the text that all persons gave their informed consent prior to their inclusion in the study.

Patient anonymity should be preserved. Photographs need to be cropped sufficiently to prevent human subjects being recognized (an eye bar must not be used because of insufficient de- identification). Images and information from individual participants will only be published where the authors have obtained the individual's free prior informed consent. Authors do not need to provide a copy of the consent form to the publisher; however, in signing the author license to publish, authors are required to confirm that consent has been obtained. Wiley has a standard patient consent form available for use. NOTE: If the consent documentation is in a language other than English, please also provide the best possible English translation of the document as a separate file, clearly marked as a translation.

Human Studies Involving Ethnic Populations

If your submission includes research on a specific ethnic population(s) or potentially vulnerable populations such as an ethnic or religious minorities, we require additional information that can be found here.

Animal Studies

A statement indicating that the protocol and procedures employed were ethically reviewed and approved, as well as the name of the body giving approval, must be included in the

Methods section of the manuscript. Authors are encouraged to adhere to animal research reporting standards, for example the ARRIVE guidelines for reporting study design and statistical analysis; experimental procedures; experimental animals and housing and husbandry. Authors should also state whether experiments were performed in accordance with relevant institutional and national guidelines for the care and use of laboratory animals:

- US authors should cite compliance with the US National Research Council's Guide for the Care and Use of Laboratory Animals, the US Public Health Service's Policy on Humane Care and Use of Laboratory Animals, and Guide for the Care and Use of Laboratory Animals.
- UK authors should conform to UK legislation under the Animals (Scientific Procedures) Act 1986 Amendment Regulations (SI 2012/3039).
- European authors outside the UK should conform to Directive 2010/63/EU.

Clinical Trial Registration

The journal requires that clinical trials are prospectively registered in a publicly accessible database and clinical trial registration numbers are included in all papers that report their results. Authors are asked to include the name of the trial register and the clinical trial registration number at the end of the Abstract. If the trial is not registered, or was registered retrospectively, the reasons for this should be explained.

Research Reporting Guidelines

Accurate and complete reporting enables readers to fully appraise research, replicate it, and use it. Authors are encouraged to adhere to recognized research reporting standards. The EQUATOR Network collects more than 370 reporting guidelines for many study types, including for:

- Randomized trials: CONSORT
- Observational studies: STROBE
- Systematic reviews: PRISMA
- Case reports: CARE
- Qualit ative research: SRQR
- Diagnostic / prognostic studies: STARD
- Quality improvement studies: SQUIRE
- Animal pre-clinical studies: ARRIVE
- Study protocols: SPIRIT
- Clini cal practice guidelines: AGREE

We also encourage authors to refer to and follow guidelines from:

- Future of Research Communications and e-Scholarship (FORCE11)
- National Research Council's Institute for Laboratory Animal Research guidelines
- The Gold Standard Publication Checklist from Hooijmans and colleagues

- Minimum Information Guidelines from Diverse Bioscience Communities (MIBBI) website
- FAIRsharing website

Species Names

Upon its first use in the title, abstract, and text, the common name of a species should be followed by the scientific name (genus, species, and authority) in parentheses. For well-known species, however, scientific names may be omitted from article titles. If no common name exists in English, only the scientific name should be used.

Genetic Nomenclature

Sequence variants should be described in the text and tables using both DNA and protein designations whenever appropriate. Sequence variant nomenclature must follow the current HGVS guidelines; see varnomen.hgvs.org, where examples of acceptable nomenclature are provided.

Human gene nomenclature should follow the standards of the HUGO Gene Nomenclature Committee (HGNC), see https://www.genenames.org/.

The GenBank reference sequence and version number for the gene(s) studied (i.e., include the decimal point following the accession number in the sequence record) must be included in the Materials and Methods and as a footnote to the relevant table(s) or figure (s).

Sequence Data

Nucleotide sequence mutation/variant data *Molecular Genetics & Genomic Medicine* supports the recommendations of the Human Variome Project (Cotton RGH et al., Nat Genet. 39: 433, 2007; https://doi.org/10.1038/ng2024). Consequently, authors are required to submit all variants included in an article to the respective Locus Specific Database (LSDB, e.g. http://www.lovd.nl) or to a central variant database prior to acceptance (e.g. db SNP or ClinVar, https://www.ncbi.nlm.nih.gov). Authors must confirm the status of database submission in their cover letter and must confirm in the manuscript (e.g., in the methods section) the LSDB(s) or central variant databases to which they have submitted their variants and provide the URL. dbSNP and ClinVar are acceptable choices. The Editors also encourage the use of widely accessible genetics databases as repositories for human gene mapping information, including loci (genes, fragile sites, DNA segments), and probes. In the case of dbSNP, the identification numbers should be used, if available, to describe the variants in the manuscript. Further information, including updates on links to Locus-Specific Databases, can be obtained from the Human Genome Variation Society (HGVS) web site http://www.hgvs.org/

Microarray data should be MIAME compliant (for guidelines see http://www.mged.org/Workgroups/MIAME/miame.html)

Nucleotide sequence data can be submitted in electronic form to any of the three major collaborative databases: DDBJ, EMBL, or GenBank. It is only necessary to submit to one database as data are exchanged between DDBJ, EMBL, and GenBank on a daily basis. The suggested wording for referring to accession-number information is: ‘These sequence data have been submitted to the DDBJ/EMBL/GenBank databases under accession number U12345’. Addresses are as follows:

- DNA Data Bank of Japan (DDBJ): www.ddbj.nig.ac.jp

- EMBL Nucleotide Archive: ebi.ac.uk/ena
- GenBank: www.ncbi.nlm.nih.gov/genbank

Protein sequence data should be submitted to either of the following repositories:

- Protein Information Resource (PIR): proteininformationresource.org/.
- SWISS-PROT: expasy.org/resources/uniprotkb-swiss-prot

Conflict of Interest

The journal requires that all authors disclose any potential sources of conflict of interest. Any interest or relationship, financial or otherwise that might be perceived as influencing an author's objectivity is considered a potential source of conflict of interest. These must be disclosed when directly relevant or directly related to the work that the authors describe in their manuscript. Potential sources of conflict of interest include, but are not limited to: patent or stock ownership, membership of a company board of directors, membership of an advisory board or committee for a company, and consultancy for or receipt of speaker's fees from a company. The existence of a conflict of interest does not preclude publication. If the authors have no conflict of interest to declare, they must also state this at submission. It is the responsibility of the corresponding author to review this policy with all authors and collectively to disclose with the submission ALL pertinent commercial and other relationships.

Funding

Authors should list all funding sources in the Acknowledgments section. Authors are responsible for the accuracy of their funder designation. If in doubt, please check the Open Funder Registry for the correct nomenclature: www.crossref.org/services/funder-registry.

Authorship

The list of authors should accurately illustrate who contributed to the work and how. All those listed as authors should qualify for authorship according to the following criteria:

1. Have made substantial contributions to conception and design, or acquisition of data, or analysis and interpretation of data;
2. Been involved in drafting the manuscript or revising it critically for important intellectual content;
3. Given final approval of the version to be published. Each author should have participated sufficiently in the work to take public responsibility for appropriate portions of the content; and
4. Agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Contributions from anyone who does not meet the criteria for authorship should be listed, with permission from the contributor, in an Acknowledgments section (for example, to recognize contributions from people who provided technical help, collation of data, writing assistance, acquisition of funding, or a department chairperson who provided general support). Prior to submitting the article all authors should agree on the order in which their names will be listed in the manuscript.

Additional Authorship Options

Joint first or senior authorship: In the case of joint first authorship, a footnote should be added to the author listing, e.g. ‘X and Y should be considered joint first author’ or ‘X and Y should be considered joint senior author.’

Wiley’s Author Name Change Policy

In cases where authors wish to change their name following publication, Wiley will update and republish the paper and redeliver the updated metadata to indexing services. Our editorial and production teams will use discretion in recognizing that name changes may be of a sensitive and private nature for various reasons including (but not limited to) alignment with gender identity, or as a result of marriage, divorce, or religious conversion. Accordingly, to protect the author’s privacy, we will not publish a correction notice to the paper, and we will not notify co-authors of the change. Authors should contact the journal’s Editorial Office with their name change request.

Publication Ethics

This journal is a member of the Committee on Publication Ethics (COPE). Read Wiley’s Top 10 Publishing Ethics Tips for Authors here. Wiley’s Publication Ethics Guidelines can be found here.

Return to the Guideline Sections

6. AUTHOR LICENSING

If a paper is accepted for publication, the author identified as the formal corresponding author will receive an email prompting them to log in to Author Services, where via the Wiley Author Licensing Service (WALS) they will be required to complete a copyright license agreement on behalf of all authors of the paper.

For authors signing the copyright transfer agreement

If the open access option is not selected the corresponding author will be presented with the copyright transfer agreement (CTA) to sign. The terms and conditions of the CTA can be previewed in the samples associated with the Copyright FAQs.

For authors choosing open access

If the open access option is selected the corresponding author will have a choice of the following Creative Commons License Open Access Agreements (OAA):

Creative Commons Attribution License (CC-BY) OAA

Creative Commons Attribution Non-Commercial License (CC-BY-NC) OAA

Creative Commons Attribution Non-Commercial -NoDerivs License (CC-BY-NC-ND) OAA

General information regarding licensing and copyright is available on the Wiley Author Services and the Wiley Open Access websites.

Note to NIH, The Wellcome Trust and the Research Councils UK Grantees

Pursuant to NIH mandate, Wiley will post the accepted version of contributions authored by NIH grant-holders to PubMed Central upon acceptance. This accepted version will be made publicly available 12 months after publication. Please click here for further information. If you select the open access option and your research is funded by The Wellcome Trust or the

Research Councils UK (RCUK) you will be given the opportunity to publish your article under a CC-BY license supporting you in compliance with The Wellcome Trust and Research Councils UK requirements.

Self-Archiving Definitions and Policies

Note that the journal's standard copyright agreement allows for self-archiving of different versions of the article under specific conditions. Please click here for more detailed information about self-archiving definitions and policies.

Return to Guideline Sections

7. PUBLICATION PROCESS AFTER ACCEPTANCE

Accepted Articles

All accepted manuscripts are subject to editing. Authors have final approval of changes prior to publication.

Proofs

Authors will receive an e-mail notification with a link and instructions for accessing HTML page proofs online. Page proofs should be carefully proofread for any copyediting or typesetting errors. Online guidelines are provided within the system. No special software is required, all common browsers are supported. Authors should also make sure that any renumbered tables, figures, or references match text citations and that figure legends correspond with text citations and actual figures. Proofs must be returned within 48 hours of receipt of the email. Return of proofs via e-mail is possible in the event that the online system cannot be used or accessed.

Publication Charges

Color figures. Figures submitted in color will be published in color free of charge.

Return to Guideline Sections

8. POST PUBLICATION

Access and Sharing

When the article is published online:

- The author receives an email alert (if requested).
- The link to the published article can be shared through social media.
- The author will have free access to the paper (after accepting the Terms & Conditions of use, they can view the article).
- The corresponding author and co-authors can nominate up to ten colleagues to receive a publication alert and free online access to the article.

For additional important information on Wiley's Article Sharing policy, click here.

Print copies of the article can now be ordered (instructions are sent at proofing stage or email chris.jones@sheridan.com.

Promoting the Article

To find out how to best promote an article, click here.

Wiley Editing Services offers professional video, design, and writing services to create shareable video abstracts, infographics, conference posters, lay summaries, and research news stories for your research – so you can help your research get the attention it deserves.

Measuring the Impact of an Article

Wiley also helps our authors measure the impact of their research through a specialist partnership with Altmetric .

Return to Guideline Sections

Author Guidelines updated February 28, 2023