

https://dspace.org/

Germline Cancer Predisposition Variants in Paediatric Rhabdomyosarcoma

Pillhofer, Gabriella Peta

2023
University of the Witwatersrand, Johannesburg
https://hdl.handle.net/10539/42989
Pillhofer, Gabriella Peta. (2024). Germline Cancer Predisposition Variants in Paediatric Rhabdomyosarcoma [Master's dissertation, University of the Witwatersrand, Johannesburg]. WireDSpace.

Downloaded from a DSpace repository. https://dspace.org/

Germline Cancer Predisposition Variants in Paediatric Rhabdomyosarcoma

Gabriella Peta Pillhofer

Student Number: 2097324

A research report submitted to the Faculty of Health Sciences, University of the Witwatersrand in partial fulfilment of the requirements for the degree of Master of Science in Medicine in the field of Genomic Medicine

November 2023

Supervisor: Dr Lindie Lamola

Co-supervisor: Dr Khuthala Mnika

Declaration by Candidate

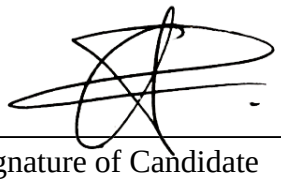
I, Gabriella Peta Pillhofer, declare that this Research Report is my own, unaided work. It is being submitted for the degree of Master of Science in Medicine in the field of Genomic Medicine at the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other University.



Gabriella Peta Pillhofer
Student Number: 2097324
27 February 2024

Contribution of Candidate

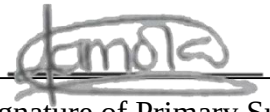
I, Gabriella Peta Pillhofer (Student Number 2097324), declare that this Research Report titled 'Germline Cancer Predisposition Variants in Paediatric Rhabdomyosarcoma' is my own work, and that I contributed significantly to the research findings presented in this paper intended for publication.



Signature of Candidate

27 November 2023

Date



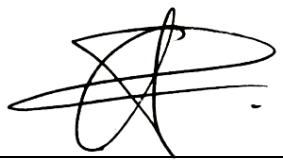
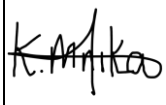
Signature of Primary Supervisor

27 November 2023

Date

Agreement by co-authors

In signing this declaration, the authors below agree to the use of the research findings intended for a published article as part of their Research Report.

Author	**Name**	**Signature**	**Date**
1st	Ms Gabriella Peta Pillhofer		27 November 2023
2nd last	Dr Khuthala Mnika		27 November 2023
last	Dr Lindie Lamola		27 November 2023

Acknowledgements

I must first express my immense thanks to my supervisors for their endless guidance and support. Dr Lindie Lamola, thank you for allowing me into this space. Your passion for the field is consistently inspiring and I can only hope to be anything like you. Dr Khuthala Mnika, your kind words and gentle direction made this work possible. It has been a privilege to learn from you.

Further thanks must be extended to all my teachers at the National Health Laboratory Service in Braamfontein. Learning from you over the last two years has been an honour – your passion and dedication never fails to excite me about genetics and genomics.

For their never-wavering love and support, I thank my parents. Mom and Dad, thank you for always encouraging me and believing in me. I am only capable of what I am because of you.

Paul, your encouragement, and love helped get me through the toughest moments. Words could never be enough to thank you.

Although you are gone, your light will not be forgotten. Phoebe, everything I do is with the knowledge that you are looking down on me. I hope I can make you proud.

Finally, to the parents and patients of the PeCan study, I extend my deep gratitude and honour. Thank you for your invaluable contribution, none of this would be possible without you.

Contents

List of Appendices

Appendix 1: Supplementary Data

Appendix 2: American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG-AMP) Guidelines for Variant Classification

Appendix 3: Human Research Ethics Committee (Medical) Clearance Certificate No. M230783

Appendix 4: Human Research Ethics Committee (Medical) Clearance Certificate No. M180855

Appendix 5: Approved Research Protocol

Appendix 6: Plagiarism Declaration and TurnItIn Plagiarism Report

Appendix 7: Human Mutation Author Guidelines

List of Tables

Table 1: *In silico* tools utilised for predicting potential effect of genetic variants

Table 2: Relevant demographic information of eight paediatric patients with rhabdomyosarcoma

Table 3: Variants retained as having un uncertain significance as cancer predisposition variants in paediatric patients with rhabdomyosarcoma

List of Abbreviations

RMS	Rhabdomyosarcoma
WES	Whole exome sequencing
ACMG-AMP	American College of Medical Genetics and Genomics and the Association for Molecular Pathology
VUS	Variant of uncertain significance
ERMS	Embryonal rhabdomyosarcoma
VCF	Variant call format
BAM	Binary alignment map
BAI	Binary alignment index
MAF	Minor allele frequency
IGV	Integrative Genomic Browser

Research report in the submissible format

To be submitted to Human Mutation

Germline Cancer Predisposition in Paediatric Rhabdomyosarcoma

Gabriella Peta Pillhofer[1], Khuthala Mnika[1,2], Lindie Lamola[1]

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

[2] Division of Human Genetics, Department of Pathology, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa

Correspondence should be addressed to Gabriella Peta Pillhofer; 2097324@students.wits.ac.za

1. Abstract

Rhabdomyosarcoma (RMS) is cancer that originates from undifferentiated skeletal muscle cells. RMS is the most common tissue sarcoma in children and adolescents and has over 50% of cases occurring in individuals under the age of 10 and thus there is reason to suspect that RMS may have a hereditary predisposition. However, much of the existing research of germline predisposition in paediatric RMS is focused on ethnically European populations, and currently very little data is available from African populations. In this study, we aimed to identify RMS predisposition variants by performing whole exome sequencing (WES) on germline DNA from eight paediatric patients diagnosed with RMS. Following WES, variant annotation and filtering was performed to identify variants in genes of interest that were potentially germline causes of malignancy. Filtered variants were then classified according to American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG-AMP) guidelines. This study identified four variants of uncertain significance (VUSs) in four patients. Two of these were in genes that have previously been associated with an RMS phenotype (*SDHB* and *BRCA1*), and two were in genes that have been associated with a hereditary cancer syndrome that is not linked to RMS (*CBL* and *CREBBP*). While the highlighted variants are not of clinical significance, this study emphasises the importance of cataloguing and reporting VUSs in research in Africa. By expanding the genomic database on African patients, the analysis of variants on the continent may be made more accurate and efficient. It is the goal that the knowledge gained from this study will contribute to the information base of hereditary paediatric cancers in Africa, and that it may encourage similar research so that the field may continue to expand.

2. Introduction

Rhabdomyosarcoma (RMS) is a malignancy that originates from undifferentiated skeletal muscle cells. Known as a paediatric cancer, RMS is the most common tissue sarcoma in children and adolescents and has over 50% of cases occurring in individuals under the age of 10 (Ognjanovic et al., 2009). There are four main subtypes of RMS: embryonal RMS (ERMS) (60-70% of cases), alveolar RMS (20-30%), pleomorphic RMS (10%) and sclerosing RMS (10%) (Kaseb et al., 2022; Zhang et al., 2015). The ERMS and alveolar RMS subtypes are the most likely to be identified in minors, with alveolar RMS in the extremities being most common in adolescents, and ERMS in hollow organs being common in children (Kaseb et al., 2022).

Rhabdomyosarcoma affects soft tissue, meaning malignancy can occur in almost any organ. The clinical presentation of RMS is varied and influenced by the affected organ systems, and specific tissue types. For instance, RMS of the bladder may result in obstructions of the urinary tract and haematuria, whereas orbital RMS may disturb vision and cause strabismus. Tumours may cause painful swellings, which can result in nausea and vomiting when occurring within the viscera (American Cancer Society, 2018; Kaseb et al., 2022). Presenting clinical complaints are often of persistent lumps, redness or swelling, with no obvious cause. Additional symptoms may include headaches, bleeding. When presenting in advanced stages, or once metastasis has begun, patients may present with bone pain and generalised weakness. (American Cancer Society, 2018).

In the United States of America (USA), RMS has an incidence of 4.58 per million per year (Martin-Giacalone et al., 2021; McEvoy et al., 2023) and makes up approximately 3% of the national paediatric cancer burden (American Cancer Society, 2020). This is similar to the annual incidence in Europe of approximately 4 per million (Ferrari et al., 2019). Incidence rates in Africa are variable, and statistics are not regularly updated. In Africa, incidence in individuals aged 0-14 has been estimated between 0.6 and 16.3 per million, with South Africa having an incidence 2.4 per million per year (Stefan et al., 2017).

There is currently no known cancer predisposition syndrome that has RMS as its major phenotype (Kaseb et al., 2022). However, approximately 18% of paediatric cancers are linked to germline pathogenic variants in cancer predisposition genes (Fiala et al., 2021), and thus there is reason to suspect that RMS may have a hereditary predisposition.

Current knowledge of variants in cancer predisposition genes being associated with RMS is limited. Previous studies have uncovered variants that may be linked to RMS in genes associated with RAS/mitogen-activated protein kinase signalling pathway syndromes (RASopathies) and RASopathy-like syndromes, as well as known cancer predisposition syndromes, such as Li-Fraumeni syndrome. Studies have identified variants in the genes associated with these syndromes in patients with RMS – suggesting the link between pathogenic variants in these cancer predisposition genes and developing RMS (Kim et al., 2021; Li et al., 2021). These previously identified pathogenic variants, along with the high proportion of paediatric cancers associated with pathogenic variants in cancer predisposition genes support the theory that there may be a hereditary component to RMS (Kim et al., 2021). However, much of the existing research of germline predisposition in paediatric RMS is focused on ethnically European populations, and currently very little data on the subject is available on African populations. Despite the high burden of disease in Africa, there is not a focus on uncovering a germline predisposition in this population.

The clinical utility of identifying germline cancer predisposition variants in paediatric cancer patients is that improved risk management for both the patient and their family can occur. Clinical management for affected patients with a germline predisposition variant may differ from sporadic cases. Additionally, a known inherited pathogenic variant can have screening and family planning implications for the parents and siblings of the patient.

This study forms part of the larger Paediatric Cancer study, which aims to investigate the genetic contribution of inherited variants in the development of cancers in children. Sub-studies of the Paediatric Cancer study, such as this one, contribute to the aim of the parent study by identifying germline variants in smaller subgroups of the larger cohort.

In this study, we aimed to identify RMS predisposition variants by performing whole exome sequencing (WES) on germline DNA from eight paediatric patients diagnosed with RMS. It is the goal that the knowledge gained from this study will contribute to the information base of hereditary paediatric cancers in Africa, and that it may encourage more similar research so that the field may continue to expand.

3. Methods

3.1. Sample selection

Eight paediatric patients with rhabdomyosarcoma were selected from the parent Paediatric Cancer (PeCan) study. Patients were prioritised for this study based on age at onset, and on

type of RMS. The cohort was comprised of six Black African patients and two Indian patients, aged 3-11. All patient information was anonymised using patient codes. Patient information other than that outlined below is only available to the principal investigator of the parent study. On request, certain clinical features were made available for genotype-phenotype correlation purposes. Only clinical data available in the patients' files was used and no additional clinal assessment was undertaken as part of this study.

3.2. Library preparation and whole exome sequencing

Library preparation was completed previously for four samples as part of the parent study. For the remaining four samples sequencing was performed as described here. Initially, library preparation was conducted according to the Ion AmpliSeq Exome RDY Library Preparation protocol from ThermoFisher Scientific (ThermoFisher Scientific, 2023). Samples were quantified prior to amplification using Qubit flourometery according to manufacturer's guidelines (ThermoFisher Scientific, 2015, 2022b). All eight prepared libraries were then quantified using the Agilent TapeStation quantification system (Agilent Technologies, 2015) and then diluted as per the Library Preparation Protocol (ThermoFisher Scientific, 2023). The TapeStation system was used specifically in library preparation as it can assess the quality of the library, and in the quantification of library size. It is imperative to determine library concentration before sequencing so that all samples can be diluted to similar ideal qualities. As this protocol pools two samples on the same sequencing chip, library dilution served to ensure that libraries being loaded together were of a similar concentration. This prevents libraries of a lower concentration having reduced coverage.

Libraries were diluted to achieve an approximate concentration of 100 picomolar and then were combined in pairs with distinct barcodes (ThermoFisher Scientific, 2023). Subsequently, the diluted and pooled libraries, along with reagents, were loaded into the IonChef system (ThermoFisher Scientific, 2019) for templating and chip loading. The IonChef system loaded two pooled samples onto an Ion 540 Chip, following the manufacturer's guidelines (ThermoFisher Scientific, 2019). Templated chips were then sequenced using the Ion S5 sequencer, adhering to the manufacturer's protocols (ThermoFisher Scientific, 2022a). The sequencing run was optimised for 200 bp fragments using a run with 500 flows.

3.3. Variant annotation and filtering

The variant call format (VCF) files from the sequencer were downloaded and used for variant annotation and to produce variant reports. This VCF file format is a text file format, allowing for easy storage and access of sequencing data (Danecek et al., 2011). Additionally, the binary alignment map (BAM) and binary alignment index (BAI) files were obtained. The BAM file stores the compressed data of the sequences aligned to the reference, whereas the BAI file contains the index required to read the BAM file (Illumina, 2017). These formats are not human-readable but are used by software tools to pinpoint exact locations of variants within the aligned sequence. These BAM and BAI files were used for variant visualisation.

Variant annotation was performed using the web-based tool from Ensembl, Variant Effect Predictor (https://www.ensembl.org/Tools/VEP). The purpose of variant annotation was to assign functional information to variants identified during sequencing. Annotation tools, such as this are used to note defining features of each variant (such as the gene in which it is located, and the type of variant) so that this information may be used to filter variants of interest for the study.

In this study, variants were filtered manually according to criteria to best assure retained variants were potentially cancer causing in RMS. Filtering variants in WES is essential to ensure resources for analysis and classification are sensibly spent on variants of interest for the highlighted phenotype.

Initially, the variant list was filtered to only include variants in the genes of interest (Supplementary Tables 1 and 2). These genes of interest were targeted for analysis in this study due to suspected associations of these genes with RMS, or their proven association with other known cancer predisposition syndromes. Associations of these genes (Supplementary Table 1) with RMS is existing in the literature (Kim et al., 2021; Li et al., 2021). The remaining genes, those associated with other cancer predisposition syndromes (Supplementary Table 2), were selected as RMS is understudied and thus increasing the number of genes analysed increases the possibility of identifying a germline pathogenic variant. These additional genes were previously selected in existing RMS research as potential targets for an RMS cancer predisposition gene (Kim et al., 2021; Li et al., 2021).

Following this, common variants were excluded by eliminating variants with a minor allele frequency (MAF) greater than 0.05. Common variants are a feature of normal variation and

are not relevant to this study. An MAF threshold of 0.05 was chosen as this is a previously established MAF associated with common variation (The International HapMap Consortium, 2005). The effect of variants was then assessed and only variants with potential protein effects were included. This involved the inclusion of missense, frameshift, and splice-site variants, while synonymous and other intronic variants were excluded. Additionally, the determination of predicted protein effect was aided using multiple *in silico* prediction tools (Table 1). For variants where multiple tools assessed a potentially deleterious effect on protein function, or if there were conflicting interpretations of pathogenicity between tools, the variant was retained for additional manual analysis. Variants with a ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/) code of pathogenic, likely pathogenic and variants of uncertain significance were prioritised. At this stage, variants with conflicting reports of pathogenicity as per ClinVar were also considered.

Variants that satisfied all the filtering criteria were visualised with the Integrative Genomic Viewer (IGV) genome browser (Robinson et al., 2011) using the previously obtained BAM and BAI files. This was to confirm sufficient coverage of the region surrounding the variant, as well as to allow the variants to be visualised in the context of the genomic region.

Table 1: *In silico* tools utilised for predicting potential effect of genetic variants

***In silico* tool**	**URL**	**Reference**
MetaLR	https://wglab.org/	Dong et al. (2015)
REVEL	https://sites.google.com/site/revelgenomics/	Ioannidis et al. (2016)
FATHMM	https://fathmm.biocompute.org.uk/	Rogers et al. (2018)
SIFT	https://sift.bii.a-star.edu.sg/	Ng & Henikoff (2001)
PolyPhen	http://genetics.bwh.harvard.edu/pph2/	Adzhubei et al. (2010)

3.4.Variant classification

Following filtering, variants were then subjected to classification using the American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG-AMP) guidelines (Richards et al., 2015). The ACMG-AMP guidelines are designed to facilitate the use of standardised, specific language when describing variants. Evidence is collected and applied to variants in order to categorise them into one of five classifications: pathogenic, likely pathogenic, uncertain significance, likely benign, and

benign. These classification groupings are intended to aid the understanding of the clinical significance of a variant, using clear terms and detailed evidence (Richards et al., 2015).

4. Results

4.1. Patient Demographics

Patients selected for this study were aged 3-11, with a varied presentation of the RMS phenotype. The cohort consisted of 5/8 patients presenting with non-specific ERMS, one patient with paratesticular RMS, one with orbital RMS, and one with a renal rhabdoid tumour. Regarding patient ethnicity, 6/8 patients were Black African, and 2/8 were Indian. Of these patients, only 2 had a family history of cancer. The specifics of these cancer histories were unavailable from the clinical files. These demographics, including ethnicity, age at onset, and cancer or tumour type are given in Table 2.

4.2.Genetic diagnosis by Whole exome sequencing

In 74 genes of interest (Supplementary Tables 1 and 2), variants were prioritised that were non-synonymous and had a MAF > 0,05. These were then manually filtered, and variants in four of the patients were considered relevant to report for this study. These variants (Table 3) were deemed VUSs based on several criteria, including literature review, classification from online tools, gene's involvement as a cancer driver, predicted effect of the variant, and novel status of the variant.

Variants in the remaining four patients (PC074, PC077, PC121 and PC127) were excluded following filtering, as they either did not satisfy the filtering criteria, or were deemed benign or likely benign in manual classification as such excluded for further analysis.

4.2.1. PC002 – Variant of uncertain significance in *SDHB* – NM_003000.3(SDHB):c.455C>T (p.Ser152Phe)

This patient is an 8-year-old Black African female, who presented with ERMS in the right nostril. A missense variant was detected in the gene of interest *SDHB* (Table 3). This variant is previously unreported in gnomAD (https://gnomad.broadinstitute.org/) (PM2). Additionally, this variant is in a region of interest where there are no benign missense variants, and all missense variants are pathogenic or VUSs (PM1). Furthermore, this variant has conflicting interpretations of pathogenicity as per ClinVar, with 4/5 submitters interpreting it as a VUS. The functional effect of this variant is considered pathogenic by

Table 2: Relevant demographic information of eight paediatric patients with rhabdomyosarcoma

Patient	**Cancer/Tumour Type**	**Age at Onset**	**Ethnicity**	Family History of Cancer?
PC002	ERMS	8	Black African	N
PC011	ERMS	3	Indian	N
PC074	ERMS	3	Black African	N
PC077	ERMS	11	Black African	N
PC083	Paratesticular RMS	3	Black African	N
PC115	Renal rhabdoid tumour	9	Black African	Y
PC121	ERMS	3	Indian	N
PC127	Orbital RMS	5	Black African	Y

RMS – rhabdomyosarcoma; ERMS – embryonal rhabdomyosarcoma; Y – yes; N – no

multiple meta-analyses of scores from i*n silico* prediction tools: MetaLR with a pathogenic supporting score of 0.892 and REVEL with a pathogenic supporting score of 0.766 (PP3). This variant is heterozygous with a coverage of over 100x. These factors considered together classify this variant as a VUS, as per the ACMG-AMP guidelines.

4.2.2. PC011 – Variant of uncertain significance in *BRCA1* – NM_007294.4(BRCA1):c.4199T>C (p.Met1400Thr)

This patient, a 3-year-old Indian male with ERMS of the bladder, has a missense variant in *BRCA1* (Table 2), a gene of interest in hereditary cancer predisposition. The variant is unreported in the global population (PM2). It has been reported as having uncertain significance by 6/6 submitters on ClinVar. Multiple lines of computational evidence

support its pathogenicity, as evidenced by a MetaLR score of 0.892 (PP3). Additionally, this variant is in a region of interest for interaction within *BRCA1*. Missense variants in this region are primarily VUSs. This variant was visualised on IGV as a heterozygous variant with a coverage of over 100x.

However, this is a missense mutation in a gene where pathogenicity is usually a result of truncating variants (BP1). Thus, the application of both benign and pathogenic criteria, as well as other supporting evidence of uncertainty, support the classification of this variant as a VUS according to ACMG-AMP guidelines.

4.2.3. PC083 – Variant of uncertain significance in *CREBBP* – NM_004380.3(CREBBP):c.293G>T (p.Gly98Val)

This patient is a 3-year-old Black African male, who presented with paratesticular RMS. A missense variant was detected in the cancer driver gene *CREBBP*, a gene of interest for RMS. The variant has an MAF of 0.004, lower than expected of the population (PM2). Scores from *in silico* prediction tools SIFT (0), PolyPhen (0.963) and FATHMM (0.9815) support this variant as having a possibly damaging effect. Furthermore, the *in silico* meta-scoring tool REVEL gives a score of 0.64, suggesting a VUS (PP3). This is a missense variant in a gene where missense variants have been previously determined to be damaging and/or have been previously associated with inherited cancers (PP2). Heterozygosity of this variant was confirmed, along with a sufficient coverage of over 100x.

However, no functional data exists to prove the effect of this variant on the protein function. Additionally, this variant is categorised as benign by multiple submitters on ClinVar (BP6). These conflicting points, and the lack of available literature and functional data on the variant result in its classification as a VUS.

4.2.4. PC115 – Variant of uncertain significance in *CBL* – NM_005188.4(CBL):c.2345C>T (p.Pro782Leu)

A missense variant was detected in *CBL* (Table 3), a gene of interest and known mutational cancer driver in this 9-year-old Black African female with a renal rhabdoid tumour. This variant is previously reported in ClinVar with conflicting interpretations of pathogenicity. The MAF for this variant is 0.002, significantly lower than is to be expected in the population (PM2). This is a missense variant in a gene where a majority of missense variants are deemed pathogenic (PP2). This variant was visualised on IGV as a heterozygous variant.

However, no functional data exists to prove the effect of this variant on the protein function. Additionally, this variant is categorised as benign by multiple *in silico* prediction tools (BP6). As criteria for both benign and pathogenic classifications are present, and are thus contradictory, this variant is as VUS as per the ACMG-AMP guidelines.

5. Discussion

This study aimed to identify RMS predisposition variants in germline DNA from eight paediatric patients diagnosed with RMS. Following variant filtering, four variants of interest were identified in four patients. These were all missense variants in cancer driver genes and were classified as having uncertain significance according to ACMG-AMP recommendations.

5.1. Target genes and genes where pathogenic variants were identified

The genes targeted for analysis in this study were selected due to previous reports of these genes having been associated with RMS. These genes are suspected to be associated with RMS or are linked to known cancer predisposition syndromes. For the genes in which variants were found, *SDHB* and *BRCA1* have not been previously linked with RMS, whereas *CREBBP* and *CBL* have (Kim et al., 2021; Li et al., 2021).

SDHB is a tumour suppressor gene that encodes for subunit B of the succinate dehydrogenase (SDH) enzyme complex. The enzyme complex functions in multiple metabolic processes within the mitochondria (OMIM #185470), and thus pathogenic variants in this gene, which are typically associated with hereditary paraganglioma-pheochromocytoma syndrome with high malignancy risk (Else et al., 2023), support the correlation between mitochondrial dysfunction and malignancy (Shi et al., 2022). Though RMS specifically has not previously been associated with variants in *SDHB*, missense mutations in the gene are implicated in the cancer predisposition syndrome leading to other sarcomas (Shi et al., 2022; Vagher et al., 2022).

BRCA1 is typically associated with hereditary breast and ovarian cancer syndrome. This gene acts as a tumour suppressor and encodes a protein necessary for retaining stability in genomic architecture (OMIM #113705). Reports of variants in this gene having pathogenicity in RMS have been limited to frameshift truncating variants. However, missense variants in *BRCA1* are typical of the cancer predisposition phenotype (Golubeva et al., 2019) and thus could contribute to the RMS phenotype.

Table 3: Variants retained as having un uncertain significance as cancer predisposition variants in paediatric patients with rhabdomyosarcoma

Patient	**Location**	**Gene**	**Transcript Name**	**rsID**	**ACMG-AMP Classification**	**ACMG-AMP Criteria**
PC002	1:17354329-17354329	*SDHB*	NM_003000.3(SDHB):c.455C>T (p.Ser152Phe)	rs200414835	VUS	PM2; PM1; PP3
PC011	17:41234579-41234579	*BRCA1*	NM_007294.4(BRCA1):c.4199T>C (p.Met1400Thr)	rs80357473	VUS	PM2; PP3; BP1
PC083	16:3900803-3900803	*CREBBP*	NM_004380.3(CREBBP):c.293G>T (p.Gly98Val)	rs141982003	VUS	PM2; PP3; PP2; BP6
PC115	11:119169161-119169161	*CBL*	NM_005188.4(CBL):c.2345C>T (p.Pro782Leu)	–	VUS	PM2; PP2; BP4

VUS – variant of uncertain significance; ACMG-AMP – American College of Medical Genetics and Genomics and the Association for Molecular Pathology Detailed explanations of the ACMG-AMP criteria can be found in Appendix 2: American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG-AMP) Guidelines for Variant Classification as per Richards et al. (2015)

Typically, *CREBBP* functions as a transcriptional activator, and its encoded protein acts to stabilise interactions within transcription complexes (OMIM #600140). Missense mutations in this gene are associated with Rubinstein-Taybi syndrome (RSTS) (Bartsch et al., 2005), a tumour predisposition syndrome. Case reports of co-occurrence of RMS and RSTS are limited and dated (Martin-Giacalone et al., 2021; Miller & Rubinstein, 1995; Sobel & Woerner, 1981), though RSTS has been associated with multiple cancer phenotypes and a variety of malignancies (Bartsch et al., 2005).

As a proto-oncogene, *CBL* acts a negative regulator in signal transduction pathways (OMIM #165360). Previously, *CBL* has typically been associated with Noonan syndrome, a RASopathy and known cancer predisposition syndrome. While malignancy in Noonan syndrome patients usually occurs as leukaemia, missense variants in *CBL* have been reported in patients with overlapping Noonan syndrome and RMS phenotypes (Ji et al., 2019). However, the patient described by Ji et al. (2019) had a clear clinical phenotype of Noonan syndrome, which was not reported in the clinical file of patient PC115 in this study. Moreover, the RMS phenotype of the described patient was of ERMS of the bladder, which does not align with the renal rhabdoid tumour seen in patient PC115 in this study. The described association of *CBL* and the RMS phenotype is limited (Ji et al., 2019; Kim et al., 2021) and thus is not a promising candidate for a cancer predisposition gene in RMS. Though further investigation into pathogenic variants in this gene in RMS patients should be undertaken to reach a definite conclusion.

These genes are notably all cancer driver genes (Martínez-Jiménez et al., 2020) and though not all previously associated with RMS – their association with cancers makes them genes of interest. Particularly when missense mutations are found occurring in paediatric patients with cancers.

5.2. Variants of uncertain significance in research in an African setting

The lack of diversity in human reference genomes presents challenge to researchers aiming to investigate genetic variation in non-European populations, especially African populations, which are the most underrepresented groups in existing references. Genomic studies of African populations are likely to encounter false positives, false negatives, inconclusive results, and VUSs (Alosaimi et al., 2021; Shriner et al., 2020).

Whereas in clinical settings these variants should only be reported with extreme caution, and guidelines can vary between conditions and clinical settings – in research in Africa it

is generally considered important for VUSs to always be reported. This allows for expansion of the knowledge base on African patients, and enough contribution of this type may eventually allow the creation of African databases to make analysis of variants on the continent more accurate and efficient (Krause, 2019).

In cancer research, non-European individuals are disproportionately likely to receive a VUS classification than European individuals. This is due to both the underrepresentation of these groups in genomic databases, and the genomic diversity of populations such as African populations, as previously discussed. This disparity results in European populations being able to benefit inequitably from the results of the genetic testing, such as a better ability to access early intervention and management (Ndugga-Kabuye & Issaka, 2019).

5.3. Clinical impact of genetic diagnosis of rhabdomyosarcoma

Although a genetic diagnosis of RMS may not directly impact the clinical management of RMS, there is still clinical utility in uncovering cancer predisposition genes linked with RMS. If pathogenic variants in syndromes associated cancer predisposition genes are linked with an RMS phenotype, there may be altered clinical management of these patients (Li et al., 2021). Children affected with these associated syndromes would have an increased risk of developing RMS, and thus would require increased screening and more rapid intervention than unaffected counterparts.

Additionally, genetic diagnosis of an inherited pathogenic variant has an impact on the siblings and other family members of the affected individual. If a cancer and/or cancer syndrome is identified as being inherited, there is an increased risk to the family. This would allow for family members, particularly siblings, to undergo screenings or potentially genetic testing to better understand their risks for cancer development. Moreover, knowledge of the cancer's status as inherited has family planning implications for the parents of the affected children. (Gomy & Estevez, 2013)

Furthermore, there is emerging evidence for the utilisation of germline pathogenic variants as a target for cancer therapeutics (Hoeben et al., 2021). Currently, RMS is treated using an individualised combination of surgical intervention, chemotherapy and radiation (Kaseb et al., 2022), all of which are invasive and inconclusive options. Personalised medicine is the future of cancer treatment and identifying pathogenic variants in cancer predisposition genes associated with a specific cancer type (Hoeben et al., 2021; Thavaneswaran et al., 2019), such as RMS, is a stepping stone towards this.

5.4. Challenges and future directions of rhabdomyosarcoma in Africa

Evaluating the role of germline variation in paediatric RMS is challenging both generally, and particularly in Africa. The relative rarity compared to other cancers (American Cancer Society, 2020), along with the small proportion of cancers caused by germline pathogenic variants (Kim et al., 2021) mean current studies are limited. This same issue is faced with RMS research in Africa and is compounded by the challenges faced by researchers in Africa attempting WES. Namely, the issue of VUSs as discussed previously.

As only approximately 8% of RMS cases can be attributed to a germline predisposition (Kim et al., 2021), 82% of the cause for cancer development is not detectable by WES of germline DNA. The low percentage of germline pathogenic variants that are causative of RMS offers explanation as to why no pathogenic results were found with these patients, and why only four had results considered to be of interest. Further studies with larger sample sizes should be conducted to increase the chances of identifying a germline pathogenic variant in an African patient with RMS. It is of particular interest to these researchers that further investigation is conducted in African cohorts, as this would further serve the aim of this research to improve the available data on RMS in paediatric, African populations.

Only a subset of genes was analysed in this study based on previous reports of their associations with RMS or known association with cancers and a suspected association with RMS. These are not the only cancer predisposition genes or genes known to be linked with a tumour predisposition syndrome. It can thus be suggested that further research into which genes are potentially cancer predisposition genes in RMS. Particularly, this research should be conducted on diverse populations and should include ethnically African participants.

Additionally, a notable limitation of this study was the minimal clinical information available. This study was conducted using only the clinical data made available through the principal investigator of the parent study. With limited clinical phenotypes available, it was challenging to comment on the potential disease-causing impact of the identified variants. Particularly for variants in genes linked with genetic conditions with specific phenotypes. Having expanded clinical information would have allowed for a more in-depth genotype-phenotype correlation to be made. Furthermore, having access to family information and potentially parent genetic information could have allowed for more thorough classification and the possibility of upgrading some of the ACMG-AMP codes.

6. Conclusion

While this study produced only VUS results, the findings described here work to serve the aim of this study which was to describe RMS variants in a paediatric African cohort, while expanding to the information base of hereditary paediatric cancers in Africa. It is the ultimate intention of the researchers that this study will facilitate further research in the field so that the understanding of germline cancer predisposition in paediatric RMS may continue to expand.

7. Declarations

7.1. Ethics

Ethics approval was granted by the Human Research Ethics Committee (HREC) Medical, University of the Witwatersrand (Certificate number: M230783). Clearance Certificates are contained within the Appendix section of this document.

7.2. Availability of data

Data is not available in public domains at present. All variants will be made available on ClinVar (https://www.ncbi.nlm.nih.gov/clinvar/).

7.3. Competing interests

The authors declare they have no competing interests.

7.4. Funding

This research project was funded under the larger PeCan project study under the principal investigator Dr Lindie Lamola. Funders include the National Research Foundation [138196] and the South African Medical Research Council [MRC SIR grant].

8. References

Adzhubei, I. A., Schmidt, S., Peshkin, L., Ramensky, V. E., Gerasimova, A., Bork, P., Kondrashov, A. S., & Sunyaev, S. R. (2010). A method and server for predicting damaging missense mutations. *Nature Methods*, *7*(4), 248–249. https://doi.org/10.1038/nmeth0410-248

Agilent Technologies. (2015). *Agilent D1000 ScreenTape System Quick Guide*.

Alosaimi, S., Van Biljon, N., Awany, D., Thami, P. K., Defo, J., Mugo, J. W., Bope, C. D., Mazandu, G. K., Mulder, N. J., & Chimusa, E. R. (2021). Simulation of African and non-African low and high coverage whole genome sequence data to assess variant calling approaches. *Briefings in Bioinformatics*, *22*(4). https://doi.org/10.1093/BIB/BBAA366

American Cancer Society. (2018). *Signs and Symptoms of Rhabdomyosarcoma*. American Cancer Society. https://www.cancer.org/cancer/types/rhabdomyosarcoma/detection-diagnosis-staging/signs-symptoms.html

American Cancer Society. (2020). *Cancer Facts & Figures 2020*. American Cancer Society. https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/cancer-facts-figures-2020.html

Bartsch, O., Schmidt, S., Richter, M., Morlot, S., Seemanová, E., Wiebe, G., & Rasi, S. (2005). DNA sequencing of CREBBP demonstrates mutations in 56% of patients with Rubinstein–Taybi syndrome (RSTS) and in another patient with incomplete RSTS. *Human Genetics*, *117*(5), 485–493. https://doi.org/10.1007/s00439-005-1331-y

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., McVean, G., & Durbin, R. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*(15), 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Dong, C., Wei, P., Jian, X., Gibbs, R., Boerwinkle, E., Wang, K., & Liu, X. (2015). Comparison and integration of deleteriousness prediction methods for nonsynonymous SNVs in whole exome sequencing studies. *Human Molecular Genetics*, *24*(8), 2125–2137. https://doi.org/10.1093/hmg/ddu733

Else, T., Greenberg, S., & Fishbein, L. (2023). Hereditary Paraganglioma-Pheochromocytoma Syndromes. In: Adam MP, Feldman J, Mirzaa GM, et al., editors. GeneReviews® [Internet]. https://www.ncbi.nlm.nih.gov/books/NBK1548/Ferrari, A., Brecht, I. B., Gatta, G., Schneider, D. T., Orbach, D., Cecchetto, G., Godzinski, J., Reguerre, Y., Bien, E., Stachowicz-Stencel, T., Ost, M., Magni, C., Kearns, P., Vassal, G., Massimino, M., Biondi, A., Bisogno, G., & Trama, A. (2019). Defining and listing very rare cancers of paediatric age: consensus of the Joint Action on Rare Cancers in cooperation with the European Cooperative Study Group for Pediatric Rare Tumors. *European Journal of Cancer*, *110*, 120–126. https://doi.org/10.1016/j.ejca.2018.12.031

Fiala, E. M., Jayakumaran, G., Mauguen, A., Kennedy, J. A., Bouvier, N., Kemel, Y., Fleischut, M. H., Maio, A., Salo-Mullen, E. E., Sheehan, M., Arnold, A. G., Latham, A., Carlo, M. I., Cadoo, K., Murkherjee, S., Slotkin, E. K., Trippett, T., Glade Bender, J., Meyers, P. A., … Walsh, M. F. (2021). Prospective pan-cancer germline testing using MSK-IMPACT informs clinical translation in 751 patients with pediatric solid tumors. *Nature Cancer*, *2*(3), 357. https://doi.org/10.1038/S43018-021-00172-1

Golubeva, V. A., Nepomuceno, T. C., & Monteiro, A. N. A. (2019). Germline Missense Variants in BRCA1: New Trends and Challenges for Clinical Annotation. *Cancers*, *11*(4), 522. https://doi.org/10.3390/cancers11040522

Gomy, I., & Estevez Diz, M. D. P. (2013). Hereditary cancer risk assessment: essential tools for a better approach. *Hereditary Cancer in Clinical Practice*, *11*(1), 16. https://doi.org/10.1186/1897-4287-11-16

Hoeben, A., Joosten, E. A. J., & van den Beuken-van Everdingen, M. H. J. (2021). Personalized Medicine: Recent Progress in Cancer Therapy. *Cancers*, *13*(2), 242. https://doi.org/10.3390/cancers13020242

Illumina. (2017). BAM File Format. Illumina. https://support.illumina.com/help/BS_App_RNASeq_Alignment_OLH_1000000006112/Content/Source/Informatics/BAM-Format.htm

Ioannidis, N. M., Rothstein, J. H., Pejaver, V., Middha, S., McDonnell, S. K., Baheti, S., Musolf, A., Li, Q., Holzinger, E., Karyadi, D., Cannon-Albright, L. A.,

Teerlink, C. C., Stanford, J. L., Isaacs, W. B., Xu, J., Cooney, K. A., Lange, E. M., Schleutker, J., Carpten, J. D., … Sieh, W. (2016). REVEL: An Ensemble Method for Predicting the Pathogenicity of Rare Missense Variants. *The American Journal of Human Genetics*, *99*(4), 877–885. https://doi.org/10.1016/j.ajhg.2016.08.016

Ji, J., Navid, F., Hiemenz, M. C., Kaneko, M., Zhou, S., Saitta, S. C., & Biegel, J. A. (2019). Embryonal rhabdomyosarcoma in a patient with a germline CBL pathogenic variant. *Cancer Genetics*, *231–232*, 62–66. https://doi.org/10.1016/j.cancergen.2018.12.006

Kaseb, H., Kuhn, J., & Babiker, H. M. (2022). Rhabdomyosarcoma. *StatPearls*. https://www.ncbi.nlm.nih.gov/books/NBK507721/

Kim, J., Light, N., Subasri, V., Young, E. L., Wegman-Ostrosky, T., Barkauskas, D. A., Hall, D., Lupo, P. J., Patidar, R., Maese, L. D., Jones, K., Wang, M., Genome Research Laboratory, C., Tavtigian, S. V., Wu, D., Shlien, A., Telfer, F., Goldenberg, A., Skapek, S. X., … Stewart, D. R. (2021). Pathogenic Germline Variants in Cancer Susceptibility Genes in Children and Young Adults With Rhabdomyosarcoma. *JCO Precision Oncology*, *5*, 75–87. https://doi.org/10.1200/PO.20.00218

Kratz, C. P., Rapisuwon, S., Reed, H., Hasle, H., & Rosenberg, P. S. (2011). Cancer in Noonan, Costello, Cardiofaciocutaneous and LEOPARD Syndromes. *American Journal of Medical Genetics. Part C, Seminars in Medical Genetics*, *157*(2), 83. https://doi.org/10.1002/AJMG.C.30300

Krause, A. (2019). New genetic testing technologies: Advantages and limitations. *South African Medical Journal*, *109*(4), 207. https://doi.org/10.7196/SAMJ.2019.v109i4.13990

Li, H., Sisoudiya, S. D., Martin-Giacalone, B. A., Khayat, M. M., Dugan-Perez, S., Marquez-Do, D. A., Scheurer, M. E., Muzny, D., Boerwinkle, E., Gibbs, R. A., Chi, Y. Y., Barkauskas, D. A., Lo, T., Hall, D., Stewart, D. R., Schiffman, J. D., Skapek, S. X., Hawkins, D. S., Plon, S. E., … Lupo, P. J. (2021). Germline Cancer Predisposition Variants in Pediatric Rhabdomyosarcoma: A Report From the Children's Oncology Group. *JNCI Journal of the National Cancer Institute*, *113*(7), 875. https://doi.org/10.1093/JNCI/DJAA204

Martínez-Jiménez, F., Muiños, F., Sentís, I., Deu-Pons, J., Reyes-Salazar, I., Arnedo-Pac, C., Mularoni, L., Pich, O., Bonet, J., Kranas, H., Gonzalez-Perez, A., & Lopez-Bigas, N. (2020). A compendium of mutational cancer driver genes. *Nature Reviews Cancer*, *20*(10), 555–572. https://doi.org/10.1038/s41568-020-0290-x

Martin-Giacalone, B. A., Weinstein, P. A., Plon, S. E., & Lupo, P. J. (2021). Pediatric Rhabdomyosarcoma: Epidemiology and Genetic Susceptibility. *Journal of Clinical Medicine 2021, Vol. 10, Page 2028*, *10*(9), 2028. https://doi.org/10.3390/JCM10092028

McEvoy, M. T., Siegel, D. A., Dai, S., Okcu, M. F., Zobeck, M., Venkatramani, R., & Lupo, P. J. (2023). Pediatric rhabdomyosarcoma incidence and survival in the United States: An assessment of 5656 cases, 2001–2017. *Cancer Medicine*, *12*(3), 3644–3656. https://doi.org/10.1002/cam4.5211

McKusick-Nathans Institute of Genetic Medicine. (n.d.). *Online Mendelian Inheritance in Man, OMIM®*. Johns Hopkins University (Baltimore, MD). Retrieved March 13, 2023, from https://www.omim.org/

Miller, R. W., & Rubinstein, J. H. (1995). Tumors in Rubinstein-Taybi syndrome. *American Journal of Medical Genetics*, *56*(1), 112–115. https://doi.org/10.1002/ajmg.1320560125

Ndugga-Kabuye, M. K., & Issaka, R. B. (2019). Inequities in multi-gene hereditary cancer testing: lower diagnostic yield and higher VUS rate in individuals who identify as Hispanic, African or Asian and Pacific Islander as compared to European. *Familial Cancer*, *18*(4), 465–469. https://doi.org/10.1007/s10689-019-00144-6

Ng, P. C., & Henikoff, S. (2001). Predicting Deleterious Amino Acid Substitutions. *Genome Research*, *11*(5), 863–874. https://doi.org/10.1101/gr.176601

Ognjanovic, S., Linabery, A. M., Charbonneau, B., & Ross, J. A. (2009). Trends in childhood rhabdomyosarcoma incidence and survival in the United States, 1975-2005. *Cancer*, *115*(18), 4218–4226. https://doi.org/10.1002/CNCR.24465

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., & Rehm, H. L. (2015).

Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genetics in Medicine*, *17*(5), 405–424. https://doi.org/10.1038/gim.2015.30

Robinson, J. T., Thorvaldsdóttir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative Genomics Viewer. *Nature Biotechnology*, *29*(1), 24. https://doi.org/10.1038/NBT.1754

Rogers, M. F., Shihab, H. A., Mort, M., Cooper, D. N., Gaunt, T. R., & Campbell, C. (2018). FATHMM-XF: accurate prediction of pathogenic point mutations via extended features. *Bioinformatics*, *34*(3), 511–513. https://doi.org/10.1093/bioinformatics/btx536

Shi, Y., Ding, L., Mo, C., Luo, Y., Huang, S., Cai, S., Xia, Y., & Zhang, X. (2022). Bladder paraganglioma, gastrointestinal stromal tumor, and SDHB germline mutation in a patient with Carney-Stratakis syndrome: A case report and literature review. *Frontiers in Oncology*, *12*. https://doi.org/10.3389/fonc.2022.1030092

Shriner, D., Adeyemo, A., & Rotimi, C. N. (2020). Implications of Genetic Distance to Reference and De Novo Genome Assembly for Clinical Genomics in Africans. *MedRxiv*, 2020.09.25.20201780. https://doi.org/10.1101/2020.09.25.20201780

Sobel, R. A., & Woerner, S. (1981). Rubinstein-Taybi syndrome andnasopharyngeal rhabdomyosarcoma. *The Journal of Pediatrics*, *99*(6), 1000–1001. https://doi.org/10.1016/S0022-3476(81)80043-1

Stefan, C., Bray, F., Ferlay, J., Liu, B., & Parkin, D. M. (2017). Cancer of childhood in sub-Saharan Africa. *Ecancermedicalscience*, *11*. https://doi.org/10.3332/ECANCER.2017.755

Thavaneswaran, S., Rath, E., Tucker, K., Joshua, A. M., Hess, D., Pinese, M., Ballinger, M. L., & Thomas, D. M. (2019). Therapeutic implications of germline genetic findings in cancer. *Nature Reviews Clinical Oncology*, *16*(6), 386–396. https://doi.org/10.1038/s41571-019-0179-3

The International HapMap Consortium. (2005). A haplotype map of the human genome. *Nature*, *437*(7063), 1299–1320. https://doi.org/10.1038/nature04226

ThermoFisher Scientific. (2015). *Qubit® dsDNA BR Assay Kits*.

ThermoFisher Scientific. (2022a). *Ion GeneStudio™ S5 Instrument*.

ThermoFisher Scientific. (2022b). *Qubit® dsDNA HS Assay Kit*.

ThermoFisher Scientific. (2023). *Ion AmpliSeq™ Exome RDY Library Preparation*.

Vagher, J., Dietz, M. S., Schiffman, J. D., Kohlmann, W., & Maese, L. (2022). Germline predisposition to soft tissue sarcoma. *Journal of Cancer Metastasis and Treatment*, *8*, 31. https://doi.org/10.20517/2394-4722.2022.31

Zhang, J., Walsh, M. F., Wu, G., Edmonson, M. N., Gruber, T. A., Easton, J., Hedges, D., Ma, X., Zhou, X., Yergeau, D. A., Wilkinson, M. R., Vadodaria, B., Chen, X., McGee, R. B., Hines-Dowell, S., Nuccio, R., Quinn, E., Shurtleff, S. A., Rusch, M., … Downing, J. R. (2015). Germline Mutations in Predisposition Genes in Pediatric Cancer. *The New England Journal of Medicine*, *373*(24), 2336. https://doi.org/10.1056/NEJMOA1508054

9. Appendices

9.1. Appendix 1: Supplementary Data

9.1.1. Supplementary Table 1: Syndrome associated germline cancer predisposition genes that have previously been linked with rhabdomyosarcoma

Gene Abbreviation	Cancer Susceptibility Syndrome	Reference
TP53	Li- Fraumeni Syndrome	1, 2
CHEK2	CHEK2 associated cancer predisposition syndrome	1, 2
NF1	Neurofibromatosis type 1	1, 2
HRAS	Costello syndrome	1
DICER1	DICER1 syndrome	1, 2
PTCH1, SUFU	Gorlin syndrome	1
SOS1, PTPN11, RAF1, CBL, KRAS, NRAS, RIT1, SHOC2	Noonan syndrome	1, 2
BRAF, MAP2K1, MAP2K2	Cardiofaciocutaneous syndrome	1
MLH1, MSH2, MSH6, PMS2	Constitutional mismatch repair deficiency	1, 2
CDKN1C	Beckwith-Wiedemann syndrome	1, 2
CREBBP	Rubenstein-Tabi syndrome	1

9.1.2. Supplementary Table 2: Cancer predisposition genes that have not previously been linked to rhabdomyosarcoma

Gene Abbreviation	Reference
ALK	1
APC	1
ATM	2
BAP1	1
BMPR1A	1
BRCA1	1
BRCA2	1
CDC73	1
CDH1	1
CDK4	1
CDKN2A	1
CEBPA	1
COL7A1	2
EPCAM	1
ERCC2	2
FH	1
GATA2	1
GBA	2
MAX	1
MEN1	1
MUTYH	2
NF2	1
PALB2	1
PAX5	1
PHOX2B	1
PRKAR1A	1
PTEN	1
RB1	1
RET	1
RUNX1	1
SBDS	2
SDHA	1
SDHAF2	1
SDHB	1
SDHC	1
SDHD	1
SMAD4	1
SMARCA4	1
SMARCB1	1
STK11	1
TMEM127	1
TSC1	1

TSC2	1
VHL	1
WT1	1

9.1.3. Supplementary References

1. Li, H., Sisoudiya, S. D., Martin-Giacalone, B. A., Khayat, M. M., Dugan-Perez, S., Marquez-Do, D. A., Scheurer, M. E., Muzny, D., Boerwinkle, E., Gibbs, R. A., Chi, Y. Y., Barkauskas, D. A., Lo, T., Hall, D., Stewart, D. R., Schiffman, J. D., Skapek, S. X., Hawkins, D. S., Plon, S. E., … Lupo, P. J. (2021). Germline Cancer Predisposition Variants in Pediatric Rhabdomyosarcoma: A Report From the Children's Oncology Group. *JNCI Journal of the National Cancer Institute*, *113*(7), 875. https://doi.org/10.1093/JNCI/DJAA204
2. Kim, J., Light, N., Subasri, V., Young, E. L., Wegman-Ostrosky, T., Barkauskas, D. A., Hall, D., Lupo, P. J., Patidar, R., Maese, L. D., Jones, K., Wang, M., Genome Research Laboratory, C., Tavtigian, S. V., Wu, D., Shlien, A., Telfer, F., Goldenberg, A., Skapek, S. X., … Stewart, D. R. (2021). Pathogenic Germline Variants in Cancer Susceptibility Genes in Children and Young Adults With Rhabdomyosarcoma. *JCO Precision Oncology*, *5*, 75–87. https://doi.org/10.1200/PO.20.00218

Author Manuscript
Author Manuscript
Author Manuscript
Author Manuscript

9.2. Appendix 2: American College of Medical Genetics and Genomics and the Association for Molecular Pathology (ACMG-AMP) Guidelines for Variant Classification

(Richards et al., 2015)

	Benign		Pathogenic			
	Strong	**Supporting**	**Supporting**	**Moderate**	**Strong**	**Very Strong**
Population Data	MAF is too high for disorder *BA1/BS1* **OR** observation in controls inconsistent with disease penetrance *BS2*			Absent in population databases *PM2*	Prevalence in affecteds statistically increased over controls *PS4*	
Computational And Predictive Data		Multiple lines of computational evidence suggest no impact on gene /gene product *BP4* Missense in gene where only truncating cause disease *BP1* Silent variant with non predicted splice impact *BP7*	Multiple lines of computational evidence support a deleterious effect on the gene /gene product *PP3*	Novel missense change at an amino acid residue where a different pathogenic missense change has been seen before *PM5* Protein length changing variant *PM4*	Same amino acid change as an established pathogenic variant *PS1*	Predicted null variant in a gene where LOF is a known mechanism of disease *PVS1*
Functional Data	Well-established functional studies show no deleterious effect *BS3*		Missense in gene with low rate of benign missense variants and path. missenses common *PP2*	Mutational hot spot or well-studied functional domain without benign variation *PM1*	Well-established functional studies show a deleterious effect *PS3*	
Segregation Data	Non-segregation with disease *BS4*		Co-segregation with disease in multiple affected family members *PP1*	Increased segregation data →		
De novo Data				*De novo* (without paternity & maternity confirmed) *PM6*	*De novo* (paternity & maternity confirmed) *PS2*	
Allelic Data		Observed in *trans* with a dominant variant *BP2* Observed in *cis* with a pathogenic variant *BP2*		For recessive disorders, detected in *trans* with a pathogenic variant *PM3*		
Other Database		Reputable source w/out shared data = benign *BP6*	Reputable source = pathogenic *PP5*			
Other Data		Found in case with an alternate cause *BP5*	Patient's phenotype or FH highly specific for gene *PP4*			

Figure 1. Evidence Framework

The following chart organizes each of the criteria by the type of evidence as well as the strength of the criteria for a benign (left side) or pathogenic (right side) assertion. Evidence code descriptions can be found in Tables 3 and 4. Abbreviations: BS, benign strong; BP, benign supporting; FH, family history; LOF, loss-of-function; MAF, minor allele frequency; path., pathogenic; PM, pathogenic moderate; PP, pathogenic supporting; PS, pathogenic strong; PVS, pathogenic very strong

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1

Population, Disease-Specific, and Sequence Databases

Population Databases	
Exome Aggregation Consortium http://exac.broadinstitute.org/	Database of variants found during exome sequencing of 61,486 unrelated individuals sequenced as part of various disease-specific and population genetic studies. Pediatric disease subjects as well as related individuals were excluded.
Exome Variant Server http://evs.gs.washington.edu/EVS	Database of variants found during exome sequencing of several large cohorts of individuals of European and African American ancestry. Includes coverage data to inform the absence of variation.
1000 Genomes http://browser.1000genomes.org	Database of variants found during low-coverage and high-coverage genomic and targeted sequencing from 26 populations. Provides more diversity compared to EVS but also contains lower quality data and some cohorts contain related individuals.
dbSNP http://www.ncbi.nlm.nih.gov/snp	Database of short genetic variations (typically 50 bp or less) submitted from many sources. May lack details of originating study and may contain pathogenic variants.
dbVar http://www.ncbi.nlm.nih.gov/dbvar	Database of structural variation (typically greater than 50 bp) submitted from many sources.
Disease Databases	
ClinVar http://www.ncbi.nlm.nih.gov/clinvar	Database of assertions about the clinical significance and phenotype relationship of human variation.
OMIM http://www.omim.org	Database of human genes and genetic conditions that also contains a representative sampling of disease-associated genetic variants.
Human Gene Mutation Database http://www.hgmd.org	Database of variant annotations published in the literature. Requires fee-based subscription for much of the content.
Locus/Disease/Ethnic/Other-Specific Databases http://www.hgvs.org/dblist/dblist.html http://www.lovd.nl	The HGVS site developed a list of thousands of different databases that provide variant annotations on specific subsets of human variation. A large percentage of databases are built in the LOVD system.
DECIPHER http://decipher.sanger.ac.uk	A molecular cytogenetic database for clinicians and researchers linking genomic microarray data with phenotype using the Ensembl genome browser.
Sequence Databases	
NCBI Genome http://www.ncbi.nlm.nih.gov/genome	Source of full human genome reference sequences.
RefSeqGene http://www.ncbi.nlm.nih.gov/refseq/rsg and Locus Reference Genomic (LRG) http://www.lrg-sequence.org	Medically relevant gene reference sequence resource
MitoMap http://www.mitomap.org/MITOMAP/HumanMitoSeq	Revised Cambridge reference sequence (rCRS) for the Human Mitochondrial DNA

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 2

In Silico Predictive Algorithms

Missense prediction	Name	Website	Basis
	ConSurf	http://consurftest.tau.ac.il	Evolutionary conservation
	FATHMM	http://fathmm.biocompute.org.uk	Evolutionary conservation
	MutationAssessor	http://mutationassessor.org	Evolutionary conservation
	PANTHER	http://www.pantherdb.org/tools/csnpScoreForm.jsp	Evolutionary conservation
	PhD-SNP	http://snps.biofold.org/phd-snp/phd-snp.html	Evolutionary conservation
	SIFT	http://sift.jcvi.org	Evolutionary conservation
	SNPs&GO	http://snps-and-go.biocomp.unibo.it/snps-and-go	Protein structure/function
	Align GVGD	http://agvgd.iarc.fr/agvgd_input.php	Protein structure/function and evolutionary conservation
	MAPP	http://mendel.stanford.edu/SidowLab/downloads/MAPP/index.html	Protein structure/function and evolutionary conservation
	MutationTaster	http://www.mutationtaster.org	Protein structure/function and evolutionary conservation
	MutPred	http://mutpred.mutdb.org	Protein structure/function and evolutionary conservation
	PolyPhen-2	http://genetics.bwh.harvard.edu/pph2	Protein structure/function and evolutionary conservation
	PROVEAN	http://provean.jcvi.org/index.php	Alignment and measurement of similarity between variant sequence and protein sequence homolog
	nsSNPAnalyzer	http://snpanalyzer.uthsc.edu	Multiple sequence alignment and protein structure analysis
	Condel	http://bg.upf.edu/condel/home	Combines SIFT, PolyPhen-2 and MutationAssessor
	CADD	http://cadd.gs.washington.edu	Contrasts annotations of fixed/nearly fixed derived alleles in humans with simulated variants
Splice site prediction			
	GeneSplicer	http://www.cbcb.umd.edu/software/GeneSplicer/gene_spl.shtml	Markov models
	Human Splicing Finder	http://www.umd.be/HSF	Position-dependent logic

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Missense prediction	Name	Website	Basis
	MaxEntScan	http://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html	Maximum entropy principle
	NetGene2	http://www.cbs.dtu.dk/services/NetGene2	Neural networks
	NNSplice	http://www.fruitfly.org/seq_tools/splice.html	Neural networks
	FSPLICE	http://www.softberry.com/berry.phtml?topic=fsplice&group=programs&subgroup=gfind	
Nucleotide conservation prediction			
	GERP	http://mendel.stanford.edu/sidowlab/downloads/gerp/index.html	
	PhastCons	http://compgen.bscb.cornell.edu/phast/	
	PhyloP	http://compgen.bscb.cornell.edu/phast/	

In-silico tools/software prediction programs used for sequence variant interpretation

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 3

Criteria for Classifying Pathogenic Variants

Very strong evidence of pathogenicity

PVS1 Null variant (nonsense, frameshift, canonical +/−1 or 2 splice sites, initiation codon, single or multi-exon deletion) in a gene where loss of function (LOF) is a known mechanism of disease

Caveats:

- Beware of genes where LOF is not a known disease mechanism (e.g. *GFAP, MYH7*)
- Use caution interpreting LOF variants at the extreme 3' end of a gene
- Use caution with splice variants that are predicted to lead to exon skipping but leave the remainder of the protein intact
- Use caution in the presence of multiple transcripts

Strong evidence of pathogenicity

PS1 Same amino acid change as a previously established pathogenic variant regardless of nucleotide change

Example: Val->Leu caused by either G>C or G>T in the same codon

Caveat: Beware of changes that impact splicing rather than at the amino acid/protein level

PS2 *De novo* (both maternity and paternity confirmed) in a patient with the disease and no family history

Note: Confirmation of paternity only is insufficient. Egg donation, surrogate motherhood, errors in embryo transfer, *etc*. can contribute to non-maternity

PS3 Well-established *in vitro* or *in vivo* functional studies supportive of a damaging effect on the gene or gene product

Note: Functional studies that have been validated and shown to be reproducible and robust in a clinical diagnostic laboratory setting are considered the most well-established

PS4 The prevalence of the variant in affected individuals is significantly increased compared to the prevalence in controls

Note 1: Relative risk (RR) or odds ratio (OR), as obtained from case-control studies, is >5.0 and the confidence interval around the estimate of RR or OR does not include 1.0. See manuscript for detailed guidance.

Note 2: In instances of very rare variants where case-control studies may not reach statistical significance, the prior observation of the variant in multiple unrelated patients with the same phenotype, and its absence in controls, may be used as moderate level of evidence.

Moderate evidence of pathogenicity

PM1 Located in a mutational hot spot and/or critical and well-established functional domain (*e.g.* active site of an enzyme) without benign variation

PM2 Absent from controls (or at extremely low frequency if recessive) (see Table 6) in Exome Sequencing Project, 1000 Genomes or ExAC

Caveat: Population data for indels may be poorly called by next generation sequencing

PM3 For recessive disorders, detected in *trans* with a pathogenic variant

Note: This requires testing of parents (or offspring) to determine phase

PM4 Protein length changes due to in-frame deletions/insertions in a non-repeat region or stop-loss variants

PM5 Novel missense change at an amino acid residue where a different missense change determined to be pathogenic has been seen before

Example: Arg156His is pathogenic; now you observe Arg156Cys

Caveat: Beware of changes that impact splicing rather than at the amino

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

	acid/protein level
PM6	Assumed *de novo*, but without confirmation of paternity and maternity
Supporting evidence of pathogenicity	
PP1	Co-segregation with disease in multiple affected family members in a gene definitively known to cause the disease
	Note: May be used as stronger evidence with increasing segregation data
PP2	Missense variant in a gene that has a low rate of benign missense variation and where missense variants are a common mechanism of disease
PP3	Multiple lines of computational evidence support a deleterious effect on the gene or gene product (conservation, evolutionary, splicing impact, etc)
	Caveat: As many *in silico* algorithms use the same or very similar input for their predictions, each algorithm should not be counted as an independent criterion. PP3 can be used only once in any evaluation of a variant.
PP4	Patient's phenotype or family history is highly specific for a disease with a single genetic etiology
PP5	Reputable source recently reports variant as pathogenic but the evidence is not available to the laboratory to perform an independent evaluation

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 4

Criteria for Classifying Benign Variants

Stand-Alone evidence of benign impact	
BA1	Allele frequency is above 5% in Exome Sequencing Project, 1000 Genomes, or ExAC
Strong evidence of benign impact	
BS1	Allele frequency is greater than expected for disorder (see table 6)
BS2	Observed in a healthy adult individual for a recessive (homozygous), dominant (heterozygous), or X-linked (hemizygous) disorder with full penetrance expected at an early age
BS3	Well-established *in vitro* or *in vivo* functional studies shows no damaging effect on protein function or splicing
BS4	Lack of segregation in affected members of a family Caveat: The presence of phenocopies for common phenotypes (*i.e.* cancer, epilepsy) can mimic lack of segregation among affected individuals. Also, families may have more than one pathogenic variant contributing to an autosomal dominant disorder, further confounding an apparent lack of segregation.
Supporting evidence of benign impact	
BP1	Missense variant in a gene for which primarily truncating variants are known to cause disease
BP2	Observed in *trans* with a pathogenic variant for a fully penetrant dominant gene/disorder; or observed in *cis* with a pathogenic variant in any inheritance pattern
BP3	In-frame deletions/insertions in a repetitive region without a known function
BP4	Multiple lines of computational evidence suggest no impact on gene or gene product (conservation, evolutionary, splicing impact, etc) Caveat: As many *in silico* algorithms use the same or very similar input for their predictions, each algorithm cannot be counted as an independent criterion. BP4 can be used only once in any evaluation of a variant.
BP5	Variant found in a case with an alternate molecular basis for disease
BP6	Reputable source recently reports variant as benign but the evidence is not available to the laboratory to perform an independent evaluation
BP7	A synonymous (silent) variant for which splicing prediction algorithms predict no impact to the splice consensus sequence nor the creation of a new splice site AND the nucleotide is not highly conserved

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 5

Rules for Combining Criteria to Classify Sequence Variants

Pathogenic

1. 1 Very Strong (PVS1) *AND*
   - **a.** ≥1 Strong (PS1–PS4) *OR*
   - **b.** ≥2 Moderate (PM1–PM6) *OR*
   - **c.** 1 Moderate (PM1–PM6) and 1 Supporting (PP1–PP5) *OR*
   - **d.** ≥2 Supporting (PP1–PP5)
2. ≥2 Strong (PS1–PS4) *OR*
3. 1 Strong (PS1–PS4) *AND*
   - **a.** ≥3 Moderate (PM1–PM6) *OR*
   - **b.** 2 Moderate (PM1–PM6) *AND* ≥2 Supporting (PP1–PP5) *OR*
   - **c.** 1 Moderate (PM1–PM6) *AND* ≥4 Supporting (PP1–PP5)

Likely Pathogenic

1. 1 Very Strong (PVS1) *AND* 1 Moderate (PM1–PM6) *OR*
2. 1 Strong (PS1–PS4) *AND* 1–2 Moderate (PM1–PM6) *OR*
3. 1 Strong (PS1–PS4) *AND* ≥2 Supporting (PP1–PP5) *OR*
4. ≥3 Moderate (PM1–PM6) *OR*
5. 2 Moderate (PM1–PM6) *AND* ≥2 Supporting (PP1–PP5) *OR*
6. 1 Moderate (PM1–PM6) *AND* ≥4 Supporting (PP1–PP5)

Benign

1. 1 Stand-Alone (BA1) *OR*
2. ≥2 Strong (BS1–BS4)

Likely Benign

1. 1 Strong (BS1–BS4) and 1 Supporting (BP1–BP7) *OR*
2. ≥2 Supporting (BP1–BP7)

*Variants should be classified as Uncertain Significance if other criteria are unmet or the criteria for benign and pathogenic are contradictory.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 6

Assessment of variant frequency in the general population for curated variant classification. Variants in some common genetic disorders with their known incidences for dominant (Kabuki syndrome and CHARGE syndrome), X linked diseases (Rett syndrome), and carrier frequencies for recessive disorders (Cystic Fibrosis, Phenylketonuria (PKU), Medium Co-Acyl Dehydrogenase Deficiency (MCADD), Autosomal Recessive Polycystic Kidney Disease (ARPKD), GJB2 associated hearing loss, and Hemochromatosis) are shown in the table. Some variants found in the ESP-GO database in these genes are listed with the variant classifications based on evidence and concordance with allele frequencies in African American (AA) and European American (EA) subpopulations. Variants that have an allele frequency greater than expected for the disorder in at least one subpopulation (AA or EA) of individuals not known to have the disorder can be considered to have strong evidence of benign impact (BS1). Variants known to be pathogenic for dominant disorders should have allele frequencies in the general population below the population disease incidence or and pathogenic variants for recessive disorders should have heterozygous frequencies consistent with their disease incidence. The allele frequencies shown in the table for certain variants are lower for certain benign and higher for certain pathogenic variants than the disease incidence/carrier frequency; therefore other data would be required to classify these variants. Variants represented in the concordance column are designated as partial when a subpopulation frequency does not conform as above or when concordance is not achieved. Variants are designated as having partial concordance when a single subpopulation frequency is inconsistent with the classification.

Disease	Gene	Inheritance	Population	Incidence	Carrier Frequency	Common Mutation	Variant Classification	ESP6500 AA MAF	ESP6500 EA MAF	ESP6500 All MAF	Concordance	Evidence to support classification
Cystic fibrosis	CFTR	AR	Caucasian	0.031%	3.6%	p.F508del	Ex24:p.F508del (Pathogenic)	n/a	n/a	n/a	n/a	Multiple publications (note variant not available in EVS)
							Ex11:c.1523T>G / p.F508C (Benign)	0.070%	0.150%	0.120%	No	Kobayashi (1990) Am J Hum Genet 47, 61
							Ex23:c.3870A>G / p.(=) (Benign)	15.090%	2.970%	7.070%	Partial	AA MAF
							5'UTR:c.-8G>C (Benign)	1.160%	5.550%	4.060%	Partial	EA MAF
							IVS6:c.743+40A>G (Benign)	0.700%	5.190%	3.670%	Partial	EA MAF
Phenylketoneuria	PAH	AR	North European	0.010%	2.0%		Ex12:c.1242C>T / p.(=) (Benign)	0.360%	1.310%	0.990%	No	PAH database
							Ex12:c.1278T>C / p.(=) (Benign)	13.550%	0.090%	4.650%	Partial	AA MAF
							IVS12:c.1316-35C>T (Benign)	0.320%	2.630%	1.850%	Partial	EA MAF
							Ex9:c.963C>T / p.(=) (Benign)	5.170%	0.000%	1.750%	Partial	AA MAF
MCADD	ACADM	AR	Not specific	0.006%	1.5%	p.K329E aka p.K304E	Ex7:c.489T>G / p.(=) (Benign)	7.010%	0.050%	2.410%	Partial	AA MAF
ARPKD	PKHD1	AR	Not specific	0.005%	1.4%		IVS20:c.1964+17G>T (Benign)	0.200%	0.810%	0.610%	No	Multiple publications
							Ex61:c.10515C>A / p.S3505R (Benign)	0.230%	1.130%	0.820%	No	Multiple publications
							Ex66:c.11738G>A / p.R3913H (Benign)	1.270%	0.000%	0.430%	No	AA MAF

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Disease	Gene	Inheritance	Population	Incidence	Carrier Frequency	Common Mutation	Variant Classification	ESP6500 AA MAF	ESP6500 EA MAF	ESP6500 All MAF	Concordance	Evidence to support classification
							Ex17:c.1587T>C / p.(=) (Benign)	1.380%	6.860%	5.010%	Partial	EA MAF
							Ex65:c.11525G>T / p.R3842L (Benign)	0.360%	2.430%	1.730%	Partial	EA MAF
							Ex61:c.10585G>C / p.E3529Q (Benign)	3.950%	0.010%	1.350%	Partial	AA MAF
Rett syndrome	MECP2	X Linked	Not specific	0.012%	de novo		Ex4:c.1161C>T / p.(=) (Benign)	0.030%	0.000%	0.010%	Partial	AA MAF
							Ex4:c.608C>T / p.T203M (Benign)	0.000%	0.060%	0.040%	Partial	Multiple publications
							Ex4:c.683C>G / p.T228S (Pathogenic)	0.830%	0.000%	0.300%	Partial	RETT database
Kabuki syndrome	KMT2D (MLL2)	AD	Not specific	0.003%	de novo		IVS31:c.8047-15C>T (Benign)	0.000%	0.020%	0.020%	Partial	EA MAF
							Ex31:c.6836G>A / p.Gly2279E (Benign)	0.000%	0.120%	0.080%	Partial	EA MAF
CHARGE syndrome	CHD7	AD	Not specific	0.010%	de novo		Ex2:c.309G>A / p.(=) (Benign)	1.460%	0.000%	0.490%	Partial	AA MAF
							Ex31:c.6478G>A / p.A2160T (Benign)	1.250%	0.000%	0.390%	Partial	AA MAF
							Ex2:c.856A>G / p.R286Gly (Benign)	0.780%	0.000%	0.250%	Partial	AA MAF
GJB2 associated hearing loss	GJB2	AR	Not specific	0.067%	2.5%	c.35delG	Ex2:c.35delG (Pathogenic)	0.090%	1.080%	0.740%	No	Multiple publications
Hemochromatosis	HFE	AR	All	0.040%	8.3%	p.C282Y	Ex4:c.845G>A / p.C282Y (Other Reportable)	1.520%	6.410%	4.750%	No	Multiple publications

9.3. Appendix 3: Human Research Ethics Committee (Medical) Clearance Certificate No. M230783

UNIVERSITY OF THE WITWATERSRAND, JOHANNESBURG	HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

Office of the Deputy Vice-Chancellor (Research and Innovation)

TO:	Ms G Pillhofer School of Pathology Division of Human Genetics Medical School University E-mail: 2097324@students.wits.ac.za
CC:	Supervisor: Drs L Lamola and K Mnika <Lindie.Lamola@wits.ac.za> and <HREC-Medical Research Office@wits.ac.za>
FROM:	Mr Iain Burns Human Research Ethics Committee (Medical) Tel: 011 717 1252 E-mail: Iain.Burns@wits.ac.za
DATE:	2023/08/18
REF:	R14/49
PROTOCOL NO:	**M230783** (This is your ethics application reference number. Please quote it in all enquiries, oral or written, relating to this study.)
PROJECT TITLE:	*Germline cancer predisposition variants in paediatric Rhabdomyosarcoma*

Please find attached the Clearance Certificate for the above project. I hope it goes well and that an article in a recognized publication comes out of it. This will reflect well on your professional standing and contribute to Government funding of the University.

MSWorks2000/Iain0007/Clearscan.wps

UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

R49 Ms G Pillhofer

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M230783

NAME: (Principal Investigator)	Ms G Pillhofer
DEPARTMENT:	School of Pathology Division of Human Genetics Medical School University
PROJECT TITLE:	*Germline cancer predisposition variants in paediatric Rhabdomyosarcoma*
DATE CONSIDERED:	Ad hoc
DECISION:	Approved unconditionally
CONDITIONS:	Sub-study under M18/08/55
NOTE:	If contact information regarding student study participants is required, please contact the Registrar's office - <Nicoleen.Potgieter@wits.ac.za>
SUPERVISOR:	Drs L Lamola and K Mnika
APPROVED BY:	Professor P Ruff, Chairperson, HREC (Medical)
DATE OF APPROVAL:	2023/08/18 **EXPIRY DATE:** 2028/08/17

This Clearance Certificate is valid for 5 years from the date of approval. An extension may be applied for.

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office secretariat on the 3rd floor, Phillip Tobias Building, Parktown, University of the Witwatersrand, Johannesburg.

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 submit details to the Committee. **I agree to submit a yearly progress report**. When a funder requires annual re-certification, the application date will be one year after the date when the study was initially reviewed. In this case, the study was initially reviewed in **July** and therefore reports and re-certification will be due in the month of **July** each year. Unreported changes to the study may invalidate the clearance given by the HREC (Medical).

Signature of Principal Investigator

2023/08/21
Date

9.4. Appendix 4: Human Research Ethics Committee (Medical) Clearance Certificate No. M180855

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG	HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)

Office of the Deputy Vice-Chancellor (Research & Post Graduate Affairs)

TO: Dr L Lamola, et al
School of Pathology
Department of Human Genetics
National Health Laboratory Service

E-mail: Lindie.Lamola@wits.ac.za

CC: Supervisor: Not applicable <>
and <HREC-Medical.ResearchOffice@wits.ac.za>

FROM: Iain Burns
Human Research Ethics Committee (Medical)
Tel: 011 717 1252

E-mail: Iain.Burns@wits.ac.za

DATE: 26/02/2019

REF: *R14/49*

PROTOCOL NO: **M180855** *(This is your ethics application study reference number. Please quote this reference number in all correspondence relating to this study)*

PROJECT TITLE: *Inherited cancer predisposing mutations in a cohort of childhood cancer patients*

Please find attached the Clearance Certificate for the above project. I hope it goes well and that an article in a recognized publication comes out of it. This will reflect well on your professional standing and contribute to the Government funding of the University.

MSWorks2000/Iain0007/Clearscan.wps

UNIVERSITY OF THE WITWATERSRAND JOHANNESBURG

R14/49 Dr L Lamola, et al

HUMAN RESEARCH ETHICS COMMITTEE (MEDICAL)
CLEARANCE CERTIFICATE NO. M180855

NAME: (Principal Investigator)	Dr L Lamola, et al
DEPARTMENT:	School of Pathology Department of Human Genetics National Health Laboratory Service
PROJECT TITLE:	Inherited cancer predisposing mutations in a cohort of childhood cancer patients
DATE CONSIDERED:	31/08/2018
DECISION:	Approved unconditionally
CONDITIONS:	
SUPERVISOR:	Not applicable
APPROVED BY:	Dr CB Penny, Chairperson, HREC (Medical)
DATE OF APPROVAL:	26/02/2019

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

DECLARATION OF INVESTIGATORS

To be completed in duplicate and **ONE COPY** returned to the Research Office Secretary on 3rd floor, Phillip V Tobias Building, Parktown, University of the Witwatersrand, Johannesburg.

I/We fully understand the conditions under which I am/we are authorised 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 to the Committee. **I agree to submit a yearly progress report**. When a funder requires annual re-certification, the application date will be one year after the date of the meeting when the study was initially reviewed. In this case, the study was initially reviewed in **August** and will therefore reports and re-certification will be due early in the month of **August** each year. Unreported changes to the application may invalidate the clearance given by the HREC (Medical).

Principal Investigator Signature

26/02/2019
Date

PLEASE QUOTE THE PROTOCOL NUMBER IN ALL ENQUIRIES

9.5. Appendix 5: Approved Research Protocol

UNIVERSITY OF THE
WITWATERSRAND,
JOHANNESBURG

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

Reference: Mrs Sandra Benn
E-mail: sandra.benn@wits.ac.za

11 August 2023
Person No: 2097324
PAG

Miss GP Pillhofer
Po Box 76356
Wendywood
2144
South Africa

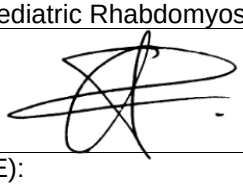
Dear Miss Gabriella Pillhofer

Master of Science in Medicine: Approval of Title

We have pleasure in advising that your proposal entitled *Germline Cancer Predisposition in Paediatric Rhabdomyosarcoma* 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

Mrs Sandra Benn
Faculty Registrar
Faculty of Health Sciences

CANDIDATE'S SURNAME: [Please print]	PILLHOFER	FIRST NAME/S:	GABRIELLA PETA	STUDENT NUMBER:	2097324
CURRENT QUALIFICATIONS: BHSc - Biomedical Sciences (WITS); BHSc (Hons) – Human Genetics (WITS)					
TEL: -	CELL: +27 79 024 8505	E-MAIL: 2097324@students.wits.ac.za		FAX: -	
DEGREE FOR WHICH PROTOCOL IS BEING SUBMITTED: Master of Science in Medicine in the field of Genomic Medicine					
PART-TIME OR FULL-TIME: Full-Time					
FIRST REGISTERED FOR THIS DEGREE:	TERM: 1	YEAR: 2023			
DEPARTMENT: Department of Human Genetics					
TITLE OF PROPOSED RESEARCH: Germline Cancer Predisposition in Paediatric Rhabdomyosarcoma					
CANDIDATE'S SIGNATURE: 		DATE: 2023/04/25			
SUPERVISOR 1 (NAME & SURNAME): Lindie Lamola		% Supervision: 60%			
SUPERVISOR'S QUALIFICATIONS: PhD					
SUPERVISOR'S DEPARTMENT: Human Genetics					
SUPERVISOR'S ADDRESS / TEL / E-MAIL: lindie.lamola@wits.ac.za					
SUPERVISOR 2 (NAME & SURNAME): Khuthala Mnika		% Supervision: 40%			
SUPERVISOR'S QUALIFICATIONS: PhD					
SUPERVISOR'S ADDRESS / TEL / E-MAIL: khuthala.mnika@wits.ac.za					

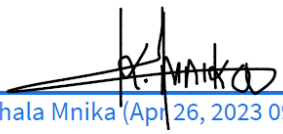
SYNOPSIS OF RESEARCH:*(Brief summary of proposed research project; between 200-300 words only; with sub-headings: an introduction and justification for study, aim/s, proposed methodology and expected outcome/s)*

Introduction: Rhabdomyosarcoma (RMS) is a soft tissue sarcoma that affects mainly paediatric patients. The prognosis of children with RMS is poor and the disease burden is high. Cancer predisposition genes (CPGs) are a mutated copy of a cancer associated in the germline, predisposing individuals with these germline mutations to developing cancer at younger ages than is expected in the general population. Despite the existing research, both RMS and mutations in its suspected CPGs are rare, and study populations are small and limited. The burden of these germline mutations in RMS cannot be definitively determined. Furthermore, in the understudied African population, the link is even less clear. This provides the basis for the research problem looking to investigate RMS germline predisposition variants in an African population.

Aim: The aim of this project is to identify predisposition variants in cancer predisposition genes in a group of children with diagnosed rhabdomyosarcoma.

Methods: Study participants have been selected from the Paediatric Cancer parent study. These are eight paediatric patients with RMS. Whole exome sequencing (WES) will be performed on four of these samples as part of this study, whereas the other four samples have already been sequenced. Exome sequencing will be performed using the Ion Torrent Ion S5 Sequencing System. The variants identified by the Ion Reporter variant analysis software will be further annotated and prioritised. Variants will be annotated using a gene based approached, and then prioritised and filtered based on variant quality, minor allele frequency and predicted pathogenicity. Following prioritisation, variants will be classified according to the American College of Medical Genetics and Genomics and Association of Medical Pathology guidelines.

Expected Outcome: This study forms part of a larger study investigating the genetics of paediatric cancers in South Africa. It is the hope that the knowledge gained from the WES data as part of this study will advance understanding of germline genetics variations that may cause RMS in our population.

<table>
<tr><td colspan="2">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</td><td>IF Y SUPPLY ETHICS CLEARANCE CERTIFICATE AS ATTACHMENT AND INCLUDE ETHICS NUMBER HERE:</td></tr>
<tr><td colspan="3">As supervisor/s, I/we confirm that I have read the protocol which has been submitted for assessment.</td></tr>
<tr><td>SIGNATURE OF SUPERVISOR/S:</td><td>
Lindie Lamola (Apr 26, 2023 09:04 GMT+2)
……</td><td>
Khuthala Mnika (Apr 26, 2023 09:23 GMT+2)
……</td></tr>
<tr><td>SIGNATURE PG OFFICE STAFF

……..</td><td>REGISTERED

YES….. NO…..</td><td>STAMP</td></tr>
</table>

a

MSc Research Proposal – HUMG7025A

Germline Cancer Predisposition Variants in Paediatric Rhabdomyosarcoma

By

Gabriella Peta Pillhofer

2097324

Master of Science in Medicine in the field of Genomic Medicine

Supervised by:

Dr Lindie Lamola

Dr Khuthala Mnika

Contents

1. Literature Review 45
1.1. Introduction 45
1.2. Cancer Predisposition Genes 46
1.3. RMS-associated Cancer Susceptibility Genes 47
2. Aims and Objectives 50
2.1. Aim 50
2.2. Objectives 50
3. Methods 50
3.1. Sample Selection 50
3.2. Quantification of DNA 51
3.3. Whole Exome Sequencing 51
3.4. Variant Annotation 52
3.5. Variant Prioritisation 52
3.5.1 Gene Selection 52
3.5.2 Variant Filtering and Minor Allele Frequencies 53
3.5.3 *In-silico* Prediction 53
3.5.4 Variant Classification 53
4. Time Frame 54
5. Budget and Funding 55
5.1. Budget 55
5.2. Funding 55
6. Ethics 55
7. Problems and Limitations 55
8. References 57
10. Supplementary Material 61
10.1. Supplementary Table 1 61
10.2. Supplementary Table 2 62
10.3. Supplementary References 63

- **Literature Review**

  - Introduction

Rhabdomyosarcoma (RMS) is a soft tissue sarcoma that originates from undifferentiated skeletal muscle cells (myoblasts). Known as a paediatric cancer, RMS is the most common tissue sarcoma in children and adolescents and has over 50% of cases occurring in individuals under the age of 10 (Ognjanovic et al., 2009). There are four main types of RMS: embryonal RMS (ERMS) (60-70% of cases), alveolar RMS (ARMS) (20-30%), pleomorphic RMS (10%) and sclerosing RMS (10%) (Kaseb et al., 2022; Zhang et al., 2015). The ERMS and ARMS subtypes are the most likely to be identified in minors, with ARMS in the extremities being most common in adolescents, and ERMS in hollow organs being common in children (Kaseb et al., 2022).

The prognosis of children with RMS is poor, with only a 63% four-year event free survival rate, and a decreased survival in individuals with RMS metastases (Hawkins et al., 2013). In the United States of America, the incidence of RMS is 4.71 per million individuals under 20 years of age (Martin-Giacalone et al., 2021). The incidence of RMS in Europe is slightly increased in comparison, with an incidence of 5.4 per million children younger than 15 years of age (Pastore et al., 2006).

Despite the severity of the RMS disease burden, much of the existing literature reports on patients of European ancestry in North America and Europe, and little to no investigation of RMS in African populations exists. This is despite an incidence rate of RMS in African populations that matches, and in some areas exceeds, the incidence in populations of majority European ancestry (between 0.6 and 16.3 cases per million in children aged 0-14 years across African countries (Stefan et al., 2017). This cancer has a large impact on the African population and yet little research exists regarding the putative genetic causes in an African population. This provides the basis for the research problem looking to investigate RMS germline predisposition variants in an African population.

- Cancer Predisposition Genes

Cancer predisposition genes (CPGs) were first discovered as a result of Alfred Knudson's two-hit hypothesis – which proposed that some individuals are at an increased risk of developing malignancies as they carry a mutated copy of a cancer associated gene (tumour suppressor or oncogene) in the germline. This means all cells would need to develop only one additional mutation to allow malignant tumour formation (Knudson, 1971). This predisposes individuals with these germline mutations to developing cancer at younger ages and in more tissues than is expected in the general population (Knudson, 1971; McGee & Nichols, 2016).

A majority of CPGs can be categorised at tumour suppressor genes or proto-/oncogenes. These are both categories of gene that have been associated with the development of cancers. Tumour suppressor genes encode proteins that function to control the cell cycle and restrict unsuitable cellular growth (Chow, 2010). For example, *TP53* is a tumour suppressor gene coding for tumour protein p53. In typical cellular growth, the p53 pathway is activated in reaction to cellular stress or DNA damage. This pathway arrests cell development until DNA repair processes can occur, or in some cases induces apoptosis. In malignancy, this pathway is inactivated – resulting in cancer cells' ability to evade DNA repair and apoptosis (Bourdon, 2007). These tumour suppressor genes are usually active in all typical cells and ensure that uncontrolled, malignant cell growth does not occur. However, mutations in these genes in cancer cells allows for carcinogenesis through sustained growth signalling and avoiding apoptosis (Chow, 2010; Hanahan & Weinberg, 2011).

Conversely, oncogenes are not active in regular cell growth. In non-malignant cells, these genes exist as protooncogenes. Protooncogenes typically function to promote cell cycle signalling, controlling cellular growth and division. When deleterious mutations occur in these genes, they are activated as oncogenes which enable uncontrolled growth of malignant cells (Chow, 2010). Genes encoding RAS proteins (e.g. *HRAS, NRAS* and *KRAS)*

act as oncogenes in cancer cells. These proteins act to hydrolyse and activate guanosine triphosphate, allowing for sustained proliferative cellular signalling. In malignancy, this RAS pathway allows cells to overcome cell signal checkpoints and continue to grow and reproduce (Chow, 2010; Pylayeva-Gupta et al., 2011).

A study by Fiala et al. (2021) suggests that 18% of paediatric cancers can be attributed to germline mutations in CPGs. This is higher than was previously thought as earlier studies estimated this to only be 8.5% (Zhang et al., 2015). It can thus be surmised that CPGs have an important role in identifying cancer risk factor in paediatric populations.

By identifying CPGs associated with specific cancer types, patients who are at an increased risk of that cancer can be earmarked for screening programmes. This allows for early treatment and intervention should malignancy be detected (Rahman, 2014a). Moreover, identifying CPGs in patients with confirmed cancer diagnoses can also be beneficial. For instance, more radical or investigative surgery techniques may be employed in an individual with a CPG mutation to increase the ability to remove all possible malignancy. Additionally, certain CPG mutations alter chemotherapeutic efficacy and so knowing an individual's CPG mutation status can guide the specific chemotherapy drug to be used. Furthermore, in certain cancers, CPG mutations alter the prognosis of an individual (McGee & Nichols, 2016; Rahman, 2014a).

- RMS-associated Cancer Susceptibility Genes

The heterogeneous genetic aetiology of RMS is difficult to categorise, and research efforts are limited. However, there are several CPG mutations that have been previously reported to underlie RMS (Supplementary Table 1), some of which are overviewed below. Additionally, many genes have been previously reported to be associated with increased cancer risk but have not yet been linked with RMS (Supplementary Table 2).

A study performed by Li et al. (2021) aimed to determine the prevalence of CPGs in a cohort of paediatric patients diagnosed with RMS in order to provide an accurate assessment of the burden of cancer susceptibility in paediatric RMS. This study investigated exome-sequencing data from an unselected cohort of European (67.7%), East Asian (4.9%), South Asian (1.4%), African American(13%), and Hispanic (13%) children who were recently diagnosed with RMS. Germline CPGs were in 7.3% of patients, and included variants in tumor protein p53 *(TP53)*, neurofibromin 1 *(NF1)* and HRas proto-oncogene *(HRAS)*. This study provides recent, substantial evidence for the association between certain CPSs (Li-Fraumeni Syndrome, Neurofibromatosis type 1 and Costello Syndrome) and RMS. While the study by Li et al. (2021) did examine multiple CPSs and RMS, the rarity of both RMS and most of the familial cancer syndromes means a majority of research performed investigates only one CPG and its association with RMS, or cancers in general.

Kratz et al. (2011) performed a retrospective review of cancer and certain CPSs to better quantify cancer risk in individuals with CPSs. Their review identified 29 patients with CS, all of which had cancer. 19 of these patients had an RMS diagnosis with a median age of 2.3 years. Most of these cases were ERMS diagnoses, once again aligning with the general population (Kaseb et al., 2022).

The link between dicer 1 ribonuclease III *(DICER1)* variations and RMS is established in a study performed by Doros et al. (2012). This study aimed to assess the link between inherited *DICER1* mutations and RMS. Initially, a cohort of DICER1 syndrome families was investigated for ERMS diagnoses. This revealed 4 patients with ERMS – all of whom had germline loss of function mutations in DICER1 confirmed with sequencing. These patients were all $\leq$ 8 years of age at the diagnosis of ERMS. Researchers then investigated 52 cases of ERMS that were previously assumed to be sporadic occurrences.

Deleterious mutations in *DICER1* were identified in 2/52 patients. Doros et al. (2012) confirm the link between familial DICER1 syndrome and paediatric onset of ERMS. Additionally, they comment on the potential link between ERMS development, and the miRNA pathway given the *DICER1* mutations identified in the 'sporadic' ERMS tumour tissue.

In a study performed by Diller et al. (1995), 33 paediatric patients with RMS were assessed for a mutation in *TP53*. Peripheral blood was analysed, and mutations in the gene were detected in 3/33 patients. The use of peripheral blood for the genomic testing, combined with the RMS may suggest these three patients had undiagnosed LFS. The authors note that all three children with *TP53* mutations were under the age of 3, and that no *TP53* mutations were identified in the population older than 3 years of age. This corroborates the idea that CPSs cause cancer to be developed at an earlier age (Rahman, 2014a).

Some CPSs are more strongly associated with a specific subtype of RMS. While NF1 patients have an increased prevalence of all RMS subtypes as compared to unaffected individuals, it is most notably associated with ERMS. In a retrospective analysis of RMS patients who had NF1, all individuals (16/16) were identified as having the embryonal subtype. The mean age at diagnosis was 2.5 years, and most patients had sarcomas in the pelvic viscera (Crucis et al., 2015), which is consistent with the population dynamics for ERMS (Kaseb et al., 2022).

Familial cancer syndromes are associated both with increased overall risk for developing associated cancers, and with the development of cancer at a younger age than is anticipated in the general population (Rahman, 2014).

Despite the existing research, both RMS and mutations in its suspected CPGs are rare, and study populations are small and limited. As a result, the burden of these germline mutations in RMS cannot be definitively determined. Furthermore, in the understudied African population, the link is even less clear. Identifying predisposition variants in CPGs in a group

of South African children with RMS could aid in informing the inherited cancer risk in the population.

- **Aims and Objectives**
  - Aim

The aim of this project is to identify germline variants in cancer predisposition genes in a group of children with diagnosed rhabdomyosarcoma.

  - Objectives

The objectives are outlined below:

i. To perform whole exome sequencing of samples from children diagnosed with RMS using the Ion Torrent platform and then analyse data using freely available tools to identify germline mutations.
ii. To annotate and prioritise putative RMS causing variants and verify presence of variants using a genome browser.
iii. To classify variants according to American College of Medical Genetics and Genomics (ACMG) and Association of Medical Pathology (AMP) guidelines.

- **Methods**
  - Sample Selection

Study participants have been selected from the Paediatric Cancer (PeCan) parent study. These are eight paediatric patients with RMS who were recruited as part of the larger study and have consented for their samples to be utilised in further research. Exome sequencing will be performed on four of these samples as part of this study, whereas the other four samples have already been sequenced. All eight samples will be annotated and prioritised once sequencing is completed, as overviewed in Figure 1 below.

Clinical, histological and demographic data of the patients was collected as part of the parent study and will be made available on the University of the Witwatersrand's RedCap database (https://redcap.core.wits.ac.za/redcap/) for phenotype correlations to be performed.

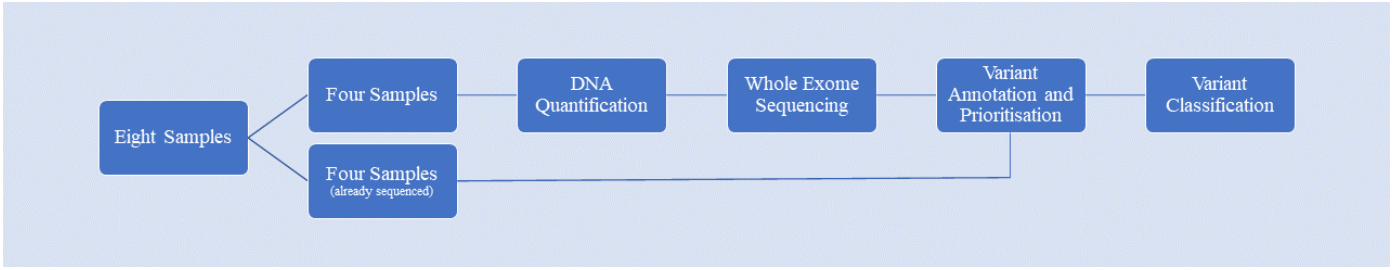


Figure 1: Ideogram representation of the proposed methods for this study. Of the eight selected samples, four will be quantified and sequenced as part of this study, whereas four have already been sequenced. All eight samples will be annotated, prioritised, and classified.

- Quantification of DNA

Quantification of DNA will be performed using both Nanodrop spectrophotometry and Qubit fluorometry. Nanodrop spectrophotometry assesses the concentration and purity of nucleic acids by measuring sample absorbance. The absorbance measured for different compounds is used to assess the concentration of specific nucleic acids, as well as to detect any contaminants, such as salts or proteins (ThermoFisher Scientific). Additionally, quantification using Qubit fluorometry will be performed to measure the quantity of double stranded DNA (dsDNA) present in the sample. This is imperative in next generation sequencing techniques such as whole exome sequencing (WES) as it is this dsDNA that is evaluated by the sequencer (ThermoFisher Scientific).

- Whole Exome Sequencing

Exome sequencing will be performed using the IonTorrent Ion S5 Sequencing System (IonTorrent & ThermoFisher Scientific). This is a next-generation sequencing system that allows for high-throughput, simplified WES. Initially, the sequencing library will be prepared. This sequencing library consists of DNA fragments with known adaptor sequences adhered to their ends, which are selected based on fragment size. Quantification of the library will then be performed using both Qubit and TapeStation.

Qubit is used to assess the quality and quantity of dsDNA in the prepared library (ThermoFisher Scientific), while TapeStation is used specifically in library preparation for assessment of the quality of the library, as well as to quantify the library size (Agilent).

Following the library set up and quantification, the templates will be prepared, and the sequencing chip loaded. These steps are automated using the Ion Chef system (ThermoFisher Scientific, 2019). Template preparation amplifies the target fragment in the sample, ensuring a strong enough signal for detection by the sequencer. Once the chip is prepared by the Ion Chef system, it can be loaded into the Ion S5 sequencer. The sequencer produces data in the form of a variant report, that lists variants identified in the test samples (ThermoFisher Scientific). The variants identified by the Ion Reporter analysis software will be further annotated and prioritised.

- Variant Annotation

Variant annotation is a process in the analysis of DNA sequencing data. Annotation of variants involves identifying and qualifying the functional impact of DNA variants (McCarthy et al., 2014). At this stage, functional effects of all variants will be annotated by the functional annotation software. Annotated variants will then be filtered and prioritised based on specific criteria for this project.

All variants identified on WES will be annotated using the Annotate Variation (ANNOVAR) software tool. This tool allows for the effects of variants on gene function to be prioritised and interpreted (Wang et al., 2010). The free web-based version of the ANNOVAR tool will be used (https://wannovar.wglab.org/).

- Variant Prioritisation
  - Gene Selection

This will involve creating a list of genes known or suspected to be CPGs implicated in the development of RMS, as per Supplementary Table 1. Following this, if no variants are identified in the initial list of genes, the list may be expanded to include other known CPGs and other regions potentially implicated in cancer predisposition, as per Supplementary Table 2.

- Variant Filtering and Minor Allele Frequencies

Following annotation, variants will be filtered, first for quality control (QC) purposes and then based on minor allele frequency (MAF). QC metrics such as the sequencing depth and allele balance will be examined and compared to current recommended standards (Pedersen et al., 2021). Variants will be excluded if they do not meet QC standards and/or if they have a MAF of >5% as determined by an allele frequency database. By excluding variants with a MAF of >5% any common variants or normal variation is excluded. At this stage variants with a limited suspected effect on protein function (e.g. synonymous and intronic variants) will also be excluded.

- *In-silico* Prediction

Scores from *in-silico* prediction tools are provided as part of annotation with the ANNOVAR tool. These scores will be used to assess variant effect on protein and transcript function (depending on the tool). These scores will allow for variants to be prioritised based on prediction of deleterious effect on protein function. Any predicted benign or likely benign variants will be filtered out at this stage.

Following prioritisation variants will be aligned to the Integrative Genomic Viewer (IGV) genome browser. Using IGV, the prioritised variants will be aligned to assess read coverage and depth in that region. Additionally, the genomic view will allow for variants to be assessed in context of their surrounding regions, and IGV will aid in confirming the variant annotations.

- Variant Classification

Once variants have been prioritised using the predictive tools, a classification of these variants will be performed. The ACMG-AMP guidelines for classification will be utilised (Richards et al., 2015). Of interest in this study will be any variants that meet the ACMG-AMP criteria for variants of uncertain significance (VUSs), pathogenic, or likely pathogenic variants.

The final stage of this study will be collating all the collected variant information and assessing which variants are of particular interest and importance as potential CPGs in RMS. A literature review will be performed to further support the genotype-phenotype correlations determined with the predictive tools, as well as to assess gene involvement in biological and molecular pathways that may contribute to the genotype. Additionally, classification will be reviewed as compared to reported classifications from the community on Varsome (Kopanos et al., 2019). This final corroboration of data from various sources will serve to strengthen the argument for any potential putative disease-causing variants for RMS.

This study forms part of the larger PeCan study investigating the genetics of paediatric cancers in South Africa. It is the goal that the knowledge gained from the WES data as part of this study will advance understanding of germline genetics variations that may cause RMS in our population.

- Patient Result Feedback

Feedback of results to patients is outside the scope of this research project. In the event a variant of interest is identified, validation using Sanger sequencing will be performed and, where relevant, results will be transmitted to the patients by the PeCan project investigators.

- **Time Frame**

Below is a Gantt chart outlining the proposed timeline for this research project, which is to take place between February and November 2023.

	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov
Literature Review										
Ethics Application										
Preparing protocol										
Protocol presentation										
Data collection										
Data analysis										

Report write-up										
Research submission										

- **Budget and Funding**

  - Budget

Item	**Cost (ZAR)**
DNA and Library Quantification	
Qubit reagents	1 500
Tape Station reagents	2 800
Exome sequencing	
Ion Ampliseq Exome RDY Kit	16 775
Ion 540 Chip Kit	12 688
Ion 540 Kit-Chef	28 900
	62 663

  - Funding

All costs are covered through the parent study, 'Inherited cancer predisposing mutations in a cohort of childhood cancer patients' (M180855). Funding has been provided by the National Research Foundation, South African Medical Research Council, Female Academic Leaders Fellowship, National Health Laboratory Service, and the Faculty of Health Sciences at the University of the Witwatersrand.

- **Ethics**

An application for ethical clearance has been submitted to the Human Research Ethics Committee (Medical) at the University of the Witwatersrand. This research is a sub-study of the parent study 'Inherited cancer predisposing mutations in a cohort of childhood cancer patients' (M180855), with principal investigator Dr Lindie Lamola, also referred to as the PeCan study.

- **Problems and Limitations**

Listed below are some problems and limitations that may be encountered:

- The patients being investigated in this study are of African descent. As there is limited data available about these populations there is a chance that potentially deleterious variants could be overlooked as no data exists to compare it to.

- Due to limited data about CPGs in RMS being available it is likely the prediction tools will identify multiple VUSs. These VUS results are not reportable or actionable and thus limit this study.
- Only approximately 8% of RMS cases arise because of a germline predisposition (Kim et al., 2021). With such a small sample size (only eight patients), the chance of finding multiple CPG variants is minimal.

- **References**

Bourdon, J.-C. (2007). p53 and its isoforms in cancer. British Journal of Cancer, 97(3), 277–282. https://doi.org/10.1038/sj.bjc.6603886

Chow, A. (2010). Cell Cycle Control by Oncogenes and Tumor Suppressors: Driving the Transformation of Normal Cells into Cancerous Cells. Nature Education, 3(9), 7.

Crucis, A., Richer, W., Brugières, L., Bergeron, C., Marie-Cardine, A., Stephan, J. L., Girard, P., Corradini, N., Munzer, M., Lacour, B., Minard-Colin, V., Sarnacki, S., Ranchere-Vince, D., Orbach, D., & Bourdeaut, F. (2015). Rhabdomyosarcomas in children with neurofibromatosis type I: A national historical cohort. Pediatric Blood & Cancer, 62(10), 1733–1738. https://doi.org/10.1002/PBC.25556

Diller, L., Sexsmith, E., Gottlieb, A., Li, F. P., & Malkin, D. (1995). Germline p53 mutations are frequently detected in young children with rhabdomyosarcoma. The Journal of Clinical Investigation, 95(4), 1606–1611. https://doi.org/10.1172/JCI117834

Doros, L., Yang, J., Dehner, L., Rossi, C. T., Skiver, K., Jarzembowski, J. A., Messinger, Y., Schultz, K. A., Williams, G., André, N., & Hill, D. A. (2012). DICER1 Mutations in Embryonal Rhabdomyosarcomas from Children With and Without Familial PPB-Tumor Predisposition Syndrome. Pediatric Blood & Cancer, 59(3), 558. https://doi.org/10.1002/PBC.24020

Fiala, E. M., Jayakumaran, G., Mauguen, A., Kennedy, J. A., Bouvier, N., Kemel, Y., Fleischut, M. H., Maio, A., Salo-Mullen, E. E., Sheehan, M., Arnold, A. G., Latham, A., Carlo, M. I., Cadoo, K., Murkherjee, S., Slotkin, E. K., Trippett, T., Glade Bender, J., Meyers, P. A., … Walsh, M. F. (2021). Prospective pan-cancer germline testing using MSK-IMPACT informs clinical translation in 751 patients with pediatric solid tumors. Nature Cancer, 2(3), 357. https://doi.org/10.1038/S43018-021-00172-1

Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. Cell, 144(5), 646–674. https://doi.org/10.1016/J.CELL.2011.02.013/ATTACHMENT/68024D79-3A9C-46C4-930B-640934F11E2E/MMC1.PDF

Hawkins, D. S., Spunt, S. L., & Skapek, S. X. (2013). Children's Oncology Group's 2013 Blueprint for Research: Soft Tissue Sarcomas. Pediatric Blood & Cancer, 60(6), 1001. https://doi.org/10.1002/PBC.24435

Kaseb, H., Kuhn, J., & Babiker, H. M. (2022). Rhabdomyosarcoma. StatPearls. https://www.ncbi.nlm.nih.gov/books/NBK507721/

Kim, J., Light, N., Subasri, V., Young, E. L., Wegman-Ostrosky, T., Barkauskas, D. A., Hall, D., Lupo, P. J., Patidar, R., Maese, L. D., Jones, K., Wang, M., Genome Research Laboratory, C., Tavtigian, S. V., Wu, D., Shlien, A., Telfer, F., Goldenberg, A., Skapek, S. X., … Stewart, D. R. (2021). Pathogenic Germline Variants in Cancer Susceptibility Genes in Children and Young Adults With Rhabdomyosarcoma. JCO Precision Oncology, 5, 75–87. https://doi.org/10.1200/PO.20.00218

Knudson, A. G. (1971). Mutation and Cancer: Statistical Study of Retinoblastoma. Proceedings of the National Academy of Sciences of the United States of America, 68(4), 820. https://doi.org/10.1073/PNAS.68.4.820

Kopanos, C., Tsiolkas, V., Kouris, A., Chapple, C. E., Albarca Aguilera, M., Meyer, R., & Massouras, A. (2019). VarSome: the human genomic variant search engine. Bioinformatics, 35(11), 1978–1980. https://doi.org/10.1093/bioinformatics/bty897

Kratz, C. P., Rapisuwon, S., Reed, H., Hasle, H., & Rosenberg, P. S. (2011). Cancer in Noonan, Costello, Cardiofaciocutaneous and LEOPARD Syndromes. American Journal of Medical Genetics. Part C, Seminars in Medical Genetics, 157(2), 83. https://doi.org/10.1002/AJMG.C.30300

Li, H., Sisoudiya, S. D., Martin-Giacalone, B. A., Khayat, M. M., Dugan-Perez, S., Marquez-Do, D. A., Scheurer, M. E., Muzny, D., Boerwinkle, E., Gibbs, R. A., Chi, Y. Y., Barkauskas, D. A., Lo, T., Hall, D., Stewart, D. R., Schiffman, J. D., Skapek, S. X., Hawkins, D. S., Plon, S. E., … Lupo, P. J. (2021). Germline Cancer Predisposition Variants in Pediatric Rhabdomyosarcoma: A Report From the Children's Oncology Group. JNCI Journal of the National Cancer Institute, 113(7), 875. https://doi.org/10.1093/JNCI/DJAA204

Martin-Giacalone, B. A., Weinstein, P. A., Plon, S. E., & Lupo, P. J. (2021). Pediatric Rhabdomyosarcoma: Epidemiology and Genetic Susceptibility. Journal of Clinical Medicine 2021, Vol. 10, Page 2028, 10(9), 2028. https://doi.org/10.3390/JCM10092028

McCarthy, D. J., Humburg, P., Kanapin, A., Rivas, M. A., Gaulton, K., Cazier, J. B., & Donnelly, P. (2014). Choice of transcripts and software has a large effect on variant annotation. Genome Medicine, 6(3), 1–16. https://doi.org/10.1186/GM543/TABLES/2

McGee, R. B., & Nichols, K. E. (2016). Introduction to cancer genetic susceptibility syndromes. Hematology: The American Society of Hematology Education Program, 2016(1), 293. https://doi.org/10.1182/ASHEDUCATION-2016.1.293

Ognjanovic, S., Linabery, A. M., Charbonneau, B., & Ross, J. A. (2009). Trends in childhood rhabdomyosarcoma incidence and survival in the United States, 1975-2005. Cancer, 115(18), 4218–4226. https://doi.org/10.1002/CNCR.24465

Pastore, G., Peris-Bonet, R., Carli, M., Martínez-García, C., de Toledo, J. S., Steliarova-Foucher, E., Petrovich, S. V., Budanov, O., Storm, H., Christensen, N., Aareleid, T., Hakulinen, T., Sankila, R., Pukkala, E., Le Gall, E., Tron, I., Lacour, B., Desandes, E., Bernard, J. L., … McDonald, A. (2006). Childhood soft tissue sarcomas incidence and survival in European children (1978-1997): report from the Automated Childhood Cancer Information System project. European Journal of Cancer (Oxford, England : 1990), 42(13), 2136–2149. https://doi.org/10.1016/J.EJCA.2006.05.016

Pedersen, B. S., Brown, J. M., Dashnow, H., Wallace, A. D., Velinder, M., Tristani-Firouzi, M., Schiffman, J. D., Tvrdik, T., Mao, R., Best, D. H., Bayrak-Toydemir, P., & Quinlan, A. R. (2021). Effective variant filtering and expected candidate variant yield in studies of rare human disease. Npj Genomic Medicine, 6(1), 60. https://doi.org/10.1038/s41525-021-00227-3

Pylayeva-Gupta, Y., Grabocka, E., & Bar-Sagi, D. (2011). RAS oncogenes: weaving a tumorigenic web. Nature Reviews Cancer, 11(11), 761–774. https://doi.org/10.1038/nrc3106

Rahman, N. (2014a). Mainstreaming genetic testing of cancer predisposition genes. Clinical Medicine, 14(4), 436. https://doi.org/10.7861/CLINMEDICINE.14-4-436

Rahman, N. (2014b). Realising the Promise of Cancer Predisposition Genes. Nature, 505(7483), 302. https://doi.org/10.1038/NATURE12981

Richards, S., Aziz, N., Bale, S., Bick, D., Das, S., Gastier-Foster, J., Grody, W. W., Hegde, M., Lyon, E., Spector, E., Voelkerding, K., & Rehm, H. L. (2015). Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation

of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genetics in Medicine, 17(5), 405–424. https://doi.org/10.1038/gim.2015.30

Stefan, C., Bray, F., Ferlay, J., Liu, B., & Parkin, D. M. (2017). Cancer of childhood in sub-Saharan Africa. Ecancermedicalscience, 11. https://doi.org/10.3332/ECANCER.2017.755

ThermoFisher Scientific. (2019). Ion 540 Kit – Chef Quick Reference.

Wang, K., Li, M., & Hakonarson, H. (2010). ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Research, 38(16), e164–e164. https://doi.org/10.1093/nar/gkq603

Zhang, J., Walsh, M. F., Wu, G., Edmonson, M. N., Gruber, T. A., Easton, J., Hedges, D., Ma, X., Zhou, X., Yergeau, D. A., Wilkinson, M. R., Vadodaria, B., Chen, X., McGee, R. B., Hines-Dowell, S., Nuccio, R., Quinn, E., Shurtleff, S. A., Rusch, M., … Downing, J. R. (2015). Germline Mutations in Predisposition Genes in Pediatric Cancer. The New England Journal of Medicine, 373(24), 2336. https://doi.org/10.1056/NEJMOA1508054

- **Supplementary Material**

  - Supplementary Table 1

Syndrome associated germline cancer predisposition genes that have previously been linked with rhabdomyosarcoma

Gene Abbreviation	Cancer Susceptibility Syndrome	Reference
TP53	Li- Fraumeni Syndrome	1, 2
CHEK2	CHEK2 associated cancer predisposition syndrome	1, 2
NF1	Neurofibromatosis type 1	1, 2
HRAS	Costello syndrome	1
DICER1	DICER1 syndrome	1, 2
PTCH1, SUFU	Gorlin syndrome	1
SOS1, PTPN11, RAF1, CBL, KRAS, NRAS, RIT1, SHOC2	Noonan syndrome	1, 2
BRAF, MAP2K1, MAP2K2	Cardiofaciocutaneous syndrome	1
MLH1, MSH2, MSH6, PMS2	Constitutional mismatch repair deficiency	1, 2
CDKN1C	Beckwith-Wiedemann syndrome	1, 2
CREBBP	Rubenstein-Tabi syndrome	1

- Supplementary Table 2

Cancer predisposition genes that have not previously been linked to rhabdomyosarcoma

Gene Abbreviation	Reference
ALK	1
APC	1
ATM	2
BAP1	1
BMPR1A	1
BRCA1	1
BRCA2	1
CDC73	1
CDH1	1
CDK4	1
CDKN2A	1
CEBPA	1
COL7A1	2
EPCAM	1
ERCC2	2
FH	1
GATA2	1
GBA	2
MAX	1
MEN1	1
MUTYH	2
NF2	1
PALB2	1
PAX5	1
PHOX2B	1
PRKAR1A	1
PTEN	1
RB1	1
RET	1
RUNX1	1
SBDS	2
SDHA	1
SDHAF2	1
SDHB	1
SDHC	1
SDHD	1
SMAD4	1
SMARCA4	1
SMARCB1	1
STK11	1
TMEM127	1
TSC1	1

TSC2	1
VHL	1
WT1	1

- Supplementary References

3. Li, H., Sisoudiya, S. D., Martin-Giacalone, B. A., Khayat, M. M., Dugan-Perez, S., Marquez-Do, D. A., Scheurer, M. E., Muzny, D., Boerwinkle, E., Gibbs, R. A., Chi, Y. Y., Barkauskas, D. A., Lo, T., Hall, D., Stewart, D. R., Schiffman, J. D., Skapek, S. X., Hawkins, D. S., Plon, S. E., … Lupo, P. J. (2021). Germline Cancer Predisposition Variants in Pediatric Rhabdomyosarcoma: A Report From the Children's Oncology Group. *JNCI Journal of the National Cancer Institute*, *113*(7), 875. https://doi.org/10.1093/JNCI/DJAA204
4. Kim, J., Light, N., Subasri, V., Young, E. L., Wegman-Ostrosky, T., Barkauskas, D. A., Hall, D., Lupo, P. J., Patidar, R., Maese, L. D., Jones, K., Wang, M., Genome Research Laboratory, C., Tavtigian, S. V., Wu, D., Shlien, A., Telfer, F., Goldenberg, A., Skapek, S. X., … Stewart, D. R. (2021). Pathogenic Germline Variants in Cancer Susceptibility Genes in Children and Young Adults With Rhabdomyosarcoma. *JCO Precision Oncology*, *5*, 75–87. https://doi.org/10.1200/PO.20.00218

9.6. Appendix 6: Plagiarism Declaration and TurnItIn Plagiarism Report

PLAGIARISM DECLARATION TO BE SIGNED BY ALL HIGHER DEGREE STUDENTS

SENATE PLAGIARISM POLICY: APPENDIX ONE

I, **Gabriella Peta Pillhofer**, (Student number: **2097324**), am a student registered for the degree **of MSc (Med) in the field of Genomic Medicine** in the academic year **2023**.
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:  Date: 29/11/2023

ContentOnly_MScResearch_GP.docx

ORIGINALITY REPORT

17%	6%	6%	13%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to University of Witwatersrand Student Paper	10%
2	iris.unige.it Internet Source	1%
3	clinvarminer.genetics.utah.edu Internet Source	1%
4	"Abstracts 25. Jahrestagung", medizinische genetik, 2014 Publication	1%
5	pericles.pericles-prod.literatumonline.com Internet Source	<1%
6	Submitted to University of Shajrah Student Paper	<1%
7	Submitted to University of Edinburgh Student Paper	<1%
8	jgeb.springeropen.com Internet Source	<1%
9	Andrea Timpanaro, Caroline Piccand, Dzhangar Dzhumashev, Stenija Anton-Joseph	<1%

et al. "CAR T cells recognizing CD276 and Dual-CAR T cells against CD276/FGFR4 promote rhabdomyosarcoma clearance in an orthotopic mouse model", Cold Spring Harbor Laboratory, 2023
Publication

10	unsworks.unsw.edu.au Internet Source	<1 %
11	Junyu Zhang, Yanyi Yao, Haixian He, Jun Shen. "Clinical Interpretation of Sequence Variants", Current Protocols in Human Genetics, 2020 Publication	<1 %
12	eprints.ncl.ac.uk Internet Source	<1 %
13	iovs.arvojournals.org Internet Source	<1 %
14	www.frontiersin.org Internet Source	<1 %
15	www.isfg.org Internet Source	<1 %
16	www.medrxiv.org Internet Source	<1 %
17	Submitted to University of Newcastle Student Paper	<1 %
18	www.mdpi.com Internet Source	<1 %

19	Deml, Brett, Linda M Reis, Emmanuelle Lemyre, Robin D Clark, Ariana Kariminejad, and Elena V Semina. "Novel mutations in PAX6, OTX2 and NDP in anophthalmia, microphthalmia and coloboma", European Journal of Human Genetics, 2015. Publication	<1 %
20	Stalbow, Lauren. "A Genotype First Approach to Diabetes", Icahn School of Medicine at Mount Sinai, 2023 Publication	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		

9.7. Appendix 7: Human Mutation Journal Author Guidelines

HUMU - For Authors | Hindawi

hindawi.com/journals/humu/guidelines

1. Publish with Hindawi

Join our community of authors and benefit from:

- An easy-to-use manuscript submission system, without manuscript formattingrequirements.
- Free of charge, full language editing report at point of submission, to help you assessand improve your manuscript prior to peer review.
- Dedicated editors who are active in their specific communities.
- High editorial standards, ensuring all published manuscripts undergo an in-depth peerreview process.
- Quick, efficient publication with full transparency on all publishing metrics andturnaround times.
- Greater impact, reach, and visibility of your research through open
- access.Retention of all ownership and copyright of your published research.
- Discount on a wide range of author services from leading providers, to help make yourmanuscript the best it can be.

2. Language editing and author services

We have partnered with a number of leading author services providers to offer our authorsan exclusive 10% discount on a wide range of manuscript preparation and post-publicationservices. This discount applies to various language editing, translation, and research communication services. Learn more about author services.

You can make use of a free artificial intelligence (AI)-based language editing tool, Writefull, atthe point of submission. Writefull will scan your manuscript and make suggestions to help
improve the quality of your writing. The tool applies machine learning, trained on millions of published scientific articles, and suggests improvements to grammar, spelling, and academiclanguage.

3. Submission

Manuscripts should be submitted by one of the authors of the manuscript through Phenom, the manuscript submission system for our journals. Only electronic PDF (.pdf) or Word (.doc,
.docx, .odt, .rtf, .txt) files can be submitted through the manuscript submission system, and
there is no page limit. Special characters should not be included in the file name of the mainmanuscript file. Submissions by anyone other than one of the authors will not be accepted.

The submitting author takes responsibility for the manuscript during submission and peerreview. For technical help, please contact help@hindawi.com.

---

4. Terms of submission

Manuscripts must be submitted on the understanding that they have not been published elsewhere and are only being considered by this journal. The submitting author is responsible for ensuring that the article's publication has been approved by all the other coauthors. It is also the submitting author's responsibility to ensure that the article has all necessary institutional approvals. Only an acknowledgment from the editorial office officiallyestablishes the date of receipt. Further correspondence and proofs will be sent to the author(s) before publication, unless otherwise indicated. It is a condition of submission that the authors permit editing of the manuscript for readability. All inquiries concerning the publication of accepted manuscripts should be addressed to help@hindawi.com. All submissions are bound by Hindawi's terms of service.

---

5. Peer review

All submitted articles are subject to assessment and peer review to ensure editorialappropriateness and technical correctness.

Research published in the journal must be:

- Scientifically valid – adhering to accepted community standards of
- research.Technically accurate in its methods and results.
- Representative of a specific advance, or replication, or null/negative result, which is
  worthy of publication.
- As reproducible as possible – sharing underlying data, code, and supporting materialswherever able.
- Ethically sound and transparent – adhering to best practice with respect to

animal andhuman studies, consent to publish, and clear declaration of potential conflicts of
interests, both real and perceived.

In the spirit of sharing findings through our open science mission, emphasis is not placed onnovelty, interest, or perceived impact. Replication studies, particularly of research publishedin this journal, are encouraged.

In order for an article to be accepted for publication, the assigned editor will first consider ifthe manuscript meets the minimum editorial standards and fits within the scope of the
journal. If an article is considered suitable for the journal, the editor will ideally solicit at leasttwo external peer reviewers (who will remain anonymous to the authors unless they chooseto disclose their identity by signing the review report) to assess the article before confirminga decision to accept. Decisions to reject are at the discretion of the editor.

Our research integrity team will occasionally seek advice outside standard peer review, for example, on submissions with serious ethical, security, biosecurity, or societal implications. We may consult experts and the editor before deciding on appropriate actions, including butnot limited to: recruiting reviewers with specific expertise, assessment by additional editors, and declining to further consider a submission.

6. Concurrent submissions

In order to ensure sufficient diversity within the authorship of the journal, authors will be limited to having three manuscripts under review at any point in time. If an author alreadyhas three manuscripts under review in the journal, they will need to wait until the review process of at least one of these manuscripts is complete before submitting another manuscript for consideration. This policy does not apply to editorials or other non-peer- reviewed manuscript types.

7. Article processing charges

The journal is open access. Article processing charges (APCs) allow the publisher to makearticles immediately available online to anyone to read and reuse upon publication.

8. Preprints

Hindawi supports the deposition of manuscripts in preprint servers, and does not consider this to compromise the novelty of the results. Articles based on content

previously made public only on a preprint server, institutional repository, or in a thesis will be considered. Thepreprint should be cited.

9. Preregistration of studies

Authors are encouraged to indicate whether the conducted research was preregistered in anindependent, institutional registry (e.g., http://clinicaltrials.gov/, https://www.socialscienceregistry.org/, http://osf.io/, https://egap.org/registry/, http://ridie.3ieimpact.org/). Preregistration of studies involves registering the study design, variables, and treatment conditions prior to conducting the research.

10. Preregistration of analysis plans

Authors are encouraged to indicate whether or not the conducted research was preregistered with an analysis plan in an independent, institutional registry (e.g., http://clinicaltrials.gov/, https://www.socialscienceregistry.org/, http://osf.io/, https://egap.org/registry/, http://ridie.3ieimpact.org/). Preregistration of studies involves registering the study design, variables, and treatment conditions. Including an analysis planinvolves specification of sequence of analyses or the statistical model that will be reported.

11. ORCID

Prior to publication, an ORCID iD must be provided for the corresponding author(s). If you already have an ORCID iD, you will be asked to provide it. If you haven't registered with ORCID yet, we'll help you create an iD at the point of submission. The ORCID is not requiredfor submission, or for peer review, but we will not be able to publish your article online until an ORCID iD is provided.

12. Article types

The journal will consider the following article types:

12.1. Research articles

Research articles should present the results of an original research study. These manuscripts should describe how the research project was conducted and provide a thorough analysis of the results of the project. Systematic reviews may be submitted asresearch articles.

12.2. Reviews

A review article provides an overview of the published literature in a particular subject area.

13. Formatting

An optional research article manuscript template can be downloaded here. We recommendthat all manuscripts include line numbers and follow the structure below:

13.1. Title and authorship information

The following information should be included:

- Manuscript title Full author names
- Full institutional mailing addresses Email addresses

Affiliations. Hindawi Limited remains neutral with regard to jurisdictional claims in institutional affiliations. Responsibility for affiliations ultimately rests with the author, althoughHindawi may request changes be made to countries listed in affiliations to ensure consistency across published output (for indexing and discovery reasons).

13.2. Abstract

The manuscript should contain an abstract. The abstract should be self-contained, citation-free, and should not exceed 300 words.

13.3. Introduction

This section should be succinct, with no subheadings.

13.4. Materials and methods

The methods section should provide enough detail for others to be able to replicate thestudy. If you have more than one method, use subsections with relevant headings, e.g.different models, in vitro and in vivo studies, statistics, materials and reagents, etc.

Hindawi journals have no space restriction on methods. Detailed descriptions of the methods(including protocols or project descriptions) and algorithms may also be uploaded as supplementary information or a previous publication that gives more details may be cited. If the method from a previous article is used then this article must

be cited and discussed. If
wording is reused from a published article then this must be noted, e.g. This study uses themethod of Smith et al. and the methods description partly reproduces their wording [1].

If a method or tool is introduced in the study, including software, questionnaires, and scales,the license this is available under and any requirement for permission for use should be stated. If an existing method or tool is used in the research, the authors are responsible for checking the license and obtaining any necessary permission. If permission was required, astatement confirming permission was granted should be included in the materials and methods section.

Publishing protocols. We encourage authors describing any methodology, in particular

laboratory-based experiments in the life sciences but also computational and bioinformaticsprotocols, to upload details of their methods to protocols.io. This is an open access websitethat allows researchers to record their methods in a structured way, obtain a DOI to allow easy citation of the protocol, collaborate with selected colleagues, share their protocol privately for journal peer review, and choose to make it publicly available. Once published, the protocol can be updated and cited in other articles.

You can make your protocol public before publication of your article if you choose, which willnot harm the peer review process of your article and may allow you to get comments about your methods to adapt or improve them before you submit your article (see also the protocols.io FAQ page).

13.5. Results and discussion

This section may be divided into subsections or may be combined.

13.6. Main text (review only)

This section may be divided into subsections or may be combined.

13.7. Conclusion

This should clearly explain the main conclusions of the article, highlighting its importance andrelevance.

13.8. Data availability

This statement should describe how readers can access the data supporting the conclusionsof the study and clearly outline the reasons why unavailable data cannot be released.

13.9. Conflicts of interest

Authors must declare all relevant interests that could be perceived as conflicting. Authors should explain why each interest may represent a conflict. If no conflicts exist, the authors should state this. Submitting authors are responsible for coauthors declaring their interests.

Conflicts of interest (COIs, also known as 'competing interests') occur when issues outside research could be reasonably perceived to affect the neutrality or objectivity of the work or itsassessment. For more information, see our publication ethics policy. Authors must declare allpotential interests – whether or not they actually had an influence – in the conflicts of interestsection, which should explain why the interest may be a conflict. If there are none, the authors should state: "The author(s) declare(s) that there is no conflict of interest regarding the publication of this article". Submitting authors are responsible for coauthors declaring their interests. Declared conflicts of interest will be considered by the editor and reviewers, and included in the published article.

Authors must declare current or recent funding (including for article processing charges) andother payments, goods or services that might influence the work. All funding, whether a conflict or not, must be declared in the funding statement. The involvement of anyone other than the authors who: i) has an interest in the outcome of the work; ii) is affiliated to an organization with such an interest; or iii) was employed or paid by a funder, in the commissioning, conception, planning, design, conduct, or analysis of the work, the preparation or editing of the manuscript, or the decision to publish must be declared.

You may be asked to make certain changes to your manuscript as a result of your declaration. These requests are not an accusation of impropriety. The editor or reviewer ishelping you to protect your work against potential criticisms.

If you are in any doubt about declaring a potential conflict, remember that if it is revealed later – especially after publication – it could cause more problems than simply declaring it at the time of submission. Undeclared conflicts of interest could lead to a corrigendum or, in themost serious cases, a retraction.

13.10. Funding statement

Authors must state how the research and publication of their article was funded, by namingfinancially supporting body(s) (written out in full) followed by associated grant number(s) insquare brackets (if applicable), for example: “This work was supported by the Engineering and Physical Sciences Research Council [grant numbers xxxx, yyyy]; the National ScienceFoundation [grant number zzzz]; and a Leverhulme Trust Research Project Grant”.

If the research did not receive specific funding, but was performed as part of the employmentof the authors, please name this employer. If the funder was involved in the manuscript
writing, editing, approval, or decision to publish, please declare this.

13.11. Acknowledgments

All acknowledgments (if any) should be included at the very end of the manuscript before thereferences. Anyone who made a contribution to the research or manuscript, but who is not a listed author, should be acknowledged (with their permission).

13.12. References

Authors may submit their references in any style. If accepted, these will be reformatted in Chicago style by Hindawi. Authors are responsible for ensuring that the information in each
reference is complete and accurate. All references should be numbered consecutively in the
order of their first citation. Citations of references in the text should be identified using numbers in square brackets e.g., “as discussed by Smith [9]”; “as discussed elsewhere [9,10]”. All references should be cited within the text and uncited references will be removed.

Citation standards. All data, program code, and other methods should be appropriately cited. Such materials should be recognized as original intellectual contributions and affordedrecognition through citation.

---

14. Date formatting

Hindawi recommends writing dates out fully to avoid confusion with different all-numeral datestyles. For example, 11/10/2018 could be 10 November 2018 or 11 October 2018 dependingon the reader, therefore, the date should be written out in full. For example, the date
September 1, 2018 should be used rather than 01/09/2018 or 09/01/2018.

15. Units of measurement

Units of measurement should be presented simply and concisely using the InternationalSystem of Units (SI).

16. Preparation of figures

Upon submission of an article, authors should include all figures and tables in the PDF file ofthe manuscript. Figures and tables should not be submitted in separate files. If the article is accepted, authors will be asked to provide the source files of the figures. Each figure should be supplied in a separate electronic file. All figures should be cited in the manuscript in a consecutive order. Figures should be supplied in either vector art formats (Illustrator, EPS, WMF, FreeHand, CorelDraw, PowerPoint, Excel, etc.) or bitmap formats (Photoshop, TIFF, GIF, JPEG, etc.). Bitmap images should be of 300 dpi resolution at least unless the resolution is intentionally set to a lower level for scientific reasons. If a bitmap image has
labels, the image and labels should be embedded in separate layers.

Maps. Hindawi Limited remains neutral with regard to jurisdictional claims in published maps. For reasons of consistency, authors are requested to use accepted standard maps asthe basis for map figure drawing, for example using the latest standard base-map of Map
Press. Responsibility for maps rests with the author and it is their responsibility to also provide any copyright or licence information when using maps that are not owned or createdby the author (e.g. Google Maps, etc.)

17. Preparation of tables

Tables should be cited consecutively in the text. Every table must have a descriptive title andif numerical measurements are given, the units should be included in the column heading.
Vertical rules should not be used.

18. Supplementary materials

Supplementary materials are the additional parts to a manuscript, such as audio files, videoclips, or datasets that might be of interest to readers. Authors can submit one file of supplementary material along with their manuscript through the manuscript submission system. If there is more than one file, they can be uploaded as a .ZIP file.

A section titled supplementary material should be included before the references list with aconcise description for each supplementary material file. Supplementary materials are not modified by our production team. Authors are responsible for providing the final supplementary material files that will be published along with the article.

19. Proofs

Corrected proofs must be returned to the publisher within two to three days of receipt. Thepublisher will do everything possible to ensure prompt publication.

20. Copyright and permissions

Authors retain the copyright of their manuscripts, and all open access articles are distributedunder the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided that the original work is properlycited.

The use of general descriptive names, trade names, trademarks, and so forth in this publication, even if not specifically identified, does not imply that these names are not protected by the relevant laws and regulations. The submitting author is responsible forsecuring any permissions needed for the reuse of copyrighted materials included in themanuscript.

While the advice and information in this journal are believed to be true and accurate on thedate of its going to press, neither the authors, the editors, nor the publisher can accept anylegal responsibility for any errors or omissions that may be made. The publisher makes nowarranty, express or implied, with respect to the material contained herein.

21. Reporting guidelines

Authors are strongly encouraged to use appropriate reporting guidelines when preparing andsubmitting manuscripts, to maximise transparency and reproducibility. Our editors and reviewers are also encouraged to use them in the review process. Completed checklists should be provided in the supplementary files on submission. We particularly encourage the use of:

---

22. Novel Variant Data

Submission of novel variant data to public databases is required prior to publication. Please read the editorial by den Dunnen (2019). *Human Mutation* recognizes the many benefits of archiving research data. We expect you to archive all the data from which your published results are derived in a public repository. The repository that you choose should offer you guaranteed preservation (see the registry of research data repositories at https://www.re3data.org/) and should help you make it findable, accessible, interoperable, and re-useable, according to FAIR Data Principles (https://www.force11.org/group/fairgroup/fairprinciples). All accepted manuscripts are required to publish a data availability statement to confirm the presence or absence of shared data. If you have shared data, this statement will describe how the data can be accessed, and include a persistent identifier (e.g., a DOI for the data, or an accession number) from the repository where you shared the data. Authors will be required to confirm adherence to the policy. *If you cannot share the data described in your manuscript, for example for legal or ethical reasons, or do not intend to share the data, then you must provide the appropriate data availability statement.* This journal notes that FAIR data sharing

allows for access to shared data under restrictions (e.g., to protect confidential or proprietaryinformation) but notes that the FAIR principles encourage you to share data in ways that areas open as possible (but that can be as closed as necessary).

22.1. Verifying Nomenclature of DNA Variants and Sharing Variant Data

Please read the editorial by den Dunnen (2019).

Documenting variation in our genomes is an important undertaking for human research andclinical care. Accuracy in the notation of DNA variants is essential for the success of this endeavor. Because of the importance of the issue and the overall

consensus on the rules, *Human Mutation* is adopting an editorial policy that requires compliance with the recommendations to describe sequence variants before manuscripts can be accepted and published. Furthermore, variants reported in manuscripts must be submitted to a public database (e.g. ClinVar https://www.ncbi.nlm.nih.gov/clinvar/ or Global Variome shared LOVDhttp://www.lovd.nl) prior to publication.

Variant descriptions should follow current recommendations of the Human Genome VariationSociety (HGVS) (https://onlinelibrary.wiley.com/doi/full/10.1002/humu.22981). Please visit https://varnomen.hgvs.org for the latest nomenclature updates, for examples of acceptable nomenclature, guidance concerning reference sequences, or if you have further questions.

Compliance with HGVS nomenclature must be verified using tools such as the Mutalyzer program (https://mutalyzer.nl/; instructions: https://github.com/mutalyzer/mutalyzer/wiki/Mutalyzer_explain.pdf) or VariantValidator (https://variantvalidator.org/ ; instructions: https://variantvalidator.org/batch_instructions/). The file resulting from this check containing each variant noted in your manuscript must beincluded in your submission (as a supplementary file for review but not publication). Thesetools are freely available on the web.

Important considerations include:

Variants should be described in the text and tables using both DNA and protein designationswhenever appropriate.

- Reference sequences defined in the HGVS nomenclature guidelines (http://varnomen.hgvs.org/bg-material/refseq/) must be used for reporting sequencevariants. Authors should always include the Accession Number of the relevant reference sequence(s), with version number where applicable (*e.g.*: RefSeq NM_003002.3, LRG_9t1 or GenBank NC_000011.10), in the Materials and Methodssection and as a footnote in any tables listing variants. Please note, RefSeq and

  Ensembl transcript reference sequences that have been denoted as the default reporting references through the Matched Annotation from the NCBI and EBI (MANE)project) *may be used once approved by the HGVS variant nomenclature working group*.

- If alternative nomenclature schemes are commonly found in the literature, they mayalso be used *in addition to approved nomenclature*, but they must be

defined clearly(e.g. *CFTR* p.Phe508del and deltaF508).

- Standard HGVS nomenclature using g. annotation and identifying the genome build must be used for non-coding variation, including those variants identified in GWAS studies (e.g., NC_000017.11:g.50201450C>T). Variants may also be described usingdbSNP genomic location identifiers, *in addition to approved nomenclature*, if the specific nucleotide change is also included.

Acceptance and/or publication may be delayed if authors do not follow these guidelines.

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. Thesuggested 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.jpEMBL 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):
- pir.georgetown.edu.SWISS-PROT: expasy.ch/sprot/sprot- top

***In Silico* Predictions of Variant Effects**

Authors must follow the journal's recommendations for the development and/or use of *insilico* prediction software, as published in the article by Vihenen (2013).

23. Ethical guidelines

In any studies on human or animal subjects, the following ethical guidelines must be observed. For any experiments on humans, all work must be conducted in accordance with the Declaration of Helsinki (1964). Manuscripts describing experimental work that carries a risk of harm to human subjects must include a statement that the experiment was conductedwith the human subjects' understanding and consent, as well as a statement that the responsible ethics committee has approved the experiments. In the case of any animal experiments, the authors must provide a full description of any

anesthetic or surgical procedure used, as well as evidence that all possible steps were taken to avoid animal suffering at each stage of the experiment.

---

24. Appeals

Authors may appeal if they feel that the decision to reject was based on: i) a major misunderstanding over a technical aspect of the manuscript; or ii) a failure to understand thescientific advance shown by the manuscript. Appeals requesting a second opinion without sufficient justification will not be considered. To lodge an appeal, please contact the journal by email, quoting your manuscript number. Appeals will only be considered from the originalsubmitting author.