CLINICAL AND MOLECULAR INVESTIGATION OF RARE GENETIC OVERGROWTH DISORDERS

By

ALISON CLAIRE FOSTER

A thesis submitted to the University of Birmingham for the degree of

DOCTOR OF PHILOSOPHY

Institute of Cancer and Genomic Sciences

College of Medical and Dental Sciences

University of Birmingham

October 2021

Abstract

Genetic overgrowth disorders are a group of rare conditions characterised by generalised and/or regional overgrowth. They are associated with a wide spectrum of clinical features including intellectual disability, developmental disorders, congenital anomalies, and other medical problems. In recent years several novel overgrowth genes have been identified but the clinical phenotypes and natural history of these emerging conditions are not yet fully understood.

The Phenotyping of Overgrowth Disorders (POD) study was established to investigate the clinical and molecular features of rare genetic overgrowth disorders. Comprehensive clinical phenotyping data was collected from 100 participants and entered in an electronic data capture system. Genomic testing was performed on a custom targeted next generation sequencing panel of overgrowth genes. Additional molecular investigation with whole exome sequencing was performed in selected participants and trios.

This work identified a molecular genetic diagnosis in over 40% of the study cohort, confirmed the genetic heterogeneity of overgrowth disorders, and identified phenotypic overlap between overgrowth disorders and other rare genetic disorders. Knowledge of the clinical phenotypes of rare genetic overgrowth disorders has been expanded, including the clinically significant discovery of vascular complications in PDGFRB-related disorders that may be amenable to targeted molecular therapy.

Acknowledgements

This work would not have been possible without the help and support of many people. I would firstly like to thank the National Institute for Health Research (NIHR) Rare Diseases Translational Research Collaboration (RD-TRC) for funding my Doctoral Fellowship.

I am indebted to my clinical and research colleagues at Birmingham Women's Hospital, Birmingham Children's Hospital, the West Midlands Regional Genetics Laboratory, the University of Birmingham, the NIHR, and recruiting sites around the country, in particular Emma Douglas, Saira Ali, Jon Hoffman, Dr Rebecca Igbokwe, Kayleigh Aston, Dr Renuka Dias, Dr Yvonne Wallis, Rachel Doak, Anna Yeung, Dr Amy Gerrish, Dr Sam Clokie, Ed Stone, Ping Yu, and many others. I would also like to thank Professor Laurence Faivre for her generous collaboration.

I am hugely grateful to my supervisors, Professor Tim Barrett, Professor Trevor Cole and Dr Derek Lim, and a special mention goes to my colleague Dr Hannah Titheradge for her encouragement and support. Most importantly, I would like to thank the patients and families who gave consent to participate in this study.

Finally, much love and thanks to my family.

Table of Contents

Chapter 1. INTRODUCTION: 1
1.1. Overgrowth disorders 1
1.1.1. Generalised overgrowth 2
1.1.2 Regional overgrowth 3
1.1.3 Clinical features of overgrowth disorders 4
1.1.4 Differential diagnosis of overgrowth disorders 4
1.2 Overgrowth disorders in the pre-NGS era 5
1.2.1 Sotos syndrome 6
1.2.2 Weaver syndrome 7
1.2.3 *PTEN* hamartoma tumour syndrome (PHTS) 8
1.2.4 Beckwith-Wiedemann syndrome (BWS) 9
1.2.5 Simpson-Golabi-Behmel (SGB) syndrome 10
1.2.6 Perlman syndrome 11
1.2.7 Proteus syndrome 11
1.3 Molecular aetiology of overgrowth disorders 11
1.3.1 Epigenetic regulation 14
1.3.2 PI3K/AKT/mTOR signalling pathway 18
1.3.3 Other molecular causes of overgrowth 22
1.4 Next Generation Sequencing 24
1.4.1 Library preparation 24
1.4.2 Sequencing by synthesis 25
1.4.3 Bioinformatic analysis and variant interpretation 26
1.4.4 NGS and diagnostic testing 27
1.5 Clinical phenotyping 28
1.6 Aims 29
1.6.1 The Phenotyping of Overgrowth Disorders (POD) study 29
Chapter 2. METHODS: The Phenotyping of Overgrowth Disorders study 32
2.1 The Phenotyping of Overgrowth Disorders Study 32
2.1.1. Inclusion criteria 34
2.1.2 Exclusion criteria 35
2.2 Phenotypic data 35
2.2.1 Phenotypic data set and Case Report Form 35

2.2.2 Measurement of growth 36
2.2.3 OpenClinica electronic data capture 36
2.3 Molecular investigations 37
2.3.1 Next Generation Sequencing panel 39
2.3.2 Whole Exome Sequencing 57
Chapter 3. RESULTS: Description of the study cohort 65
3.1 Flowchart of participant recruitment 65
3.2 Method of data collection 66
3.3 Participant demographics 66
3.3.1 Sex 66
3.3.2 Age 67
3.3.3 Trios 68
3.3.4 Participants with molecular diagnoses at recruitment 70
3.3.5 Type of overgrowth in undiagnosed participants 71
3.4 Total cohort growth parameters 72
3.4.1 Height 72
3.4.2 Head circumference 73
3.4.3 Height vs head circumference 74
3.4.4 Birthweight 75
3.5 Comparison of participants with and without a molecular diagnosis 76
3.5.1 No significance difference in height 76
3.5.2 No significant difference in head circumference 79
3.5.3 No association between developmental delay and presence of molecular diagnosis 82
3.5.4 There is a relationship between autism spectrum disorder and absence of molecular diagnosis 84
3.5.5 No association between dysmorphic features and presence of molecular diagnosis 86
3.6 Comparison of participants with and without developmental delay 88
3.6.1 No significant difference in height 88
3.6.2 No significant difference in head circumference 91
3.7 Comparison of participants with and without autism 93
3.7.1 No significant difference in height 93
3.7.2 No significant difference in head circumference 95
3.8 Summary of study cohort 97
Chapter 4. RESULTS: Expanding the phenotypes of overgrowth disorders 98
4.1 Overgrowth disorders 98

4.1.1 Beckwith-Wiedemann syndrome 99
4.1.2 *CHD8* overgrowth syndrome 101
4.1.3 *DNMT3A*: Tatton-Brown-Rahman syndrome (TBRS) 105
4.1.4 *EZH2*: Weaver syndrome 110
4.1.5 *NFIX*: Malan syndrome 114
4.1.6 *NSD1*: Sotos syndrome 115
4.1.7 *PDGFRB*: Kosaki overgrowth syndrome (KOGS) 125
4.1.8 *PIK3CA*: PIK3CA-related overgrowth spectrum (PROS) 165
4.1.9 *PPP2R5D* -related neurodevelopmental disorder 170
4.1.10 *PTEN:* PTEN-hamartoma tumour syndrome (PHTS) 173
4.1.11 *SUZ12:* Imagawa-Matsumoto syndrome (SUZ12-related overgrowth syndrome) 185
4.2 Other single gene disorders with phenotypes overlapping overgrowth disorders. 211
4.2.1 *FBN1:* Marfan syndrome 211
4.2.2 *FOXP2*-related speech and language disorder 216
4.2.3 *GLI3*: Grieg cephalopolydactyly syndrome (GCPS) 219
4.2.4 *HIST1H1E* syndrome 222
4.2.5 *MAGED2*: transient antenatal Bartter syndrome 226
4.2.6 KMT5B syndrome 228
4.2.7 *PTCH1*: Nevoid basal cell carcinoma syndrome (BCNS; Gorlin syndrome) 231
4.2.8 *TMCO1:* Cerebro-facio-thoracic dysplasia (CFTD) 234
4.2.9 *THRA:* Resistance to thyroid hormone alpha (RTH-alpha) 238
4.3 Microdeletion and microduplication syndromes 240
4.3.1 16p13.11 microduplication syndrome 240
4.3.2 15q11.2 BP1-BP2 microdeletion syndrome (Burnside-Butler syndrome) 245
4.3.3 Participants with other findings on microarray 249
Chapter 5. RESULTS: Genomic analysis of overgrowth disorders 256
5.1 Molecular data generated outside the study 256
5.1.1 Clinical diagnostic testing 256
5.1.2 Molecular diagnoses made through other studies 260
5.2 NGS panel of overgrowth genes 261
5.2.1 POD 068.0 *DNMT3A* 263
5.2.2 POD 002.0 *NFIX* 265
5.2.3 POD 013.0 *PIK3CA* 267
5.2.4 POD 028.0 *NSD1* 269

5.2.5 POD 064.0 *PDGFRB* 271
5.2.6 POD 030.0 *PPP2R5D* 273
5.2.7 POD 035.0 *PTCH1* 275
5.3 Whole Exome Sequencing 276
5.3.1 POD 077.0 *DNMT3A* 278
5.3.2 POD 008.0 *FBN1* 279
5.3.3 POD 085.0 *KMT5B* 280
5.3.4 POD 016.0 *MAGED2* 281
5.3.5 POD 009.0 *CHD8* 283
5.3.6 POD 005.0 *FBN1* 285
5.3.7 POD 038.0 *FOXP2* 287
5.4 Summary of molecular diagnoses 289
5.4.1 Diagnoses made on testing in the POD study 289
5.4.2 Diagnostic rate according to testing strategy 292
5.4.3 All molecular diagnoses 294
Chapter 6. DISCUSSION: 296
6.1 Limitations of the study 296
6.2 Overgrowth disorders are a heterogenous group of conditions 297
6.3 Definition of overgrowth and overgrowth disorders 298
6.4 Expanding the phenotypes of overgrowth disorders 300
6.4.1 Future work: Targeted molecular treatment 301
6.5 Molecular diagnosis in overgrowth disorders 303
6.5.1 Future work: DNA methylation signature analysis 306
Chapter 7. CONCLUSION: 310
List of References 312
Appendix 337
A. ACMG and AMP standards and guidelines for the interpretation of sequence variants 337
B. Documentation for the POD study 338
B.1 IRAS application form 338
B.2 Participant information 376
B.3 Consent forms 402
B.4 Clinical Record Form (CRF) 422
C. Human Phenotype Ontology terms in the OpenClinica database 434
D. Sample sheets 449

D.1 Example of sample sheet for upload to MiSeq (QXT) 449
D.2 Example of sample sheet for upload to MiSeq (TSCA) 449
E. Lists of variants 450
E.1 Example of Agilent SureCall test sample 14 451
E.2 Example of Illumina VariantStudio test sample 14 453
F. Variant confirmation form 455
G. Summary of participants 456
H. Output arising from the study 460
H.1 Publications 460
H.2 Poster presentations 461
H.3 Presentations 461

List of Illustrations

Figure 1: PI3K/AKT/mTOR pathway 19

Figure 2: An overview of the POD study 33

Figure 3: Pathway for molecular testing for generalised overgrowth in the POD study 38

Figure 4: UCSC Genome Browser view of SETD2 with custom track 'Missed Regions' 42

Figure 5: UCSC Genome Browser view of NSD1 with custom track 'Missed Regions' 43

Figure 6: Example of D1000 ScreenTape image of libraries with fragment sizes in the optimal range 45

Figure 7: Example of High Sensitivity D1000 ScreenTape image of libraries with fragments in the optimal range 47

Figure 8: UCSC Genome Browser view of MTOR with custom track 'GapTrack' 50

Figure 9: Example of pedigree information for upload to Congenica 61

Figure 10: Summary of participant recruitment. 65

Figure 11: Total cohort - sex of participants 66

Figure 12: Total cohort - participant ages 67

Figure 13: Probands recruited as singletons, parent-child duos and parent-child trios 69

Figure 14: Individuals with a molecular diagnosis of an overgrowth disorder at entry to the study. 70

Figure 15: Type of overgrowth in participants without a molecular diagnosis of an overgrowth disorder at entry to the study 71

Figure 16: Total cohort height in SD compared to mean for age and sex 72

Figure 17: Total cohort - OFC in SD compared to mean for age and sex 73

Figure 18: Total cohort - height against OFC 74

Figure 19: Total cohort - birthweight in SD 75

Figure 20: Participants with a single gene molecular genetic diagnosis - height in SD compared to mean for age and sex 77

Figure 21: Participants without a single gene molecular diagnosis - height in SD compared to mean for age and sex 77
Figure 22: Participants with a molecular genetic diagnosis - OFC in SD compared to mean for age and sex 79
Figure 23: Participants without a diagnosis - OFC in SD compared to mean for age and sex 80
Figure 24: Presence or absence of developmental delay in the group with a molecular genetic diagnosis and without a molecular genetic diagnosis 82
Figure 25: Presence or absence of autism in the group with a molecular genetic diagnosis and without a molecular genetic diagnosis 84
Figure 26: Presence or absence of dysmorphic features (of face, hands or feet) in the group with a molecular genetic diagnosis and without a molecular genetic diagnosis 86
Figure 27: Participants with developmental delay – height in SD 89
Figure 28: Participants without developmental delay – height in SD 89
Figure 29: Participants with developmental delay – OFC in SD 91
Figure 30: Participants without developmental delay – OFC in SD 92
Figure 31: Participants with autism – height in SD 93
Figure 32: Participants without autism – height in SD 94
Figure 33: Participants with autism – OFC in SD 95
Figure 34: Participants without autism – OFC in SD 96
Figure 38: Pedigree POD 017.0 111
Figure 39: Clinical photographs of POD 064.0: facial features 128
Figure 40: Clinical photographs of POD 064.0 showing tall stature, dermatological and skeletal features. A scar on the abdomen is the result of VP shunt insertion. 128
Figure 41: Clinical photograph of POD 064.0 showing hyperpigmented dermal and subcutaneous nodules 129

Figure 42: Clinical photographs of POD 064.0 showing camptodactyly and lax skin on the palms. Scars on both wrists are from carpal tunnel release surgery. 129
Figure 43: POD 064.0 CT head age 18 months showing almost complete fusion of the sagittal, coronal and metopic sutures 130
Figure 44: POD 064.0 Hand x-ray age three years showing widening of the metacarpals and phalanges with carpal crowding 130
Figure 45: POD 064.0 MRI head age three years showing generalised parenchymal loss and severe cystic changes 131
Figure 46: POD 064.0 CT head age five years showing copper beaten skull resulting from chronic raised intrcranial pressure 131
Figure 47: POD 064.0 MRI head age five years showing foci in the right midbrain occluding the third ventricle and cerebral aqueduct, enlargement of the lateral ventricles and third ventricle, and desent of the brainstem and cerebellar tonsils 132
Figure 48: Adult patient with KOGS - childhood photograph 138
Figure 49: Adult patient with KOGS – clinical photograph showing facial features 139
Figure 50: Adult patient with KOGS – clinical photograph showing hyperpigmented atrophic skin and abnormal blood vessels 139
Figure 51: Adult patient with KOGS – clinical photographs showing camptodactyly, scarring, and dystrophic nails 140
Figure 52: Clinical photographs of patient with novel PDGFRB variant: facial features age nine, 11 and 20 144
Figure 53: Clinical photographs of patient with novel PDGFRB variant: side profile age nine, 11 and 20 144
Figure 54: Clinical photograph of patient with novel PDGFRB variant: yellow ring around the iris 145
Figure 55: Clinical photographs of patient with novel PDGFRB variant: thin hyperelastic skin with wrinkling of the palms and soles 145

Figure 56: Clinical photograph of patient with novel PDGFRB variant: haematoma on the lower limb 146
Figure 57: Patient with novel PDGFRB variant: CT angiogram images showing prominent frontal bone and thrombosis of a large fusiform aneurysm of the basilar artery 146
Figure 58: Patient with novel PDGFRB variant: MRA head scan images showing large fusiform aneurysm of the basilar artery indenting the brain stem including an area of thrombosis; multiple foci of signal change in the white matter consistent with small vessel ischaemia; and areas of dural ectasia with thinning of the skull vault and skull base 147
Figure 59: PDGFRB and location of known disease-causing variants 160
Figure 60: Pedigree POD 087.0, 089.0, and 089.1 174
Figure 61: Pedigree POD 088.0 and 088.2 181
Figure 62: Pedigree of family with seven individuals with deletions of SUZ 186
Figure 63: POD 103.0 - clinical photographs of face and hands 187
Figure 64: POD 103.1- Clinical photographs of face and hands 188
Figure 65: POD 103.3 - Clinical photographs of face, scalp hair, and hands 191
Figure 66: POD 104.0 – Clinical photographs of face, hands and feet 192
Figure 67: POD 104.1 - Clinical photographs of face, hands, right leg and feet 195
Figure 68: POD 104.3 - Clinical photographs of face, hands and feet 197
Figure 69: POD 104.4 – clinical photographs of face, hands, feet, and birthmark left calf 200
Figure 70: Height vs Age in participants with deletion of SUZ12 202
Figure 71: Weight vs Age in participants with deletion of SUZ12 203
Figure 72: Height vs OFC in participants with deletion of SUZ12 204
Figure 73: Scatterplot of height vs OFC according to type of SUZ12 variant in this study and in previously reported individuals 206

Figure 74: Head circumference and type of variant in SUZ12 in this study and in previously reported individuals 206
Figure 75: 17q24.2q24.3 duplication in DECIPHER database browser 252
Figure 76: Screenshot of DNMT3A c.499C>T in Alamut Visual 264
Figure 77: Screenshot of NFIX c.248T>G in Alamut Visual and PolyPhen-2 266
Figure 78: Screenshots of PIK3CA c.2740G>A in Alamut and PolyPhen-2 268
Figure 79: Screenshots of NSD1 c.5791T>C in Alamut and Polyphen2 270
Figure 80: Screenshots of PDGFRB c.1751C>G in Alamut and PolyPhen-2 272
Figure 81: Screenshots of PPP2R5D c.598G>A in Alamut and PolyPhen-2 274
Figure 82: Screenshot of PTCH1 c.2611_2624del in Alamut Visual 275
Figure 83: Screenshots of DNMT3A variant in Congenica 278
Figure 84: Screenshots of FBN1 variant in Congenica 279
Figure 85: Screenshot of KMT5B variant in Congenica 280
Figure 86: Screenshot of MAGED2 variant in Congenica 282
Figure 87: Screenshots of POD 009.0 CHD8 c.716delA; p.(Lys239ArgfsTer22) in Congenica 284
Figure 88: Screenshots of POD 005.0: FBN1 c.247+1G>A in Congenica 286
Figure 89: Screenshot from Alamut Visual showing effect of the FBN1 c.247+1G>A variant on the splice site 287
Figure 90: Screenshots of POD 038.0: FOXP2 c.982C>T in Congenica 288
Figure 91: Number of tests performed and number of diagnoses made according to testing strategy 292
Figure 92: Diagnostic rate of different testing strategies 293

List of Tables

Table 1: Examples of causes of macrosomia without childhood overgrowth 5

Table 2: Molecular causes of Beckwith-Wiedemann syndrome, frequency, recurrence risk and tumour risk 10

Table 3: Examples of epigenes associated with overgrowth disorders 15

Table 4: PI3K/AKT/mTOR genes and associated disorders 20

Table 5: Other selected single gene disorders associated with overgrowth 22

Table 6: Chromosomal translocations, deletions and duplications associated with overgrowth 23

Table 7: Genes on 20 gene panel 39

Table 8: Samples with known variants for validation of the gene panel 40

Table 9: Genes included on panel version 1 and version 2 54

Table 10: Mean height in SD compared between group with molecular diagnosis and group without molecular diagnosis 78

Table 11: Two sample T-test comparing height in SD between the group with molecular diagnosis and group without molecular diagnosis 78

Table 12: Mean OFC in SD compared between group with molecular diagnosis and group without molecular diagnosis 81

Table 13: Two sample T-test comparing OFC in SD between the group with molecular diagnosis and group without molecular diagnosis 81

Table 14: Number of individuals with developmental delay in the group with a molecular genetic diagnosis and group without a molecular genetic diagnosis 83

Table 15: Chi-Square tests of presence of developmental delay and molecular genetic diagnosis 83

Table 16: Number of individuals with autism in the group with a molecular genetic diagnosis and group without a molecular genetic diagnosis 85

Table 17: Chi-Square tests of presence of autism and molecular genetic diagnosis 85

Table 18: Number of individuals with dysmorphic features in the group with a molecular genetic diagnosis and group without a molecular genetic diagnosis 87
Table 19: Chi-Square tests of presence of dysmorphic features and molecular genetic diagnosis 87
Table 20: Mean height in SD compared between group with developmental delay and group without developmental delay..... 90
Table 21: Two sample T-test comparing height in SD between the group with developmental delay and group without developmental delay 90
Table 22: Mean OFC in SD compared between group with developmental delay and group without developmental delay..... 92
Table 23: Two sample T-test comparing OFC in SD between the group with developmental delay and group without developmental delay 92
Table 24: Mean height in SD compared between group with autism and group without autism 94
Table 25: Two sample T-test comparing height in SD between the group with features of autism and group without autism..... 94
Table 26: Mean OFC in SD compared between group with autism and group without autism..... 96
Table 27: Two sample T-test comparing OFC in SD between the group with features of autism and group without autism..... 96
Table 28: Summary of clinical features of participants with Sotos syndrome 123
Table 29: Clinical features of the first six patients with Kosaki overgrowth syndrome 134
Table 30: All known individuals with KOGS, vascular findings and complications..... 150
Table 31: All known individuals with PDGFRB-associated disorders and vascular abnormalities ... 155
Table 32: PDGFRB associated disorders 158
Table 33: Summary of clinical features of PDGFRB activating disorders (excluding vascular)..... 161
Table 34: Proposed nomenclature for PDGFRB-activating variants 164
Table 35: Summary of participants with PROS 170

Table 36: Summary of clinical features of participants with SUZ12 deletion 209

Table 37: Participants with a pathogenic variant in FBN1 215

Table 38: OMIM disease genes in the 17q24.224.3 duplication 253

Table 39: Participants with pathogenic variants and likely pathogenic variants on CGH microarray 258

Table 40: Participants with molecular genetic diagnoses made on diagnostic testing 259

Table 41: Participants with molecular diagnoses made through other studies 260

Table 42: Molecular diagnoses made on 20 gene panel (panel v.1) 262

Table 43: Molecular diagnoses made on 44 gene panel (panel v.2) 262

Table 44: Molecular diagnoses made on exome sequencing 277

Table 45: Summary of NGS performed in the POD study 290

Table 46: POD participants with a molecular diagnosis 294

List of Abbreviations

AA	amino acid
ACMG	American College of Medical Genetics and Genomics
AD	autosomal dominant
AR	autosomal recessive
ADHD	attention deficit hyperactivity disorder
ASD	autism spectrum disorder
ASQ	Ages & Stages Questionnaire
AVM	arteriovenous malformation
BCH	Birmingham Children's Hospital
bp	base pair
BP1-BP2	breakpoint 1 – breakpoint 2
BWH	Birmingham Women's Hospital
BRRS	Bannayan-Riley-Ruvalcaba syndrome
BWCH	Birmingham Women's and Children's Hospital
BWS	Beckwith-Wiedemann syndrome
BWSp	Beckwith-Wiedemann spectrum
CAS	childhood apraxia of speech

CATSHL	camptodactyly, tall stature and hearing loss
CCDS	Consensus Coding Sequence project
CCTN	cerebriform connective tissue naevus
CFTD	cerebrofaciothoracic dysplasia
CGH	comparative genomic hybridisation
CLOVES	congenital lipomatous overgrowth, vascular malformations, epidermal naevi, and spinal/skeletal anomalies/scoliosis
CNV	copy number variation
COG	Childhood Overgrowth study
CPMS	Central Portfolio Management System
CRF	Case Report Form
CS	Cowden syndrome
CSF	cerebrospinal fluid
dbSNP	single nucleotide polymorphism database
DDD	Deciphering Developmental Disorders study
DDG2P	Developmental Disorders Genotype-to-Phenotype database
DECIPHER	Database of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources
DNA	deoxyribonucleic acid

DGV	Database of Genomic Variants
EDC	electronic data capture
EEG	electroencephalogram
EMG	electromyography
ESP	Exome Sequencing Project
ExAC	Exome Aggregation Consortium
FIL	Facial Infiltrating Lipomatosis
GAGs	glycosaminoglycans
GCPS	Grieg cephalopolysyndactyly syndrome
GCSE	General Certificate of Secondary Education
GLOW	global developmental delay, lung cysts, overgrowth, Wilms tumour
GI	gastrointestinal
gnomAD	Genome Aggregation Database
GOF	gain of function
HIE	hypoxic-ischaemic encephalopathy
HGNC	HUGO (Human Genome Oraganisation) Gene Nomenclature Committee
HPO	Human Phenotype Ontology
IC1	imprinting centre 1
IC2	imprinting centre 2

ID	intellectual disability
IGF-1	insulin-like growth factor 1
IGF-2	insulin-like growth factor 2
IM	infantile myofibromatosis
IQ	intelligence quotient
IRAS	Integrated Research Application System
kbp	kilo base pairs
KOGS	Kosaki overgrowth syndrome
KTS	Klippel-Trenaunay syndrome
LDD	Lhermitte-Duclos disease
LOF	loss of function
LRTI	lower respiratory tract infection
MAF	minor allele frequency
Mb	megabase
MCAD	medium-chain acyl-CoA dehydrogenase
MCAP	megalencephaly-capillary malformation syndrome
MDEM	Mendelian disorders of the epigenetic machinery
MDT	multidisciplinary team
MLPA	multiplex ligation-dependent probe amplification

MPPH	megalencephaly-polymicrogyria-polydactyly-hydrocephalus syndrome
MRA	magnetic resonance angiography
MRI	magnetic resonance imaging
NBCCS	naevoid basal cell carcinoma syndrome
ng	nanogram
NGS	next generation sequencing
NHS	National Health Service
NIHR	National Institute for Health Research
NIHR CRF	National Institute for Health Research Clinical Research Facility
NRES	National Research Ethics Service
OC	OpenClinica
OFC	occipitofrontal circumference
OGID	overgrowth with intellectual disability
OMIM	Online Mendelian Inheritance in Man
OPDKD	ocular pterygium-digital keloid dysplasia
PAVS1	*PDGFRB* activating spectrum disorder-1
PAVS2	*PDGFRB* activating spectrum disorder-2
PFBC	primary familial brain calcification
PIL	patient information leaflet

PCR	polymerase chain reaction
PHTS	PTEN hamartoma tumour syndrome
pLI	probability of being loss of function intolerant
POD	Phenotyping of Overgrowth Disorders
Polyphen-2	Polymorphism Phenotyping v2
PRC2	Polycomb repressive complex 2
PROS	PIK3CA-related overgrowth spectrum
PS	Penttinen syndrome of premature ageing
PUJ	pelviureteric junction
RCC	renal cell carcinoma
RCPCH	Royal College of Paediatrics and Child Health
RD-TRC	Rare Diseases Translational Research Collaboration
REC	Research Ethics Committee
RNA	ribonucleic acid
RTK	receptor tyrosine kinase
SBS	sequencing by synthesis
SD	standard deviation
SGB	Simpson-Golabi-Behmel
SIFT	Sorting Intolerant from Tolerant

SKS	Smith-Kingsmore syndrome
SNOMED-CT	Systematized Nomenclature of Medicine - Clinical Terms
SNP	single nucleotide polymorphism
SOLAMEN	segmental overgrowth, lipomatosis, arteriovenous malformation and epidermal naevus
SOS	Segmental Overgrowth Study
TBRS	Tatton-Brown-Rahman syndrome
TSCA	TruSeq Custom Amplicon
ul	microlitre
UPD	uniparental disomy
URTI	upper respiratory tract infection
USS	ultrasound scan
UTI	urinary tract infection
UTR	untranslated region
VATER	vertebral defects, anal atresia, tracheoesophageal fistula with oesophageal atresia, and radial or renal dysplasia
VCF	variant call format
VEGA	Vertebrate Genome Annotation database
VEP	variant effect predictor

VP	ventriculoperitoneal
WES	whole exome sequencing
WGS	whole genome sequencing
WMRGL	West Midlands Regional Genetics Laboratory
100KGP	The 100,000 Genomes Project
1000G	1000 Genomes Project

Chapter 1. INTRODUCTION:

Overgrowth disorders are a group of rare genetic conditions, with each disorder having a prevalence of 1 in 2,000 or lower. These disorders are genetically and phenotypically heterogeneous and are associated with a wide spectrum of clinical features, most frequently intellectual disability. Other features may include congenital anomalies, facial dysmorphism, and many different medical complications, depending on the disorder. Although some disorders have distinct clinical features, there can be considerable phenotypic variability within disorders and a large degree of clinical overlap between disorders, leading to diagnostic difficulty. Achieving a molecular diagnosis is important for informing clinical management, for example the need for tumour surveillance in disorders associated with increased risk of cancer, and for genetic counselling in the family. The rate of discovery of novel overgrowth genes has increased dramatically in recent years with the advent of next generation sequencing technologies. There is a need to characterise the clinical features and natural history of these emerging syndromes, and to ensure that diagnostic testing is available for novel genes, to translate our increasing knowledge of the genetic aetiology of overgrowth into improving clinical care for patients with these rare disorders.

1.1. Overgrowth disorders

Overgrowth is defined as the excessive growth of an individual compared to other individuals of the same peer group (age, sex, and population). It can be further classified into generalised overgrowth, with increased height and/or head circumference, and regional overgrowth, causing asymmetrical overgrowth of one or more discrete areas of the body. Regional overgrowth is also known as segmental, somatic or mosaic overgrowth[1]. Regional overgrowth

also encompasses lateralised overgrowth, the term now used to refer to hemihyperplasia and/or hemihypertrophy[2]. The terms generalised overgrowth and regional overgrowth will be used hereafter.

1.1.1. Generalised overgrowth

Generalised overgrowth refers to increased height (or length in children under the age of two), known as tall stature, and/or increased occipitofrontal circumference (OFC), known as macrocephaly. Macrocephaly may result from increased brain size (megalencephaly) or increased size of fluid spaces within the brain. Tall stature and macrocephaly are variably defined as more than two standard deviations (SD) above the mean (approximately 98th centile)[3], or above the 99.6th centile (the uppermost of the nine centile lines on the standard UK growth chart[4,5], equivalent to more than +2.67 SD above the mean), compared to the age-related peer group.

Although the definitions of tall stature and macrocephaly are clear, it is considerably more difficult to define an 'overgrowth disorder'. The definition of an overgrowth disorder in every individual with a height or head circumference >2 SD above the mean would include many individuals are simply at the upper end of the normal variation between individuals in the population. These individuals generally have tall parent(s), or parent(s) with large head circumference, and have constitutional tall stature or familial macrocephaly respectively.

Equally, not all individuals who do have an overgrowth disorder meet the above definition of tall stature or macrocephaly, not only because of the considerable variability in phenotype seen in these disorders, but also individual factors such as familial short stature, chronic

disease or coexisting growth retarding genetic condition. There is also the question of at what stage of life tall stature and/or macrocephaly presents, which can vary between different overgrowth disorders, and sometimes between individuals with the same disorder. Some individuals may have prenatal overgrowth and high birthweight (macrosomia or large for gestational age) that does not persist into childhood. Other individuals may have a birth weight within the normal range, but overgrowth becomes apparent in early childhood. Overgrowth that presents in childhood may persist into adult life, or growth may slow in later childhood, resulting in adult growth parameters within the normal range. The definition of an overgrowth disorder cannot be made on growth parameters alone.

The definition of a generalised overgrowth disorder must therefore include other additional clinical features characteristic of these disorders. A working definition of a generalised overgrowth disorder could be proposed as follows: height or OFC >2 SD, present at any age beyond the neonatal period, in association with an additional clinical feature typical of an overgrowth disorder, such as facial dysmorphism, developmental delay, congenital anomaly, or embryonal tumour.

1.1.2 Regional overgrowth

Assessment of regional overgrowth is more subjective, as there are no accepted criteria to distinguish a normal degree of asymmetry from regional overgrowth and it is therefore reliant on the judgement of the examining clinician[2]. An area of regional overgrowth may be non-progressive, growing at a similar rate to the rest of the child, or progressive, growing more quickly and becoming more disproportionate with age.

1.1.3 Clinical features of overgrowth disorders

In addition to general and/or regional overgrowth, these disorders are associated with a wide spectrum of medical problems. Most individuals with generalised overgrowth have intellectual disability and the term 'overgrowth-intellectual disability' (OGID) has been used to describe this group of conditions[6]. Developmental and behavioural disorders, particularly autism spectrum disorder (ASD), are also often associated with these disorders.

Congenital anomalies are commonly associated with overgrowth disorders, for example congenital cardiac and renal anomalies in Sotos syndrome.

An increased risk of tumours is also associated with some overgrowth disorders, for example an increased risk of embryonal tumours in some molecular subtypes of Beckwith-Wiedemann syndrome (BWS), and several different tumour types in PTEN hamartoma tumour syndrome (PHTS).

1.1.4 Differential diagnosis of overgrowth disorders

Some conditions are associated with tall stature and/or macrocephaly but are not considered to be primary overgrowth disorders. This is either because the genetic disorder is characterised by other features that are more distinctive than the growth phenotype, such as the connective tissue abnormalities in Marfan syndrome; the increased growth is hormonally mediated, for example caused by an endocrine tumour; or another non-genetic non-endocrine factor has

resulted in an abnormal growth parameter; for example macrocephaly due to hydrocephalus secondary to meningitis.

There is also a group of disorders that present with a pattern of growth of macrosomia/large for gestational age (birth weight >2 SD for gestational age, sex and population) because of excessive growth in fetal life, but subsequently demonstrate normal or reduced growth in infancy and childhood (see Table 1).

Table 1: Examples of causes of macrosomia without childhood overgrowth

Condition	Aetiology
Infant of a diabetic mother	Non-genetic
Cantu syndrome	Gain of function variants in *ABCC9* and *KCNJ8*[7]
Familial hyperinsulinemic hypoglycaemia	Heterogenous; most commonly homozygous loss of function variants in *ABCC8* and *KCNJ11*[8]
Marshall-Smith syndrome	Specific frameshift and splice variants in *NFIX* that escape nonsense-mediated decay[9]
Costello syndrome	Gain of function variants in *HRAS*[10]
Pallister-Killian syndrome	Mosaic tetrasomy 12p[11]

1.2 Overgrowth disorders in the pre-NGS era

Overgrowth syndromes were first described on a clinical basis when groups of patients were recognised to share similar features. These conditions included Sotos syndrome, Weaver syndrome, BWS, Bannayan-Riley-Ruvalcaba syndrome (BRRS, now part of the PTEN hamartoma syndrome PTHS), Simpson-Golabi-Behmel (SGB) syndrome, Perlman syndrome, and Proteus syndrome.

1.2.1 Sotos syndrome

First described in 1964 by Juan Sotos[12], Sotos syndrome (OMIM *117550) is one of the commonest overgrowth disorders, accounting for one-third of all cases of overgrowth with intellectual disability in a recent series[6]. It comprises a triad of characteristic facial features, overgrowth (height and head circumference) and developmental delay[13]. There are a wide range of associated medical problems. In the neonatal period, hypotonia, jaundice, and poor feeding are often present and there may be an antenatal history of maternal pre-eclampsia[14]. Congenital anomalies including cardiac, renal and brain are also associated with Sotos syndrome[14]. Other medical problems include seizures, hypermobility and scoliosis[14]. Advanced bone age is also often present[13]. The classical facial phenotype of Sotos syndrome consists of a prominent forehead, narrow at the temples, full cheeks and a pointed chin[15]. The face may be rounder in infancy but the usual phenotype is easily recognised by early childhood[15]. With increasing age, the face lengthens and the chin becomes more prominent[15]. Cranial magnetic resonance imaging (MRI) findings include ventricular anomalies such as prominence of the trigone, prominent occipital horns, and ventriculomegaly, increased supratentorial extracerebral fluid spaces, and abnormalities of the midline structures most commonly thinning of the corpus callosum[16,17]. Behavioural characteristics of Sotos syndrome include a high prevalence of symptoms of autism spectrum disorder[18,19]. Cognitive profiling has identified better visuospatial memory and verbal ability compared to quantitative reasoning and non-verbal reasoning ability[20].

Sotos syndrome is caused by haploinsufficiency of the histone methyltransferase NSD1 (nuclear receptor binding SET domain protein 1). In 2002, a patient with Sotos syndrome and a chromosomal translocation with a breakpoint at 5q35 was identified by Kurotaki et al.[21].

The gene at the breakpoint was identified as *NSD1*, and they went on to describe 42 other individuals with Sotos syndrome with point mutations or microdeletions of this gene[21].

Somatic re-arrangements or mutations in *NSD1* occur in haematological and other malignancies, and although tumours have been reported in Sotos syndrome (e.g. sacrococcygeal teratoma) the risk of malignancy appears to be low for these patients.

1.2.2 Weaver syndrome

An overgrowth syndrome first clinically recognised by Weaver in 1974[22], Weaver syndrome comprises tall stature, increased head circumference, developmental delay, and advanced bone age[23]. There is some phenotypic overlap with other syndromes, particularly Sotos syndrome, but clinical features of soft doughy skin, camptodacyly and other joint contractures, umbilical hernia and a low-pitched cry in infancy are characteristic[23,24]. The facial gestalt can be subtle[24] but typically consists of hypertelorism, broad forehead, almond shaped palpebral fissure and a pointed chin with horizontal crease[23,24]. In infancy the face is rounded with retrognathia, small distinctive chin, long prominent philtrum, and large ears[23,24]. Tumours including lymphoma and neuroblastoma occur rarely in individuals with Weaver syndrome[24].

Following the identification of *NSD1*, there were reports that some cases of Weaver syndrome were also due to mutations in this gene[25,26]. However mutations in *EZH2* (enhancer of zeste homolog 2) were subsequently discovered to be the cause of Weaver syndrome through whole exome sequencing by two independent groups[27,28]. There is no overlap in the molecular aetiology of these two syndromes. The majority of mutations identified in individuals with Weaver syndrome are missense and there are no early truncating mutations, suggesting

haploinsufficiency is not the mechanism of pathogenesis[24]. Functional work indicates that a partial loss of function may be responsible[29]. EZH2 is a histone methyltransferase that forms a subunit of the polycomb repressive complex 2 (PRC2)[30].

1.2.3 *PTEN* hamartoma tumour syndrome (PHTS)

PTEN hamartoma tumour syndrome (PHTS) encompasses the spectrum of clinical entities caused by mutations in *PTEN*. Cowden syndrome (CS) was first identified in 1963 by Lloyd and Dennis who described a patient with mucocutaneous features, skeletal anomalies, craniofacial features, benign growths, and ductal breast carcinoma[31]. Additional cutaneous features, macrocephaly and a predisposition to a range of benign and malignant tumours including breast, non-medullary thyroid cancer and Lhermitte-Duclos disease (LDD; dysplastic gangliocytoma of the cerebellum)[32]. Bannayan-Riley-Ruvalacaba syndrome (BRRS) encompasses several previously named disorders including Riley-Smith (macrocephaly, papilloedema, and hemagiomata)[33], Bannayan-Zonana syndrome (macrocephaly, angiomatosis and lipomatosis)[34], and Ruvalcaba-Myhre (macrocephaly, intestinal polyposis and pigmentary changes of the genitalia)[35]. In 1992 Gorlin proposed that these disorders were in fact a single entity and expanded the phenotype to include joint hypermobility, developmental delay, and high birth weight[36]. The key features of BRRS are macrocephaly, developmental delay, autism, lipomas and genital freckling.[37]

PTEN is a tumour suppressor gene and a negative regulator of the PI3K/AKT/mTOR signalling pathway[38]. Individuals with PHTS have a high lifetime risk of cancer and tumour surveillance is recommended[39].

Biallelic variants in *PTEN*, with a germline variant and a second mosaic variant resulting in mosaic *PTEN* nullizygosity, has been reported to cause a regional phenotype. This has been described as a 'Proteus-like disorder'[40] and is termed SOLAMEN (segmental overgrowth, lipomatosis, arteriovenous malformation and epidermal nevus[41].

1.2.4 Beckwith-Wiedemann syndrome (BWS)

Beckwith-Wiedemann syndrome is characterised by generalised and regional overgrowth and congenital anomalies including abdominal wall defects, organomegaly, macroglossia, ear creases and/or ear pits[42]. Intellectual disability is not a typical feature[43]. There is an increased risk of embryonal tumours in early childhood associated with specific molecular subtypes (Table 2).

The molecular aetiology of BWS is complex, involving epigenetic or genomic mechanisms leading to abnormal gene expression in the BWS critical region at 11p15.5[44]. This region contains many imprinted genes involved in regulation of growth. Imprinted genes are differentially expressed depending on whether they are maternally inherited or paternally inherited.

The BWS critical region contains two domains, imprinting centre 1 (IC1) that regulates the expression of genes including *IGF2* (paternally expressed) and *H19* (maternally expressed long noncoding RNA growth suppressor); and imprinting centre 2 (IC2), that regulates the expression of *CDKN1C*, *KCNQ1* and *KCNQ1OT1*. Disruption of the finely balanced expression of imprinted genes promoting and suppressing growth can result in

an overgrowth phenotype. The epigenetic or genomic mechanism responsible is often mosaic and a wide range of features termed Beckwith-Wiedemann spectrum (BWSp) can be seen.[43]

Table 2: Molecular causes of Beckwith-Wiedemann syndrome, frequency, recurrence risk and tumour risk

Molecular finding	Frequency	Recurrence risk	Tumour risk
Loss of methylation of IC2 on the maternal allele	50%[45]	<1% (if no causative genomic anomaly; can be up to 50%)[46]	Low risk Wilms tumour[47,48]
Paternal uniparental isodisomy 11p15.5	20%[45]	<1%[46]	High risk Wilms tumour and hepatoblastoma[47,48]
Gain of methylation of IC1 on the paternal allele	5%[45]	<1% (if no causative genomic anomaly; can be up to 50%)[46]	High risk Wilms tumour[47,48]
Pathogenic variant in *CDKN1C* (maternal allele)	5%[45]	50% (if maternally inherited)[49]	Low risk Wilms tumour[47,48]
11p15.5 duplication	2-4%[50]	50% (if paternally inherited)[50] Risk of Silver-Russell syndrome (if maternally inherited)[51]	Variable depending on precise size and location
11p15 deletion	1-2%[50]	variable[50]	Variable depending on precise size and location

1.2.5 Simpson-Golabi-Behmel (SGB) syndrome

Simpson-Golabi-Behmel syndrome (SGB) was first described in 1975 by Simpson et al[52]. It is an X-linked condition characterised by generalised overgrowth of prenatal onset, coarse facial features, congenital abnormalities, variable intellectual disability and an increased risk of embryonal tumours[53]. The associated congenital abnormalities include supernumerary nipples, umbilical hernia, diaphragmatic hernia, genitourinary anomalies, congenital heart disease and skeletal abnormalities including polydactyly[53]. Embryonal tumours include Wilms tumour and hepatoblastoma. In 1996 Pilia et al identified that variants in *GPC3*, a member of the Glypican family, are the cause of many cases of SGB[54]. This gene is thought to have a role in the regulation of the Hedgehog signalling pathway[55].

1.2.6 Perlman syndrome

This autosomal recessive syndrome was first described in 1973 by Perlman et al. and is characterised by macrosomia, organomegaly, learning disability, a high neonatal mortality rate and a high risk of Wilms tumour[56]. In 2012 Astuti et al identified germline mutations in *DIS3L2* to be the cause of Perlman syndrome[57]. This gene encodes an exoribonuclease and lack of this protein has been shown to cause abnormalities of mitosis[57].

1.2.7 Proteus syndrome

Proteus syndrome is a regional overgrowth disorder characterised by progressive and distorting overgrowth, cerebriform connective tissue naevus (CCTN), linea verrucous epidermal naevus, and lipomatous overgrowth[58]. It was first recognised as a clinical entity by Cohen et al. in 1979[59] and later given the name 'Proteus syndrome' by Wiedemann et al[60]. In 2011, Lindhurst et al. demonstrated Proteus syndrome is caused by a somatic activating variant c.49G>A p.(Glu17Lys) in *AKT* [58].

1.3 Molecular aetiology of overgrowth disorders

Normal human growth is a complex process largely controlled by endocrine factors: predominantly insulin and insulin-like growth factors in fetal life, growth hormone and insulin-like growth factor 1 (IGF-1) in childhood, and sex hormones in puberty[61]. Environmental factors, in particular placental function in the fetal period and nutrition during infancy, also play a crucial role[61].

Normal fetal growth is significantly influenced by insulin-like growth factor 2 (IGF-2), which is highly expressed in the placenta, affecting both placental size and the functional capacity to transfer nutrients to the fetus[62]. Thus in Beckwith-Wiedemann syndrome (section 1.2.4), with overexpression of IGF-2, prenatal overgrowth often results in features such as a high birth weight and/or placentomegaly[43].

The rapid rate of growth in the fetus decreases postnatally, with linear growth in children largely the result of growth plate chondrogenesis. Chondrogenesis is regulated by hormonal factors including growth hormone and IGF-1, with IGF-2 becoming relatively less important[63]. Other factors affecting growth plate chondrogenesis include nutrition (largely mediated through altered hormone levels); inflammatory cytokines (such as TGF-beta); paracrine growth factors (such as C-natriuretic peptide); the extracellular matrix and intracellular proteins[63].

At puberty, oestrogen stimulates secretion of growth hormone by the pituitary, and androgens have a direct stimulating effect on the growth plate as well as being converted to oestrogen, causing the pubertal growth spurt. Ultimately oestrogen accelerates senescence of the growth plate. Epiphyseal fusion results from age-related local changes in gene expression, with consequent cease in cell proliferation[64].

Both IGF-1 and IGF-2 bind to the insulin-like growth factor receptor 1 (IGF1) on the cell surface of targeted tissues, leading to activation of pathways that are crucial for growth, including the PI3K/AKT/mTOR signalling pathway[65]. Variants in genes encoding component proteins in this pathway have been identified as a key cause of overgrowth

disorders (see section 1.3.2). In the NGS era, several novel disorders have been identified in addition to the previously described Proteus syndrome and PTEN hamartoma tumour syndrome (see section 1.2.3 and 1.2.7).

The other main group of overgrowth disorders, which has also expanded through the identification of novel syndromes by NGS, is that of the epigenetic regulators. The previously described disorders Sotos syndrome (see 1.2.1) and Weaver syndrome (see 1.2.2) are included in this group.

There are complex inter-relationships between the molecular pathways involved in growth and the pathophysiology of overgrowth disorders. In the fetus, mTOR has been reported to control the secretion of insulin like growth factor binding protein 1 (IGFBP-1), thereby regulating IGF-1 and IGF-2[66]. In individuals with Sotos syndrome, hypomethylation of regions including the IGF2-DMR (differentially methylation region) has been identified, suggesting *NSD1* plays a role in the establishment or maintenance of the IGF2-DMR[67]. Overexpression of *NSD1* results in decreased mTOR signalling, suggesting that loss of *NSD1* would cause activation of the mTOR pathway[68]. *EZH2* appears to repress negative regulators of the mTOR pathway[69]. There is also interaction between epigenetic regulators, with *NSD1* variants shown to deregulate PRC2, of which *EZH2* encodes a subunit[70]. These shared molecular mechanisms between the epigenetic regulators and the PI3K/AKT/mTOR pathways could explain some of the growth similarities between the overgrowth disorders. Equally, differences in the stage of life that overgrowth becomes apparent, prenatal and/or postnatal, could be due to differing effects of perturbations of these pathways on IGF-2 and IGF-1 respectively.

1.3.1 Epigenetic regulation

Normal human development requires specialisation of cells into many different tissues. This is achieved through cell type specific gene expression[71].

The long molecules of DNA in eukaryotic cells are complexed with histone proteins to form a compact structure, chromatin. The nucleosome, the basic unit of chromatin, is composed of two turns of DNA wound around an octamer of the four core histone proteins H2A, H2B, H3 and H4[72]. Other histone proteins are known as linker histones and are important in the higher order structure of chromatin[72]. Chromatin can be open, allowing transcription factors to bind and gene expression to occur, or closed, repressing transcription and preventing gene expression.

Epigenetic marks are modifications to DNA, chromatin, and other related proteins that do not change the DNA sequence itself, but affect chromatin structure and thus regulate gene expression[73]. The epigenetic machinery is composed of proteins that write, erase, read, and remodel the epigenetic marks on DNA or histone proteins, or remodel chromatin[74]. The three main mechanisms are post-translational modification of histone tails, DNA methylation, and chromatin remodelling[71].

The methylation of lysine residues in histone tails is a key type of post-translational modification. Lysine methylation activates or represses transcription depending on the site of the residue on the histone tail (e.g. K4, K9, K27, K36) and methylation status (mono-,di-, or trimethylation)[75].

The genes that encode components of the epigenetic machinery are known as epigenes. It is increasingly apparent that many overgrowth disorders are caused by pathogenic variants in epigenes, and fall in the wider group of genetic conditions called 'Mendelian disorders of the epigenetic machinery' (MDEMs) (Table 3)[76].

Table 3: Examples of epigenes associated with overgrowth disorders

Gene	Protein	Associated overgrowth disorder
NSD1	H3K36 histone methyltransferase	Sotos syndrome
EZH2	H3K27 histone methyltransferase (catalytic subunit of PRC2 complex)	Weaver syndrome
EED	Subunit of PRC2 complex	Cohen-Gibson syndrome
SUZ12	Subunit of PRC2 complex	Imagawa-Matsumoto syndrome
DNMT3A	DNA methyltransferase	Tatton-Brown-Rahman syndrome
CHD8	Chromatin remodelling protein	CHD8 overgrowth syndrome

Pathogenic variants in these genes are associated with generalised overgrowth and most commonly occur de novo in the germline. There are some reports of familial cases, where the disorder is inherited from an affected parent with mild clinical features in an autosomal dominant pattern.

1.3.1.1 Polycomb repressive complex 2 (PRC2) genes: EED (Cohen-Gibson syndrome) and SUZ12 (Imagawa-Matsumoto syndrome)

The proteins encoded by *EED*, *SUZ12*, and *EZH2* are members of the polycomb repressive complex 2 (PRC2). PRC2 methylates K27 on histone H3 and represses transcription[77]. Variants in *EZH2* are associated with Weaver syndrome as discussed in Chapter 1.2.2. Investigation of individuals with Weaver-like phenotypes but no identifiable *EZH2* variants first identified variants in *EED* in 2015[78–80] and in *SUZ12* in 2018[81,82]. The phenotypes of Cohen-Gibson syndrome and Imagawa-Matsumoto syndrome overlap considerably with that

of Weaver syndrome, including similar facial features, variable overgrowth and intellectual disability[81,83–87]. It has been suggested that there may be differences in the incidence of some clinical characteristics between the disorders, for example intellectual disability being universal in Cohen-Gibson syndrome, occurring in the majority of individuals with Weaver syndrome, but present in just over half of individuals with Imagawa-Matsumoto syndrome[83–85].

1.3.1.2 DNMT3A: Tatton-Brown-Rahman syndrome (TBRS)

Variants in a DNA methyltransferase, *DNMT3A*, were reported to be associated with an overgrowth disorder in 2014, causing a phenotype of overgrowth, intellectual disability and characteristic facial features[88]. Dysmorphic features in TBRS typically consist of low-set, broad, horizontal eyebrows; macrodontia of the upper central incisors; and sometimes short, widely spaced toes[89]. Other clinical features include joint hypermobility, obesity, hypotonia, autism spectrum disorder, scoliosis, and seizures[89].

Somatic variants in *DNMT3A* are commonly associated with haematological malignancies, and the same variants can be seen in individuals with TBRS in the germline[90]. To date two individuals with TBRS are known to have developed acute myeloid leukaemia and one individual has developed medulloblastoma[89,91,92].

1.3.1.3 CHD8 overgrowth syndrome

In 2007 Zahir et al. reported three patients with de novo 14q11.2 deletions with developmental delay, intellectual disability and similar facial features (widely-spaced eyes, short nose with flat nasal bridge, long philtrum, prominent Cupid's bow, full lower lip and ear

anomalies)[93]. Two genes in this region, *SUPT16H* and *CHD8*, were identified as potential candidate genes for this phenotype. Subsequently, O'Roak et al. performed whole exome sequencing in a cohort of trios of children with autism spectrum disorder and unaffected parents and found recurrent protein-altering variants in *CHD8*[94]. They also identified that CHD8 related autism spectrum disorder was associated with a macrocephaly phenotype[95].

The phenotype was expanded to include gastrointestinal (GI) problems by Bernier et al.[96] who sequenced CHD8 in a large cohort of children with ASD or developmental delay. Loss of CHD8 was described as a distinct neurodevelopmental syndrome with ID, ASD, characteristic facial features, macrocephaly, GI problems and sleep problems by Yasin et al.[97] and as an OGID syndrome by Tatton-Brown et al[6]. Further phenotype studies have identified an increased male to female rate, tall stature, regression of speech, seizures and hypotonia[98]. Additional features include pes planus, scoliosis, fifth finger clinodactyly, umbilical hernia and glabellar haemangioma[99].

The CHD (chromodomain, helicase, DNA binding) family of proteins are ATP-dependent chromatin remodelling enzymes[100]. CHD8 inhibits beta-catenin mediated transcriptional activation by promoting the association of beta-catenin and histone H1[100,101].

1.3.2 PI3K/AKT/mTOR signalling pathway

The phosphatidylinositol-3-kinase PI3K/AKT/mTOR signalling pathway is a key molecular pathway that regulates cell growth, metabolism, and survival. Human diseases including overgrowth disorders and malignancy are associated with activating variants in genes encoding proteins in the pathway including *AKT1*[58], *AKT2*[102], *AKT3*[103], *PDGFRB*[104], *PIK3CA*[103,105,106], *PIK3R2*[103], and *MTOR*[107] (Figure 1). Loss of function variants in the tumour suppressor gene *PTEN* are also associated with overgrowth and malignancy. Overgrowth disorders associated with the pathway include Proteus syndrome, PIK3CA-related overgrowth spectrum (PROS), Kosaki overgrowth syndrome (KOGS), Smith-Kingsmore syndrome (SKS), and *PTEN* hamartoma tumour syndrome (PTHS) (Table 4).

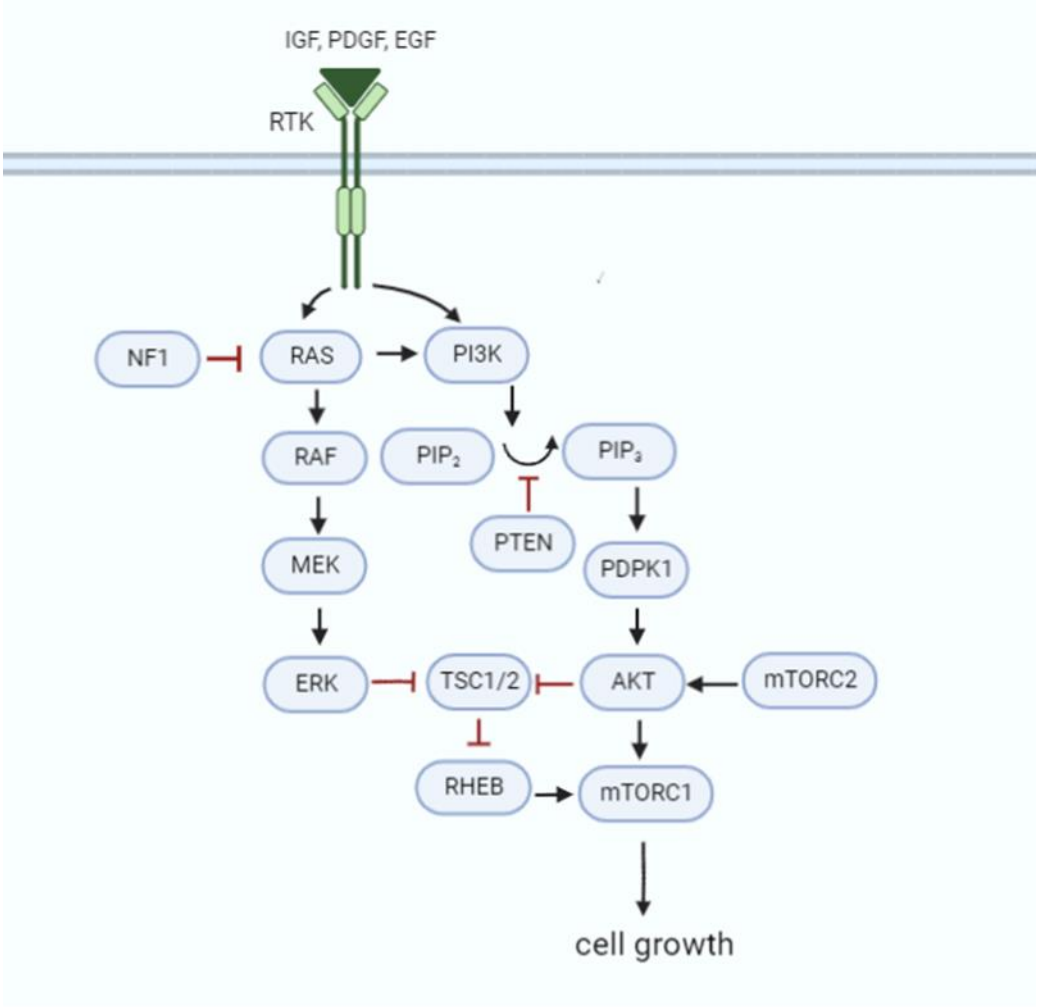


Figure 1: PI3K/AKT/mTOR pathway. AKT – protein kinase B, EGF – epidermal growth factor, ERK – extracellular signal-regulated kinase, IGF – insulin-like growth factor, MEK – mitogen-activated protein kinase, mTORC1 – mechanistic target of rapamycin (MTOR) complex 1, mTORC2 – mechanistic target of rapamycin (MTOR) complex 2, NF1 – neurofibromin, PDGF – platelet-derived growth factor, PDPK1- 3-phosphoinositide-dependent protein kinase-1, PIP_2 – phosphatidylinositol (4,5)-bisphosphate, PIP_3 – phosphatidylinositol (3,4,5)-triphosphate, PI3K – phosphoinositide 3-kinase, PTEN – phosphatase and tensin homolog, RAF – rapidly accelerated fibrosarcoma kinase, RAS – rat sarcoma virus GTPase, RAF – rapidly accelerated fibrosarcoma kinase, RTK – receptor tyrosine kinase, TSC1/2 – hamartin-tuberin complex. Created with BioRender.com

Table 4: PI3K/AKT/mTOR genes and associated disorders

Gene	Protein	Types of pathogenic variants	Phenotypes
AKT1	AKT kinase	GOF mosaic	Proteus syndrome[58]
AKT2	AKT kinase	GOF germline/ mosaic LOF germline	Hypoinsulinaemic hypoglycaemia and asymmetric overgrowth[108] AD type 2 diabetes[109]
AKT3	AKT kinase	GOF germline GOF mosaic LOF germline	MPPH (megalencephaly-polymicrogyria-polydactyly-hydrocephalus syndrome)[103] Megalencephaly; hemimegalencephaly; dysplastic megalencephaly; focal cortical dysplasia[110] Microcephaly[111]
MTOR	mTOR kinase	GOF germline/mosaic	Smith-Kingsmore syndrome[107]
PDGFRB	Receptor tyrosine kinase	GOF germline/mosaic	Numerous, including Kosaki overgrowth syndrome[104]
PIK3CA	PI3K catalytic subunit	GOF germline GOF mosaic	Megalencephaly[110] PROS[112]
PIK3R1	PI3K regulatory subunit	LOF germline	SHORT syndrome, syndromic insulin resistance with lipoatrophy[113]
PIK3R2	PI3K regulatory subunit	GOF germline GOF mosaic	MPPH[103] Bilateral perisylvian polymicrogyria[114]
PTEN	Tumour suppressor	LOF germline/mosaic LOF germline and mosaic (nullizygous)	PTEN hamartoma tumour syndrome SOLAMEN[41]

Pathogenic variants in these genes are often mosaic and are often associated with regional overgrowth. Some disorders such as *AKT1* are only found in the mosaic form, most likely because they would be lethal if present in the germline.

1.3.2.1 PIK3CA: PIK3CA-related overgrowth spectrum (PROS)

PIK3CA related overgrowth spectrum (PROS) was defined as a disorder by Keppler et al. in 2014[115], and encompasses a large number of clinical entities including fibroadipose

dysplasia[106], CLOVES (congenital lipomatous overgrowth, vascular malformations, epidermal nevi, scoliosis/skeletal and spinal abnormalities)[105], megalencephaly-capillary malformation-polymicrogyria syndrome (MCAP)[116], facial infiltrating lipomatosis (FIL)[117], Klippel-Trenaunay syndrome (KTS)[118], isolated macrodactyly[119], and other phenotypes. These disorders are clinically overlapping, and many individuals have features that do not fit entirely within a single disorder but fall within the wider spectrum of PROS[120]. Clinical features of PROS include congenital or early childhood onset of regional overgrowth of adipose, muscle, nerve, or skeletal tissue; and vascular malformations[112]. Height in individuals with PROS is usually within the normal range but macrocephaly due to megalencephaly is common[115].

The spectrum is caused by activating variants in *PIK3CA*, usually mosaic, with the cell type, timing of the mutational event, and severity of the variant determining the resulting phenotype[120,121].

Activating variants in *PIK3CA* are also the most common somatic mutations in human malignancies[122]. It has been suggested that cancer hot-spot variants are generally the cause of PROS without megalencephaly, and variants with weaker oncogenic activity are the cause of MCAP[121]. However this genotype phenotype correlation has not yet been fully established, with other studies finding that all PROS phenotypes can be caused by a wide range of *PIK3CA* variants[123].

1.3.3.2 PDGFRB: Kosaki overgrowth syndrome (KOGS)

PDGFRB (OMIM #173410) encodes platelet-derived growth factor receptor beta, a tyrosine kinase receptor involved in activation of multiple signalling pathways including the PI3K/AKT/mTOR pathway. Activating missense mutations are associated with a number of conditions including infantile myofibromatosis (IM)[124–127], premature ageing syndrome, Penttinen type [128], and Kosaki overgrowth syndrome (KOGS)[104,129]. Loss-of-function variants are associated with primary familial brain calcification (PFBC)[130]. Somatic rearrangements with a t(5;12) translocation leading to a ETV6-PDGFRB fusion gene are associated with myeloproliferative disorder with eosinophilia (OMIM#131440)[131].

KOGS was first described in 2015 and less than ten patients have been described to date[104,129,132,133].

1.3.3 Other molecular causes of overgrowth

In addition to the two main groups of epigenetic regulator genes and genes in the PI3K/AKT/mTOR pathway, a number of other molecular causes of overgrowth disorders have been identified. These include single gene disorders (Table 5), the imprinting disorder BWS (discussed in section 1.2.4), and chromosomal disorders (Table 6). The single gene disorder Malan syndrome is described in this section. Simpson-Golabi-Behmel syndrome has been described in Section 1.2.5.

Table 5: Other selected single gene disorders associated with overgrowth

Gene	Disorder	Protein
GPC3	Simpson-Golabi-Behmel syndrome	Extracellular proteoglycan[54]
NFIX	Malan syndrome	Transcription factor[9]

1.3.3.1 NFIX: Malan syndrome

In 2010, *NFIX* was identified as the cause of Marshall-Smith syndrome, a disorder that has some features of an overgrowth disorders such as increased birth length, macrocephaly and advanced skeletal maturation but subsequently results in failure to thrive and life-threatening respiratory difficulties[9]. A separate overgrowth phenotype, now termed Malan syndrome [134] was simultaneously found to be caused by a separate spectrum of variants in *NFIX* [9]. It is thought that simple haploinsufficiency results in Malan syndrome, whereas Marshall-Smith syndrome is associated with variants that escape nonsense-mediated decay and have a dominant-negative effect[9]. However, these conditions have some overlapping features and it may therefore be most appropriate to describe these as *NFIX*-related overgrowth disorders. *NFIX* is thought to play a key role in chondrocyte differentiation[9].

1.3.3.3 Chromosomal disorders associated with overgrowth

Table 6: Chromosomal translocations, deletions and duplications associated with overgrowth

	Putative gene(s)/genomic element	Associated single gene overgrowth disorder
Translocation 2q37[135,136]	*NPPC*	
3q13.31 deletion[137]	*ZBTB20*	Primrose syndrome[138]
4p16.3 duplication[139]	*FGFR3*[140]	CATSHL syndrome[141,142]
9q22.3 deletion	? *PTCH1* and regulatory element(s)	Gorlin syndrome
13q14.2q14.3 deletion[143]	?multiple genes	
13q31.3 duplication[144]	*MIR17HG* (miR-17~92 cluster)[145]	
15q26 duplication[146]	? *LRRK1*[147]	
19p13.13 deletion[148]	*NFIX*[9]	
22q13.3 deletion (Phelan-McDermid syndrome)[149]	? *PARVB*[150]	

1.4 Next Generation Sequencing

Sanger sequencing was developed in 1977 by Frederick Sanger and provided a robust and accurate method of determining the sequence of DNA bases[151,152]. This technology enabled the entire human genome to be sequenced in a huge decade-long international undertaking, the Human Genome Project, with the draft human genome published in 2001[153,154]. However, the traditional method of identifying novel human disease genes through positional cloning remained a laborious and time consuming process[155]. The advent of next generation sequencing (NGS), enabling large scale high throughput sequencing at ever decreasing costs, dramatically increased the rate of gene discovery. In 2010, Ng et al. published the first example of NGS being used to identify a novel human disease gene[156]. Several novel overgrowth genes were identified using NGS prior to the commencement of this study in 2015.

NGS is known as massively parallel sequencing technology, with many millions of reads (pieces of DNA sequence) being produced per run. There are a number of different NGS sequencing technologies. The Illumina platform (sequencing by synthesis, SBS) is a popular and economic option[157].

1.4.1 Library preparation

NGS requires the DNA (or RNA) of interest to be prepared into a library that is compatible with the desired sequencing platform[157]. The preparation of a high quality library is essential

to obtain high quality sequencing data with maximal genomic coverage[158] As sequencing platforms have improved, the methods for constructing NGS libraries have also improved and there are now a wide variety of NGS library preparation protocols[157,159].

All library preparation methods involve fusing fragments of DNA to adaptors that contain sequences necessary for immobilisation on a solid surface for sequencing[159]. The core steps in preparing DNA libraries for NGS analysis are: 1) fragmentation or sizing of DNA (by physical, enzymatic or chemical methods) to the required length 2) end repairing and ligation to oligonucleotide adaptors 3) amplification by PCR (in most protocols) and 4) quantitation of the DNA libraries[160].

1.4.2 Sequencing by synthesis

The DNA libraries are loaded into a flow cell where the fragments bind to primers complementary to the adaptors. Bridge amplification with PCR cycles of denaturation and synthesis produces clonal clusters.

After cluster amplification, the sequencing stage begins. Reagents for DNA synthesis and fluorescently labelled nucleotides flow over the cell. There is a different colour fluorophore to distinguish each of the four bases and a reversible chain terminator is attached so only one nucleotide can be added per cycle. The flow cell is imaged and the emission wavelength and intensity used to identify which base is added to each cluster in each cycle. The flow cell is washed, removing the chain terminator, and the next cycle begins. The cycles are repeated to build up the DNA reads at each cluster.

1.4.3 Bioinformatic analysis and variant interpretation

The short DNA reads are then aligned and mapped to a reference sequence and variants are called and annotated. Bioinformatic software programmes are used to construct this 'bioinformatic pipeline' resulting in a list of variants for analysis.

To filter pathogenic variants from the thousands of benign variants present in each individual, a number of strategies are used. Databases such as dbSNP, 1000 Genomes, ESP, ExAC and gnomAD provide the frequency of variants in the general population[161–164]. A Minor Allele Frequency (MAF) of <0.05% is often specified to identify de novo dominant pathogenic variants, and <1% if a recessive disorder is suspected[165]. Disease or gene specific databases, such as UMD-FBN1 for Marfan syndrome, are also available for some conditions[166].
Trio analysis, comparing the sequence of an individual to that of both parents, is a powerful technique for identifying pathogenic variants. This is particularly useful for identifying de novo dominant variants[165].

Bioinformatic predictive software algorithms, such as SIFT, PolyPhen-2, and MutationTaster, provide *in silico* analysis of whether a variant is likely to be benign or pathogenic[167–169]. These predictive tools use data such as conservation between species, the physical and chemical properties of amino acids, and the likely effect on the resulting protein structure and function, and are most useful for missense variants. The pLI (probability of being loss of function intolerant) score is useful for interpreting loss of function variants. A gene with a high pLI score indicates that haploinsufficiency is not tolerated[163]. There are also a number of

software programmes for predicting the disruption or creation of splice site consensus sequences[170].

Standards and guidelines for classification of variants as ‘pathogenic’, ‘likely pathogenic’[, ‘uncertain significance’, ‘likely benign’, and ‘benign’ have been developed for use in the clinical diagnostic laboratory by the American College of Medical Genetics and Genomics and the Association for Molecular Pathology[171] (see Appendix).

1.4.4 NGS and diagnostic testing

At the beginning of this study, clinical testing for overgrowth disorders in NHS regional genetics laboratories in the UK was limited to a small number of single genes (including *NSD1*, *PTEN*, and *GPC3*), BWS testing, and CGH microarray. Sequential testing of single genes was a lengthy and expensive process and the provision of diagnostic clinical testing of single genes had not kept pace with the exponential increase in genes identified by NGS. Pursuing a molecular diagnosis was therefore a difficult, and sometimes fruitless task.

Over the past five years, large scale NGS has been introduced to the clinical testing arena. This has included clinical exome sequencing (sequencing of all genes with a known association with a human disease phenotype), whole exome sequencing (WES), whole genome sequencing (WGS) through the 100,000 Genomes Project, and finally WGS as a clinical test for NHS patients with rare disease and cancer. As the scale of sequencing increases, the diagnostic yield is likely to be greater. This advantage must be weighed against the need for huge amounts of computing power and requirement for bioinformatics expertise to store and analyse such large quantities of complex data, and the practical implications of

taking consent and feeding back results to patients. Even within a relatively small defined panel of overgrowth disorders, the clinical features, prognosis and genetics may vary considerably. WES and WGS are very likely to produce many more results of uncertain significance, and incidental findings. Taking informed consent for a test with these possible outcomes poses a significant challenge to the clinician and raises a number of ethical issues that must be addressed. The consenting process for these investigations is time consuming and requires additional resources for each patient in terms of clinic time.

In a diagnostic setting, an NGS gene panel that includes only known genes of relevance to the phenotype being investigated overcomes many of these difficulties. Panels also have the advantage that due to the smaller size it is possible to obtain better coverage of the specific genomic regions of interest, a greater number of samples can be run on a single sequencing run, and a greater depth of sequencing coverage can be achieved at a lower cost than with larger designs. In addition, as panel can be custom designed, it is possible to include newly discovered genes that are not present in 'off the shelf' kits for clinical exome sequencing. At the outset of this study, a panel-based approach as a first line investigation followed by large-scale sequencing for individuals who remain undiagnosed following this test was chosen as the strategy.

1.5 Clinical phenotyping

Deep phenotyping refers to 'the complete and detailed understanding of the spectrum of phenotypic abnormalities associated with each disease entity' [172]. Accurate and comprehensive phenotyping is essential to identify clinically relevant sub-classifications within disease groups. However, phenotyping is often imprecise and incomplete in the medical literature[172]. Sequencing projects have successfully identified novel genes causing

overgrowth, however these novel disorders are often identified in only a small number of individuals and the associated phenotypic information is unlikely to represent the full spectrum of the condition. The phenotypic information provided by recruiting clinicians is often brief. Detailed phenotyping of overgrowth disorders is therefore required to understand the natural history, spectrum of clinical features and genotype-phenotype correlations of these rare diseases and thus enable accurate diagnosis and optimal management for patients.

Comprehensive and accurate recording of phenotypes is enabled by using a resource such as The Human Phenotype Ontology (HPO) project, which describes over 10,000 terms and over 13,000 subclasses [173]. Using a well-defined set of terms for clinical phenotyping is increasingly important as clinical practice begins to utilise large scale genomic sequencing and moves towards personalised medicine.

1.6 Aims

1.6.1 The Phenotyping of Overgrowth Disorders (POD) study

The introduction summarises what is currently known on overgrowth disorders and the gaps in our knowledge. Areas in which knowledge is lacking include:

- the phenotypic spectrum of novel overgrowth disorders recently identified by NGS
- the phenotypic spectrum of 'pre-NGS era' overgrowth disorders diagnosed by genomic testing instead of clinical diagnosis
- the underlying genomic aetiology in undiagnosed individuals with overgrowth
- the proportion of individuals in whom current molecular testing techniques identify a diagnosis

- the optimal diagnostic strategy for investigating individuals with overgrowth
- genotype-phenotype correlations in overgrowth disorders
- clinical management strategies for each rare overgrowth disorder including screening for potential complications

My hypothesis is that addressing these areas will improve our ability to diagnose, to screen for complications, to offer genetic counselling, and to provide personalised care for patients.

Therefore, the following aims were developed:

1) To recruit and phenotype a large cohort of individuals with rare genetic overgrowth disorders
2) To design an overgrowth panel to provide diagnostic testing to individuals without a molecular diagnosis
3) To identify the genomic aetiology of overgrowth in the cohort
4) To explore the role of deep phenotyping in diagnosis of overgrowth disorders
5) To identify genotype-phenotype correlations in overgrowth disorders
6) To identify medical complications in overgrowth disorders to guide clinical care

The Phenotyping of Overgrowth Disorders (POD) study was devised in order to achieve these aims. The principal objective of this study was to characterise a cohort of individuals with rare genetic overgrowth disorders to increase our knowledge of the phenotypic spectrum, natural history and genotype-phenotype correlations in these conditions. The secondary objectives were to provide a cohort for further study into the pathophysiology of overgrowth disorders and to support future multicentre studies of therapeutic intervention in overgrowth disorders.

Chapter 2. METHODS: The Phenotyping of Overgrowth Disorders study

2.1 The Phenotyping of Overgrowth Disorders Study

The Phenotyping of Overgrowth Disorders (POD) study (REC 15-YH-0252; IRAS 161520, CPMS ID: 19361) was funded by an NIHR Rare Disease Translational Research Collaboration (RD-TRC) Doctoral Fellowship.

The study protocol and patient facing documentation were written in accordance with NIHR guidelines. An overview of the study is provided in Figure 2. Patient information leaflets (PILs) and consent forms were devised for participant groups including competent adults, parents, children, young people, and adults with intellectual disability not competent to give consent (consultee forms). An application for ethical approval was submitted through the Integrated Research Application System (IRAS). The study Sponsor was the University of Birmingham. A favourable ethical opinion was given by the Research Ethics Committee (NRES Committee Yorkshire and the Humber – Leeds East) on the 3/7/2015. The first participant was recruited on the 13/8/2015 and the study was adopted onto the NIHR Clinical Research Network (CRN) Portfolio.

Initially participants were identified from clinical genetics outpatient clinics at Birmingham Women's Hospital (BWH) and paediatric endocrinology outpatient clinics at Birmingham Children's Hospital (BCH). Recruitment took place in the clinical genetics outpatient clinic at BWH and in the NIHR Clinical Research Facility (CRF) at BCH. In April 2016 the study was approved under the Musketeer's Memorandum (NIHR Rare Genetic Disease Research

Consortium) for recruitment from all regional clinical genetics centres in the United Kingdom.

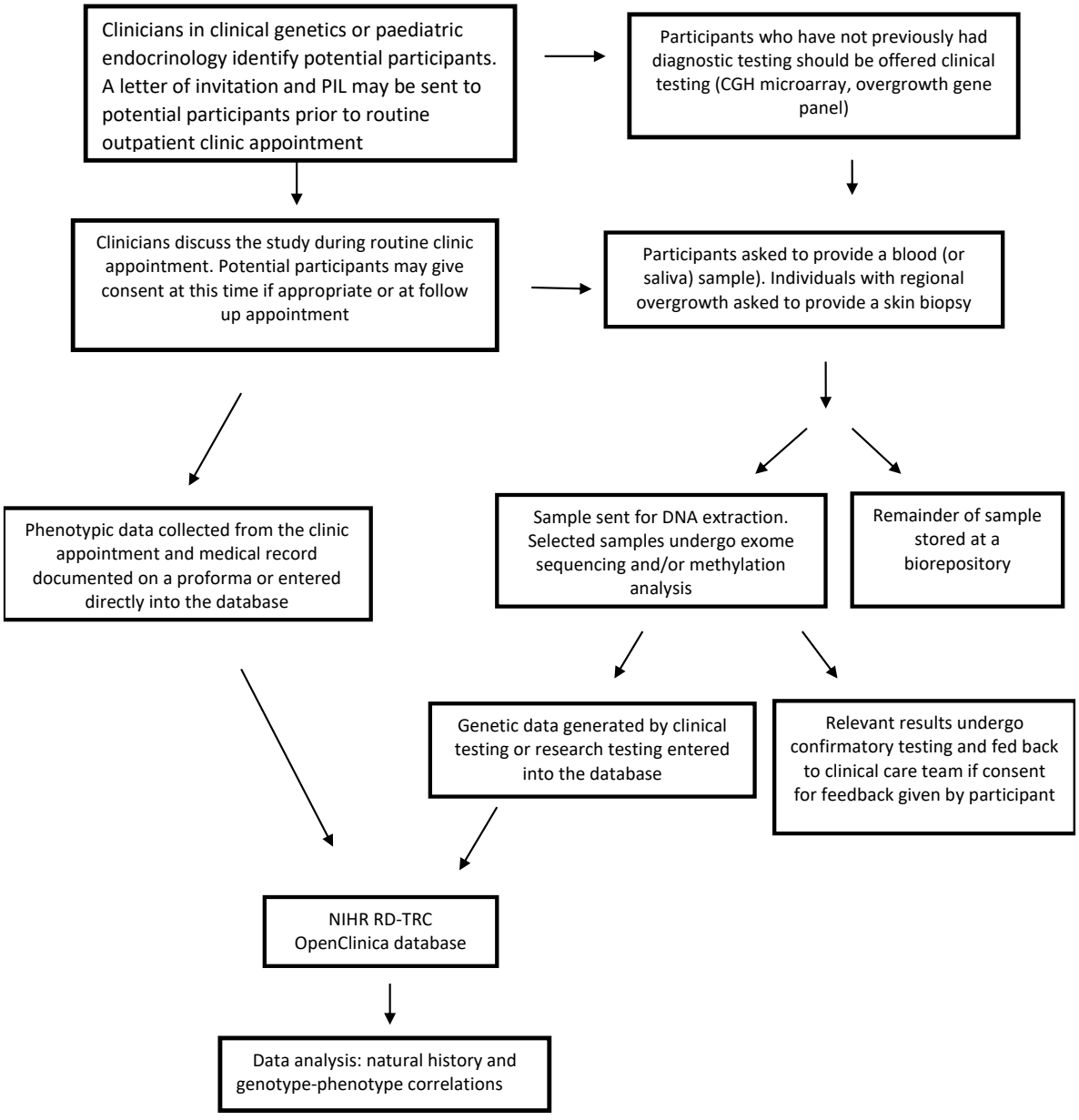


Figure 2: An overview of the POD study

2.1.1. Inclusion criteria

The inclusion criteria for the POD study were:

Participants may be of any age. Participants must meet one of the following criteria:

1. Height and/or head circumference greater than two standard deviations above the mean in association with one or more of the following:
    a. Dysmorphic facial features
    b. Developmental delay/intellectual disability
    c. Congenital anomaly
    d. Childhood tumour

Or

2. Height and/or head circumference more than three standard deviations above the mean

Or

3. Regional overgrowth

Or

4. A genetic or genomic variant associated with overgrowth

Or

5. Be a parent of an individual meeting one of the above criteria (unaffected parent participant)

2.1.2 Exclusion criteria

An individual may not enter the study if any of the following apply:

1. The individual has a clinical diagnosis of a genetic condition causing tall stature or increased head circumference that is not considered to be a primary overgrowth disorder (for example, a connective tissue disorder such as Marfan syndrome).
2. The individual has tall stature or increased head circumference that is solely due to an alternative diagnosis (for example an acquired endocrine condition such as acromegaly)
3. The individual does not give consent for participation in the study

2.2 Phenotypic data

2.2.1 Phenotypic data set and Case Report Form

A Case Report Form (CRF) was designed (See Appendix) to collect comprehensive 'deep' phenotyping data about every participant. This included sections on general information such as age, genetic investigations, pregnancy and neonatal history, clinical features at birth, growth data, medical history (including endocrine, cardiac, respiratory, neurological, renal, gastrointestinal, musculoskeletal, dermatological, ophthalmic, audiological, immunological, malignancy and psychiatric), development and behaviour, features of adult life, family history, and examination findings including dysmorphology. Specific data fields relevant to

our knowledge of the features of known overgrowth conditions (e.g. macroglossia and omphalocoele in BWS) were included.

2.2.2 Measurement of growth

Height (or length in participants under the age of two), weight and occipitofrontal circumference (OFC) were measured in accordance with Royal College of Paediatrics and Child Health (RCPCH) guidelines. Centiles and standard deviation values were generated using the LMSgrowth Excel add-in (program version 2.77, compiled on 15 September 2012, authors Huiqi Pan and Tim Cole copyright © 2002-12 Medical Research Council, UK) which uses 1990 British growth data. The LMS method uses power (L), mean (M), and coefficient of variance (S) curves to convert measurements into exact SD scores[174].

2.2.3 OpenClinica electronic data capture

In addition to the paper CRF, an electronic data capture (EDC) plan was designed for use on the OpenClinica (OC) platform (https://openclinica.com). To facilitate the collection of a clean data set, data items were coded with Human Phenotype Ontology (HPO), OMIM and SNOMED CT (Systematized Nomenclature of Medicine – Clinical Terms) wherever possible. 'Closed' data collection fields were designed, for example requiring 'yes/no' answers, or the selection of a term from a drop-down list. Additional data fields remained hidden unless a participant was recorded to have specific phenotypic features. For example, if a participant had no dysmorphic features, no further data fields would appear. However if a participant did have dysmorphic features, many more data fields would then appear. This meant data could

be entered in a time efficient manner but also enabled the collection of 'deep' phenotypic data when needed.

The electronic data capture plan was modelled on the paper CRF, with sections on general information, perinatal, growth, development, medical, family history, pedigree and dysmorphology. The data fields and coding terms were supplied to the NIHR RD-TRC data team. *The NIHR RD-TRC data team constructed the database in OpenClinica.*

2.3 Molecular investigations

Participants with generalised overgrowth followed the molecular testing strategy in Figure 3. *Clinical diagnostic testing by the West Midlands Regional Genetics Laboratory (WMRGL) included CGH microarray, single gene testing, and BWS testing, as indicated by the clinical presentation.*

Participants with regional overgrowth and a clinical suspicion of a mosaic PI3K/AKT/mTOR pathway disorder were offered buccal swab and/or skin biopsy. *The WMRGL extracted DNA from buccal samples and cultured fibroblasts (and paired blood samples) and provided clinical diagnostic testing of the PIK3CA gene on a validated cancer panel.*

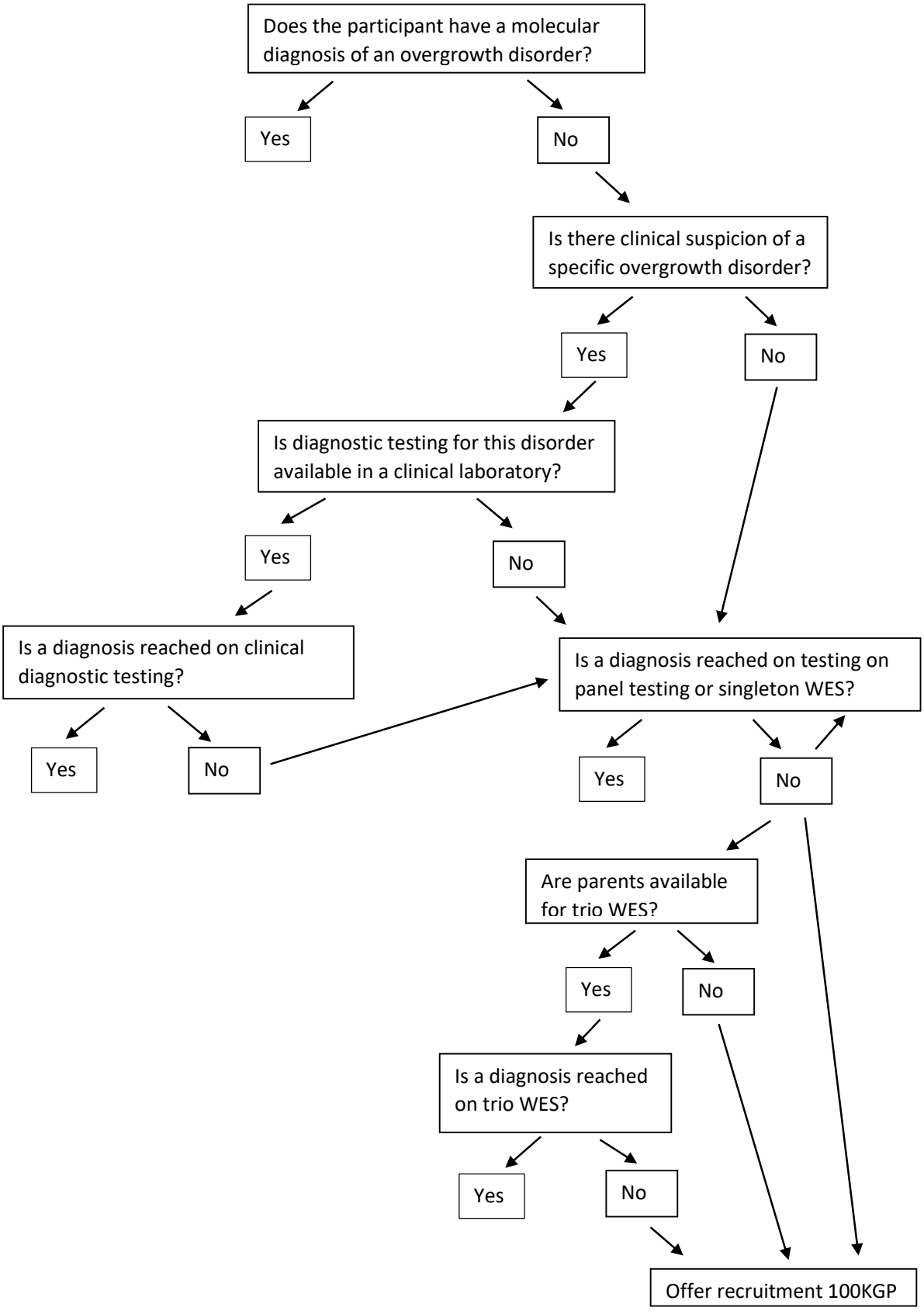


Figure 3: Pathway for molecular testing for generalised overgrowth in the POD study

Participants who remained undiagnosed following panel and/or exome sequencing in the POD study were offered recruitment to The 100,000 Genomes Project, a national clinical transformation project established to sequence genomes from NHS patients with rare disease or cancer (https://www.genomicsengland.co.uk).

2.3.1 Next Generation Sequencing panel

A PubMed database search was performed in October 2014 for known genes in which variants have been shown to cause a generalised (height and/or OFC) or regional overgrowth phenotype in humans. Search terms were overgrowth AND gene AND syndrome. 20 genes were identified in the literature and included on the overgrowth gene panel (see Table 7).

Table 7: Genes on 20 gene panel

Genes	Overgrowth disorder
NSD1	Sotos syndrome[21]
EZH2	Weaver syndrome[27]
PTEN	PTEN hamartoma tumour syndrome[175,176]
DIS3L2	Perlman syndrome[57]
GPC3	Simpson-Golabi-Behmel syndrome[54]
NFIX	Malan syndrome[9]
RNF135	RNF135 related overgrowth[177]
NPR2	NPR2 related overgrowth[178]
DNMT3A	Tatton-Brown-Rahman syndrome[88]
DICER1	GLOW syndrome (global developmental delay, lung cysts, overgrowth and Wilms tumour)[179]
SETD2	SETD2 related overgrowth/Luscan-Lumish syndrome[180]
AKT1	Proteus syndrome[58]
AKT2	Hypoinsulinaemic hypoglycaemia with asymmetrical overgrowth[108]
AKT3	AKT3 related megalencephaly[103]
PIK3CA	PIK3CA-related overgrowth spectrum[106]
PIK3R2	Megalencephaly-polymicrogyria-polydactyly-hydrocephalus (MPPH) syndrome[103]
CDKN1C	Beckwith-Wiedemann syndrome (BWS)[49]
MTOR	Smith-Kingsmore syndrome[107]
ZBTB20	Primrose syndrome[138]

Stored samples from patients with known pathogenic variants in overgrowth genes were chosen for validation of the panel (Table 8). These were clinical samples that had undergone diagnostic genetic testing in the WMRGL. Prior consent was given by the families of these patients for the use of stored DNA samples in developing new tests. The samples were selected to represent a range of pathogenic variants, including insertions and deletions of varying sizes, duplications, missense variants, and splice site variants. The number of genes represented was limited by the small number of genes for which diagnostic testing was currently available in the WMRGL.

Table 8: Samples with known variants for validation of the gene panel

Sample	Gene	Variant	Variant category	DNA conc (ng/ul)
1	*NSD1*	c.4378+1delGTGA intron 9	Splice site	241
2	*NSD1*	AC insertion nucleotide 1730 exon 5	Insertion 2bp	291
3	*NSD1*	TG deletion nucleotide 3464-66 exon 5	Deletion 2bp	740
4	*NSD1*	T insertion nucleotide 5744	Insertion 1bp	456
5	*DIS3L2*	Homozygous deletion exon 6 c.367-41553_602+40962	Deletion 82.8kb	40.1
6	*CDKN1C*	Heterozygous deletion c.301delG	Deletion 1bp	660
7	*PTEN*	Duplication 24 bases c.22_234dup24 exon 5	Duplication	28.3
8	*PTEN*	Heterozygous missense c.389 G>A exon 5	Missense	58.7
9	*GPC3*	Hemizygous missense mutation c.254 C>T exon 3	Hemizygous missense	45.2
10	*NSD1*	c.6004_6007delGACA	Deletion 4bp	580
11	*NSD1*	4bp insertion exon 5	Insertion 4bp	5.65
12	*NSD1*	c.6376delG exon 22	Deletion 1bp	600
13	*PTEN*	Missense c.202T>G exon 3	Missense	466
14	*NSD1*	C1905R exon 18	Missense	295
15	*NSD1*	R1473X exon 10	Missense	47.4

Commercially available options for library construction including Agilent Haloplex, Agilent Nextera, Agilent SureSelect QXT, and Illumina TSCA were compared. The factors taken into consideration were: the size of the genomic regions of interest to be covered by the panel, the level of coverage achieved by each panel, and the cost of the kits. Due to the relatively small size of the panel, Agilent Haloplex and Agilent Nextera were immediately excluded as these

are marketed for large-scale designs. Agilent SureSelect QXT and Illumina TSCA were compared on a cost basis and were equal in price. These two technologies were therefore selected for comparison.

2.3.1.1 Agilent SureDesign

The Agilent SureSelect QXT panel was designed using the Agilent online SureDesign wizard.

Step1: The design name (*overgrowth)* and species *(H. sapiens*) were entered.

Step 2: The target regions for capture were entered by gene name (e.g. *nsd1*) and the genome annotation databases for obtaining genomic coordinate information for these targets were selected (RefSeq, Ensembl, CCDS, Gencode, VEGA, SNP, CytoBand). The regions of interest were specified as coding exons plus 25bp 3'UTR and 25bp 5'UTR.

Step 3: The target summary was reviewed and the genomic regions identified by SureDesign were viewed in the UCSC Genome Browser.

Step 4: Parameters for probe selection were entered: stringency set as 'most stringent' (to mask repetitive regions).

Step 5: The program's algorithms selected the probe sequences for the design

Steps 1 through 5 were then repeated twice, changing the stringency setting to 'moderately stringent' then 'least stringent' in order to cover regions missed by the 'most stringent' setting. The probe groups were then combined to create the final design consisting of 4125 probes covering 111.659kbp. The panel contains two gaps in coverage that are both in intronic regions (of *NSD1* and *SETD2*, see Figure 4and Figure 5). Variants in these regions have not been reported to be associated with a disease phenotype.

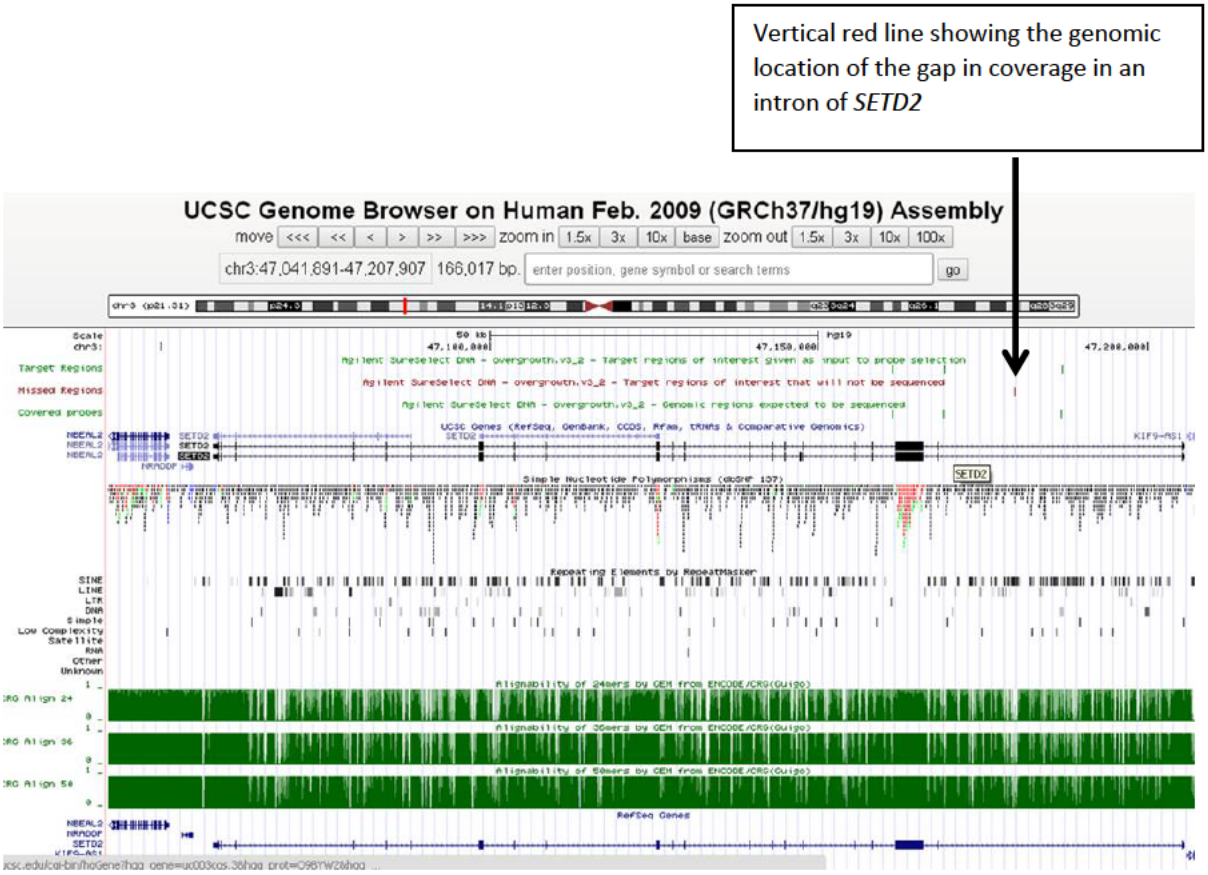


Figure 4: UCSC Genome Browser view of SETD2 with custom track 'Missed Regions'

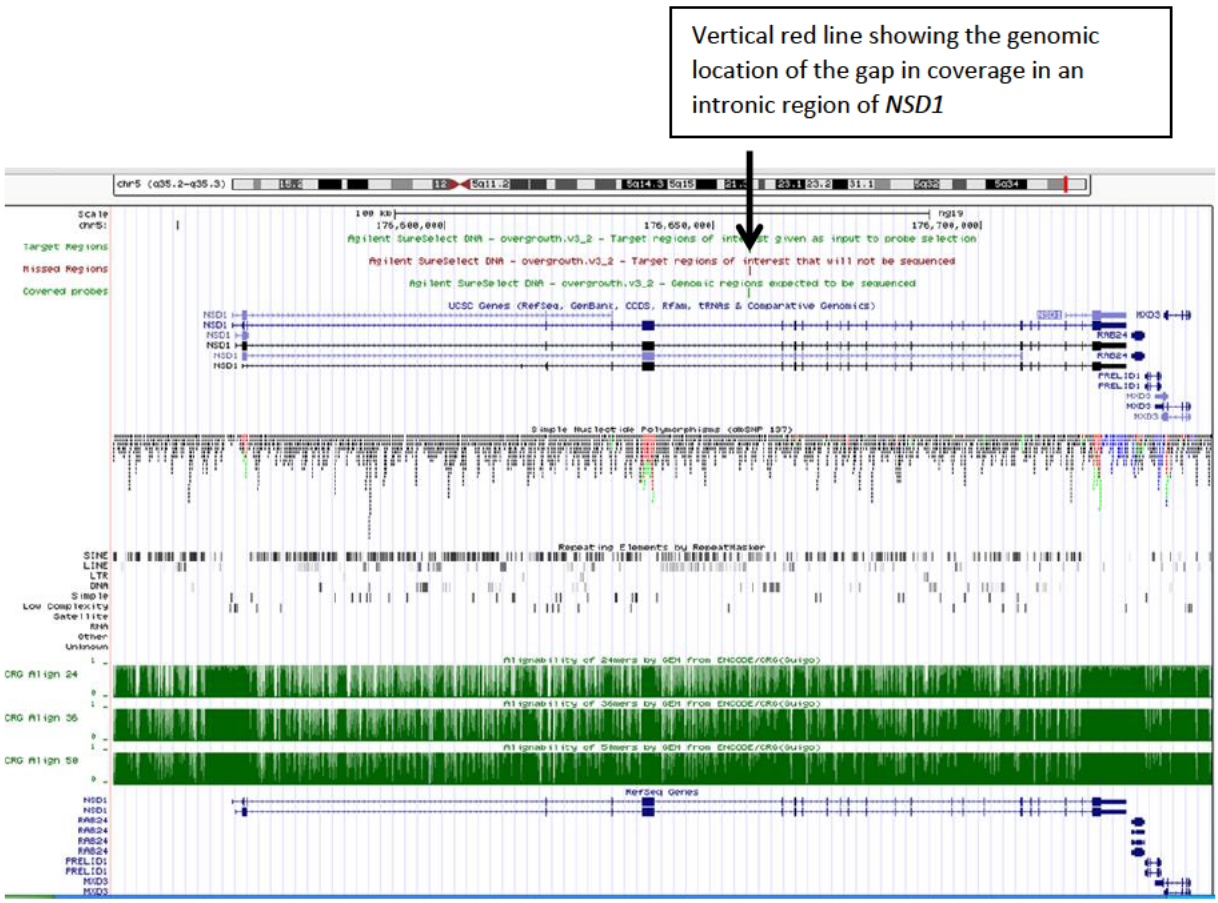


Figure 5: UCSC Genome Browser view of NSD1 with custom track 'Missed Regions'

2.3.1.2 Agilent SureSelect QXT Target Enrichment Protocol

DNA was extracted from study samples by an NHS laboratory technician in the WMRGL according to local SOPs.

Following extraction of DNA from study samples, the SureSelect QXT Target Enrichment for Illumina Multiplexed Sequencing version C0, January 2015 was followed. The protocol consisted of 4 stages: 1) DNA quantification, 2) sample preparation, 3) hybridisation and capture, and 4) indexing and sample processing for multiplexed sequencing.

1) DNA quantification: DNA samples were quantified using the Qubit according to standard protocols in the WMRGL. Samples were diluted using two serial fluorometric assays to a final concentration of 25ng/ul.

2) Sample preparation: The DNA samples were enzymatically fragmented and adaptors added to the ends of the fragments in a single reaction. The adaptor-tagged libraries were purified using AMPure beads, amplified using PCR, and again purified using AMPure beads. At this stage the DNA library quantity and quality were assessed using the Agilent 2200 TapeStation and D1000 Screentape; the peak DNA fragment sizes were identified on the electropherogram and the concentration of each library measured by integrating under the peak. Libraries with peak fragment sizes in the optimal range of 245-325bp were taken forward to the next stage. (see Figure 6)

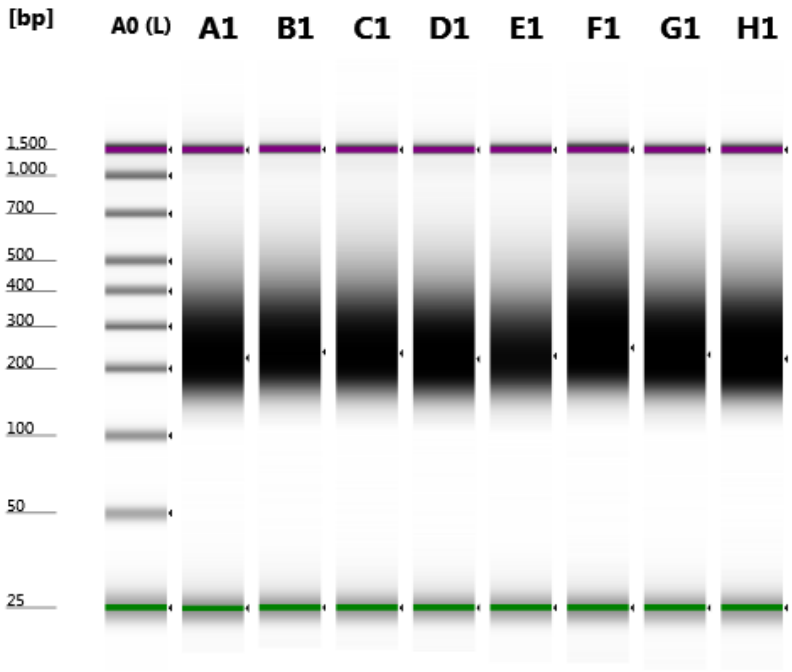


Figure 6: Example of D1000 ScreenTape image of libraries with fragment sizes in the optimal range

3) Hybridisation and capture: The prepared DNA libraries were hybridised to the SureSelect Capture Library. Streptavidin-coated magnetic beads were used to capture the hybridised DNA.

4) Indexing and sample processing for multiplexed sequencing: The captured libraries were PCR amplified using the appropriate pair of dual indexing primers to add index tags. The amplified captured libraries were purified using AMPure beads then assessed using the Agilent 2200 TapeStation and High Sensitivity D1000 ScreenTape. The peak DNA fragment sizes were identified on the electropherogram and the concentration of each library measured by integrating under the peak, checking that the peak fragment sizes were in the optimal 325-450bp range (see Figure 7). The samples were then pooled for multiplexed sequencing.

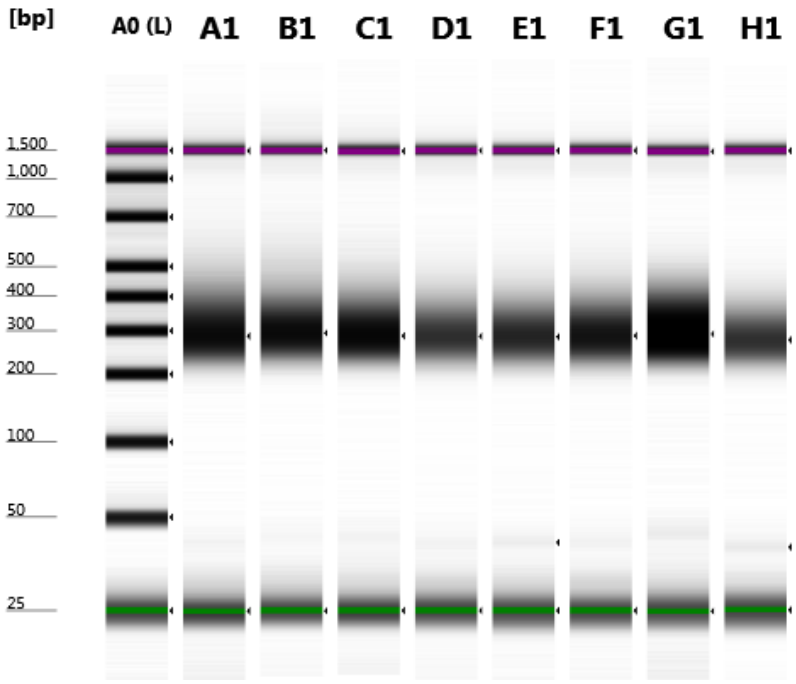


Figure 7: Example of High Sensitivity D1000 ScreenTape image of libraries with fragments in the optimal range

2.3.1.3 Sequencing on the Illumina MiSeq

Sequencing was performed under the supervision of an NHS laboratory technician in the WMRGL.

Libraries were prepared for sequencing according to the Illumina protocol 'Preparing DNA libraries for Sequencing on the MiSeq'. Libraries were denatured and diluted before a PhiX control spike-in was added. The prepared libraries were loaded onto a MiSeq reagent cartridge.

A MiSeq sample sheet with the information required for setting up, performing, and analysing a sequencing run was completed in Excel according to WMRGL protocols. The workflow was specified as 'GenerateFQ', Chemistry as 'Amplicon', number of cycles in Read 1 as 150 and number of cycles in Read 2 as 150. Index 1 and Index 2 were entered as specified in the Agilent QXT user guide.

The library mix was loaded onto the MiSeq reagent cartridge. A washed and dried flow cell, reagent cartridge, PR2 bottle and waste bottle were loaded into the MiSeq and sequencing commenced.

Following completion of the sequencing run, FASTQ data files were downloaded by the WMRGL technical team.

2.3.1.4 Data analysis using Agilent SureCall

Agilent's SureCall software was used for NGS data analysis. This software removes the adaptor sequences, aligns the reads to the reference genome, and calls variants. An example of the variant data produced by SureCall can be seen in the Appendix. This spreadsheet shows the variants identified in test sample 14, including the pathogenic missense variant c.202T>G in *PTEN*.

2.3.1.5 Illumina DesignStudio

The Illumina online design tool DesignStudio was used to design the TSCA panel. Species information (*Homo sapiens*), Source (UCSC) and Genome Build (hg19) were entered. The SNP source was set as 1000 Genomes and an amplicon length of 250bp was specified. The names of the genes of interest were entered and target selection set to all coding regions with 25bp into the 3' and 5' UTR. Probe design was automatically performed by the programme, taking into account GC content, specificity, probe interaction and coverage. The best design produced by DesignStudio contained gaps in coverage of exons in *PIK3CA*, *PIK3R2*, and *RNF125*, due to surrounding repetitive regions. *It was not possible to improve the design using DesignStudio and therefore the Illumina custom design concierge service was contacted to produce a custom design. This design covered coding regions only with no coverage of UTRs. The final design covered all regions except for a gap in coverage over an exon of MTOR (see Figure 8).*

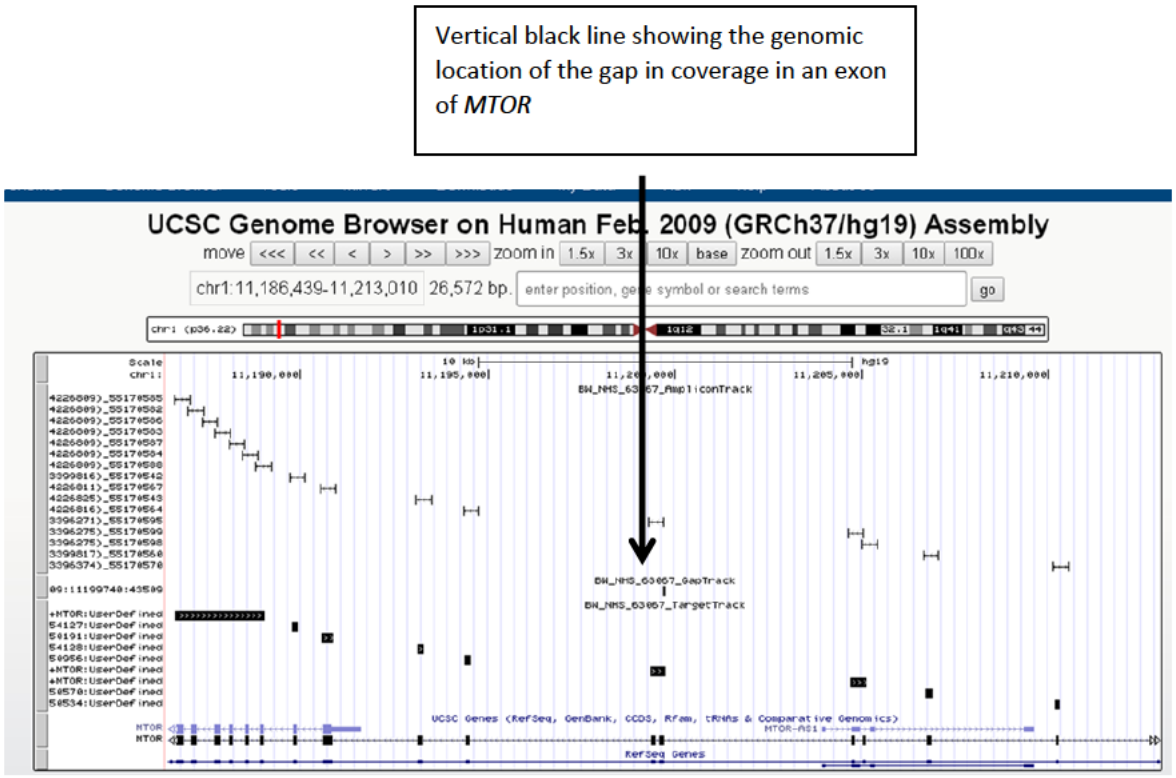


Figure 8: UCSC Genome Browser view of MTOR with custom track 'GapTrack'

2.3.1.6 Illumina TruSeq Custom Amplicon protocol

DNA was extracted from study samples by an NHS laboratory technician in the WMRGL according to local SOPs.

Following extraction of DNA from study samples, the protocol described in the Illumina 'TruSeq Custom Amplicon v1.5 Reference Guide' was followed. Stages included 1) DNA quantification, 2) Hybridisation of oligo pools, removal of unbound oligos, and extension and ligation of bound oligos, and 3) Amplification, clean up, normalisation, and pooling of libraries.

1) DNA quantification: DNA samples were quantified using a Qubit to check that the minimum concentration was at least 50ng/ul.

2) Hybridisation of oligo pools, removal of unbound oligos, and extension and ligation of bound oligos: A custom oligo pool containing upstream and downstream oligos was hybridised to the specific regions of interest. Unbound oligos were then removed using a size specific filter. Bound oligos were connected using a DNA polymerase and DNA ligase to extend from the upstream oligo through the targeted region and ligate to the 5' end of the downstream oligo.

3) Amplification, clean up, normalisation, and pooling of libraries: The extension-ligation products were amplified and adapters and sequences required for cluster formation were added in a PCR. Library quality was assessed by gel electrophoresis of an aliquot of the library and control on a 4% agarose gel. AMPure beads were used to purify the PCR products. The libraries were normalised, pooled, diluted and denatured prior to sequencing.

2.3.1.7 Sequencing on the Illumina MiSeq

Sequencing was performed under the supervision of an NHS laboratory technician in the WMRGL.

A MiSeq sample sheet with the information required for setting up, performing, and analysing a sequencing run was completed in Excel according to WMRGL protocols (see Appendix). The workflow was specified as 'Custom Amplicon', Chemistry as 'Amplicon', number of cycles in Read 1 as 150 and number of cycles in Read 2 as 150. Index 1 and Index 2 were entered as specified in the Illumina TruSeq Custom Amplicon user guide.

The library mix was loaded onto the MiSeq reagent cartridge. A washed and dried flow cell, reagent cartridge, PR2 bottle and waste bottle were loaded into the MiSeq and sequencing commenced.

Following completion of the sequencing run, VCF files were downloaded by the WMRGL technical team.

2.3.1.8 Data analysis using Illumina VariantStudio

Illumina VariantStudio software was used to annotate the variant data. An example of the data produced by VariantStudio software can be seen in the Appendix. This spreadsheet shows the variants identified in test sample 14, including the pathogenic missense variant c.202T>G in *PTEN*.

2.3.1.9 Selection of Agilent SureSelect QXT for overgrowth panel

Both strategies successfully called the pathogenic variants in 14 out of the 15 test samples. The large homozygous deletion in *DIS3L2* was not identified. Given the equal performance of the two strategies, comparisons of cost and efficiency were made. The Agilent platform was less expensive and had a faster workflow and was therefore chosen for samples undergoing panel testing in the POD study.

2.3.1.10 Redesign of panel and updated SureCall software

Several novel overgrowth genes were published in the literature during the course of the study. A further literature search was performed and a total of 24 additional genes were identified (Table 9). A redesign of the panel with 44 overgrowth genes using the Agilent online SureDesign wizard was therefore performed. The design strategy was the same as that for the 20 gene panel, with the exception of the regions of interest being specified as coding exons plus 10bp 3'UTR and 10bp 5'UTR instead of plus 25bp 3'UTR and 25bp 5'UTR. The non-coding exon 1 of *NSD1* was also included in the 44 gene design. The final design consisted of 2533 probes covering 174.874kbp.

Table 9: Genes included on panel version 1 and version 2

Genes on 20 gene panel	Overgrowth disorder
NSD1	Sotos syndrome[21]
EZH2	Weaver syndrome[27]
PTEN	PTEN hamartoma tumour syndrome[175,176]
DIS3L2	Perlman syndrome[57]
GPC3	Simpson-Golabi-Behmel syndrome[54]
NFIX	Malan syndrome[9]
RNF135	RNF135 related overgrowth[177]
NPR2	NPR2 related overgrowth[178]
DNMT3A	Tatton-Brown-Rahman syndrome[88]
DICER1	GLOW syndrome (global developmental delay, lung cysts, overgrowth and Wilms tumour)[179]
SETD2	SETD2 related overgrowth/Luscan-Lumish syndrome[180]
AKT1	Proteus syndrome[58]
AKT2	Hypoinsulinaemic hypoglycaemia with asymmetrical overgrowth[108]
AKT3	AKT3 related megalencephaly[103]
PIK3CA	PIK3CA-related overgrowth spectrum[106]
PIK3R2	Megalencephaly-polymicrogyria-polydactyly-hydrocephalus (MPPH) syndrome[103]
CDKN1C	Beckwith-Wiedemann syndrome (BWS)[49]
MTOR	Smith-Kingsmore syndrome[107]
ZBTB20	Primrose syndrome[138]
Additional genes on 44 gene panel	
RNF125	Tenorio syndrome[181]
CCND2	Megalencephaly-polymicrogyria-polydactyly-hydrocephalus (MPPH) syndrome [182]
APC2	APC2 related overgrowth[183]
EED	Cohen-Gibson syndrome[81]
PPP2R5B	PPP2R5B related neurodevelopmental/overgrowth disorder[184]
PPP2R5C	PPP2R5C related neurodevelopmental/overgrowth disorder[184]
PPP2R5D	PPP2R5D related neurodevelopmental/overgrowth disorder[185]
HERC1	HERC1 overgrowth disorder[186]
PDGFRB	Kosaki overgrowth syndrome[104]
PIGA	Simpson-Golabi-Behmel type 2[187]
CHD8	CHD8 overgrowth syndrome[188]
HIST1H1E	HIST1H1E syndrome[6]
BRWD3	BRWD3 syndrome[189]
TCF20	TCF20 syndrome[190]
FIBP	FIBP syndrome[191]
SUZ12	Imagawa-Matsumoto syndrome[81]
NLGN2	NLGN2 syndrome[192]
CHD4	CHD4-related syndrome Sifrim-Hitz-Weiss (SIHIWES)[193]
NFIA	NFIA syndrome[194]
KPTN	KPTN sydnrome[195]
PTCH1	Gorlin syndrome[196]
SUFU	Gorlin syndrome[197]
MED12	MED12-related disorders[198]
TBC1D7	TBC1D7 syndrome[199]
FGFR3	Camptodactyly, tall stature and hearing loss (CATSHL)[141]

2.3.1.11 NGS panel variant interpretation

Variants identified by SureCall software (example seen in the Appendix) were filtered as follows:

Read depth: > 30

MAF: < 0.0005 if heterozygous; <0.01 if homozygous

Filtered variants were then entered in Alamut Visual (see https://www.interactive-biosoftware.com/alamut-visual/. This software application integrates information from public databases (dbSNP, ESP, gnomAD etc), nucleotide and amino acid conservation across species, missense variant pathogenicity prediction tools (including SIFT, MutationTaster, and PolyPhen-2), and assessment of impact on splicing. Examples of the Alamut Visual window and analysis of variants are found in Chapter 5 Results; 5.2 NGS panel of overgrowth genes.

A PubMed literature search was performed to investigate if an identified variant had previously been published. The participant's deep phenotyping data was reviewed to assess if the putative diagnoses was consistent with the patient's phenotype.

Variants were classified as likely pathogenic or pathogenic according to the '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[200].

A variant confirmation form was submitted to the WMRGL to request confirmation (see Appendix) of pathogenic and likely pathogenic variants by Sanger sequencing. *Variant confirmation by Sanger sequencing was performed by NHS laboratory staff in the WMRGL to enable a diagnostic report to be issued to the participant's clinical team. Sanger sequencing of parental samples to establish segregation was performed when possible.*

2.3.2 Whole Exome Sequencing

As the study progressed, the ever decreasing cost of whole exome sequencing allowed the initial panel based strategy to evolve into a whole exome sequencing (WES) approach with a 'virtual panel'. This strategy had the advantage that the virtual panel could be easily updated with novel overgrowth genes. Participant data could be reviewed at any point for variants in novel genes without the need for resequencing.

Some participants who had already undergone panel testing underwent second line testing with WES. Later in the study, participants recruited to the project underwent WES as a first line investigation. Selected participants with generalised overgrowth, who remained undiagnosed after panel testing or singleton WES, underwent trio WES, with samples from the participant and both parent participants. This allowed for a 'gene agnostic' approach in addition to a virtual panel approach.

DNA was extracted from study samples by the West Midlands Regional Genetics Laboratory (WMRGL) technicians according to local SOPs.

Samples for WES were sent to the Genomics Core Facility next-generation sequencing service provider within the Biosciences Institute at Newcastle University in the UK. Sequencing libraries were prepared using the Twist Bioscience 'Twist Human Core Exome kit' with additional probes from the RefSeq gene panel. The DNA samples were sheared enzymatically and libraries sequenced to an average depth of 90x on an Illumina NovaSeq 6000 running a

2x100 bp flow cell. FASTQ files were sent to the WMRGL bioinformatic team for upload to the Congenica Sapientia genome analytics software platform (https://www.congenica.com).

Later WES samples were sequenced in house in the WMRGL on the Nonacus exome sequencing platform. This platform is validated for diagnostic panel testing of clinical samples from NHS patients.

2.3.2.1 WES variant interpretation using Congenica Sapientia

Congenica Sapientia genome analysis software was used for interpretation of variants in WES data. The HPO terms for each participant undergoing WES were supplied to the WMRGL bioinformatic team. *The WMRGL bioinformatic team uploaded these terms to Congenica*. Examples of the Sapientia interpretation window are shown in Chapter 5 Results; 5.3 Whole Exome Sequencing.

2.3.2.2 Singleton WES

Two virtual panels were applied to singleton WES data. Firstly, a custom virtual 44 gene overgrowth panel, and secondly, the Developmental Disorder Genotype-Phenotype Database (DDG2P). This database was created as part of the DDD study and is a list of curated genes that are associated with developmental disorders.

The Configuration in Congenica was set as follows for the 44 gene overgrowth panel:

Genes location: gene panel overgrowth 44

Population frequency: MAF ExAC, UK10K, 1000G <0.01

VEP Consequence: transcript ablation, splice donor variant, splice acceptor variant, stop gained, frameshift variant, stop lost, start lost, initiator codon variant, inframe insertion, inframe deletion, missense variant, protein altering variant, splice region variant

Zygosity: homozygous, heterozygous, hemizygous

Inheritance: not selected

If a causative variant was not identified in the 44 gene overgrowth panel, the Configuration was reset for DDG2P genes.

Genes location: DDG2P

Population frequency: MAF ExAC, UK10K, 1000G <0.01

VEP Consequence: transcript ablation, splice donor variant, splice acceptor variant, stop gained, frameshift variant, stop lost, start lost, initiator codon variant

Zygosity: homozygous, heterozygous, hemizygous

Inheritance: not selected

The information available in Congenica from databases (dbSNP, ExAC, UK10K, 1000G, ClinVar), haploinsufficiency score, missense variant pathogenicity prediction tools (SIFT and PolyPhen), the medical literature and the participant's phenotype were reviewed in relation to the lists of variants identified. Potential spliceogenic variants were assessed in AlamutVisual

Variants were classified as likely pathogenic or pathogenic according to the '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[200].

A variant confirmation form was submitted to the WMRGL to request confirmation (see Appendix) of pathogenic and likely pathogenic variants by Sanger sequencing. *Variant confirmation by Sanger sequencing was performed by NHS laboratory staff in the WMRGL to enable a diagnostic report to be issued to the participant's clinical team. Sanger sequencing of parental samples to establish segregation was performed when possible.*

2.3.2.3 Trio WES

For trio analysis, pedigree information was provided to the WMRGL bioinformatics team (Figure 9). *The WMRGL bioinformatics team uploaded this information to Congenica.*

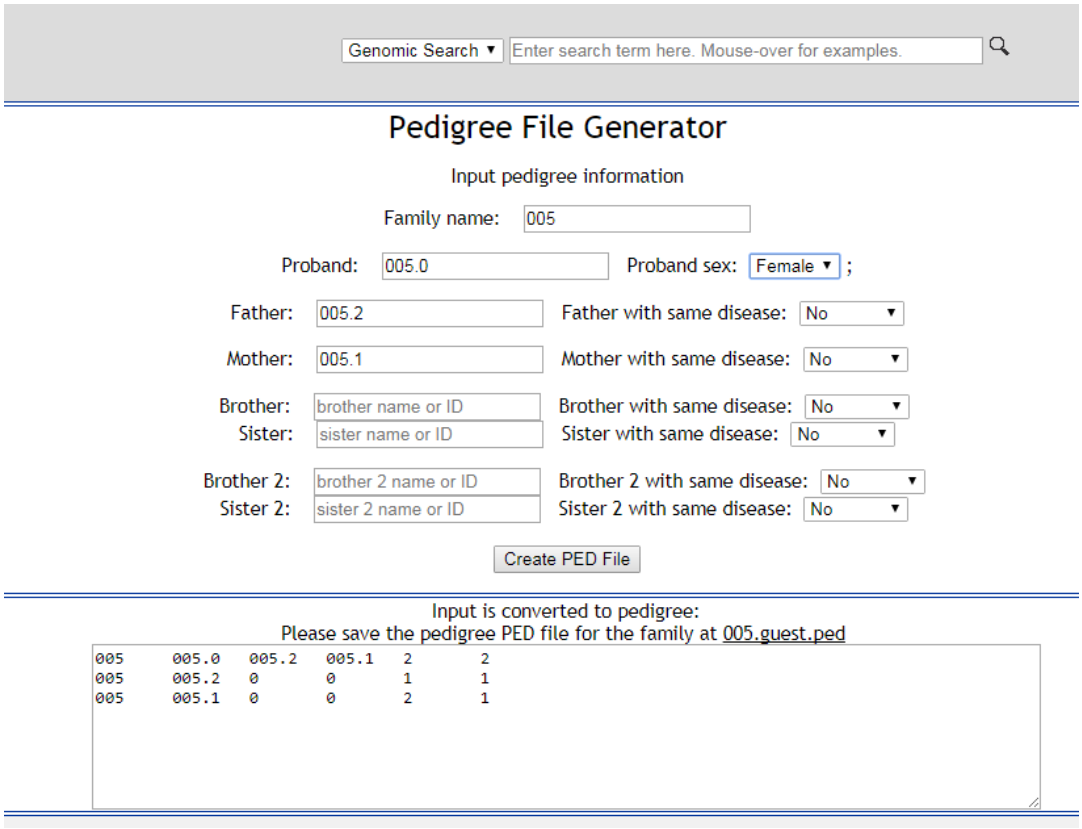


Figure 9: Example of pedigree information for upload to Congenica

Initial analysis was performed of virtual panels as per the singleton exome approach. If no causative variants were identified, a gene agnostic strategy was used. Three separate configurations were set to identify de novo dominant disorders, recessive disorders, and X-linked disorders.

De novo dominant disorders:

Genes location: not selected

Population frequency: MAF ExAC, UK10K, 1000G <0.01

VEP Consequence: transcript ablation, splice donor variant, splice acceptor variant, stop gained, frameshift variant, stop lost, start lost, initiator codon variant, inframe insertion, inframe deletion, missense variant, protein altering variant, splice region variant

Zygosity: heterozygous

Inheritance: de novo

Recessive disorders:

Genes location: not selected

Population frequency: MAF ExAC, UK10K, 1000G <0.01

VEP Consequence: transcript ablation, splice donor variant, splice acceptor variant, stop gained, frameshift variant, stop lost, start lost, initiator codon variant, inframe insertion, inframe deletion, missense variant, protein altering variant, splice region variant

Zygosity: homozygous

Inheritance: biparental, de novo

X-linked disorders

Genes location: not selected

Population frequency: MAF ExAC, UK10K, 1000G <0.01

VEP Consequence: transcript ablation, splice donor variant, splice acceptor variant, stop gained, frameshift variant, stop lost, start lost, initiator codon variant, inframe insertion, inframe deletion, missense variant, protein altering variant, splice region variant

Zygosity: hemizygous

Inheritance: maternal, de novo

The information available in Congenica from databases (dbSNP, ExAC, UK10K, 1000G, ClinVar), haploinsufficiency score, missense variant pathogenicity prediction tools (SIFT and PolyPhen), the medical literature and the participant's phenotype were reviewed in relation to the lists of variants identified. Potential spliceogenic variants were assessed in AlamutVisual.

Variants were classified as likely pathogenic or pathogenic according to the '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[200].

A variant confirmation form was submitted to the WMRGL to request confirmation (see Appendix) of pathogenic and likely pathogenic variants by Sanger sequencing. *Variant confirmation by Sanger sequencing was performed by NHS laboratory staff in the WMRGL to enable a diagnostic report to be issued to the participant's clinical team. Sanger sequencing of parental samples to establish segregation was performed when possible.*

Chapter 3. RESULTS: Description of the study cohort

100 participants recruited between August 2015 and December 2019 were included in the data analysis.

3.1 Flowchart of participant recruitment

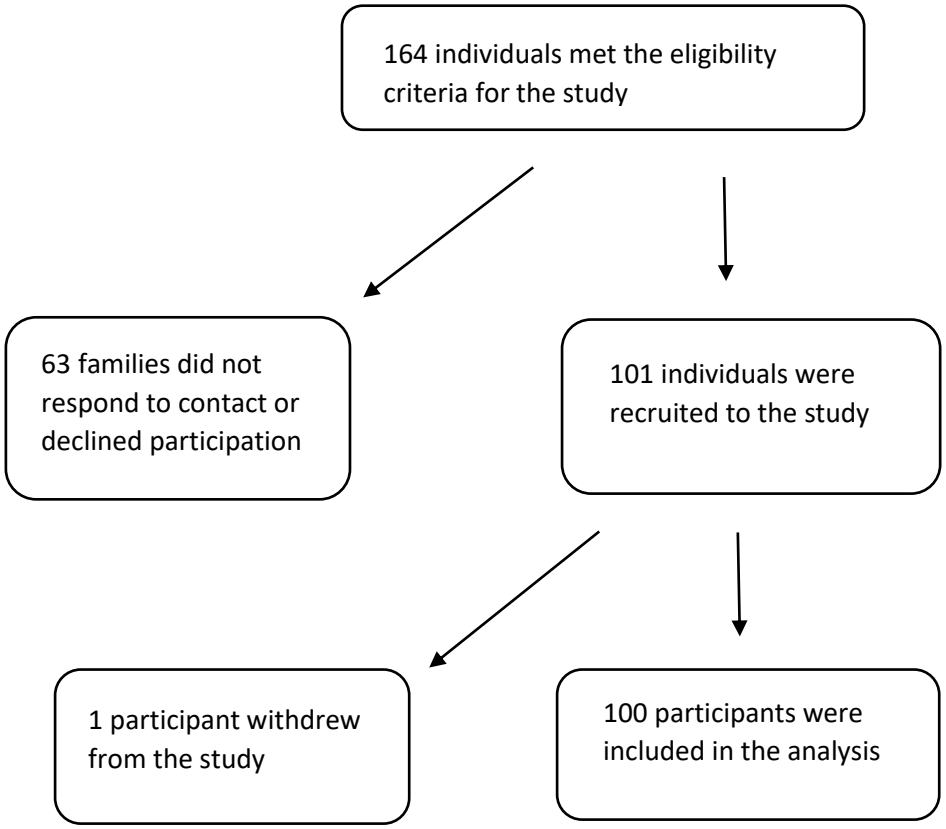


Figure 10: Summary of participant recruitment.

An additional 123 unaffected parents were also recruited.

See Appendix for summary table of all participants.

3.2 Method of data collection

Phenotypic data collection was performed at a face to face clinic appointment (68 participants), by other recruiting clinicians at face to face clinic appointments (20 participants), or from the medical records (12 participants).

3.3 Participant demographics

3.3.1 Sex

65 participants were male and 35 participants were female (Figure 11).

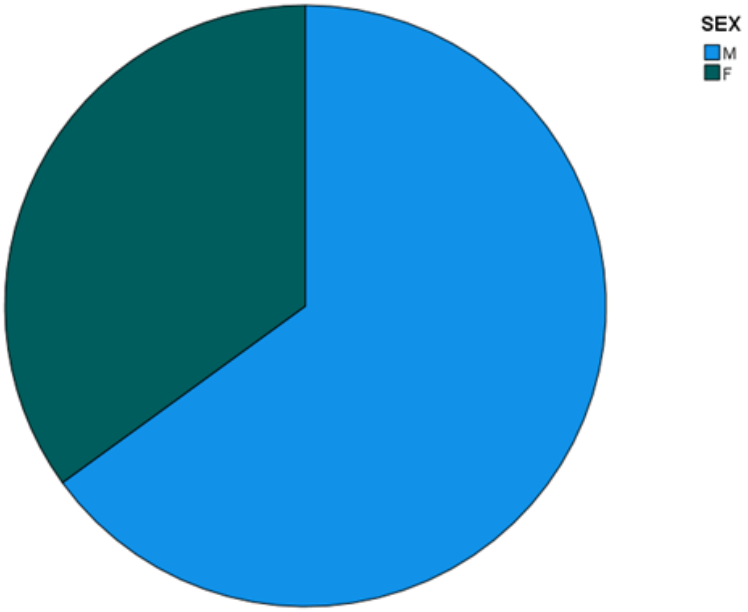


Figure 11: Total cohort - sex of participants

The higher proportion of male participants could be attributed to the excess of males with intellectual disability in the general population, generally thought to be due to X-linked intellectual disability syndromes[201]. However, no diagnoses of X-linked ID were made in any participants in the study.

3.3.2 Age

Participants ranged in age from 1-55 years with a mean of 11 years and median of 9 years (Figure 12).

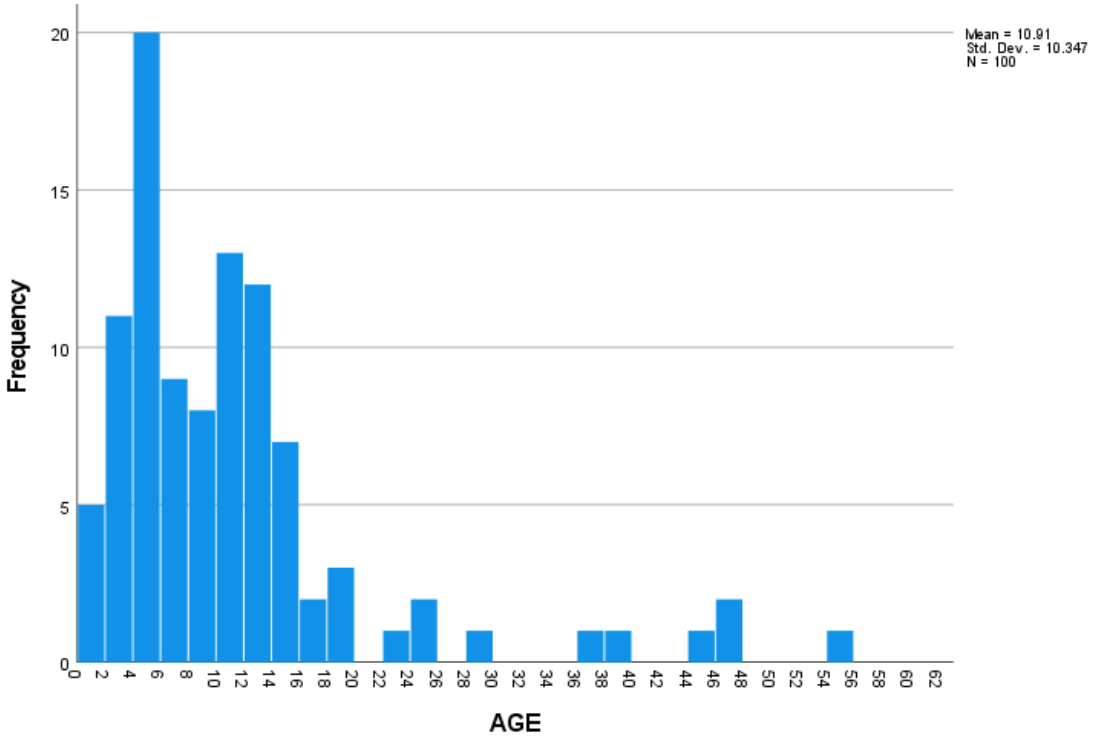


Figure 12: Total cohort - participant ages

This distribution with heavy skewing towards the paediatric population is likely to reflect referral patterns to clinical genetics. Children with increased growth and developmental delay are followed up by paediatricians who often refer to clinical genetics, while adults with intellectual disability may not be seen regularly by a clinician. The two most common age brackets for recruitment, age 4-6 and age 10-12, are the ages at which children start primary school and secondary school respectively. This pattern may reflect parents' and teachers' concerns about growth and development at these critical stages in education, prompting referral to services.

3.3.3 Trios

49 probands were recruited as trios with two unaffected parents and 26 were parent-child duos (23 mother-child duos and three father-child duos), giving a total of 123 unaffected relatives (the mother of siblings 99.0 and 99.3 counted once) (Figure 13).

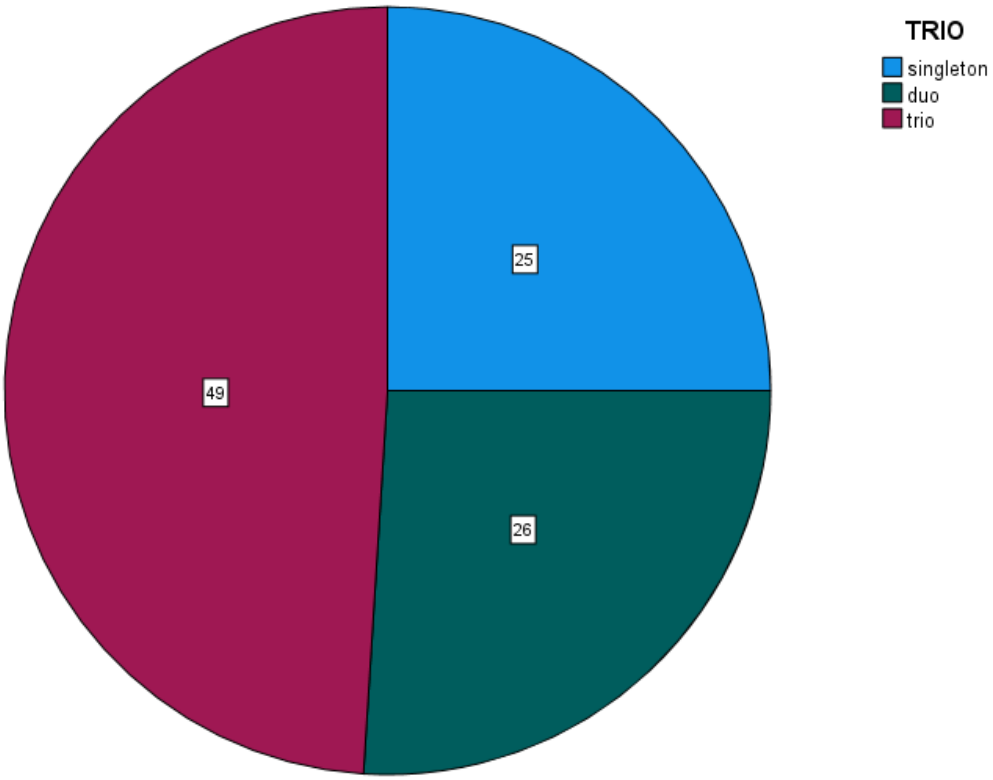


Figure 13: Probands recruited as singletons, parent-child duos and parent-child trios

3.3.4 Participants with molecular diagnoses at recruitment

At entry to the study, 18/100 had a molecular diagnosis of an overgrowth disorder and 82/100 did not (see Figure 14). The five individuals with PTHS are from two families and the seven individuals with deletions of *SUZ12* are from a single family. Three of the four individuals with Sotos syndrome had intragenic variants and the fourth had a deletion including *NSD1*. Another individual had a molecular diagnosis of Gorlin syndrome (*PTCH1*) (not included here as not usually considered to be an overgrowth disorder).

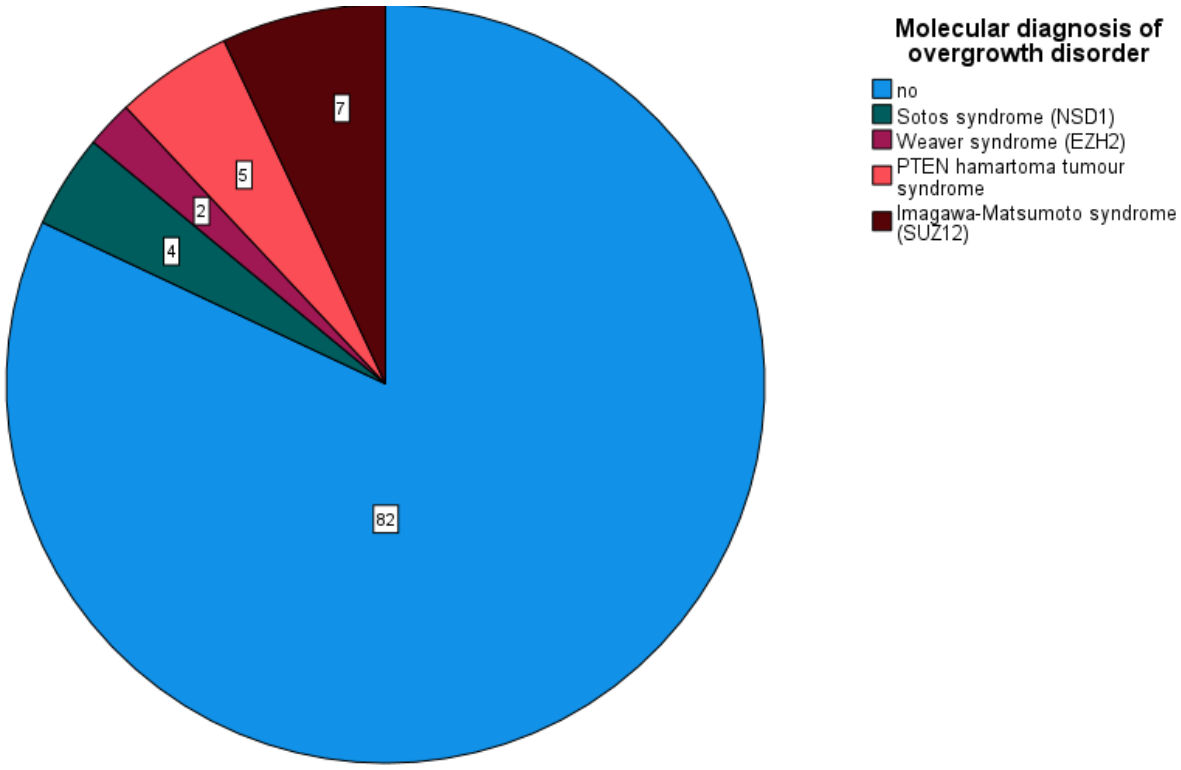


Figure 14: Individuals with a molecular diagnosis of an overgrowth disorder at entry to the study.

The molecular diagnoses in participants at entry to the study reflect the diagnostic testing available at the start of the study, with single gene testing of *NSD1*, *EZH2* and *PTEN* being performed by Sanger sequencing in regional NHS genetics laboratories. The deletion of *SUZ12* was identified on CGH microarray.

3.3.5 Type of overgrowth in undiagnosed participants

Of the 82 participants without a molecular diagnosis of an overgrowth diagnosis at entry to the study, ten had regional overgrowth and 72 did not have regional overgrowth (Figure 15).

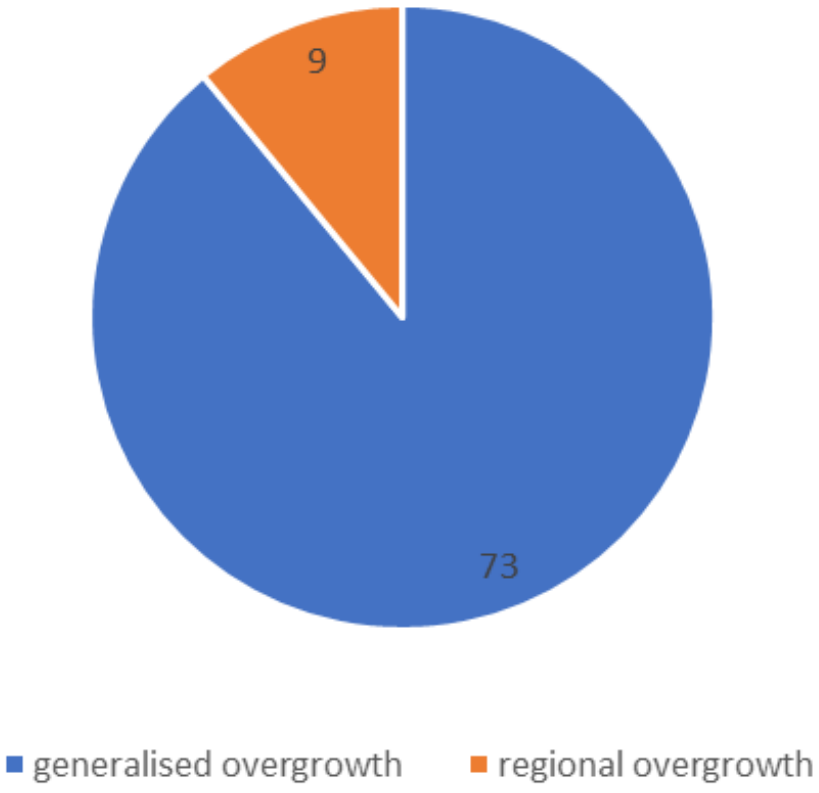


Figure 15: Type of overgrowth in participants without a molecular diagnosis of an overgrowth disorder at entry to the study

3.4 Total cohort growth parameters

3.4.1 Height

59/100 (59%) of participants had tall stature (height > 2 SD above the mean for age and sex). Height ranged from -3.4 SD to +7.0 SD with a mean of +2.2 SD and standard deviation 1.7 (Figure 16).

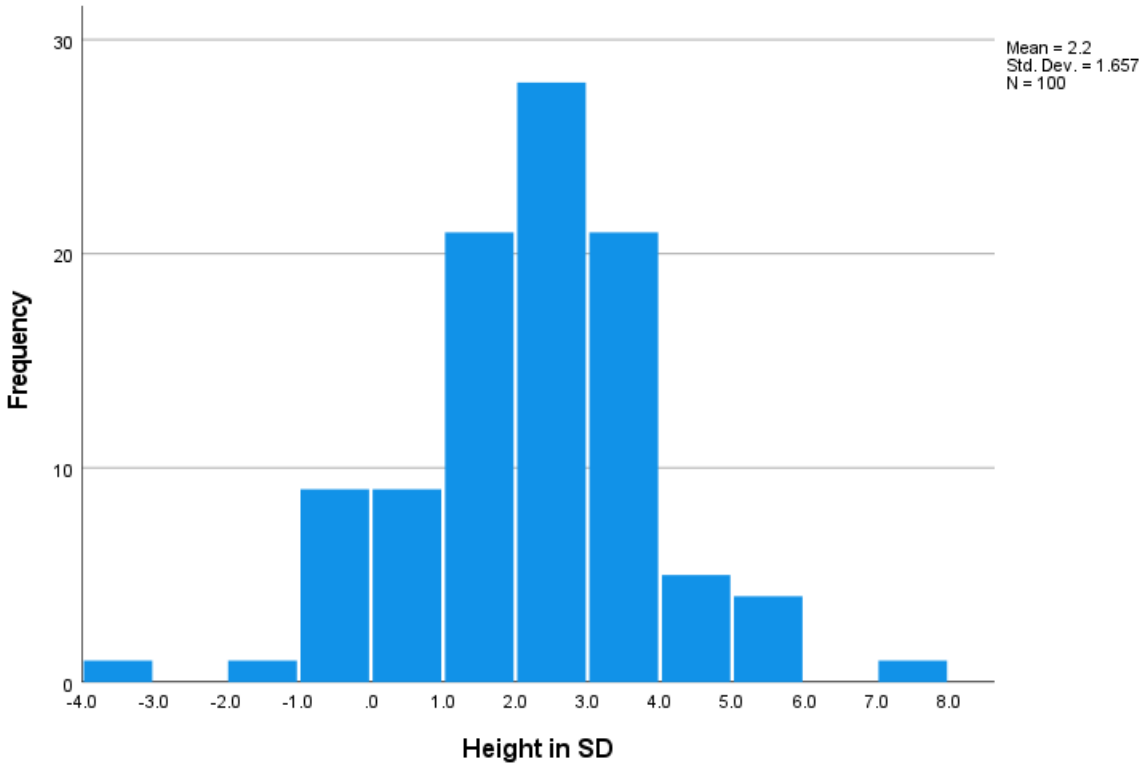


Figure 16: Total cohort height in SD compared to mean for age and sex

3.4.2 Head circumference

38/96 (40%) of participants had macrocephaly (OFC > 2SD above the mean for age and sex). Head circumference ranged from -1.8 SD to +6.5 SD with a mean of +1.6 SD and standard deviation 1.6 (Figure 17.)

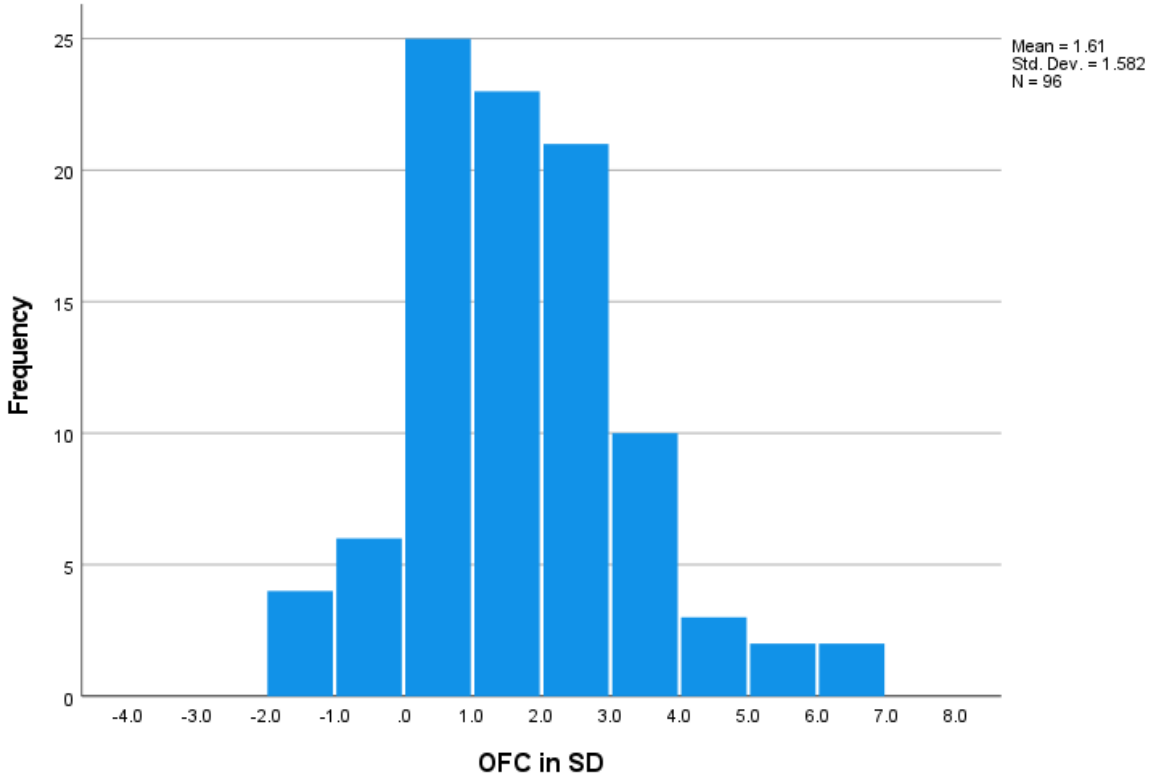


Figure 17: Total cohort - OFC in SD compared to mean for age and sex

3.4.3 Height vs head circumference

17 participants had both tall stature and macrocephaly (Figure 18).

19 participants had neither tall stature nor macrocephaly at the time of entry to the study. Eight of these had a growth parameter >2 SD at a younger age documented by a clinician (five with previous tall stature and three with previous macrocephaly); seven participants were eligible for the study because they had regional overgrowth; and four were eligible on the basis of a molecular diagnosis of an overgrowth disorder (1 participant with a variant in *EZH2*, one participant with a variant in *PTEN*, and two participants with a deletion of *SUZ12*).

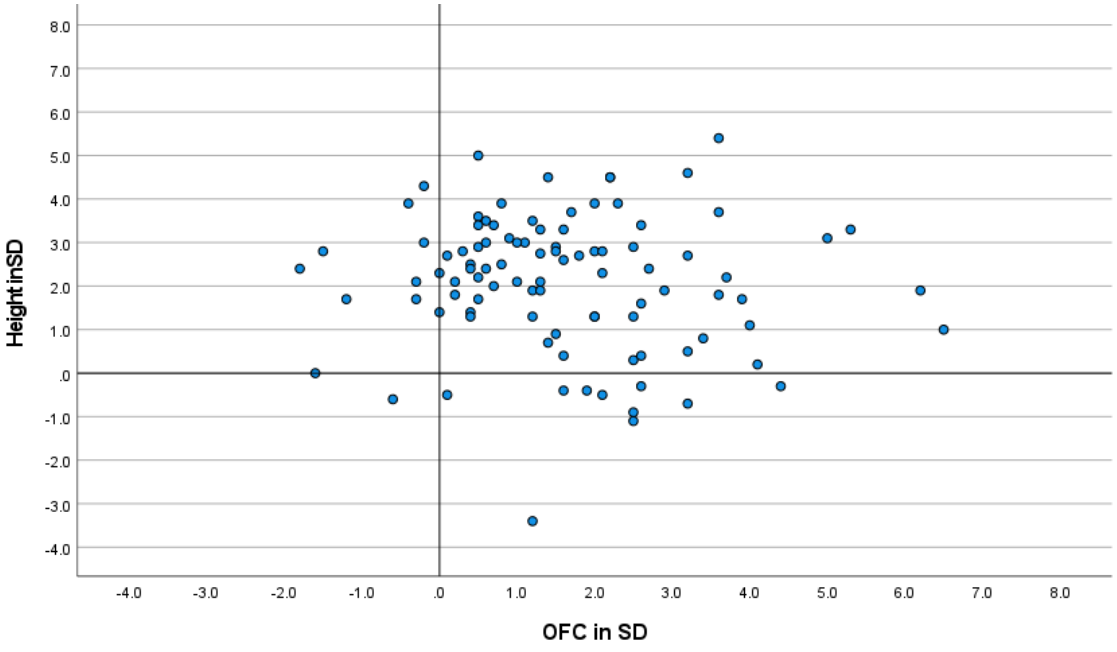


Figure 18: Total cohort - height against OFC

3.4.4 Birthweight

Of the 92 participants for whom information on birthweight was available, 19 had macrosomia (large for gestational age) with a birthweight > 2 SD above the mean for gestation and sex. The mean birthweight was + 0.9 SD (Figure 19).

Information on birth length was available for nine participants only. Four of these had a birth length > 2 SD above the mean.

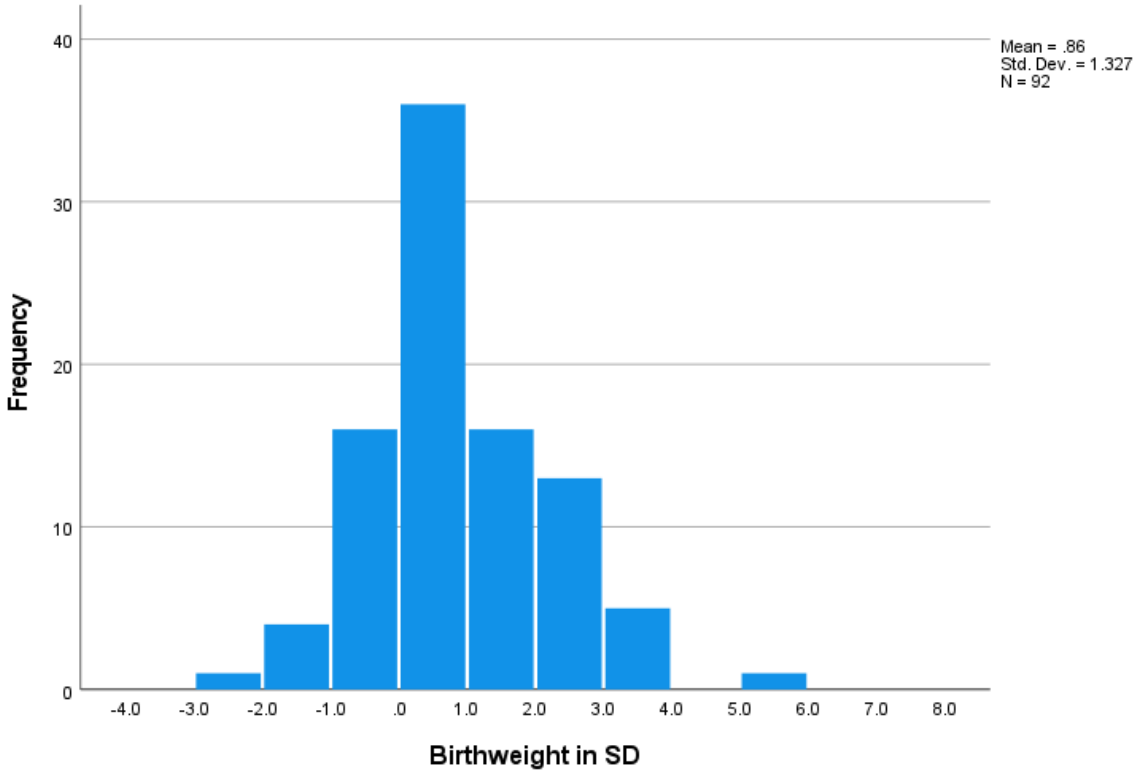


Figure 19: Total cohort - birthweight in SD

3.5 Comparison of participants with and without a molecular diagnosis

At the time of data analysis, 42 participants had a molecular diagnosis of a single gene or imprinting disorder (not including the neurosusceptibility loci 16p13.11 and 15q11.2) (see Chapter 5 for details of molecular investigations) and 58 participants did not.

In the following sections, the cohort is divided into those with and without a molecular diagnosis to investigate if there were any differences in height, head circumference, developmental delay, autism spectrum disorder, or dysmorphic features between the two groups. It would seem plausible that each of these characteristics could be associated with an increased chance of identifying a molecular diagnosis and/or a specific subgroup of diagnoses.

3.5.1 No significance difference in height

21/42 (50%) of participants with a molecular diagnosis had tall stature (> 2 SD above the mean for age and sex.), with a mean of +2.0 SD (Figure 20). 38/58 (66%) number of participants without a molecular diagnosis of a single gene disorder had tall stature (> 2 SD above the mean for age and sex), with a mean of +2.3 SD (Figure 21).

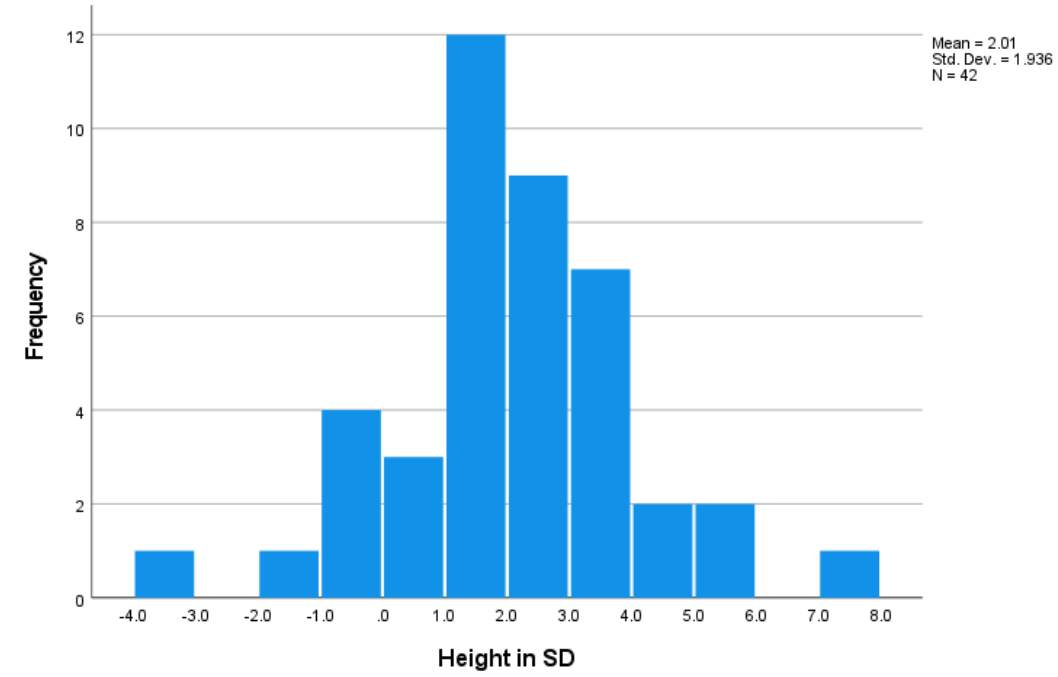


Figure 20: Participants with a single gene molecular genetic diagnosis - height in SD compared to mean for age and sex

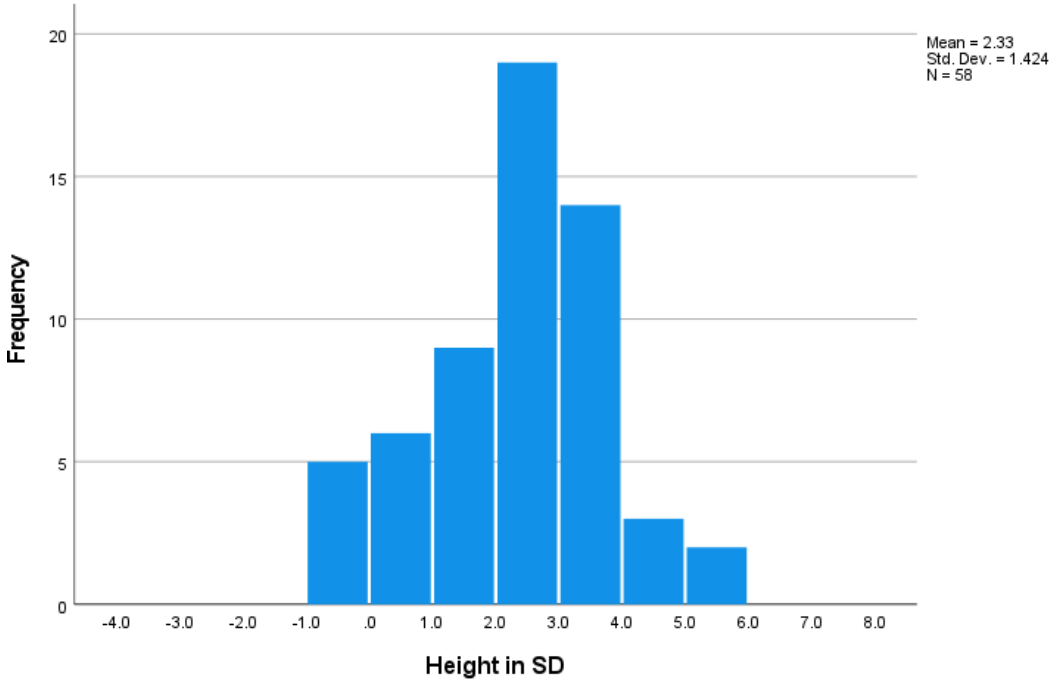


Figure 21: Participants without a single gene molecular diagnosis - height in SD compared to mean for age and sex

There was no significant difference in mean height in SD (p = .346) between the group with a molecular genetic diagnosis and without a molecular genetic diagnosis (

Table 10 and Table 11).

Table 10: Mean height in SD compared between group with molecular diagnosis and group without molecular diagnosis

Group Statistics

	SINGLE GENE CONDITION PRESENT	N	Mean	Std. Deviation	Std. Error Mean
HEIGHT_SD	no	58	2.330	1.4245	.1870
	yes	42	2.012	1.9360	.2987

Table 11: Two sample T-test comparing height in SD between the group with molecular diagnosis and group without molecular diagnosis

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance		Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p			Lower	Upper
HEIGHT_SD	Equal variances assumed	2.898	.092	.948	98	.173	.346	.3183	.3359	-.3483	.9848
	Equal variances not assumed			.903	71.540	.185	.370	.3183	.3524	-.3844	1.0209

3.5.2 No significant difference in head circumference

16/39 (41%) of participants with a molecular diagnosis of a single gene disorder or imprinting disorder had macrocephaly (> 2 SD above the mean for age and sex), with a mean of +1.5 SD (Figure 22). 22/57 (39%) number of participants without a molecular diagnosis of a single gene disorder had macrocephaly (> 2 SD above the mean for age and sex), with a mean of +1.7 SD (Figure 23).

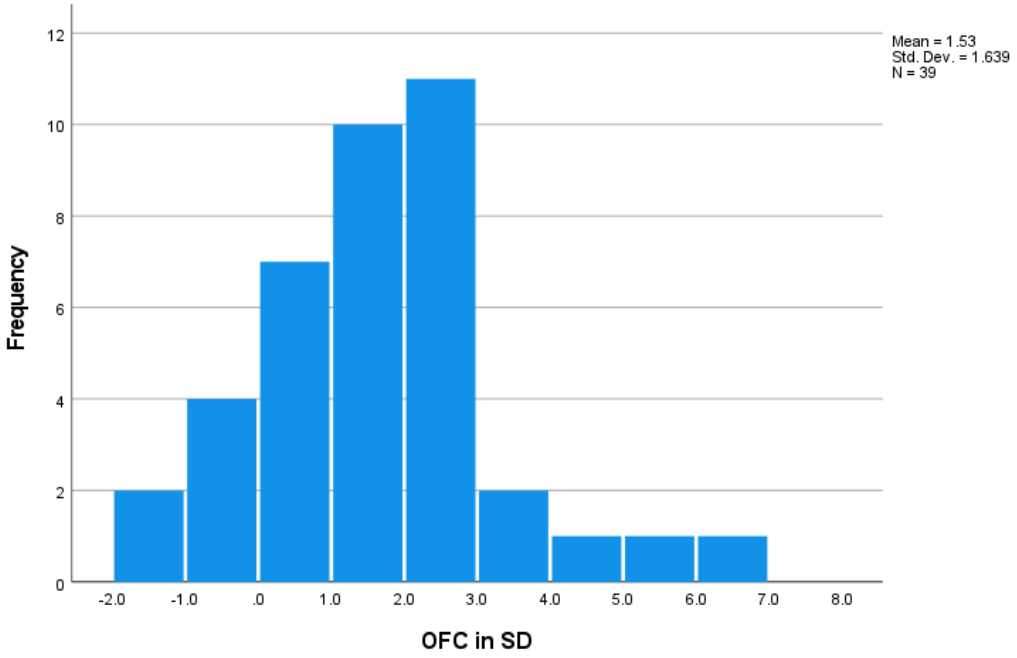


Figure 22: Participants with a molecular genetic diagnosis - OFC in SD compared to mean for age and sex

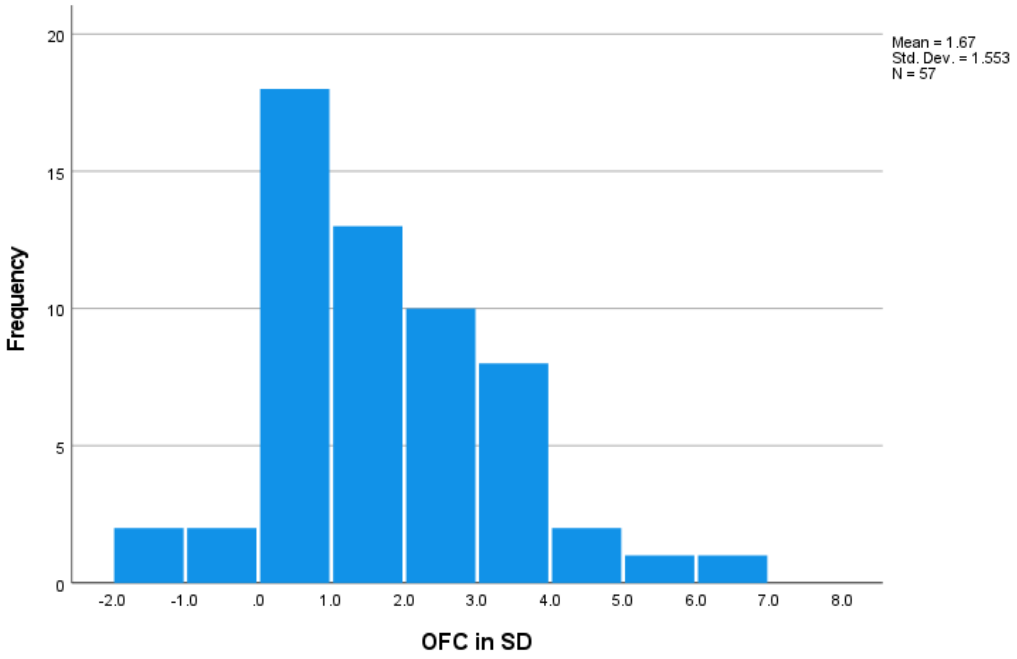


Figure 23: Participants without a diagnosis - OFC in SD compared to mean for age and sex

There was no significant difference (p = .655) in mean OFC in SD between the group with a molecular genetic diagnosis and the group without a molecular genetic diagnosis (Table 12 and Table 13).

Table 12: Mean OFC in SD compared between group with molecular diagnosis and group without molecular diagnosis

Group Statistics

	SINGLE_GENE_CONDITION_PRESENT	N	Mean	Std. Deviation	Std. Error Mean
OFC_SD	no	57	1.674	1.5531	.2057
	yes	39	1.526	1.6388	.2624

Table 13: Two sample T-test comparing OFC in SD between the group with molecular diagnosis and group without molecular diagnosis

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
OFC_SD	Equal variances assumed	.001	.975	.449	94	.327	.655	.1480	.3301	-.5073	.8034
	Equal variances not assumed			.444	78.847	.329	.658	.1480	.3334	-.5157	.8117

3.5.3 No association between developmental delay and presence of molecular diagnosis

There was no statistically significant association between developmental delay and a molecular genetic diagnosis (p=.806) (Figure 24; Tables 14 and 15).

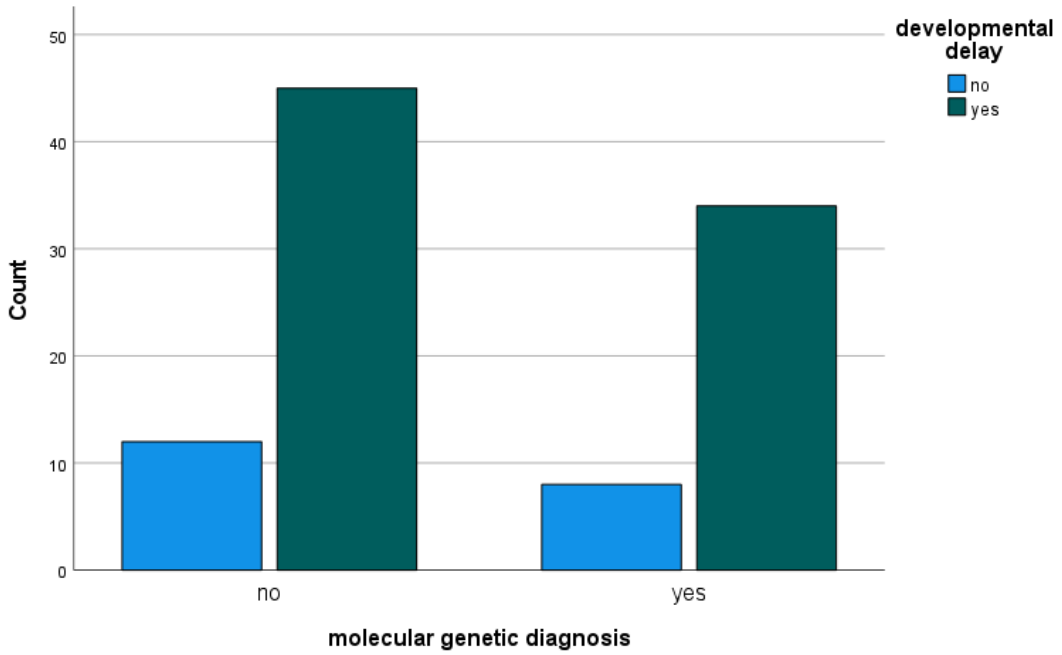


Figure 24: Presence or absence of developmental delay in the group with a molecular genetic diagnosis and without a molecular genetic diagnosis

Table 14: Number of individuals with developmental delay in the group with a molecular genetic diagnosis and group without a molecular genetic diagnosis

			DEV_DELAY		
			no	yes	Total
CONDITION_PRESENT	no	Count	12	45	57
		% within CONDITION_PRESENT	21.1%	78.9%	100.0%
		% within DEV_DELAY	60.0%	57.0%	57.6%
		% of Total	12.1%	45.5%	57.6%
	yes	Count	8	34	42
		% within CONDITION_PRESENT	19.0%	81.0%	100.0%
		% within DEV_DELAY	40.0%	43.0%	42.4%
		% of Total	8.1%	34.3%	42.4%
Total		Count	20	79	99
		% within CONDITION_PRESENT	20.2%	79.8%	100.0%
		% within DEV_DELAY	100.0%	100.0%	100.0%
		% of Total	20.2%	79.8%	100.0%

Table 15: Chi-Square tests of presence of developmental delay and molecular genetic diagnosis

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.060[a]	1	.806		
Continuity Correction[b]	.000	1	1.000		
Likelihood Ratio	.061	1	.806		
Fisher's Exact Test				1.000	.506
Linear-by-Linear Association	.060	1	.807		
N of Valid Cases	99				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 8.48.
b. Computed only for a 2x2 table

3.5.4 There is a relationship between autism spectrum disorder and absence of molecular diagnosis

Participants with a molecular diagnosis were less likely to have autism spectrum disorder, and participants without a molecular diagnosis were more likely to have autism spectrum disorder (Figure 25).

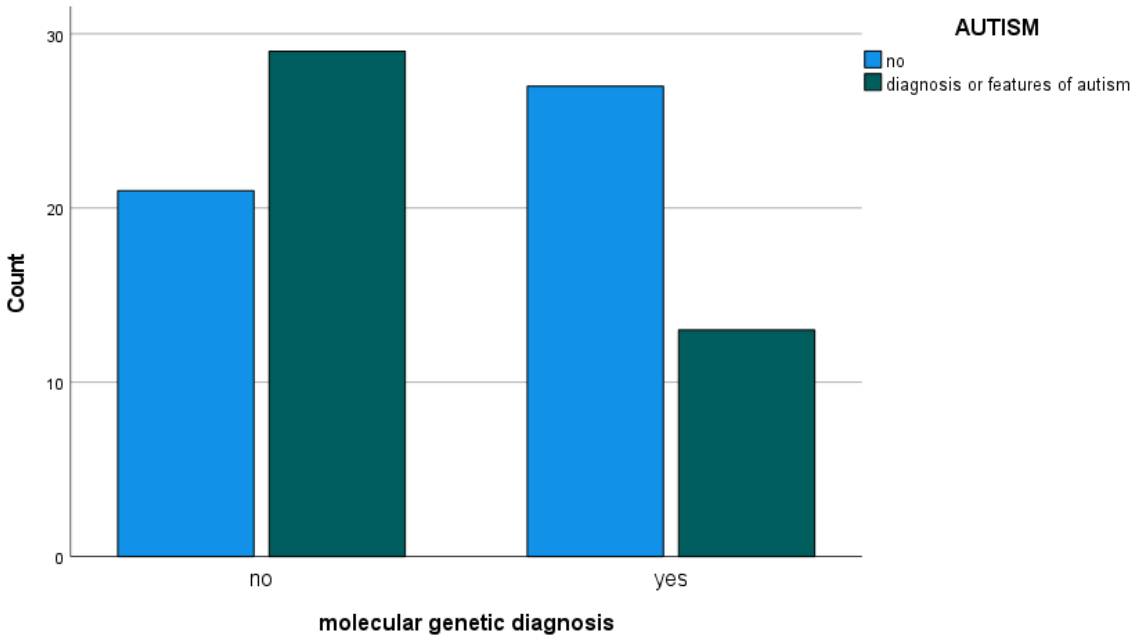


Figure 25: Presence or absence of autism in the group with a molecular genetic diagnosis and without a molecular genetic diagnosis

A significant association (p = .016) was identified between the presence of autism spectrum disorder and absence of molecular genetic diagnosis (Table 16 and Table 17).

Table 16: Number of individuals with autism in the group with a molecular genetic diagnosis and group without a molecular genetic diagnosis

			AUTISM diagnosis or features of autism		
			no		Total
CONDITION_PRESENT	no	Count	21	29	50
		% within CONDITION_PRESENT	42.0%	58.0%	100.0%
		% within AUTISM	43.8%	69.0%	55.6%
		% of Total	23.3%	32.2%	55.6%
	yes	Count	27	13	40
		% within CONDITION_PRESENT	67.5%	32.5%	100.0%
		% within AUTISM	56.3%	31.0%	44.4%
		% of Total	30.0%	14.4%	44.4%
Total		Count	48	42	90
		% within CONDITION_PRESENT	53.3%	46.7%	100.0%
		% within AUTISM	100.0%	100.0%	100.0%
		% of Total	53.3%	46.7%	100.0%

Table 17: Chi-Square tests of presence of autism and molecular genetic diagnosis

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	5.806[a]	1	.016		
Continuity Correction[b]	4.826	1	.028		
Likelihood Ratio	5.891	1	.015		
Fisher's Exact Test				.020	.014
Linear-by-Linear Association	5.741	1	.017		
N of Valid Cases	90				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 18.67.
b. Computed only for a 2x2 table

3.5.5 No association between dysmorphic features and presence of molecular diagnosis

There was no statistically significant association between the presence of dysmorphic features and a molecular genetic diagnosis (p = .452) (see Figure 26; Tables 18 and 19).

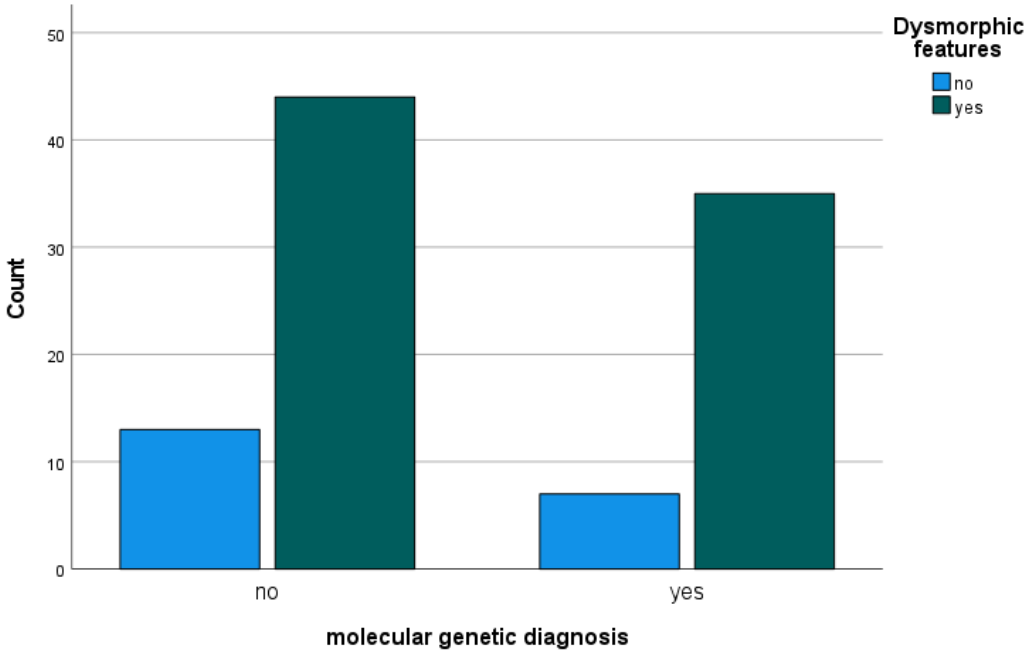


Figure 26: Presence or absence of dysmorphic features (of face, hands or feet) in the group with a molecular genetic diagnosis and without a molecular genetic diagnosis

Table 18: Number of individuals with dysmorphic features in the group with a molecular genetic diagnosis and group without a molecular genetic diagnosis

			DYMORPHOLOGY no	DYMORPHOLOGY yes	Total
CONDITION_PRESENT	no	Count	13	44	57
		% within CONDITION_PRESENT	22.8%	77.2%	100.0%
		% within DYMORPHOLOGY	65.0%	55.7%	57.6%
		% of Total	13.1%	44.4%	57.6%
	yes	Count	7	35	42
		% within CONDITION_PRESENT	16.7%	83.3%	100.0%
		% within DYMORPHOLOGY	35.0%	44.3%	42.4%
		% of Total	7.1%	35.4%	42.4%
Total		Count	20	79	99
		% within CONDITION_PRESENT	20.2%	79.8%	100.0%
		% within DYMORPHOLOGY	100.0%	100.0%	100.0%
		% of Total	20.2%	79.8%	100.0%

Table 19: Chi-Square tests of presence of dysmorphic features and molecular genetic diagnosis

Chi-Square Tests

	Value	df	Asymptotic Significance (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	.566[a]	1	.452		
Continuity Correction[b]	.249	1	.618		
Likelihood Ratio	.574	1	.449		
Fisher's Exact Test				.613	.311
Linear-by-Linear Association	.560	1	.454		
N of Valid Cases	99				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 8.48.
b. Computed only for a 2x2 table

3.6 Comparison of participants with and without developmental delay

In the following section, participants are divided into those with and without developmental delay. It could be hypothesised that variants in genes affecting neurodevelopment could cause both developmental delay and abnormal brain development in the form of macrocephaly. If a relationship between macrocephaly and developmental delay were identified, further analysis of this subgroup of participants could identify a different molecular aetiology to the larger cohort. Data analysis was performed to investigate potential associations between developmental delay and macrocephaly in comparison to developmental delay and tall stature.

3.6.1 No significant difference in height

There was no significant difference (p=.238) in mean height in SD between the group with developmental delay and the group without developmental delay (Figures 27 and 28; Tables 20 and 21).

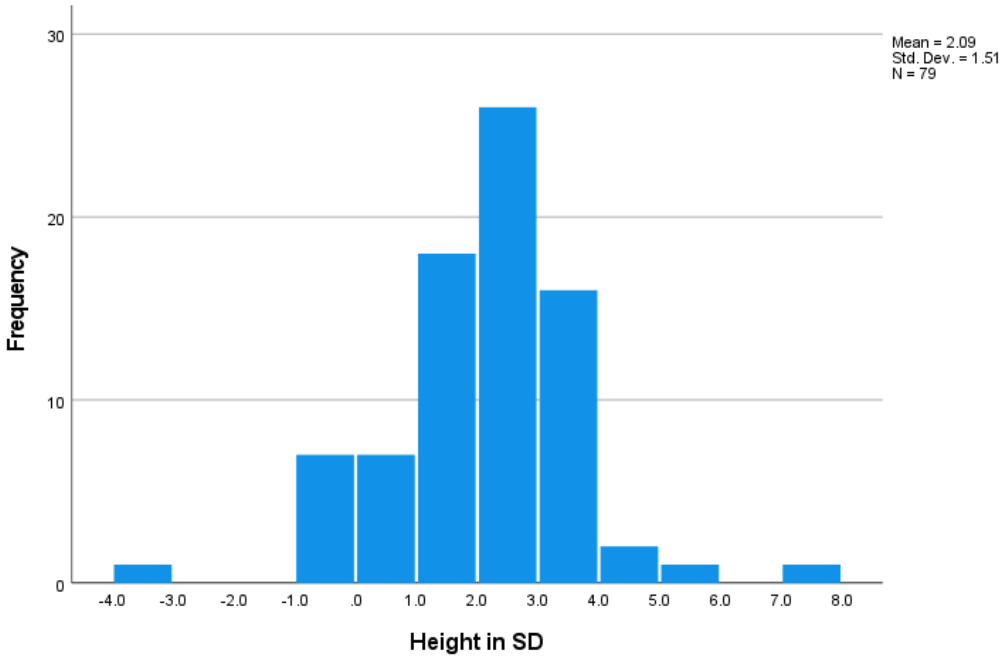


Figure 27: Participants with developmental delay – height in SD

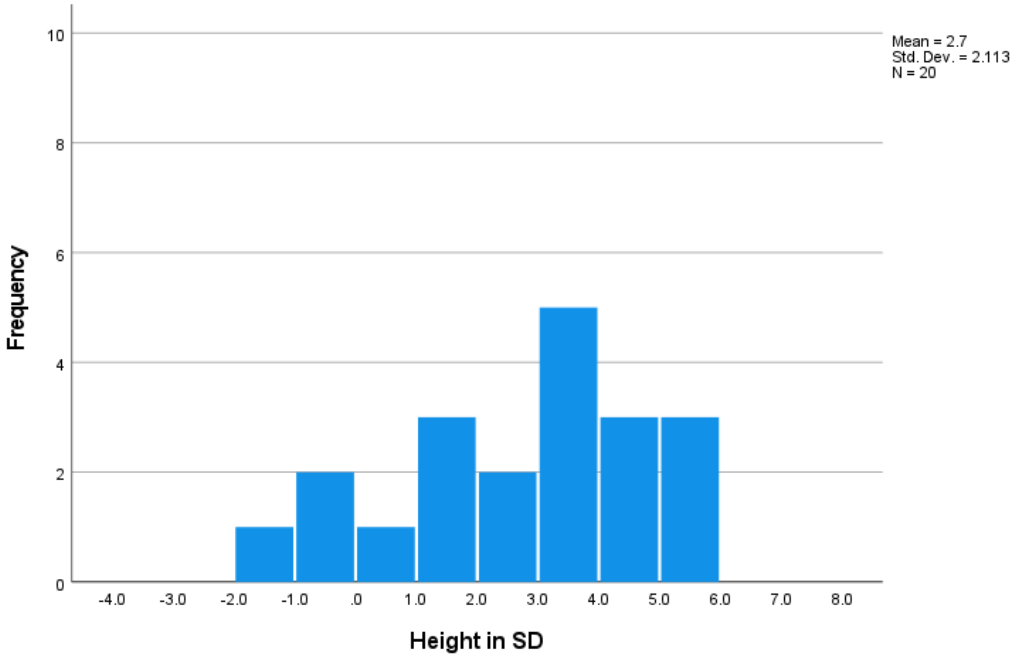


Figure 28: Participants without developmental delay – height in SD

Table 20: Mean height in SD compared between group with developmental delay and group without developmental delay

Group Statistics

	DEV_DELAY	N	Mean	Std. Deviation	Std. Error Mean
HEIGHT_SD	no	20	2.700	2.1134	.4726
	yes	79	2.092	1.5097	.1698

Table 21: Two sample T-test comparing height in SD between the group with developmental delay and group without developmental delay

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
						Significance				95% Confidence Interval of the Difference
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
HEIGHT_SD	Equal variances assumed	7.608	.007	1.477	97	.071	.143	.6082	.4119	-.2092	1.4257
	Equal variances not assumed			1.211	24.128	.119	.238	.6082	.5022	-.4279	1.6443

3.6.2 No significant difference in head circumference

There was no significance difference (p=.204) in the mean OFC in SD between the group with developmental delay and the group without developmental delay (Figures 29 and 30; Tables 22 and 23).

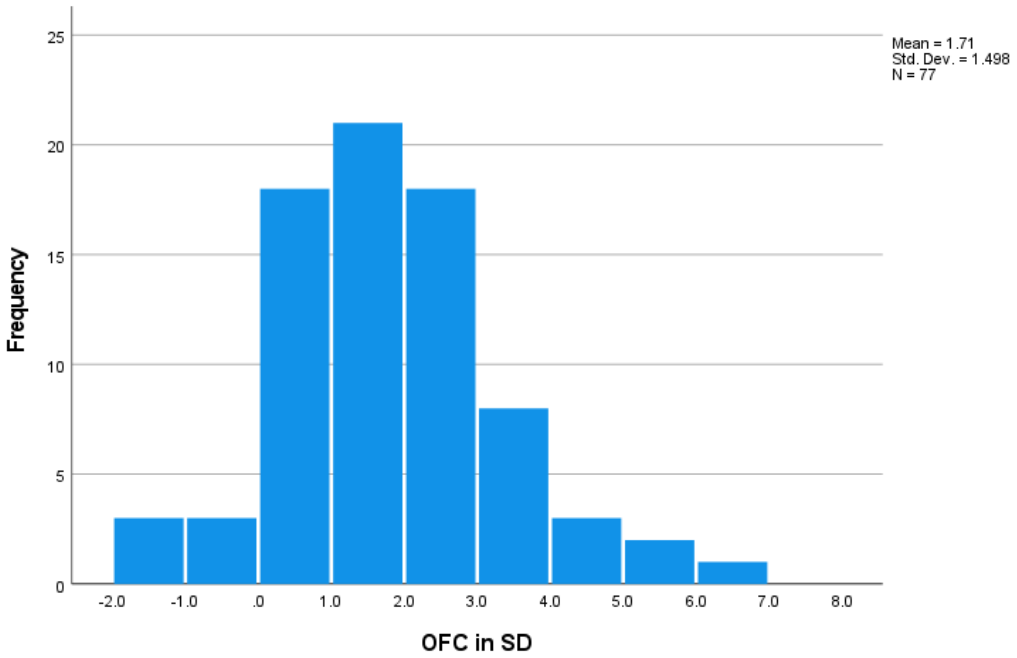


Figure 29: Participants with developmental delay – OFC in SD

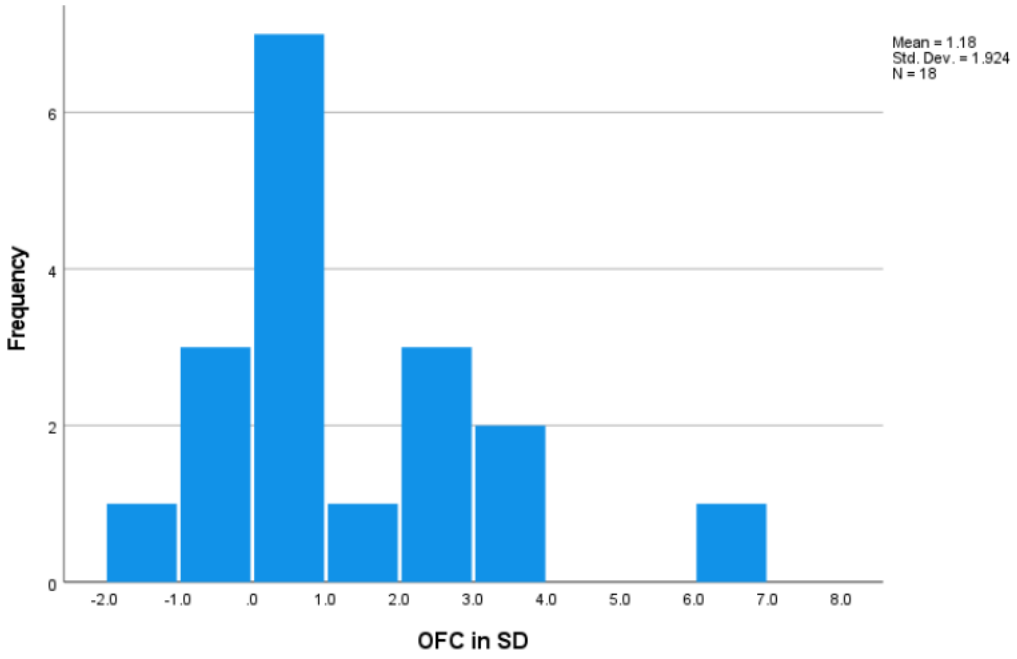


Figure 30: Participants without developmental delay – OFC in SD

Table 22: Mean OFC in SD compared between group with developmental delay and group without developmental delay

Group Statistics

	DEV_DELAY	N	Mean	Std. Deviation	Std. Error Mean
OFC_SD	no	18	1.183	1.9242	.4535
	yes	77	1.714	1.4981	.1707

Table 23: Two sample T-test comparing OFC in SD between the group with developmental delay and group without developmental delay

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
						Significance				95% Confidence Interval of the Difference
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
OFC_SD	Equal variances assumed	1.641	.203	-1.280	93	.102	.204	-.5310	.4149	-1.3548	.2929
	Equal variances not assumed			-1.096	22.060	.143	.285	-.5310	.4846	-1.5358	.4739

3.7 Comparison of participants with and without autism

In the following section, participants are divided into those with and without autism. It could be hypothesised that variants in genes affecting neurodevelopment could cause both autism and abnormal brain development in the form of macrocephaly. If a relationship between macrocephaly and autism were identified, further analysis of this subgroup of participants could identify a different molecular aetiology to the larger cohort. Data analysis was performed to investigate potential associations between autism and macrocephaly in comparison to autism and tall stature.

3.7.1 No significant difference in height

There was no significant difference (p=.220) in mean height between the group with features of autism and the group without autism (Figures 31 and 32; Tables 24 and 25).

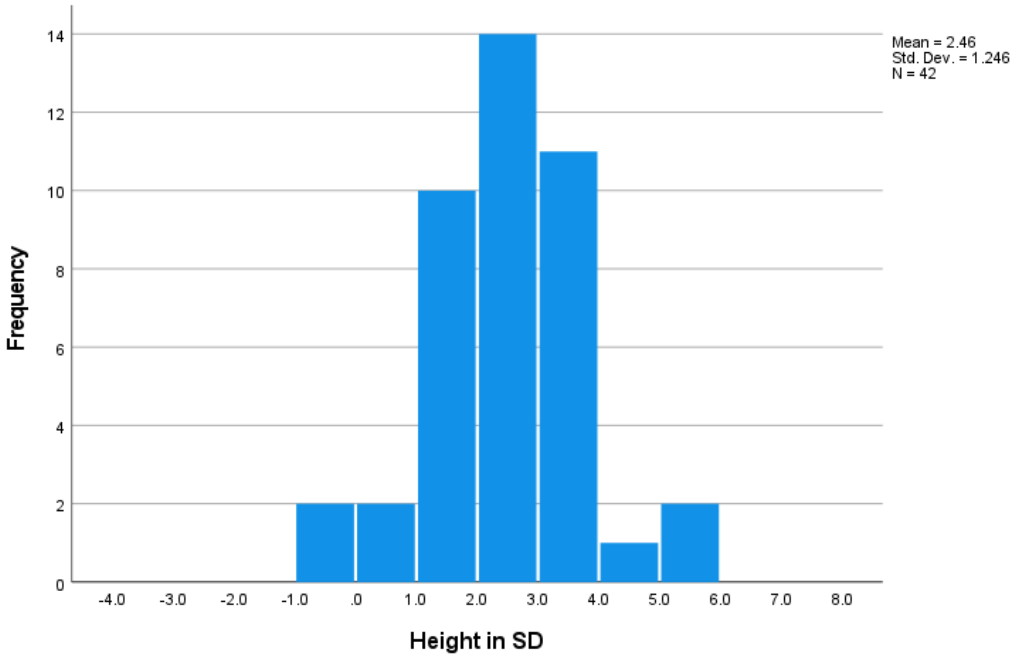


Figure 31: Participants with autism – height in SD

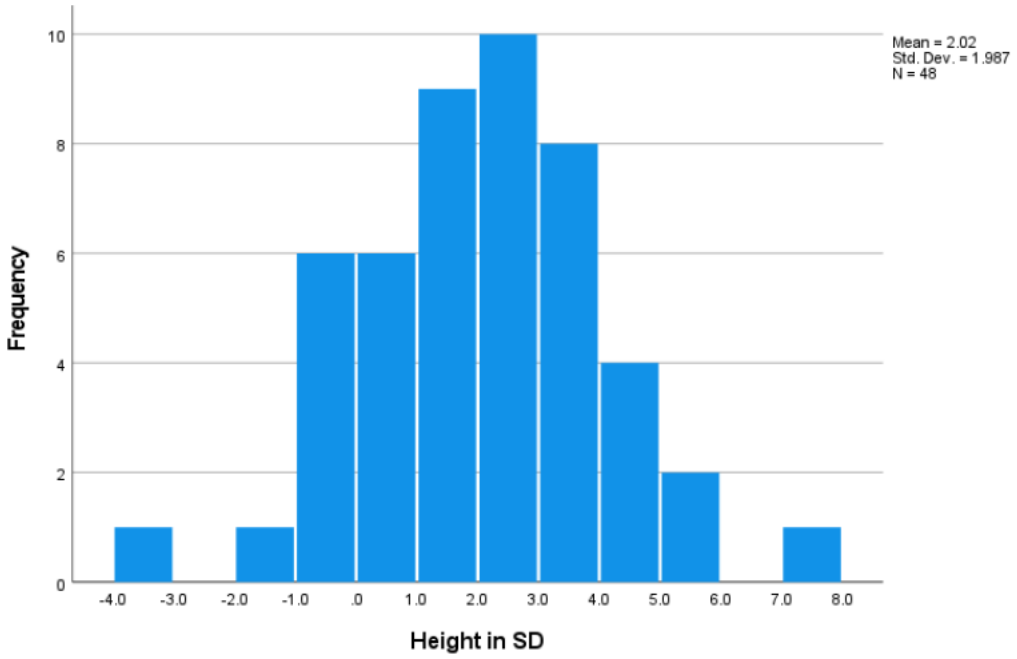


Figure 32: Participants without autism – height in SD

Table 24: Mean height in SD compared between group with autism and group without autism

Group Statistics

	AUTISM	N	Mean	Std. Deviation	Std. Error Mean
HEIGHT SD	No	48	2.018	1.9872	.2868
	features or diagnosis	42	2.457	1.2461	.1923

Table 25: Two sample T-test comparing height in SD between the group with features of autism and group without autism

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
						Significance				95% Confidence Interval of the Difference
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
HEIGHT_SD	Equal variances assumed	7.034	.009	-1.236	88	.110	.220	-.4394	.3556	-1.1461	.2672
	Equal variances not assumed			-1.273	80.174	.103	.207	-.4394	.3453	-1.1266	.2477

3.7.2 No significant difference in head circumference

There was no significant difference (p=.768) in mean OFC between the group with features of autism and the group without autism (Figures 33 and 34; Tables 26 and 27).

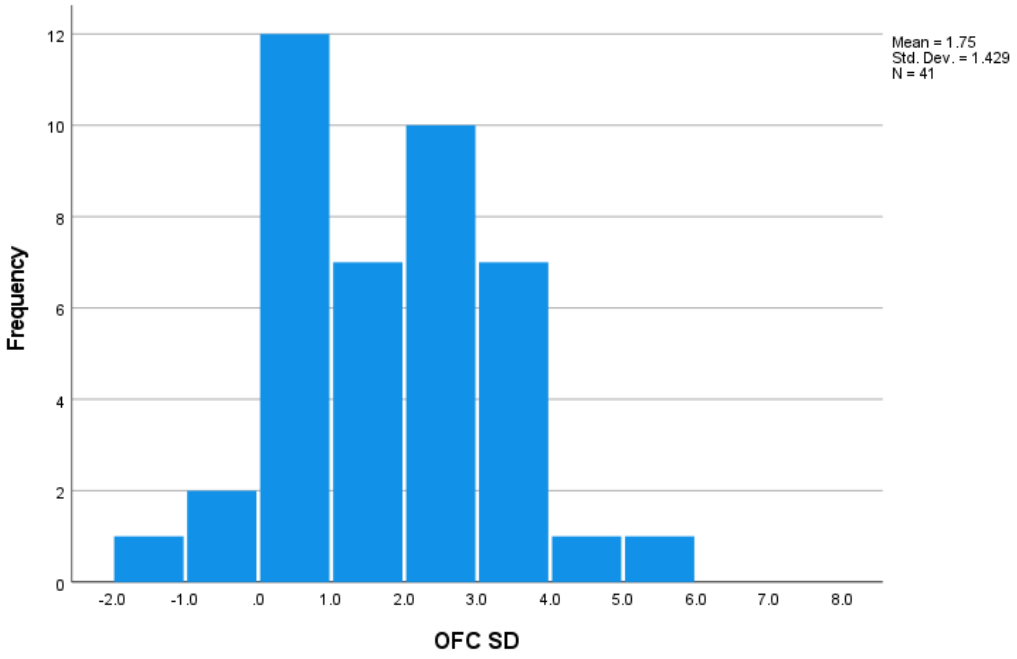


Figure 33: Participants with autism – OFC in SD

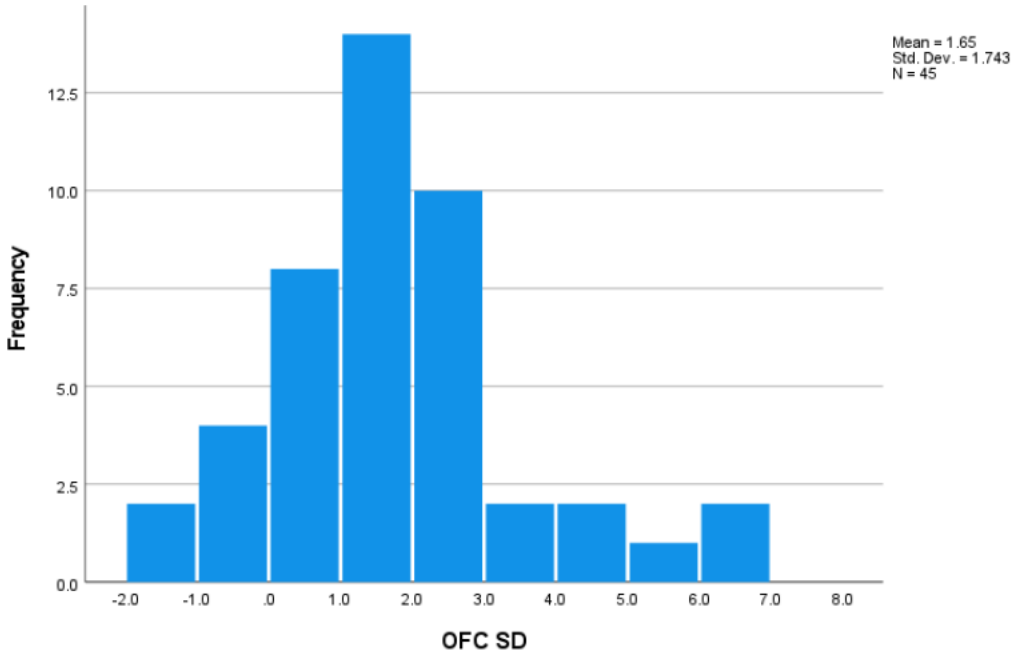


Figure 34: Participants without autism – OFC in SD

Table 26: Mean OFC in SD compared between group with autism and group without autism

Group Statistics

	AUTISM	N	Mean	Std. Deviation	Std. Error Mean
OFC_SD	No	45	1.649	1.7427	.2598
	features or diagnosis	41	1.751	1.4288	.2231

Table 27: Two sample T-test comparing OFC in SD between the group with features of autism and group without autism

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means							
						Significance				95% Confidence Interval of the Difference	
		F	Sig.	t	df	One-Sided p	Two-Sided p	Mean Difference	Std. Error Difference	Lower	Upper
OFC_SD	Equal variances assumed	.181	.671	-.296	84	.384	.768	-.1023	.3456	-.7897	.5850
	Equal variances not assumed			-.299	83.112	.383	.766	-.1023	.3425	-.7835	.5788

3.8 Summary of study cohort

In summary, the cohort was largely paediatric and there was a slight excess of males. 18 participants had a molecular diagnosis at entry to the study, with only copy number variants and three single gene disorders (Sotos syndrome, Weaver syndrome, and PTEN hamartoma tumour syndrome) identified. This reflects the diagnostic testing available at commencement of the study. Most of the 82% of participants without a diagnosis had generalised overgrowth and a minority had regional overgrowth.

59% of the cohort had tall stature, 40% had macrocephaly, 17% had both tall stature and macrocephaly, and 21% had high birth weight.

By the end of the study, 42 participants had achieved a molecular diagnosis. Analysis of the study data did not identify any significant associations between presence or absence of genetic diagnosis and height, head circumference, developmental delay, or dysmorphic features. No significant associations were identified between developmental delay and height or head circumference, or between macrocephaly and height or head circumference.

A significant association ($p=0.016$) was identified between the presence of autism spectrum disorder and the absence of a molecular genetic diagnosis. The reason for this is unclear but might be explained by a polygenic rather than monogenic aetiology for autism spectrum disorder in these participants.

Chapter 4. RESULTS: Expanding the phenotypes of overgrowth disorders

4.1 Overgrowth disorders

Phenotypic descriptions of the subset of participants with a confirmed molecular diagnosis of an overgrowth disorder, either at entry to the study or identified through molecular investigations as part of the study, are given in the following chapter. Confirmed molecular diagnoses included BWS (one individual), *CHD8* overgrowth syndrome (one individual), TBRS (two individuals), Weaver syndrome (two individuals), Malan syndrome (one individual), Sotos syndrome (five individuals), Kosaki overgrowth syndrome (one individual), PROS (four individuals), *PPP2R5D*-related neurodevelopmental disorder (one individual), PHTS (three individuals from one family and two individuals from a second family), and Imagawa-Matsumoto syndrome (seven individuals from one family). Descriptions are also given of two other individuals with Kosaki overgrowth syndrome who were not identified through the POD study.

A summary of the relevant clinical history is given for each participant. Growth parameters at birth and at recruitment to the study are given for each participant unless unknown. Parental height and weight are also listed if known. Unless otherwise specified, the following clinical history is true for each participant: the pregnancy was a naturally conceived, singleton pregnancy with no medical problems or medication in pregnancy, the antenatal scans were unremarkable and the birth was by a spontaneous cephalic vaginal delivery; in the neonatal period, neonatal resuscitation and admission to NNU were not required, no abnormalities on newborn examination were identified; there was no significant medical or developmental

history, parents were non-consanguineous, there was no family history of an overgrowth disorder and there were no dysmorphic features. Any deviation from this description is noted in the phenotypic information.

Each clinical description is followed by a discussion of whether the participant's phenotype information is consistent with our current knowledge of the overgrowth disorder and how it expands the known phenotype.

4.1.1 Beckwith-Wiedemann syndrome

4.1.1.1 Results POD 061.0

POD 061.0: mosaic hypermethylation at H19/IGF2:IG-DMR within 11p15.5 absent in blood and identified on repeat buccal sampling.

POD 061.0 was born at 40 weeks and three days. He was macrosomic with a birthweight of 5.04 kg (+2.7 SD), required resuscitation at birth, and was treated on the neonatal unit for 14 days for neonatal hypoxic-ischaemic encephalopathy (HIE). His OFC was 35.9 cm (+0.4 SD). He had seizures, hypotonia, and hypoglycaemia. A naevus flammeus was noted on newborn examination. An MRI head scan was normal.

At the age of one his weight was 11.5 kg (+0.7 SD), OFC 49.0 (+0.4 SD) and length 82 cm (+1.4 SD). He had right sided regional overgrowth affecting the face, upper limb, and lower limb. He was also noted to have macroglossia. Renal ultrasound scan (USS) screening was performed at three monthly intervals.

His mother's birthweight was 3.6 kg (+0.4 SD) with an adult height of 172.7 cm (+1.5 SD) and his father's birthweight was 4.8 kg (+2.6 SD) with an adult height of 182 cm (+0.7 SD). His four year old sibling had a birthweight of 3.35 kg (-0.5 SD).

4.1.1.2 Discussion POD 061.0

The regional overgrowth and macroglossia present in this participant are cardinal features of BWS. In combination with the suggestive features of a birth weight >2 SD and facial naevus, a clinical diagnosis of BWS was made in this participant in line with the scoring system presented by Brioude et al. 2018[202]. A clinical diagnosis of classical BWS is consistent with a score of at least 4 points. In this participant, two cardinal features (2 points each) and two suggestive features (1 point each) give a score of 6 points.

Other cardinal phenotypic features of BWS are exomphalos, multifocal and/or bilateral Wilms tumour or nephroblastomatosis, hyperinsulinism lasting more than 7 days requiring escalated treatment, and findings on pathological examination (adrenal cortex cytomegaly, placental mesenchymal dysplasia or pancreatic adenomatosis). Other suggestive features of BWS are polyhydramnios and/or placentomegaly, ear creases and/or pits, transient hypoglycaemia lasting less than seven days, typical tumours associated with BWS, and umbilical hernia or

diastasis recti[202]. This participant had hypoglycaemia in the neonatal period which could be related to BWS but might also be caused by HIE. The unexpectedly difficult delivery and HIE could have been related to the BWS-associated high birth weight that was not suspected antenatally. The other cardinal and suggestive features were either not present in this individual, or in the case of pathology findings, information was not available.

Specific clinical features of BWS are more common in some molecular subgroups. Hypermethylation at H19/IGF2:IG-DMR is associated with a low frequency of exomphalos and a higher risk of Wilms tumour[43]. This participant did not have an exomphalos or other abdominal defect, in keeping with the molecular findings. Screening for Wilms tumour will continue until he is seven. This participant had no additional phenotypic features outside the cardinal and suggestive features of BWS previously identified in the literature.

4.1.2 *CHD8* overgrowth syndrome

4.1.2.1 Results POD 008.0

POD 009.0: *CHD8* de novo c.716delA; p.(Lys239ArgfsTer22)

POD 009.0 was born at 40 weeks and three days gestation weighing 4.0 kg (+0.9 SD) with a birth length of 58 cm (+3.5 SD). Neonatal problems included feeding difficulties, hypotonia and anal stenosis requiring treatment with dilatation. Age four his height was +3.0 SD, weight was 18.2 kg (+0.7 SD), and OFC 52 cm (-0.2 SD). He had a mildly advanced bone age of 3.5 years at three years old. He also had a borderline raised testosterone level of 10.9. He had global developmental delay and sat age 12 months, walked age 18 months, and spoke his first word age 24 months. Developmental and behavioural issues included features of autism. He

had facial flushing and long slender fingers and toes with slightly flattened deep set nails. His mother was 165.7cm (+0.3 SD) tall with an OFC of 54.6 cm (-0.7 SD), and his father was 171cm (-0.9 SD) tall with an OFC of 58.8 cm (1.2 SD).

4.1.2.2 Discussion POD 008.0

Data on birth length has not previously been reported for *CHD8* overgrowth syndrome. This participant had an extremely long length at birth (+3.5 SD) and this finding may represent a key recognisable feature of *CHD8* overgrowth syndrome. Further study is needed to establish the mean and range birth weights in this condition. The normal birthweight (+0.9 SD) is consistent with the previously reported findings, with only a small proportion of individuals with *CHD8* overgrowth syndrome being large for gestational age[203]. Neonatal hypotonia is a recognised feature of this condition, being described in approximately one-third of individuals in two separate studies[98,203].

This participant's tall stature of +3.0 SD is in keeping with heights described in *CHD8* overgrowth syndrome, with a recent review of 27 patients finding a mean height of +2.8 SD, a range of +0.2 SD to +6.3 SD, and over three-quarters of individuals having a height >2.0 SD[203]. Douzgou et al. found only 47% of 25 individuals had tall stature but this included four individuals with height >3 SD. The normal head circumference in POD 009.0 is uncommon, with macrocephaly present in between 62.5% and 96% of three reported cohorts[96,98,203], 53/66 individuals (80%) in total. The finding of an OFC of -0.2 SD in this participant demonstrates that even a below average head circumference can be present in individuals with *CHD8* overgrowth syndrome.

The congenital anomaly of anal stenosis has not previously been reported in association with *CHD8* overgrowth syndrome however other gastrointestinal issues have been described. 40% of one cohort were reported to have gastrointestinal problems, mostly commonly constipation and alternating constipation and diarrhoea. The majority (80%) of individuals reported by Bernier et al.[96] had significant gastrointestinal issues, particularly chronic constipation.

Studies using a zebrafish model system found that mutants with disruption of *CHD8* had an average of 50% fewer enteric neurons in the hindgut[96]. This would explain a phenotype of reduced gut motility. It is not clear whether this mechanism might also account for the presence of anal stenosis. Hirschsprung's disease (characterised by absence of enteric ganglions) and anorectal malformation are not commonly associated but there is an increased risk of anorectal malformation in individuals with Hirschsprung's compared to the general population[204]. The anal stenosis in POD 009.0 could be part of a spectrum of gut-related complications of *CHD* overgrowth syndrome.

The advanced bone age present in POD 009.0 has not been documented in previous literature. Advanced bone age is a feature of several overgrowth syndromes and it is possible that it is also common in *CHD8* overgrowth syndrome. However as next generation sequencing becomes a first line investigation, using bone age as a diagnostic investigation has become less relevant. It is therefore unlikely that data on bone age will become available for many individuals with *CHD8* overgrowth syndrome on a clinical basis, and further knowledge of

bone age in *CHD8* would only be possible through inclusion of this investigation in a future research study.

CHD8 was initially identified as an autism susceptibility gene rather than an overgrowth gene. The presence of autistic features in POD 008.0 is consistent with this association between *CHD8* and autism. Depending on the cohort, estimates of ASD rates in individuals with *CHD8* overgrowth syndrome vary between 56% of individuals with features of ASD[203] up to over 85% with diagnosed ASD[96,98]. Developmental delay and/or intellectual disability are extremely common, with cohorts reporting between 80 and 100% of individuals affected[98,203].

POD 008.0 is not obviously facially dysmorphic. *CHD8* overgrowth syndrome is described to have a subtle facial appearance with features including a long face, high hairline, frontal bossing, pronounced supra-orbital ridges, arched and sparse eyebrows, wide spaced eyes, down-slanted palpebral fissures, broad nose with full nasal tip, thick ear helix with a notch in the upper third, full cheeks and a prominent/pointed chin[96,98,203]. On review of photographs following the molecular diagnosis, a slightly long face with pointed chin could be appreciated. The facial flushing in POD 009.0 has not previously been commented on but is known to be a feature of Sotos syndrome, an overgrowth disorder with overlapping clinical features. The long fingers and toes with deep set nails in this individual are described in *CHD8* overgrowth syndrome[98].

In summary, POD 008.0 shares many of the known phenotypic features of *CHD8* overgrowth syndrome including tall stature, hypotonia, developmental delay and autistic features. Anal

stenosis has not previously been reported and may be a rare congenital anomaly in this disorder. Extremely long birth weight, advanced bone age, and facial flushing may also be features of *CHD8* overgrowth syndrome.

4.1.3 *DNMT3A*: Tatton-Brown-Rahman syndrome (TBRS)

4.1.3.1 Results POD 068.0

POD 068.0: *DNMT3A* c.499C>T; p.(Gln167*). Parental samples not available to confirm de novo status.

POD 068.0 was born at 37 weeks gestation by instrumental delivery weighing 2.6 kg (-0.5 SD) with a birth length of 51 cm (+1.8 SD). The pregnancy was complicated by pre-eclampsia and her mother was treated with anti-hypertensives. There were feeding difficulties in the neonatal period. Age eight she was 155.5 cm tall (+5.0 SD) with a weight of 60.6 kg (+3.7 SD) and an OFC 53.7 cm (+0.5 SD). Her bone age was advanced. Medical problems included febrile convulsions, vacant episodes, recurrent urinary tract infections treated with prophylactic antibiotics, scoliosis, seborrheic dermatitis, myopia, recurrent upper respiratory tract infections, sleep apnoea, tonsillectomy, and slow wound healing. She was also noted to have an abnormality of temperature regulation with absent sweating. She had facial erythema, full cheeks, and tapering fingers bilaterally. She had global developmental delay and walked at 18 months. Behavioural difficulties included aggression, emotional lability, temper tantrums and features of autism. She attended a special school. Her mother was 160.2 cm (-0.6 SD) tall with an OFC of 55.3 cm (-0.2 SD) and her father was 174.5 cm (-0.4 SD) tall with an OFC of 58.5 cm (+0.9 SD).

4.1.3.2 Discussion POD 068.0

Pregnancy complications are not often reported in the literature on TBRS and it is not clear whether the pre-eclampsia in this case is associated with the condition. However, a series of six patients included two individuals with an antenatal history of pre-eclampsia[205].

Pre-eclampsia is described in about 15% of pregnancies in Sotos syndrome, an overgrowth syndrome with shared clinical features and a similar mechanism of haploinsufficiency of a methyltransferase. It is therefore feasible that TBRS may have a similar association. This participant's below average birth weight (-0.5 SD) is lower than the mean birth weight of 1.3 SD reported in the literature, but is within the described range of -1.2 SD to 4.0 SD[89]. There is less information available for birth length in TBRS but this participant's length of +1.8 SD is similar to the reported mean of +1.6 SD[89]. Feeding difficulties in the neonatal period are again not reported in the literature, but not unexpected given the high frequency of hypotonia in this condition, reported in more than half of individuals[89].

This participant has extremely tall stature at +5 SD. Tall stature of >2 SD is common in TBRS, present in at least 80% of individuals, but of the greater than 50 reported individuals only one other individual in the literature has a height of +5.1 SD[89]. That individual was age ten and reported to have precocious puberty, which is likely to have accelerated their growth even further than their peers, who are unlikely to have started their pubertal growth spurt. Obesity (+2 SD) is also common in TBRS (approximately two thirds of individuals) although in this participant the height SD is more extreme compared to weight (+5.0 vs +3.7 SD). An OFC of 0.5 SD is within the reported range of -0.8 to +4.0 SD[89].

Scoliosis is a well-recognised feature of TBRS, present in about one third of reported individuals[89]. Afebrile seizures are also a known association and the absent episodes in this participant may represent seizure episodes. Epilepsy is often preceded by a history of febrile convulsions as reported in this individual. Recurrent upper and lower respiratory tract infections and myopia have been described in TBRS[206,205]. Recurrent urinary tract infection in two individuals. Sleep apnoea has been reported in three individuals possibly as part of a wider pattern of autonomic dysfunction including postural orthostatic hypotension and episodic vasomotor instability[205]. Three other individuals have been reported to have malar flushing[207]. The abnormality of temperature regulation, absent sweating, and facial flushing in this participant would be consistent with dysautonomia. Slow wound healing, seborrheic dermatitis and tapering fingers have not previously been reported in TBRS.

Intellectual disability is universal in TBRS and developmental and behavioural difficulties including autism are common[89]. Aggressive behaviour and temper tantrums have been reported in a small number of individuals[89,207].

In summary, this participant shares many of the common features of TBRS, including tall stature, increased weight, scoliosis, seizures, intellectual disability and autism. She also has some of the less common phenotypes of recurrent upper and lower respiratory infections and myopia. Recurrent urinary tract infections may also be a feature of TBRS. A height of +5 SD demonstrates that extreme tall stature can occur in TBRS. Her sleep apnoea, abnormal temperature regulation, sweating, and facial flushing add to the evidence that autonomic

dysfunction is part of the phenotype of TBRS. Finally, slow wound healing, seborrheic dermatitis and tapering fingers are possible novel features of TBRS.

4.1.3.3 Results POD 077.0

POD077.0: *DNMT3A* c.993delC; p.(Phe331Leufs14*). Parental samples not available to confirm de novo status.

POD077.0 was born at 38 weeks and five days gestation by instrumental delivery with a birthweight of 4.3 kg (+2.2 SD). His mother had gestational diabetes during pregnancy and was treated with insulin. At the anomaly scan at 20 weeks gestation, he was found to have dilated renal pelvices and pelviureteric junction (PUJ) obstruction. He had neonatal hypotonia, feeding difficulties, jaundice and an umbilical hernia. Age eleven he was 160.5 cm tall (+2.6 SD) with a weight of 56.8 kg (+2.3 SD) and an OFC of 57.3 cm (+1.6 SD). Medical problems included hypotonia, febrile convulsions, horseshoe kidney with left hydronephrosis, PUJ dysfunction and pyeloplasty age two, hypermobility and pes planus. He also had an abnormality of temperature regulation with increased sweating and pain insensitivity. He had global developmental delay and sat age eight months and walked at 22 months. He put two words together at 25 months. He had behavioural issues including anxiety and short attention span and he attended a special school. He had a broad forehead, deeply set eyes, horizontal eyebrows, a short philtrum, overgrowth of the alveolar ridges, and a high narrow palate. He had long tapering fingers and clinodactyly of the 4th and 5th toes. His mother was 167.8 cm (+0.7 SD) tall and his father was 177.8 cm (+0.1 SD) tall.

4.1.3.4 Discussion POD 077.0

Gestational diabetes has not previously been described in association with TBRS and is a common complication of pregnancy. It may have contributed to the high birth weight in this participant although +2.2 SD is within the range reported in TBRS (-1.2 SD – 4.0 SD)[89]. His growth parameters age 11 are also within the reported range for this condition. Jaundice and feeding difficulties in the neonatal period are likely to be related to hypotonia, a common feature of TBRS. Hypermobility is also a common feature of TBRS. Umbilical hernia and pes planus have been described in a number of individuals with this condition[89,91,206]. Horseshoe kidney and PUJ obstruction have not previously reported in TBRS however there is one individual with bilateral hydroureteronephrosis and left ureteral ectasia and one individual with multiple renal cysts[89]. This participant also had autonomic features with abnormal temperature regulation with increased sweating.

Developmental delay is present in all individuals with TBRS and behavioural issues of anxiety and short attention span have been previously reported. This participant has the characteristic facial features of TBRS of horizontal eyebrows and narrow palpebral fissures. A high narrow palate has also been described in other individuals[206]. Tapering fingers and toe clinodactyly have not previously been reported.

Many of the known features of TBRS, including high birth weight, tall stature, hypotonia, umbilical hernia, pes planus, hypermobility, intellectual disability and behavioural issues are present in this participant. His abnormal temperature regulation, increased sweating and pain insensitivity confirm autonomic dysfunction is part of the TBRS phenotype. The congenital

renal anomalies of horseshoe kidney and PUJ obstruction in this participant are novel features of TBRS and expand the phenotype of this overgrowth disorder.

4.1.4 *EZH2*: Weaver syndrome

4.1.4.1 Results POD 017.0

POD 017.0: *EZH2* maternally inherited c.1876 G>A; p.(Val626Met)

POD 017.0 was born at 38 weeks gestation by elective LSCS for breech presentation. Her mother had gestational diabetes and preeclampsia during the pregnancy. She was macrosomic with a birth weight of 4.1kg (+2.5 SD) and had neonatal jaundice that did not meet the threshold for treatment.

Age five she was 123.4 cm tall (+1.9 SD), weighed 27.8 kg (+1.7 SD) and her OFC was 53.6 cm (+1.2 SD). She had seizures and poor coordination. Her MRI head scan was reported to be normal. She had mild scoliosis, camptodactyly of her fingers and toes, flat fingernails and toenails, and ingrown toenails. She also had strabismus and cataracts. Her facial features included a round face, horizontal crease in her chin and long ears.

She had global developmental delay and sat at the age of 12 months and walked at 24 months. Behaviour issues included polyphagia and poor sleep. She attended mainstream school with assistance.

Her mother also had a confirmed molecular diagnosis of Weaver syndrome but declined participation in the study. POD 017.0 met the criteria for inclusion in the study following her molecular diagnosis.

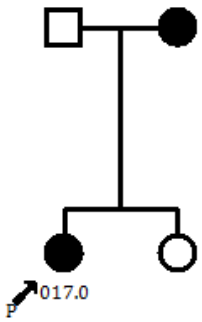


Figure 35: Pedigree POD 017.0

4.1.4.2 Discussion POD 017.0

The high birth weight of +2.5 SD in this participant is within the range reported in Weaver syndrome (-1.6 SD to +4.6 SD)[24]. Her height of +1.9 SD is on the lower side for Weaver syndrome, as the large majority of individuals have a height over +2 SD and a proportion have a height over +4 SD[24]. Her head circumference is in keeping with reported OFCs ranging from -0.9 to +5.5 SD with a median value of 1.8 SD in one study[24].

Poor coordination, scoliosis, camptodactyly, and deep-set nails are frequently reported in Weaver syndrome[23,24,27,208]. Seizures, strabismus and polyphagia have also been described in a small number of individuals[24,27]. A round face, long ears and a horizontal chin crease are characteristic facial features of Weaver syndrome in early childhood[23,24]. Developmental

delay and mild intellectual disability are common in Weaver syndrome. Poor sleep has not been specifically reported as a feature of Weaver syndrome but is not uncommon in young children with developmental delay. Cataracts have not previously been reported in Weaver syndrome. This child is undergoing further ophthalmic assessment and further investigations may identify an alternative explanation for the development of cataracts.

This participant inherited Weaver syndrome from her mother. As might be expected with an overgrowth disorder in which mild intellectual disability is more common than moderate or severe ID, a proportion of individuals in the literature are known to have inherited Weaver syndrome from a parent (14 out of 48 individuals in one series)[24].

In summary, the phenotype of this individual confirms that seizures, strabismus and polyphagia are rare features of Weaver syndrome. Cataracts may be a novel feature of Weaver syndrome.

4.1.4.3 Results POD 046.0

POD046.0: *EZH2* c.1299C>T p.(Tyr733*). Parental samples unavailable to confirm de novo status.

POD046.0 was born at 40 weeks gestation by emergency LSCS and was large for gestational age with a birthweight of 5.5 kg (+3.8 SD). He had an umbilical hernia, naevus flammeus, and neonatal jaundice requiring treatment with phototherapy.

Age eight he was 160.1 cm tall (+7.0 SD) with a weight of 53.3 kg (+4.2 SD). Previous OFC measurement was +2.8 SD. His bone age was advanced at four months of age to 12 months. His medical problems included distal arthrogryposis of the hands, hypermetropia, recurrent otitis media with conductive hearing impairment. He had global development delay. He walked at 18 months. His first word was at 24 months with two word phrases at 30 months. Behavioural issues include aggression, emotional lability, temper tantrums, short attention span, polyphagia and sleep difficulties. He attended a special school. His mother was 174.3 cm (+1.8 SD) tall with an OFC 56.7 cm (+0.9 SD) and his father was 193 cm (+2.2 SD) tall with an OFC of 60 cm (+1.8 SD).

4.1.4.4 Discussion POD 046.0

This participant's high birth weight and umbilical hernia are recognised features of Weaver syndrome[208]. He has exceptionally tall stature at +7.0 SD. Although height in Weaver syndrome is often over +4 SD, only two other individuals have a height over +6 SD and these measurements have been from young children under the age of five[24]. This participant's extreme tall stature is likely to have resulted from a familial tendency to tall stature, with a predicted midparental height of 190cm +- 8.5 cm.

Neonatal jaundice and naevus flammeus have not been reported in association with Weaver syndrome. Advanced bone age, joint contractures, hypermetropia, hearing loss, intellectual disability, and behavioural issues including temper tantrums and attention deficit have all been reported in Weaver syndrome[24]. Polyphagia has been rarely reported[27].

This participant's phenotype confirms polyphagia is a behavioural issue that may occur in Weaver syndrome. Extreme tall stature can occur in individuals with Weaver syndrome, and is likely to result from having tall parents and genetic background of tendency to constitutional tall stature.

4.1.5 *NFIX*: Malan syndrome

4.1.5.1 Results POD 002.0

POD 002.0: *NFIX* de novo c.248T>G; p.(Ile83Ser)

POD 002.0 was born at 40 weeks and 14 days gestation by instrumental delivery weighing 3.6 kg (-0.4 SD). Her bone age was advanced to 7.1 years at five years and five months of age. Age ten she was 150.4 cm tall (+1.3 SD), 40 kg in weight (+0.8 SD) and her OFC was 56.5 cm (+2.0 SD). Her medical problems included hypotonia, ingrown toenails, reduced visual acuity, mild-moderate conductive hearing loss, and anxiety. She had global developmental delay and was able to sit independently at six months and walk independently at 21 months. Her first word was at 24 months. She had developmental and behavioural issues including features of autism and pain insensitivity and she attended a special school. Facial features included a long triangular face, prominent forehead, a tall chin, deeply set eyes, and a thin upper lip vermilion. She also had long palms. Her mother was 153 cm (-1.8 SD) tall with an OFC of 54.5 cm (-0.7 SD) and her father was 187.5 cm (+1.5 SD) tall with an OFC of 60.5 cm (+2.0 SD).

4.2.5.2 Discussion POD 002.0

This participant's birth weight at -0.4 SD is unusual for Malan syndrome, as approximately 90% of reported individuals have an above average birth weight[209]. Her height is within the normal range at +1.3 SD but her head circumference is +2 SD, consistent with a pattern described in this condition where height is often above +2 SD in early childhood but with increasing age tends to normalise, and head circumference remains macrocephalic[209]. Advanced bone age, hypotonia, reduced visual acuity, features of autism spectrum disorder, and especially anxiety are common reported features of Malan syndrome. Intellectual disability is present in all reported individuals with Malan syndrome[209].

POD 002.0 was included in an international collaboration of 45 individuals with Malan syndrome and clinical details were contributed to the publication 'Further delineation of Malan syndrome' (see Appendix). This participant was patient number 13 in the publication[209].

4.1.6 *NSD1*: Sotos syndrome

Five participants in the study (POD 028.0, 065.0, 074.0, 080.0 and 095.0) have intragenic variants in *NSD1* causing Sotos syndrome. Another participant has a large deletion including *NSD1* and is described in section 4.3 under 'Other findings on microarray'.

4.1.6.1 Results POD 028.0

POD 028.0: *NSD1* de novo c.5791T>C; p.(Cys1931Arg)

POD 028.0 was born at 42 weeks gestation weighing 4.2 kg (+0.2 SD). During the pregnancy, an anomaly scan at 20 weeks gestation had detected a cystic dysplastic kidney. At birth he required admission to NICU for three days for feeding difficulty and hypoglycaemia.

Age four, his height was 121.8 cm (+4.5 SD), weight 26.3 kg (+3.5 SD) and OFC 55.5 cm (+2.2 SD). His medical problems included febrile convulsions, generalised seizures, and myoclonus, with an electroencephalogram (EEG) consistent with primary generalised epilepsy. He had a left sided multicystic dysplastic kidney confirmed on USS and a phimosis that required circumcision age four. He also had craniosynostosis of the metopic and sagittal sutures. Conductive hearing impairment was treated with grommets.

He had global developmental delay with motor milestones including sitting independently age 18 months and walking at 22 months. His first word was spoken age 30 months and he put two words together age 36 months. He attended mainstream school with assistance. His facial features included a thick vermilion of the upper lip and large ears.

4.1.6.2 Discussion POD 028.0

Characteristic facies are a cardinal feature of Sotos syndrome[14]. Interestingly, participant 28.0 does not have the typical facial features and Sotos syndrome was therefore not clinically suspected despite the presence of other major features of this condition (poor feeding in the neonatal period, congenital renal anomalies and seizures). This is likely to be because craniosynostosis of the metopic and sagittal sutures affected his facial appearance. Craniosynostosis is not a major feature of Sotos syndrome although it has previously been

reported in an individual with Sotos syndrome[14]. This finding confirms craniosynostosis is a rare association of Sotos syndrome and importantly demonstrates how it may affect other phenotypic features.

This participant has the two other cardinal features of Sotos syndrome, overgrowth and ID, and several of the major features, including poor feeding in the neonatal period, seizures, and congenital renal anomaly. Neonatal hypoglycaemia is not considered a major feature of this condition, but transient hyperinsulinaemic hypoglycaemia has been reported in a few individuals[210]. Hypoglycaemia is this participant could have been caused by feeding difficulty and/or hyperinsulinism.

Phimosis and conductive hearing loss are not considered major features of this condition but have previously been reported. Their presence in this participant confirms the association with Sotos syndrome.

4.1.6.3 Results POD 065.0

POD 065.0: *NSD1* de novo intragenic duplication exons 11-22

POD 065.0 was born at 34 weeks gestation by emergency LSCS because of maternal preeclampsia. She required resuscitation and ventilation for respiratory distress and was admitted to NICU for 42 days. She also had feeding difficulties, jaundice and hypotonia. Her birthweight was 2.5 kg (+1.1 SD) and on examination she had an umbilical hernia and a strawberry naevus on her scalp.

Age 15 her height was 181.5 cm (+3.1 SD), weight 66.3 kg (+1.4 SD), and OFC 61.9 cm (+5.0 cm). Her bone age was advanced to five years two months at the age of three years and three months. She entered puberty at 11 years and 10 months. She was hypotonic and had an abnormal MRI head scan at the age of three years and five months showing dilated asymmetric lateral ventricles, enlarged extra-axial spaces, and a thin corpus callosum. She had recurrent urinary tract infections (UTIs) however a renal USS was reported as normal. She complained of back and leg pain. Slight scoliosis and pes planus were present on examination. A history of early eruption of teeth and overcrowded teeth requiring extraction was noted.

Global developmental delay was present, with age of first sitting at seven months and walking at 23 months. Behavioural difficulties included aggression, emotional lability, anxiety, features of autism, short attention span, obsessive-compulsive behaviour and pain insensitivity. She also had increased sweating with abnormal temperature regulation. She attended mainstream school with a statement of special educational needs. Facial features included a long face, broad forehead, tall pointed chin and a high palate. She had long palms. Her mother was 170 cm (+1.1 SD) tall with an OFC of 57.8 cm (+1.6 SD) and her father was 173 cm (-0.6 SD) tall with an OFC of 58.8 cm (+1.1 SD).

4.1.6.4 Discussion POD 065.0

Maternal pre-eclampsia is a major feature of Sotos syndrome. This participant also had the neonatal complications of hypotonia, feeding difficulties, and jaundice, which are likely to be multifactorial in this individual because of both prematurity and Sotos syndrome. Umbilical hernia is not a major feature of Sotos syndrome, unlike in Beckwith-Wiedemann syndrome,

but has previously been reported in the literature. Strawberry naevi have not been described in association with Sotos syndrome but are not uncommon in the general neonatal population.

This participant has the cardinal features of characteristic facies, overgrowth, and intellectual disability. Her head circumference of 5 SD above the mean for age and sex is strikingly increased, demonstrating that extreme macrocephaly can be a feature of Sotos syndrome.

4.1.6.5 Results POD 074.0

POD 074.0: *NSD1* c.5279_5282delTCTC; p.(Val1760Glyfs*2). Parental samples not available to confirm de novo status.

POD 074.0 was born at 37 weeks and three days gestation with a birthweight of 2.93 kg (+0.1 SD), birth length 50 cm (+1.0 SD), and OFC 33cm (0 SD). There was a history of maternal hypertension in pregnancy. She required resuscitation at birth and was admitted to NICU for 21 days for CPAP to treat respiratory distress and suspected sepsis. She had hypotonia, feeding difficulties, hypoglycaemia and jaundice requiring treatment with exchange transfusion. She was noted to have micrognathia.

Age four her height was +1.3 SD, weight -0.7 SD, and OFC +2.0 SD. Her medical problems included recurrent upper and lower respiratory tract infections and otitis media, adenoidectomy and insertion of grommets, hypotonia, discoloured worn teeth, and astigmatism. She had global developmental delay and was able to sit independently at 11 months and walk independently at 18 months. Behavioural difficulties included aggression,

temper tantrums, anxiety, short attention span, and pain insensitivity. She had increased sweating and abnormal temperature regulation. She attended mainstream school. Her facial features included a high anterior hairline, long triangular face, facial erythema and frontal bossing. Her mother was 153 cm (-1.7 SD) tall and her father was 189 cm (+1.7 SD) tall. Her two year old brother did not have an overgrowth disorder.

4.1.6.6 Discussion POD 074.0

The absence of tall stature in this participant is not typical of Sotos, but not all individuals with Sotos have a height > 2 SD. Micrognathia is not known to be a feature of Sotos syndrome. Otherwise this patient is typical of Sotos syndrome, with characteristic facial features, overgrowth in the form of macrocephaly, and developmental delay.

4.1.6.7 Results POD 080.0

POD 080.0: *NSD1* c.1187delC; p.(Pro396Leufs*23). Parental samples not available to confirm de novo status.

POD 080.0 was conceived using IVF and was born at 39 weeks and 5 days gestation by elective LSCS weighing 3.3kg (-0.5 SD). He required resuscitation and was admitted to NICU for 28 days. He required ventilation for respiratory distress, phototherapy for jaundice and phenobarbitone for neonatal seizures. He also had hypotonia and feeding difficulties.

Age 12 he was 179 cm tall (+3.9 SD), weighed 65.6 kg (+2.4 SD), and his OFC was 58.5 cm (+2.0 SD). He entered puberty age 11 years. His medical problems included recurrent lower respiratory tract infections and asthma, hypermobility, abnormal dentition with loss of

enamel, and an episode of diplopia. He had global developmental delay with milestones including age of first word at 48 months. He had features of autism and attended a mainstream school with assistance. He had a prominent forehead, anterior earlobe creases, and broad long palms. His mother was 167 cm tall (+0.6 SD) and his father 179 cm tall (+0.2 SD).

4.1.6.8 Discussion POD 080.0

The anterior earlobe creases in this participant are not a feature of Sotos syndrome, and are often seen in BWS, but can also be a familial feature. The episode of diplopia is unexplained and it is uncertain if this is clinically significant. Other features in this participant are typical of Sotos syndrome.

4.1.6.9 Results POD 095.0

POD 095.0: *NSD1* c.4833T>G; p.(Cys1611Trp). Parental samples not available to confirm de novo status.

POD 095.0 was born at 35 weeks and six days gestation weighing 3.135 kg (+1.5 SD). She was admitted to NICU for four days for respiratory distress, hypotonia, feeding difficulties, hypoglycaemia and jaundice. She was noted to have low set ears, a sacral dimple, turricephaly, high arched palate, pectus excavatum, widely spaced nipples, 2,3 toe syndactyly bilaterally and peripheral oedema.

At 1 year she weighed 10.4 kg (+0.8 SD), his OFC was 49.7 cm (+2.6 SD), and her length was 83.0 cm (+3.4 SD). Her MRI head showed benign enlargement of the subarachnoid space

and borderline myelination. She had pectus excavatum, hearing loss, and global developmental delay. She was able to sit independently at 12 months.

4.1.6.10 Discussion POD 095.0

The neonatal history of hypotonia, feeding difficulties and jaundice is commonly seen in Sotos syndrome. This participant's tall stature, macrocephaly and global developmental delay are cardinal features of this condition. Pectus excavatum, 2,3 toe syndactyly, neonatal hypoglycaemia and hearing loss have also been reported[14]. Abnormal findings on neuroimaging, particularly enlargement of CSF spaces, have also been described[211].

Turricephaly is not described in Sotos syndrome, with the typical head shape being dolichocephalic. The unusual head shape was not reported to be present at 12 months of age and it is possible the appearance of the head could have been due to moulding at birth instead of true turricephaly. Her apparently low set ears, again not described as a feature of Sotos syndrome, may have appeared low set because of the head shape at birth. A high arched palate, sacral dimple, and widely spaced nipples are also not generally seen in Sotos syndrome.

The neonatal peripheral oedema described in this participant is also not typical of Sotos syndrome. There are however reports of adults with Sotos syndrome rarely developing lymphoedema[212]. There are several causes of oedema in the neonatal period including congenital heart disease. There would be a high index of suspicion for congenital heart disease in a neonate with Sotos syndrome but this was not the cause of oedema in this participant.

The features in POD 095.0 demonstrate the phenotypic variability that can be seen in individuals with Sotos syndrome.

4.1.6.11 Summary of participants with Sotos syndrome

Table 28: Summary of clinical features of participants with Sotos syndrome

	028.0	**065.0**	**074.0**	**080.0**	**095.0**
Age	4	15	4	12	1
Cardinal features					
Facial features	-	+	+	+	nk
Overgrowth > 2 SD					
Birth weight	+	-	-	-	-
Birth length	nk	nk	-	nk	nk
Birth OFC	nk	nk	-	nk	nk
Height	+	+	-	+	+
Weight	+	-	-	+	-
OFC	+	+	+	+	+
Intellectual disability	+	+	+	+	+
Major features					
Maternal pre-eclampsia	-	+	+	-	-
Neonatal hypotonia	nk	+	+	+	+
Neonatal poor feeding	+	+	+	+	+
Neonatal jaundice	-	+	+	+	+
Advanced bone age	nk	+	nk	nk	nk
Cardiac anomalies	-	-	-	-	-
Renal anomalies	+	-	-	-	-
MRI head anomalies	-	+	nk	nk	+
Seizures	+	-	-	+	-
Scoliosis	-	+	-	-	-
Hypermobility	-	-	-	+	-
Other features					
	Neonatal hypo glycaemia Phimosis Craniosyno-stosis Conductive hearing impairment	Umbilical hernia Strawberry naevus Recurrent UTIs Pes planus Overcrowded teeth Behavioural issues Autonomic dysfunction	Neonatal hypo glycaemia Recurrent URTI and LRTI Worn teeth Astigmatism Behavioural issues Autonomic dysfunction	Recurrent LRTIs Loss of dental enamel Behavioural issues	Pectus excavatum Hearing impairment

The clinical features of Sotos syndrome in childhood are well described[13,14]. The cardinal features present in >90% of individuals are characteristic facial features, overgrowth, and

intellectual disability. The major features of Sotos syndrome present in >15% of individuals are neonatal hypotonia, poor feeding and jaundice, congenital cardiac anomalies, congenital renal anomalies, abnormalities on MRI brain, joint hypermobility (including pes planus), scoliosis, seizures, and advanced bone age.

Other less common features include craniosynostosis, pectus excavatum, 2,3 toe syndactyly, phimosis, astigmatism, and umbilical hernia. The development of craniosynostosis in POD 28.0 meant the characteristic facial features of Sotos syndrome were not present in this individual and despite the presence of other major clinical features, Sotos syndrome was not suspected until a molecular diagnosis was made on panel testing.

4.1.7 *PDGFRB*: Kosaki overgrowth syndrome (KOGS)

4.1.7.1 Results POD 064.0

POD 064.0: *PDGFRB* de novo c.1751C>G; p.(Pro584Arg)

Participant POD 064.0 was suspected to have fetal overgrowth on antenatal ultrasound scans. He was born by normal vaginal delivery at 37 weeks gestation weighing 3.7kg (+1.9 SD) following induction of labour for prolonged rupture of the membranes and oligohydramnios. At 48 hours of age, he was suspected to have testicular torsion. He had unilateral cryptorchidism and a bilateral orchidopexy was performed.

At the age of 14 months, he was diagnosed with craniosynostosis affecting the sagittal, coronal, and metopic sutures (Figure 43). An MRI brain also showed bilateral periventricular cystic foci (Figure 45). Molecular genetic analysis for craniofacial disorders with sequencing of *FGFR1* exon 7, *FGFR2* exons IIIa and IIIc, *FGFR3* exons 7 and 10, *TWIST1* exon 1, *FGFR2* exons 3,5,11,14,15,16,17, and 18, and MLPA of *TWIST1*, *RUNDX2*, *ALX1*, *ALX3*, *ALX4*, *MSX2*, and *EFNB1*, did not identify any pathogenic variants. Microarray analysis did not identify any clinically significant copy number variants.

He began to lose his primary teeth early at the age of three instead of the usual six or seven years. He developed obstructive ventriculomegaly with tonsillar descent at the age of six (Figure 47) and required emergency surgery to insert a ventriculoperitoneal (VP) shunt. Imaging showed changes consistent with chronic raised intracranial pressure (Figure 46). The combination of raised intracranial pressure and enlargement of mid brain cysts was thought to explain his loss of vision at this time. Visual assessment with the Kay Picture Test Linear

Crowded Book found visual acuities of perception of light only in the right eye and 0.6 at 10 cm in the left eye.

At age six, he had tall stature with a height of 132.5cm (+2.8 SD), head circumference 51.0 cm (-1.5 SD) and weight 27.8 kg (+1.7 SD). His relatively small head size was attributed to his craniosynostosis. Other medical problems included hearing loss requiring hearing aids. He had equinovarus of the feet and ankles managed with orthotic boots and x-rays showed widening of the metatarsals. He had finger contractures of both hands and widening of the metacarpals and phalanges. A history of gnawing of his fingers was investigated and he was subsequently found to have severe carpal tunnel syndrome. Bilateral carpal tunnel release was performed for severe median nerve compression aged eight years.

There were no concerns about early developmental milestones and he walked independently at the age of 14 months. His development at the age of four was assessed to be normal by an Ages and Stages Questionnaire (ASQ). He started mainstream school but moved to special education age six because of his disabilities including visual impairment, hearing impairment, and fine motor difficulties because of flexion contractures of his fingers. He also missed periods of education because of frequent hospital admissions with recurrent VP shunt blockages.

His parents were unrelated white British. He had two elder maternal half siblings and a younger sibling, all of whom were healthy.

On clinical examination at the age of six, he was noted to have dysmorphic facial features consisted of brachycephaly, a sloping forehead, prominent supraorbital ridges, widely spaced eyes, proptosis, downslanting palpebral fissures, a wide nasal bridge, nasal base, and tip; malar flattening, midface retrusion, a smooth philtrum, thin vermilion of the upper lip, everted vermilion of the lower lip, widely spaced teeth and cupped ears (Figure 39). He had a number of dermatological findings including an extremely tanned appearance (out of keeping with ethnic background and other family members and with no history of sun exposure), hyperelastic soft thin skin, lax skin on the palms, and pigmented lesions on the thighs. Skeletal features were also noted, including pectus excavatum, reduced extension at the elbows, long and broad palms, and camptodactyly of the fingers and toes (Figure 40 and Figure 42).

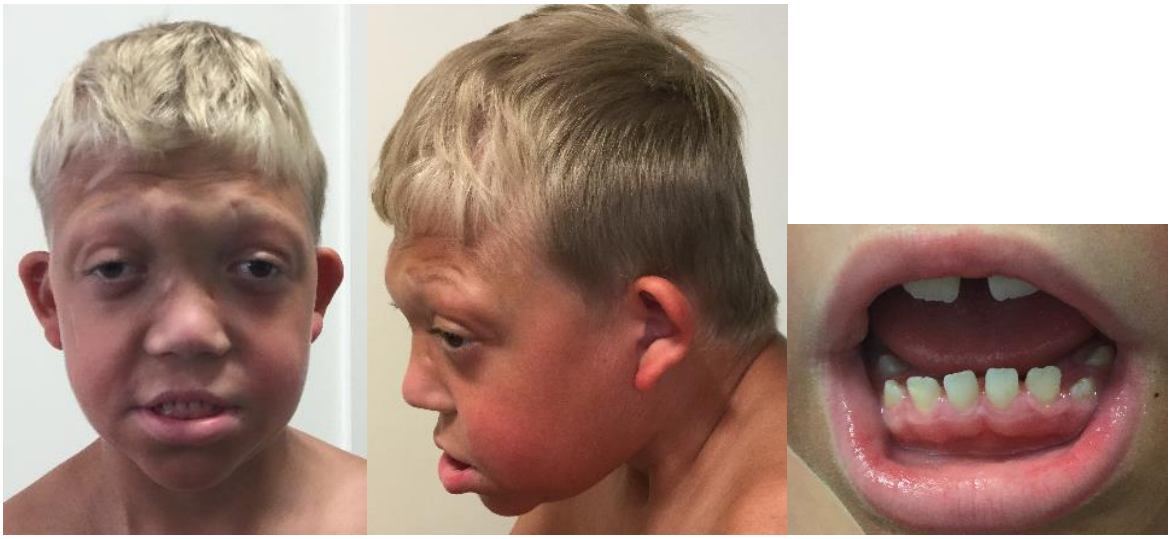


Figure 36: Clinical photographs of POD 064.0: facial features

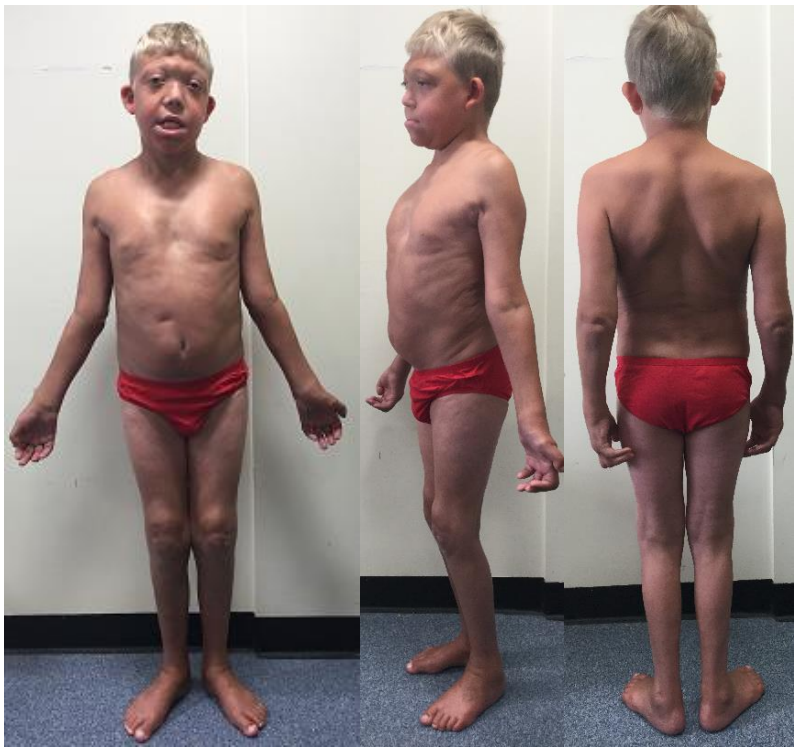


Figure 37: Clinical photographs of POD 064.0 showing tall stature, dermatological and skeletal features. A scar on the abdomen is the result of VP shunt insertion.

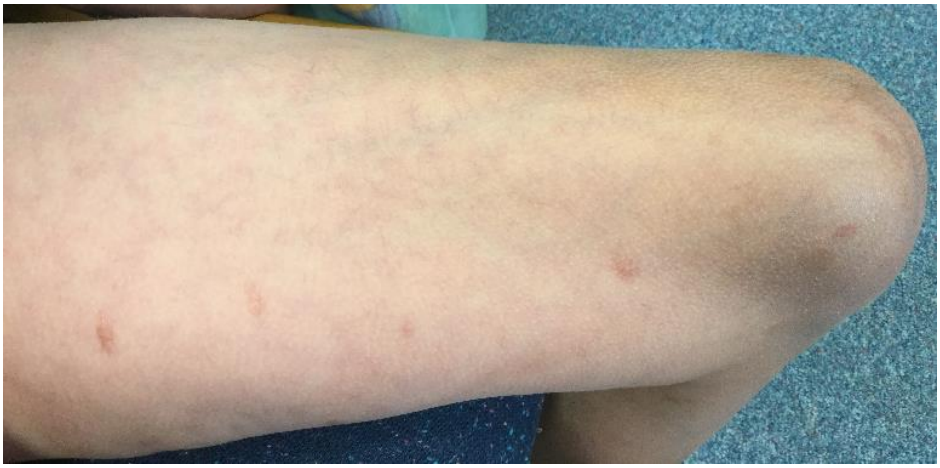


Figure 38: Clinical photograph of POD 064.0 showing hyperpigmented dermal and subcutaneous nodules

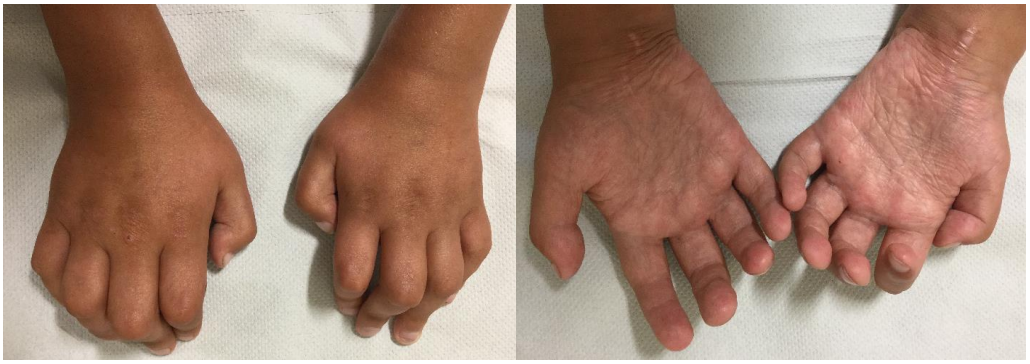


Figure 39: Clinical photographs of POD 064.0 showing camptodactyly and lax skin on the palms. Scars on both wrists are from carpal tunnel release surgery.

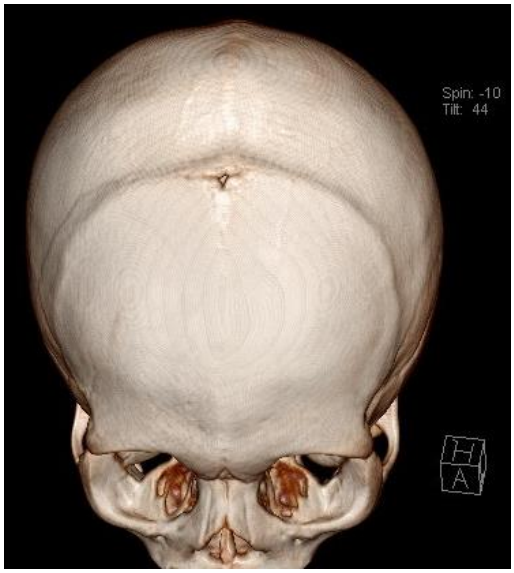


Figure 40: POD 064.0 CT head age 18 months showing almost complete fusion of the sagittal, coronal and metopic sutures

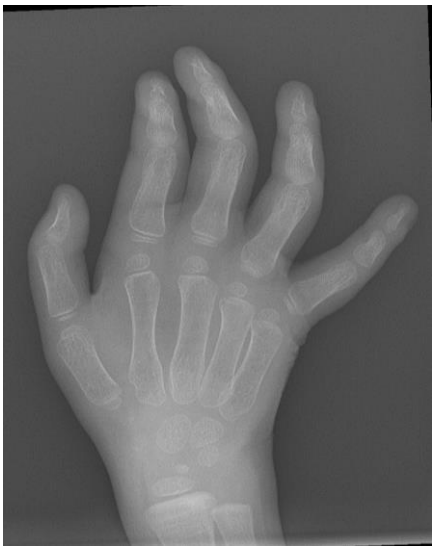


Figure 41: POD 064.0 Hand x-ray age three years showing widening of the metacarpals and phalanges with carpal crowding

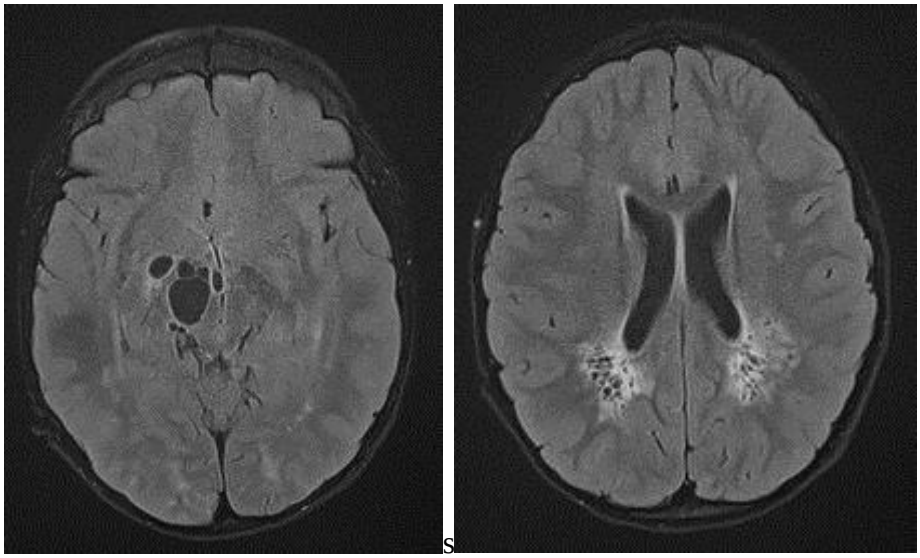


Figure 42: POD 064.0 MRI head age three years showing generalised parenchymal loss and severe cystic changes

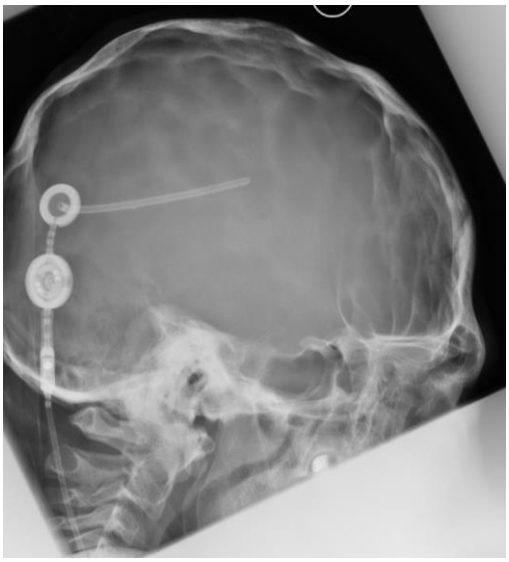


Figure 43: POD 064.0 CT head age five years showing copper beaten skull resulting from chronic raised intrcranial pressure

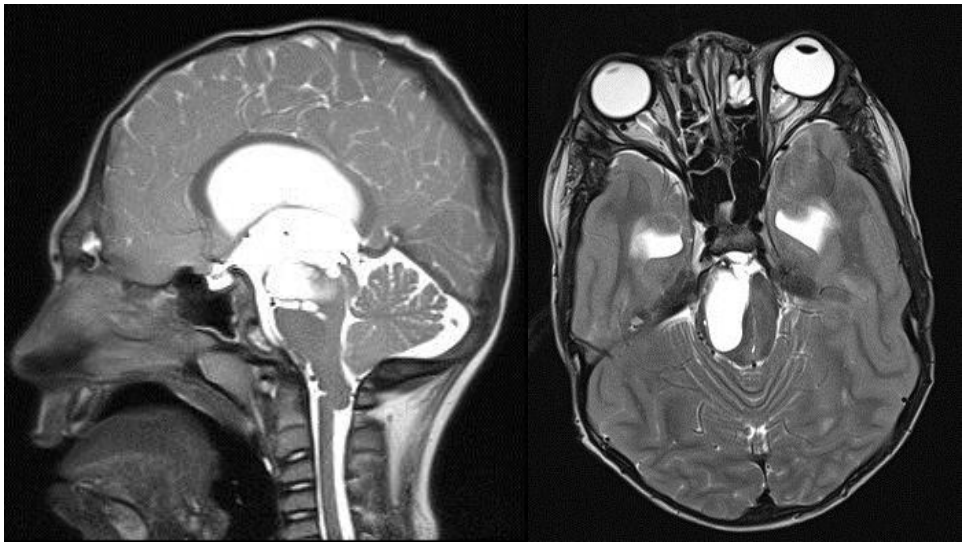


Figure 44: POD 064.0 MRI head age five years showing foci in the right midbrain occluding the third ventricle and cerebral aqueduct, enlargement of the lateral ventricles and third ventricle, and desent of the brainstem and cerebellar tonsils

Testing on the 44 gene overgrowth panel identified a pathogenic variant c.1751C>G p.(Pro584Arg) in *PDGFRB* consistent with a diagnosis of Kosaki overgrowth syndrome (KOGS).

4.1.7.2 Discussion POD 064.0

At the time POD 064.0 was diagnosed with KOGS, only five patients worldwide had been identified with this ultra-rare disorder[104,129,132]. Takenouchi et al. described two girls in 2015 with overgrowth, distinctive facial features, thin and hyperelastic skin, scoliosis, abnormal cranial shape, and hyperintense lesions in the white matter on MRI head scan[104]. The first patient also developed a myofibroma age eight and psychiatric symptoms of depression, anxiety and schizophrenia age 14[104]. The second patient experienced progressive intellectual disability, with achievement of developmental milestones in infancy but an IQ of 73 age five and IQ <40 at age 13[104].

Two further patients were reported by Minatogawa et al. in 2017. Additional features of multiple myofibromas, constriction bands of the fingers, 2,3,4 toe syndactyly, arachnoid cysts and ventricular enlargement on MRI brains scan were described in the first patient[129]. The second patient also had hydrocephalus, Dandy-Walker variant, pectus excavatum, bifid uvula, progressive lipodystrophy, a prematurely aged appearance, advanced bone age and findings on echocardiogram of mitral valve bowing, mild pulmonary stenosis and post stenotic dilatation[129].

Gawlinski et al. described the fifth patient in 2017, proposing that the KOGS phenotype also included macrocephaly, camptodactyly, elbow contractures, cryptorchidism, strabismus and widely spaced teeth [132]. Table 29 summarises the features of the first five reported patients and POD 064.0.

Table 29: Clinical features of the first six patients with Kosaki overgrowth syndrome

	Patient 1 Takenouchi et al. 2015	Patient 2 Takenouchi et al. 2015	Patient 3 Minatogawa et al. 2017	Patient 4 Minatogawa et al. 2017	Patient 5 Gawlinski et al. 2018	POD 064.0 Foster et al. 2020
Pathogenic variant	c.1751C>G p.(Pro584Arg)	c.1751C>G p.(Pro584Arg)	c.1696T>C p.(Trp566Arg)	c.1696T>C p.(Trp566Arg)	c.1751C>G p.(Pro584Arg)	c.1571C>G p.(Pro584Arg)
Sex	F	F	F	F	M	M
Age (years)	14	17	3	15	10	6
Core features						
Tall stature	+	+	+	+	+	+
Hyperelastic, thin, fragile skin	+	+	+	+	+	+
Periventricular white matter lesions	+	+	+	+	+	+
Large hands	+	+	+	+	+	+
Facial features	+	+	+	+	+	+
Additional features						
Hypotonia/delay in infancy	-	-	+	+	+	-
Scoliosis	+	+	-	+	+	-
Sparse hair	-	-	-	+	+	(+)
Ptosis	+	-	(+)	(+)	+	(+)
Lipodystrophy	-	-	-	+	+	+
Prematurely aged appearance	-	-	-	+	+	+
Progressive ID	-	+	nk	+	-	-
Psychiatric symptoms	+	-	-	+	-	-
Myofibroma	+	-	+	-	-	-
Abnormal cranial shape	+	+	-	-	+	+
Large ventricles /hydrocephalus	-	-	+	+	+	+
Arachnoid cysts	-	-	+	-	+	+
Proposed features						
Macrocephaly	nk	nk	nk	-	+	-
Strabismus	-	-	-	-	+	-
Cryptorchidism	na	na	na	na	+	+
Camptodactyly	-	-	-	-	+	+
Elbow contractures	-	-	-	-	+	+
Widely spaced teeth	nk	nk	nk	nk	+	+
Constriction bands	-	-	+	-	-	-
Dandy-Walker malformation	-	-	-	+	-	-
Novel features in POD 060.0						
Craniosynostosis	-	-	-	-	-	+
Increased pigmentation	-	-	-	-	-	+
Premature loss primary dentition	nk	nk	nk	nk	nk	+
Hearing loss	-	-	-	-	-	+

- Absent + present (+) mild presentation nk not known na not applicable

POD 064.0 shares the core features of KOGS, namely tall stature, hyperelastic thin fragile skin, periventricular white matter lesions on MRI brain, large hands, and characteristic facial features, with all of the five previously reported patients. The characteristic facial features of KOGS consist of prominent forehead and supraorbital ridges, hypertelorism, downslanted palpebral fissures, wide nasal bridge, wide nasal base, malar flattening and pointed chin[104,129,132].

Several of the additional features previously reported in more than one patient with KOGS, including sparse hair, ptosis, lipodystrophy, prematurely aged appearance, enlarged ventricles and arachnoid cysts, are also present in POD 064.0, confirming these features are frequently seen in this condition. POD 064.0's phenotype also confirms that the proposed features of camptodactyly, elbow contractures, cryptorchidism, widely spaced teeth, and pectus excavatum, each previously reported in a single individual,[104,129,132] are part of the phenotypic spectrum of KOGS. The camptodactyly in POD 064.0 is however more severe and disabling than seen in the patient reported by Gawlinski et al.

Although abnormal cranial shape has previously been described, POD 064.0 is the first patient with KOGS known to have craniosynostosis. The sudden onset and subsequent progression of visual impairment likely caused by raised intracranial pressure secondary to craniosynostosis, development of obstructive ventriculomegaly and enlargement of mid brain cysts, is unique to POD 064.0. The clinical consequences of craniosynostosis make this an important addition to the phenotypic spectrum of KOGS. Increased pigmentation, carpal

tunnel syndrome, premature loss of primary dentition, and hearing loss are also novel features of KOGS in POD 064.0.

The expansion of the KOGS phenotype was presented at a poster at the Manchester Dysmorphology Meeting in 2018. Following this presentation, two further patients with phenotypes overlapping that of POD 064.0 and variants in *PDGFRB* were identified. These three individuals were published in a paper by Foster et al. in the journal 'Clinical Genetics' in 2020, with POD 064.0 being Patient 2 in the publication[213].

4.1.7.3 Adult patient with Kosaki overgrowth disorder

Another patient with KOGS (Patient 1 in Foster et al. 2020) was identified by a colleague in France. Clinical information was shared with the patient's consent.

He was born following an uneventful pregnancy and had no neonatal problems. At 15 months he had tall stature, dolichocephaly, and increased pigmentation and elasticity of the skin. He developed a corneal dystrophy at the age of five. Age eight, his skin continued to be hyperpigmented and hyperelastic, and he had dystrophic scars. Skin biopsy identified increased melanic pigmentation of the epidermis, atrophy of the epidermis and dermis, and increased elastic fibres in the dermis. He also had global hypotonia, and musculoskeletal features including hyperostosis of metopic and sagittal sutures, contractures of the knee and ankle joints, camptodactyly of the fingers, pes cavus and claw toes. Tall stature of +4 SD was noted (see Figure 48). He later developed severe scoliosis. He also had dysmorphic facial features including widely spaced eyes, wide nasal bridge and wide nasal base, long philtrum,

micrognathia, a high narrow palate and thin and fragile gums. Investigations with skeletal survey showed an advanced bone age, large phalanges, and global demineralization of the bones. He developed a pterygium of the right eye age ten and underwent keratoplasty but it later recurred. His intellectual ability was normal, with an IQ of 105. His parents were unrelated and he had four healthy older brothers. He was suspected to have Shprintzen-Goldberg or atypical Ehlers-Danlos syndrome, and was published by Stoll et al. in 1974[214].

As an adult, at the age of 32 he developed a pterygium of the left eye and again underwent keratoplasty but recurrence resulted in visual impairment. He was found to have osteoporosis age 37 but could not be treated with bisphosphonates because he had plantar calcifications.

He had an ischaemic stroke at the age of 53 and cerebral imaging identified the cause as thrombosis of a basilar artery aneurysm. Other abnormalities on MRI imaging included dolichoectasia of the cerebral arteries, subarachnoid cysts of the anterior temporal lobes, and Dandy-Walker malformation (hypoplasia of the cerebellum with large posterior fossa cyst in continuity with the fourth ventricle). Further vascular investigations with echocardiogram and magnetic resonance angiography (MRA) identified a sinuous aorta but no aortic dilatation and no other abnormalities of major blood vessels.

His phenotype in adulthood included progression of dermatological features, with hyperpigmented, atrophic, and fragile skin, and hypertrophic, retractile, calcified scars. The previous hyperelasticity was no longer present. The nails were dystrophic and pterygia were present on two toenails. Body hair was sparse. Skeletal features included disproportionate tall

stature (188cm) with reduced upper-lower segment ratio, severe scoliosis, mild pectus carinatum, joint contractures at the elbow and knee, genu valgum and camptodactyly of the 2^{nd}-5^{th} fingers. Muscle wasting, macropenis, lipodystrophy, abnormal blood vessels, and Ainhum circumferential constriction of the second and third toes of the left foot were also present. Facial features included prominent supraorbital ridges and wide nasal base and nasal bridge (see Figure 49). His intellectual ability was normal.

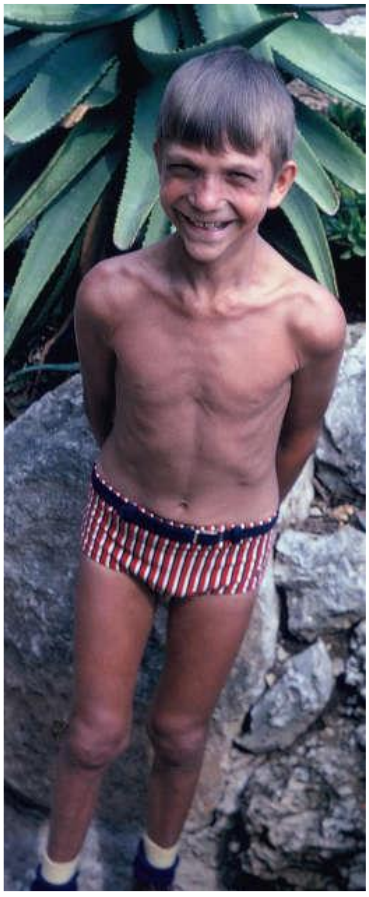


Figure 45: Adult patient with KOGS - childhood photograph

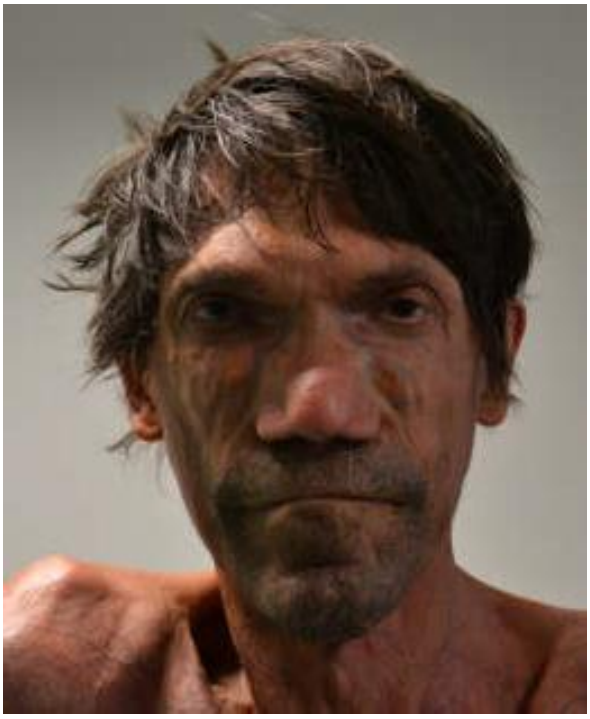


Figure 46: Adult patient with KOGS – clinical photograph showing facial features

Figure 47: Adult patient with KOGS – clinical photograph showing hyperpigmented atrophic skin and abnormal blood vessels

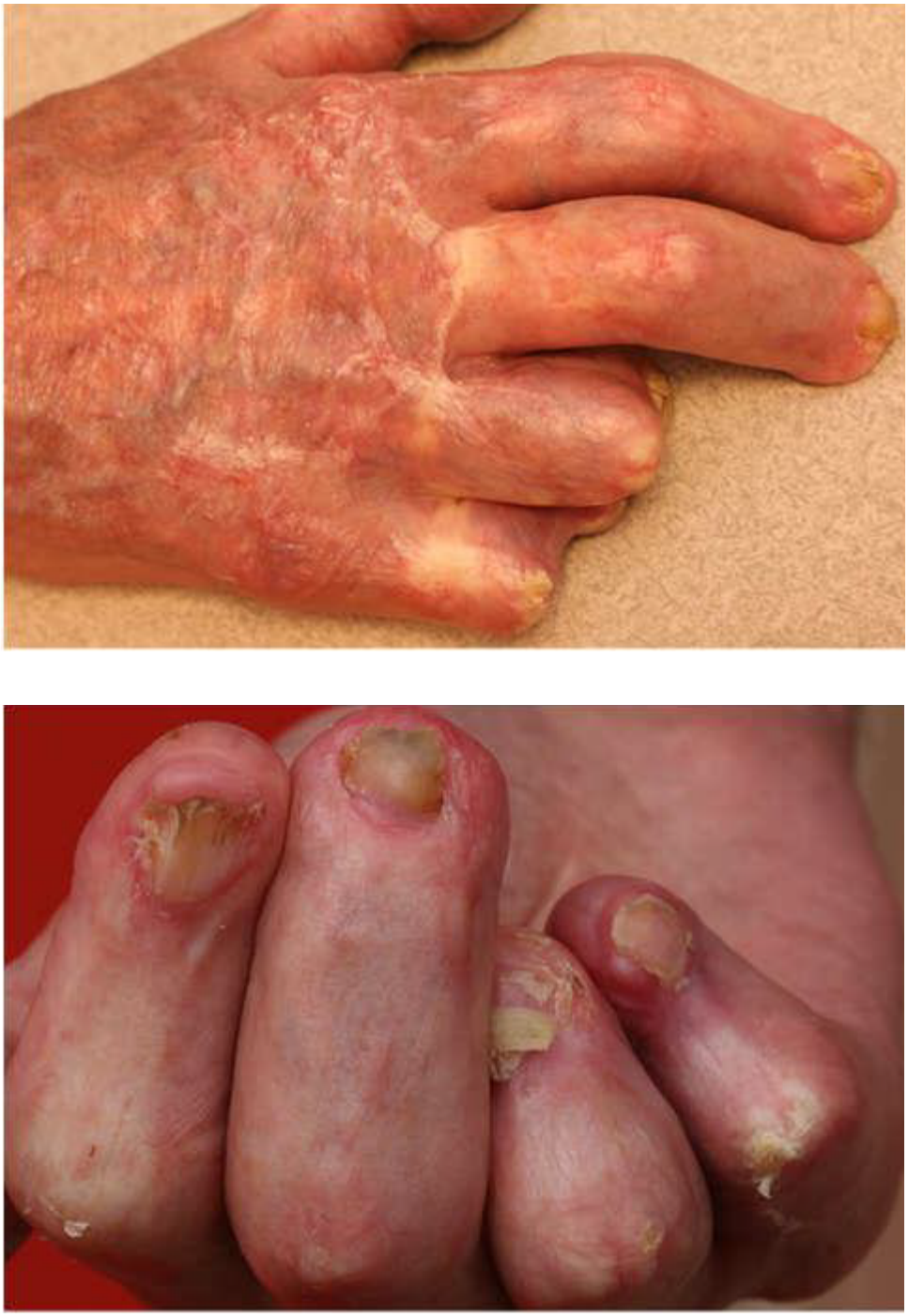


Figure 48: Adult patient with KOGS – clinical photographs showing camptodactyly, scarring, and dystrophic nails

He underwent whole exome sequencing at the Centre National de Genotypage and a c.1751C>G, p.(Pro584Arg) variant in *PDGFRB* was identified, confirming a diagnosis of Kosaki overgrowth syndrome.

This patient is the oldest individual known to have KOGS, with all other reported patients being under the age of 20 at the time of publication. This patient provides a valuable insight into the progressive nature of the dermatological and skeletal complications of KOGS. The fact that his intellectual ability and mental health remained good is important information about the natural history of this condition, given the increasing intellectual disability and psychiatric issues described in the teenage years in the first two reported patients.

4.1.7.4 Patient with a novel PDGFRB variant

A third patient with a variant in *PDGFRB* (Patient 3 in Foster et al. 2020) was identified following recruitment to the 100,000 Genomes Project and subsequent discussion of the WGS findings at the regional West Midlands clinical and laboratory genetics 100KGP MDT meeting. Phenotypic information for this patient was obtained from the medical record with parental consent.

Patient 3 was born at term weighing 3.26kg following induction of labour for hypertension. The pregnancy was unremarkable apart from the hypertension from 34 weeks gestation. At birth he appeared very thin with cutis laxa despite his normal birth weight. He also had diastasis recti. His hair and irides were noted to be very pale.

Between the ages of five months and three years, he had recurrent hypoglycaemia episodes that were ascribed to ketotic hypoglycaemia following extensive metabolic investigations. He had joint, leg and foot pain from early childhood. Following a fractured tibia treated with plaster and subsequent muscle wasting age four, there was asymmetry of the lower limbs, right leg and foot being smaller than the left. Age ten, he was noted to have a yellow ring around the iris and found to have a minor Rieger-like anterior chamber cleavage syndrome with normal vision.

On examination at the age of 11, he had a tanned appearance, soft skin with wrinkling on the palms and soles, and numerous non-atrophic scars. Other ectodermal features included minimal body hair and advanced dentition with prominent gums and widely spaced teeth. He had prominent blood vessels on his arms and legs and easy bruising. A pulsatile mass on the occiput was later surgically removed at the age of 12 and confirmed to be a vascular malformation, and a 3 cm soft lump that developed age 14 was removed and reported to be a benign angiofibromatous lesion with unusual histology. He experienced recurrent haematomas and cellulitis on his legs following minimal trauma and noted to have abnormal blood vessels when surgical treatment was attempted. Platelets, bleeding time and a clotting screen were normal. He also had little subcutaneous fat, a muscular build, and divarication of the rectus muscles. Large hands and feet and hypermobility of the fingers, knee and spine were noted. He later dislocated his shoulder age 15. He did not have tall stature, with a height on the 50th centile, weight +1.1 SD and OFC +1.0 SD. Age 20, his height was +1.2 SD and weight +1.9 SD. He had facial features including prominent supraorbital ridges, widely

spaced eyes, and a broad nasal bridge and nasal base. He attended a mainstream school with some additional support but did not need a formal statement of special education needs and proceeded to degree level education at university.

Genetic investigations including a CGH microarray and molecular analysis of *COL3A1* and *COL5A1* were normal. Fibroblast analysis of collagen type I, III and V and a metabolic screen (urinary GAGs and AA, oligosaccharides, sialic acid, MCAD) were also unremarkable. A skeletal survey reported slightly increased bone density, undermodelling of the distal femur and proximal tibia, possible mild elongation of the metacarpals and metatarsals, and an unaerated and prominent frontal bone. Brain MRI showed bilateral subcortical high signal areas, hypoplasia of the superior and inferior cerebellar vermis, and prominent subarachnoid spaces in the posterior fossa. He underwent electromyography (EMG) and nerve conduction studies that showed no evidence of myotonia and normal motor and sensory conduction in the upper and lower limbs.

He died suddenly age 21 of a stroke caused by thrombosis and dissection of a fusiform aneurysm of the basilar artery and subarachnoid haemorrhage. Imaging also showed the presence of multiple foci of signal change in the supratentorial white matter consistent with chronic ischaemia of the small vessels. The vertebral arteries and aorta were normal however the coronary arteries had moderate calcified atheroma.

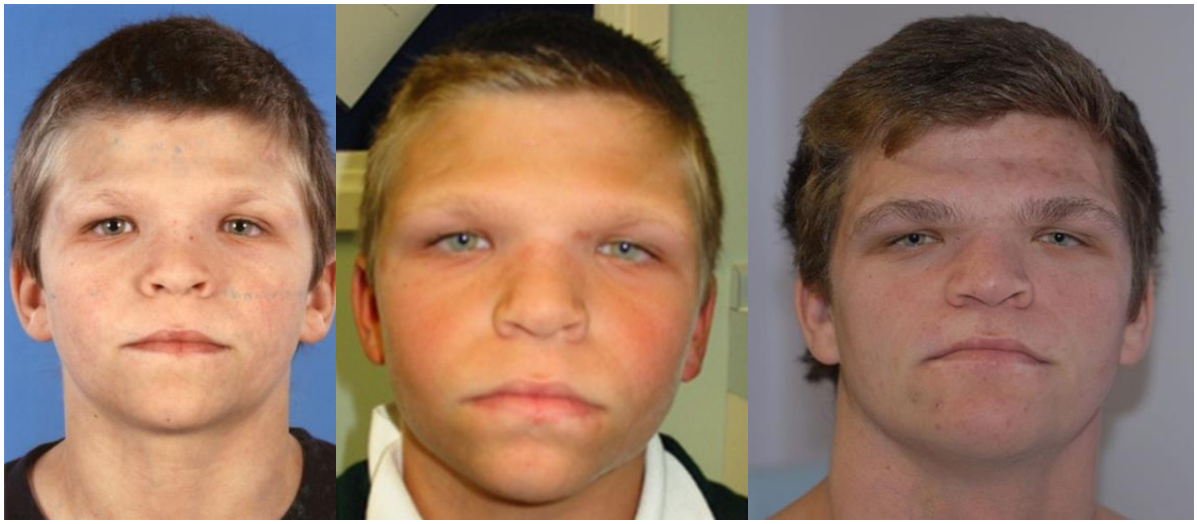


Figure 49: Clinical photographs of patient with novel PDGFRB variant: facial features age nine, 11 and 20

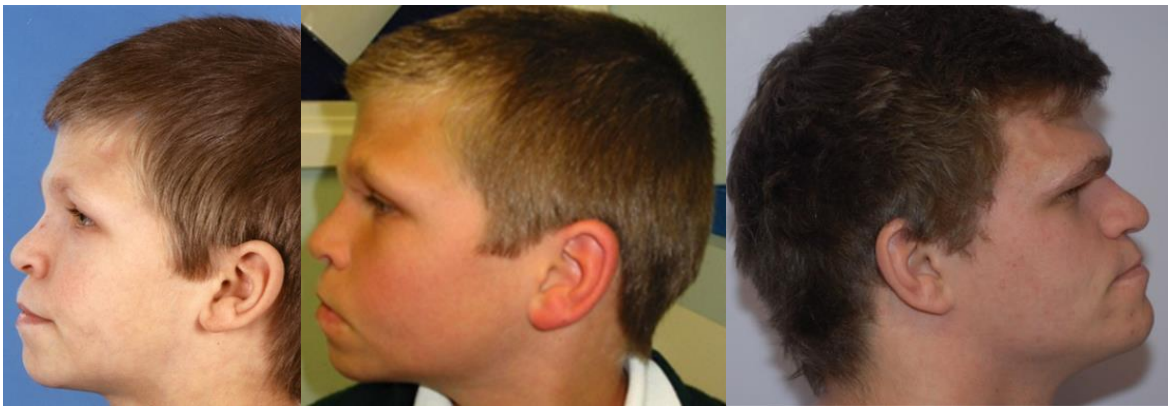


Figure 50: Clinical photographs of patient with novel PDGFRB variant: side profile age nine, 11 and 20

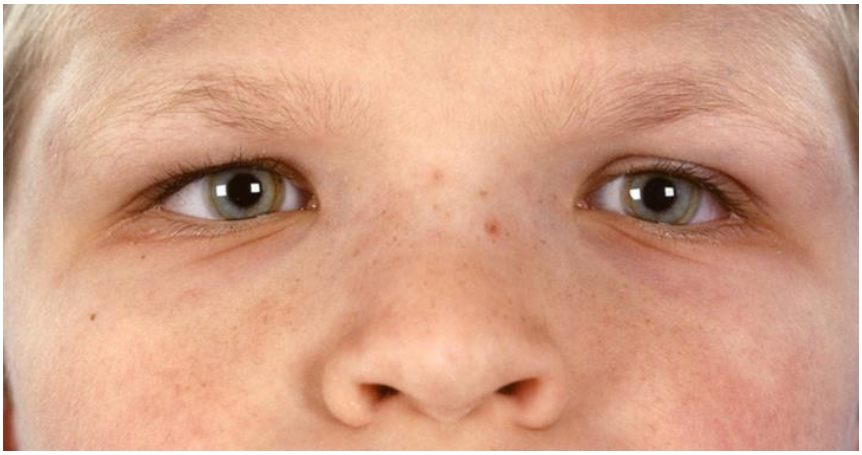


Figure 51: Clinical photograph of patient with novel PDGFRB variant: yellow ring around the iris

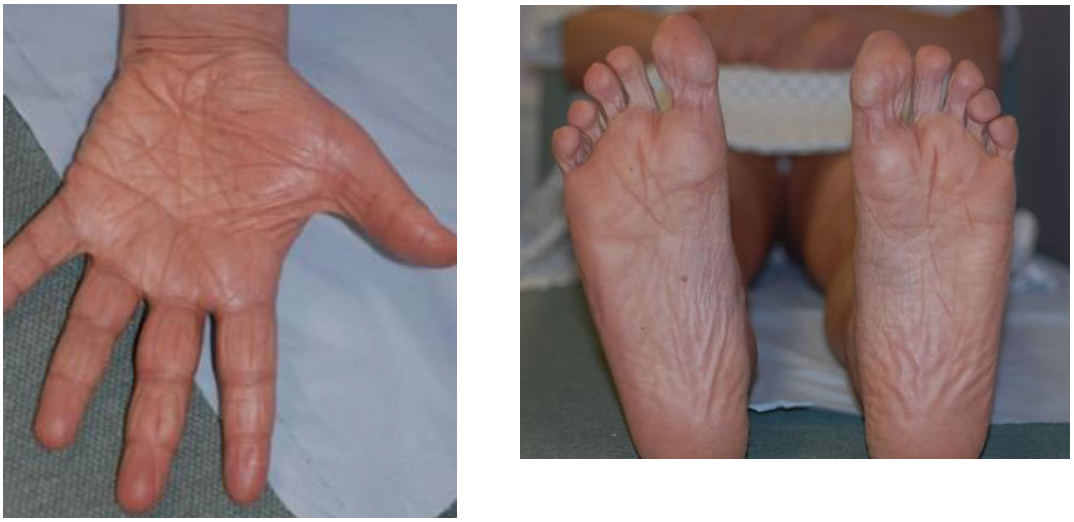


Figure 52: Clinical photographs of patient with novel PDGFRB variant: thin hyperelastic skin with wrinkling of the palms and soles

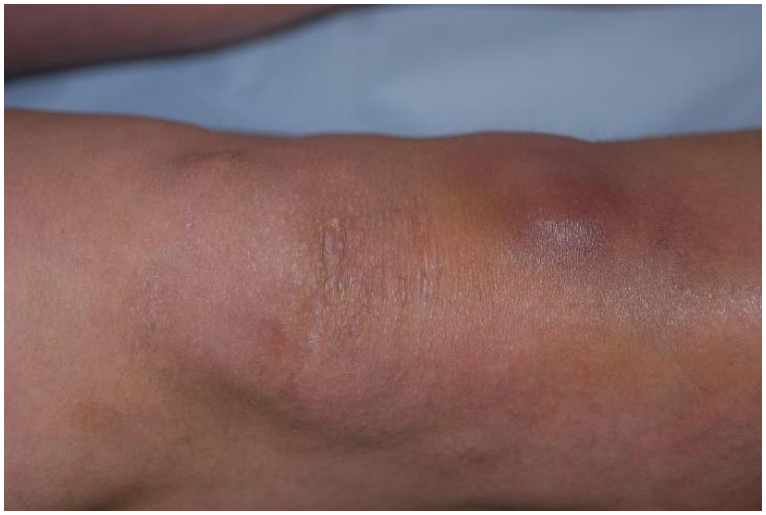


Figure 53: Clinical photograph of patient with novel PDGFRB variant: haematoma on the lower limb

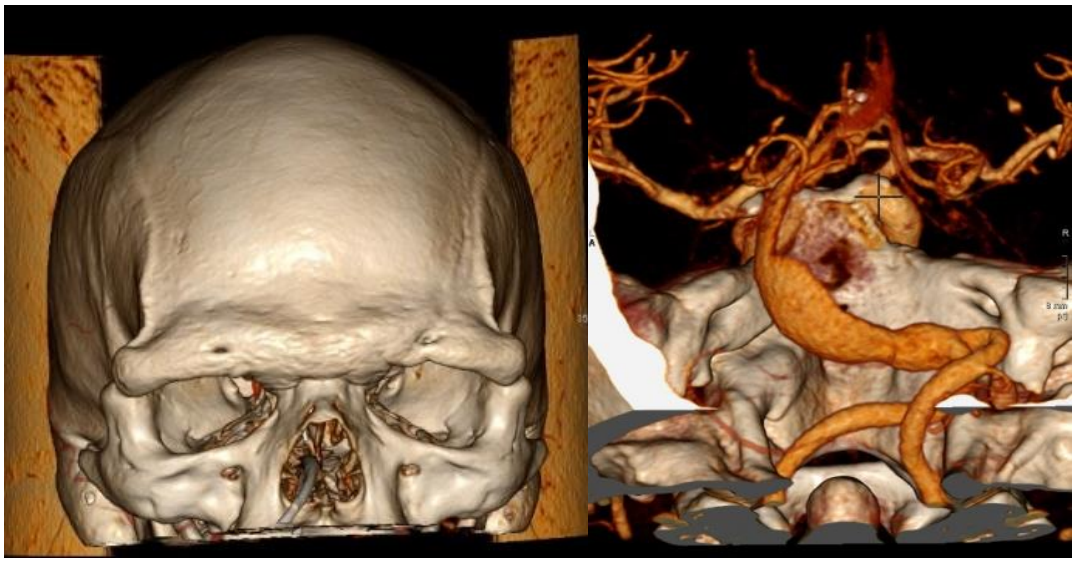


Figure 54: Patient with novel PDGFRB variant: CT angiogram images showing prominent frontal bone and thrombosis of a large fusiform aneurysm of the basilar artery

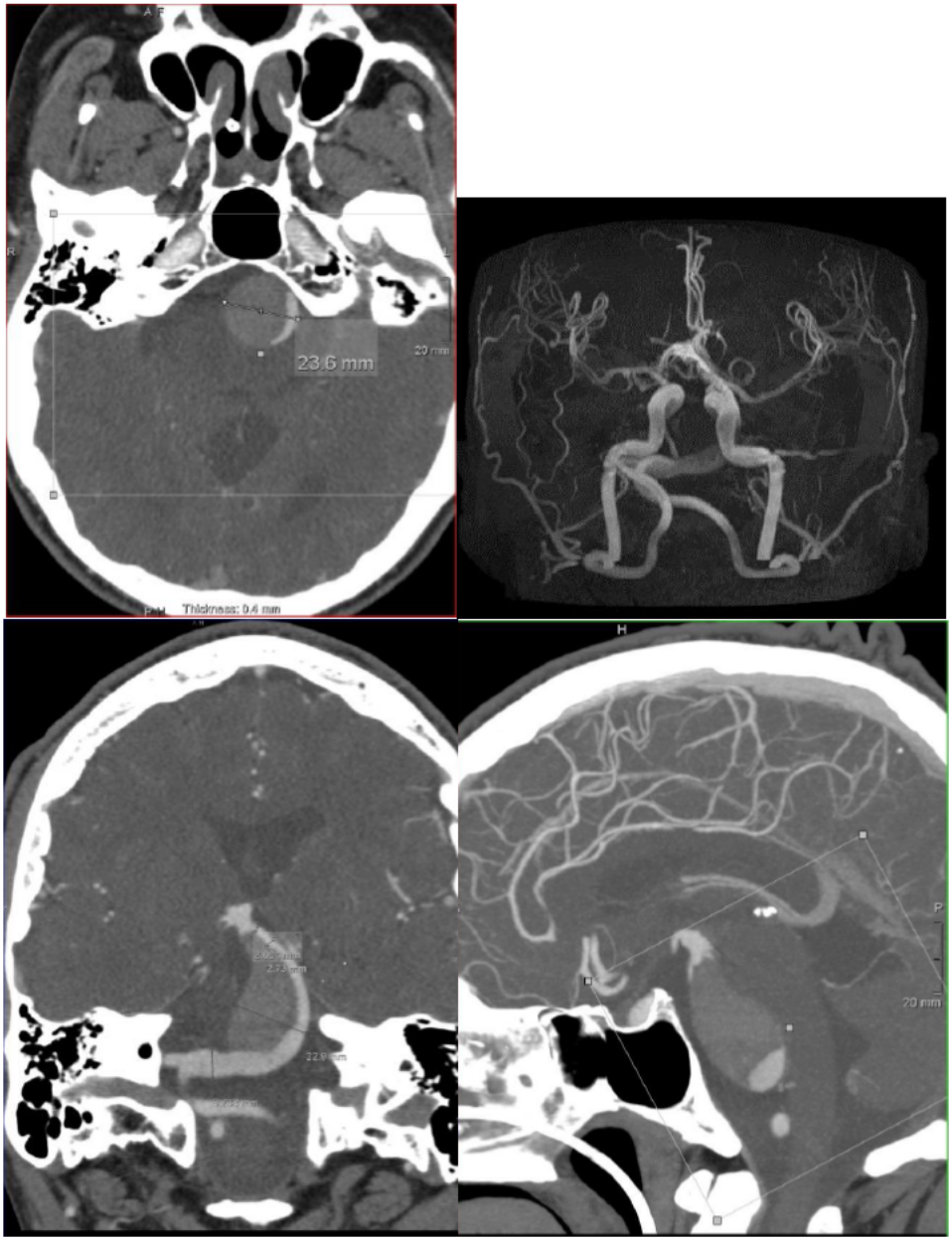


Figure 55: Patient with novel PDGFRB variant: MRA head scan images showing large fusiform aneurysm of the basilar artery indenting the brain stem including an area of thrombosis; multiple foci of signal change in the white matter consistent with small vessel ischaemia; and areas of dural ectasia with thinning of the skull vault and skull base

With the consent of his family, whole genome sequencing was performed as part of the 100,000 Genomes Project. He was found to have a novel c.1477A>T p.(Ser493Cys) variant in *PDGFRB*. This variant has not previously been reported in association with KOGS, but he shares many phenotypic features, including the typical facial features, large hands, lipodystrophy, skin features, and MRI brain anomalies. He does not have the tall stature reported in all previous individuals with KOGS, but increased height may not be universal in KOGS. Parallels can be drawn with Penttinen syndrome, another PDGFRB related disorder, in which two of the five known individuals have tall stature and three do not. He also has several novel features including recurrent hypoglycaemia in infancy, anterior chamber cleavage syndrome, recurrent haematomas, prominent musculature, joint dislocation and splenomegaly. These are not previously reported in KOGS, but given the small number of known individuals, it is possible that these are rarer features of KOGS and represent an expansion of the phenotype. Consideration must be also given however to the precise variant identified in this patient. The recurrent KOGS variants, Trp566Arg and Pro584Arg, are both in the juxtamembrane domain of the protein, but this patient's variant, Ser493Cys, is located in the transmembrane domain. Until further understanding is gained of the genotype phenotype correlations in PDGFRB-associated disorders, it remains unclear whether this patient has KOGS or a separate PDGFRB-associated 'KOGS-like' disorder.

4.1.7.5 Vascular complications in KOGS

The most striking and clinically significant feature of this patient's phenotype is his sudden death at a young age. The presence of basilar artery aneurysm with consequent thrombosis in both the French adult patient and the patient with the novel Ser493Cys variant immediately prompts consideration of the possibility of a serious vascular phenotype in KOGS. A second

review of the literature was performed, and a case report of a third individual with KOGS with vascular complications was identified, published by Zarate et al. in 2019[133]. This patient with a c.1696T>C p.(Trp566Arg) variant in *PDGFRB* had an echocardiogram suggestive of dilatation of the proximal left main coronary artery on echocardiogram, and proceeded to CT angiography which confirmed a 9.6mm saccular aneurysm in the left main coronary artery and an 8 mm saccular aneurysm in the right coronary artery[133]. She died suddenly at the age of nineteen. No post-mortem examination was performed but it would seem likely that her death was the result of a sudden vascular event.

The identification of a clinically significant vascular phenotype through this work was published in 'Kosaki overgrowth syndrome: A novel pathogenic variant in PDGFRB and expansion of the phenotype including cerebrovascular complications' by Foster et al. in the journal Clinical Genetics in 2020. Subsequently, Takenouchi et al. reported that imaging of the original two patients with KOGS identified that both individuals had similar vascular findings, with progressive dilatation of the basilar, vertebral, and coronary arteries[215]. Another patient with the KOGS-associated variant Trp566Arg and dilation of the right and left coronary arteries was reported by Wenger et al.in 2020[216]. Imaging did not identify any vascular abnormalities in POD 064.0, however it is possible that this might develop in view of his relatively young age and the apparently progressive nature of this condition. All known patients with KOGS-associated variants Trp566Arg and Pro584Arg, associated vascular findings, and morbidity and mortality are listed in Table 30.

Table 30: All known individuals with KOGS, vascular findings and complications

	Variant	Vasculature involved (age of diagnosis)	Complications (age of occurrence)
Minatogawa et al. 2015	Trp566Ar	Unknown (3)	
Minatogawa et al. 2015	Trp566Arg	Mitral valve bowing, mild pulmonary stenosis and post stenotic pulmonary artery dilation (15) Cerebral vessels unknown	
Zarate et al. 2019	Trp566Arg	Saccular aneurysms right and left coronary arteries (13) Subtle tortuosity vertebral arteries (13)	Death (19) (no post-mortem)
Wenger et al. 2020	Trp566Arg	Unknown (8)	Death (8) recurrent apnoeic episodes
Wenger et al. 2020	Trp566Arg	Dilation right and left coronary arteries (6 m) Dysplastic mitral valve with focal area of prolapse (13)	
Rustad et al. 2021	Trp566Arg	Normal imaging (9)	
Takenouchi et al. 2015; 2021	Pro584Arg	Normal imaging (12) Tortuosity and dilation of coronary arteries; occlusion middle segment L descending artery (20) Semi-fusiform aneurysm at the bifurcation of the left internal carotid and anterior choroidal artery (21)	
Takenouchi et al. 2015; 2021	Pro584Arg	Normal imaging (15) Tortuous and dilated/dolichoectasia basilar and vertebral arteries (23) Tortuous bilateral coronary arteries with multiple calcified aneurysms (23)	Compression of brainstem and cranial nerves VII and VIII – right sided hearing loss and left sided visual loss (23)
Gawlinski et al. 2018	Pro584Arg	Unknown (10)	
Foster et al. 2020	Pro584Arg	Fusiform basilar artery aneurysm (53) Dolichoectasia of cerebral arteries (53) Sinuous thoracic aorta (53)	Thrombosis basilar artery aneurysm and stroke (53)
Foster et al. 2020 (POD 064.0)	Pro584Arg	Normal imaging (9)	

Of the 11 individuals now reported with KOGS worldwide, eight have undergone vascular imaging and three have not had vascular imaging. Six out of the eight that had imaging were found to have abnormal vasculature (75%). The most common blood vessels involved were the cerebral (vertebral, basilar and carotid) arteries (4 individuals) and coronary arteries (4 individuals). Blood vessels were described as stenotic/occluded in two cases, tortuous in four cases, and dilated/dolichoectatic in five cases. Aneurysms were present in four cases. Two individuals also had valvular changes. Serious clinical complications of vascular abnormalities were confirmed in two individuals (compression of cranial nerves leading to hearing and visual loss age 23, and thrombosis of a basilar artery aneurysm age 53). A third individual died suddenly age 19, which may have been secondary to known coronary artery aneurysms, however an alternative cause of death cannot be excluded. A fourth individual died age 8 with recurrent apnoeic episodes; she had not previously undergone vascular imaging but was known to have other severe congenital brain anomalies. The incidence of serious morbidity or death resulting from vascular abnormalities is therefore at least 2/11 (18%) but could be as high as 4/11 (36%).

Vascular abnormalities were detected in individuals ranging in age from 6 months to 53 years, with a mean age of 21 years and median of 17.5 years. Normal vasculature was identified in individuals between the ages of nine and fifteen, with a mean age of 11. Of note, the two patients reported by Takenouchi et al. had normal imaging at the ages of 12 and 15 respectively, and abnormalities of both cerebral and coronary vessels age 21 and 23. A pattern of progression from normal blood vessels in childhood, to dilated, tortuous, aneurysmal

coronary and cerebral arteries in the late teen and early 20s is emerging as the common pattern. However, the presence of dilated coronary arteries in a six month old suggests that abnormalities can develop at any age.

Six of the reported individuals with KOGS have the Trp566Arg variant and five have the Pro584Arg variant. Although these are small numbers, abnormal vasculature appears to be equally associated with the Trp566Arg or Pro584Arg variant, being identified in three out of six and three out of five individuals respectively.

4.1.7.6 Vascular complications in the wider PDGFRB spectrum

The presence of vascular complications in the patient with the novel Ser493Cys variant, which as previously discussed may be a third genotype associated with KOGS or may cause a novel PDGFRB-associated disorder, suggests that a review of individuals with other PDGFRB-associated disorders is warranted.

Of note, Zufferey et al. reported a patient with Penttinen syndrome of premature ageing who had major dilatation of the left coronary artery and an aneurysm of the basilar artery leading to stroke at the age of nine[217]. Interestingly this patient had many features overlapping the KOGS phenotype, including craniosynostosis, hydrocephalus, arachnoid cysts, white matter lesions, thin transparent skin, prematurely aged appearance, sparse hair, ocular pterygia, conductive hearing loss, tall stature (+3 SD), and progressive contractures of the fingers[217]. However, he had the progressive acro-osteolysis that is typical of Penttinen syndrome and has not been described in any patient with KOGS. He also developed extreme progression of joint

contractions, lipomyoatrophy, and skin ulcerations that affected his mobility and breathing and he ultimately died from respiratory insufficiency at the age of 20. This very severe phenotype has not been seen in KOGS. This patient was reported prior to the identification of the Val665Ala variant in *PDGFRB* as the cause of Penttinen syndrome and the precise variant causing his disorder is unknown.

There is also evidence that in addition to being associated with germline variants in PDGFRB, fusiform cerebral aneurysms and other vascular complications are also associated with somatic variants in *PDGFRB*. Karasozen et al. reported an individual with a mosaic Tyr562Cys variant who had a dissecting fusiform paraclinoid internal carotid artery aneurysm at the age of nine and a giant dissecting fusiform aneurysm of the right vertebral artery at the age of 23. He later developed large aneurysms of the distal left main coronary artery and right distal radial artery[218]. He had a mosaic distribution of other clinical features affecting the upper right half of the body, comprising thin, fragile, hyperelastic skin with haemangioma-like discoloration, prominent tortuous veins on the right arm, overgrowth with a longer right arm and hand, clinodactyly of the 3rd, 4th and 5th fingers of the right hand, and absence of subcutaneous adipose tissue[218]. The variant was present in aneurysmal and other abnormal tissues and absent in normal tissue. Further work was performed on 50 samples of cerebral aneurysmal tissue taken at the time of surgery; mosaic *PDGFRB* variants were identified in three out of five samples of fusiform aneurysms, but none in the more common saccular aneurysms, suggesting mosaic variants in *PDGFRB* are associated with the development of the rare fusiform subtype of aneurysms[218].

A review of all reported individuals with vascular abnormalities and a clinical or molecular diagnosis of a PDGFRB associated disorder (IM, Penttinen syndrome, KOGS, novel disorder), and resulting clinical sequelae, is given in Table 31.

Table 31: All known individuals with PDGFRB-associated disorders and vascular abnormalities

Reported by	Clinical diagnosis	Variant	Vasculature involved (age of diagnosis)	Complications
Wright et al. 2004	IM	unknown	Progressive multifocal aneurysmal dilatation of the thoracic and abdominal aorta, carotid, iliac, and lower limb arteries, (5 months)	Thrombosis iliac and femoral arteries, shortening right limb, ischemic limb (5 m); high output cardiac failure, death (2)
Brasseur et al. 2010	IM	unknown	Aneurysm renal arteries and R common iliac artery (4m) Aneurysm internal iliac artery (14m) Stenosis proximal renal artery (14m)	Renal failure and dialysis
Frezin et al. 2015	IM	unknown	Left renal artery stenosis Aneurysms right renal artery and right common iliac artery (2)	Renal failure and successful renal transplant (3)
Zufferey et al. 2013	Penttinen?	unknown	Major dilatation left coronary arteries (7) Basilar artery aneurysm (7)	Haemorrhagic stroke (9)
Wenger et al. 2021	Penttinen	Val665Ala	Wide-spread vessel tortuosity and ectasia (14) Renal artery aneurysm (14)	
Takenouchi et al. 2015; 2021	KOGS	Pro584Arg	Semi-fusiform aneurysm at the bifurcation of the left internal carotid and anterior choroidal artery (21) Tortuosity and dilation of coronary arteries; occlusion middle segment L descending artery (20)	
Takenouchi et al. 2015; 2021	KOGS	Pro584Arg	Tortuous and dilated/dolichoectasia basilar and vertebral arteries (23) Tortuous bilateral coronary arteries with multiple calcified aneurysms (23)	Compression of brainstem and cranial nerves VII and VIII – right sided hearing loss and left sided visual loss (23)
Foster et al. 2020	KOGS	Pro584Arg	Fusiform basilar artery aneurysm (53) Dolichoectasia of cerebral arteries (53) Sinuous thoracic aorta (53)	Thrombosis basilar artery aneurysm and stroke (53)
Minatogawa et al. 2015	KOGS	Trp566Arg	Mitral valve bowing, mild pulmonary stenosis and post stenotic pulmonary artery dilation (15); cerebral vessels unknown	
Zarate et al. 2018	KOGS	Trp566Arg	Saccular aneuryms right and left coronary arteries (13) Subtle tortuosity vertebral arteries (13)	Death (19) (no postmortem)
Wenger et al. 2021	KOGS	Trp566Arg	Dilation coronary arteries (6m) Dysplastic mitral valve with focal area of prolapse (13)	
Foster et al. 2020	KOGS-like novel	Ser493Cys	Vascular malformation occiput (12) Fusiform aneurysm basilar artery (21)	Thrombosis and dissection of basilar artery aneurysm; death (21)
Chenbhanich et al. 2021	Novel mosaic Tyr562Cys	Tyr562Cys mosaic	Fusiform aneurysms left internal carotid, vertebrobasilar junction, left cerebral artery, and right middle cerebral artery (26) Fusiform aneuryms coronary arteries (26)	Compression of optic nerve and visual loss (25); rupture of left internal carotid aneurysm and stroke (27)
Wenger et al. 2021	Novel mosaic Tyr562Cys	Tyr562Cys mosaic	Fusiform aneurysm internal carotid artery (26) Cutaneous vascular malformations	Compression optic nerve (26)
Karasozen et al. 2019	Novel mosaic Tyr562Cys	Tyr562Cys mosaic	Paraclinoid internal carotid artery aneurysm (9) Giant fusiform aneurysm right vertebral artery (23) Coronary artery aneurysms Radial artery aneurysms	Dissection vertebral artery (23)

It is apparent that several variants in *PDGFRB* are associated with abnormal vasculature, including the germline variants Pro584Arg (KOGS), Trp566Arg (KOGS), Ser493Cys (KOGS-like), and mosaic variants in Tyr562Cys. Only four individuals with mosaic Tyr562Cys variants have been reported in the literature[216,218,219], and three have had fusiform aneurysms affecting the cerebral and/or coronary arteries, resulting in serious clinical complications (compression of optic nerve in two individuals and stroke in two individuals) in early adult life. The high incidence, location (cerebral and coronary artery) and type of vascular abnormality (aneurysm) reported in mosaic Tyr562Cys is therefore very similar to that reported in the KOGS and KOGs-like associated variants.

Of the two individuals reported to have Penttinen syndrome and vascular abnormalities, the patient with a clinical diagnosis reported by Zufferey et al. has the same phenotype of involvement of the coronary and cerebral arteries. However, given the lack of molecular data it is not certain this patient had Penttinen syndrome and it cannot be confirmed this vascular phenotype is consistent with this diagnosis. The second individual, first reported by Johnston et al. and then Wenger et al., has the Val665Ala variant associated with Penttinen syndrome. MRA of the chest and body demonstrated widespread vessel ectasia and a renal artery aneurysm, but there is no description of involvement of the cerebral or coronary arteries. Penttinen syndrome remains an exceptionally rare disorder. Of the three other reported individuals in the literature[220,221], one had normal coronary artery imaging at the age of 29 and the other two have not had vascular imaging. It can be concluded that there is an association between the Val665Ala variant and vascular abnormality, as demonstrated by the features in the individual reported by Johnston and Wenger, but it is currently not possible to draw any

further inferences about the incidence, location or type of vascular lesions characteristic of Penttinen syndrome.

Of note, there are three individuals in the literature with clinical diagnoses of IM and vascular abnormalities, but no molecular genetic information is available[222–224]. In contrast to the pattern of cerebral and coronary artery involvement seen in KOGS and mosaic Tyr562Cys, two of the patients with IM had aneurysms affecting the iliac and renal arteries. The third had widespread aneurysmal dilatation of the thoracic and abdominal aorta, carotid, iliac, and lower limb arteries. These distributions are reminiscent of the patient with Pettinen syndrome reported by Johnston et al. and Wenger et al., with widespread vessel ectasia and renal artery aneurysm. The IM patients had a young age of onset ranging between four months and three years, and serious clinical complications including ischaemic limb, renal failure, and death. IM is much commoner than KOGS or Penttinen syndrome, with over 100 families in the literature reported to have familial IM. The vascular abnormalities associated with IM therefore appear to be much less common, occur at a younger age, and affect different blood vessels than the vascular abnormalities associated with KOGS, KOGs-like disorder, and mosaic Tyr562Cys.

The vascular abnormalities associated with variants in PDGFRB may result from its function in pericytes, where it is highly expressed[218,219]. These cells form part of the smooth muscle of the tunica media in arteries including cerebral vessels[218,219].

4.1.7.7 Genotype-phenotype correlations in the PDGFRB spectrum

The overlapping phenotypes in the individuals with PDGFRB associated disorders already discussed make it clear that a detailed analysis of genotype-phenotype correlations is indicated.

Table 32 details the clinical diagnoses, number of individuals described, associated variants, and type of variant (germline or mosaic; gain or loss of function) of all known PDGFRB associated disorders.

Table 32: PDGFRB associated disorders

Clinical diagnosis	Number of reported patients	Report	Associated variants	Germline/ mosaic	Mechanism
Kosaki overgrowth syndrome (KOGS)	11	Takenouchi et al. 2015[104] Minatogawa et al. 2017[129] Gawlinski et al. 2018[132] Zarate et al. 2019[133] Foster et al. 2020[213] Rustad et al. 2021[225] Wenger et al. 2021[216]	Trp566Arg Pro584Arg	Germline	Gain of function
KOGS-like syndrome	1	Foster et al. 2020[213]	Ser493Cys	Germline	Suspected gain of function
Penttinen syndrome of premature ageing	5	Penttinen et al. 1997[226] (Zufferey et al. 2013[217]) Johnston et al. 2015[227] Zhang et al. 2018[221]	Val665Ala	Germline	Gain of function
Severe Penttinen -like syndrome	2	Bredrup et al. 2019[228]	Asn666Ser	Germline	Gain of function
Ocular pterygium-digital keloid dysplasia (OPDKD)	4	Abarca et el. 2014[229] Bredrup et al. 2021[230]	Asn666Tyr	Germline	Gain of function
Familial infantile myofibromatosis	47	Hettmer et al. 2020[231] Wenger et al. 2021[216]	Pro560Leu Arg561Cys (recurrent) Arg561Ser Lys567Glu Pro660Thr	Germline	Gain of function

Sporadic multifocal IM Solitary IM		Art et al. 2017[232] Al Qawahmed et al. 2019[233]	Arg561Cys Asn666Lys Asp850Val Ile538_539insArg	Mosaic	Gain of function
Cerebral aneurysms	3	Karasozen et al. 2019[218]	Arg849_Lys860delins Asp850Tyr Tyr562_Arg565del	Mosaic	Gain of function
Novel variant reported by Pond et al.	1	Pond et al. 2018[234]	Asn666His	Germline	Gain of function
Novel variant reported by Guimier et al.	1	Guimier et al. 2019[235]	Arg561_Tyr562delinsHis	Mosaic	Unknown
Novel variant reported by Zhong et al.	1	Zhong et al. 2018[236]	Asn666Lys	Mosaic	Suspected gain of function
Novel syndrome mosaic Tyr562Cys	4	Karasozen et al. 2019 Chenbhanich et al. 2020[219] Wenger et al. 2021[216]	Tyr562Cys	Mosaic	Gain of function
Unicentric Castleman disease	7	Li et al. 2019[237]	Asn666Ser	Mosaic	Gain of function
Primary familial brain calcification (PFBC)	>100 families	Hertz et al. 2019[238]	Gly612Arg Leu658Pro Arg695Cys Asp737Asn Asp844Gly Arg987Trp Glu1071Val	Germline	Loss of function
Myeloid neoplasm with eosinophilia		Golub et al. 1994 [239] Savage et al. 2013[240]	*PDGFRB* rearrangement;> 20 known gene fusion partners	Somatic	Gain of function

Within the missense variants, there is clustering at amino acid position 561/562 and at position 665/666. A representation of the location of variants is shown in Figure 59.

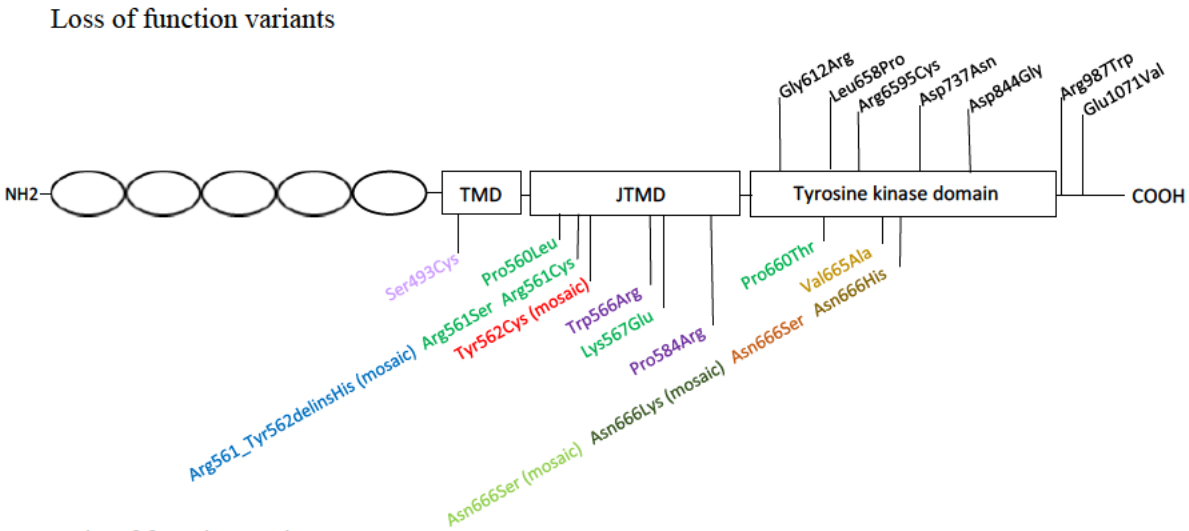


Figure 56: PDGFRB and location of known disease-causing variants

A summary of the clinical features of PDGFRB associated disorders is given in Table 33. Unicentric Castleman disease (lymphoproliferative disorder), PFBC (neurodegenerative disorder caused by calcification of the basal ganglia and other areas of the brain), and myeloid neoplasm with eosinophilia are excluded from the table as these disorders have distinct phenotypes.

Table 33: Summary of clinical features of PDGFRB activating disorders (excluding vascular)

Disorder	Growth	Face	Myo-fibroma	Eye	Skeletal	Dermatology/ tissue	Brain
Kosaki overgrowth syndrome (KOGS)	Tall	Supra-orbital ridges Wide spaced eyes Wide nasal bridge Maxillary hypoplasia	+	Ocular pterygia	Scoliosis Large hands Contractures Carpal tunnel syndrome Craniosyno-stosis	Hyperelastic thin fragile skin Sparse hair Lipodystrophy Hyper-pigmentation Constriction bands	Periventricular white matter lesions Arachnoid cysts Large ventricles /hydrocephalus
KOGS-like syndrome	-	Supra-orbital ridges Wide spaced eyes Wide nasal base	?	Anterior chamber cleavage	Large hands Dislocations	Soft skin Hyper-pigmentation Wrinkled palms and soles Sparse hair	Cerebellar hypoplasia
Penttinen syndrome of premature ageing	Average-tall	Proptosis Narrow nasal bridge Maxillary hypoplasia Midface retraction Closely or widely spaced eyes	-	Corneal clouding	Progressive brachydactyly Acro-osteolysis Contractures Scoliosis	Thin translucent skin Hypertrophic scar-like lesions Thin sparse hair Lipoatrophy Aged appearance Delayed tooth eruption	Hydrocephalus Arachnoid cyst Periventricul;ar white matter changes
Severe Penttinen-like syndrome	Average-tall	Maxillary hypoplasia Midface retraction Proptosis Narrow nasal bridge	-	Corneal neo-vascular-isation Ocular ptetygia	Progressive brachydactyly Acro-osteolysis Contractures Scoliosis	Thin translucent skin Ulceration Lipoatrophy Hyper- and hypo-pigmentation Aged	Intracranial haemangioma Hydrocephalus

						appeanance Delayed tooth eriuption	
Ocular pterygium-digital keloid dysplasia (OPDKD)	-	Narrow nasal bridge	-	Corneal neo-vascular-isation	Camptodactyly scoliosis, hammertoes	Keloids, hyper-pigmented lesions Truncal fat deposit	nk
Familial infantile myofibromatosis	-	-	+	-	-	-	-
Sporadic multifocal IM Solitary IM	-	-	+	-	-	-	-
Cerebral aneurysms	-	-	-	-	-	-	-
Novel variant reported by Pond et al.	Macro-cephaly	'Coarse'	?	-	Progressive brachydactyly Acro-osteolysis Sagittal craniosynostosis Contractures Carpal tunnel syndrome	Gingival hypertrophy Sparse hair	Retrocerebellar cyst Arachnoid cyst
Novel variant reported by Guimier et al.	Short Macro-cephaly	Proptosis Malar hypoplasia	+	-	Severe scoliosis Limitation of finger flexion Prominent metopic ridge Persistent open fontanelle	Thin skin with cutis marmorata Visible veins Wrinkled palms Sparse hair Umbilical hernia	Massive interhemispheric cysts Bilateral periventricular and basal ganglia calcifications Severe delay
Novel variant reported by Zhong et al.	-	-	+	-	-	Reticulated vascular skin changes Subcutaneous atrophy	-
Novel syndrome mosaic Tyr562Cys	Regional over-growth	-	-	Visual loss	Scoliosis Large hands	Hyperelastic skin Wrinkled skin Sparse hair	-

A degree of phenotypic overlap can be seen between many of these disorders. However, sporadic myofibroma, familial infantile myofibromatosis, and isolated cerebral aneurysms, can be delineated as having single phenotypes (myofibroma, multiple myofibromas, and aneurysm, respectively).

Penttinen syndrome of premature ageing (Val665Ala) and severe Penttinen-like syndrome (Asn666Ser) are distinctive in being associated with a devastating progressive phenotype resulting in severe disability, with progressive shortening of fingers, acro-osteolysis, hypertrophic/keloid scarring, dermal atrophy and ulceration. These individuals also experience severe progressive midface retrusion and have a narrow nasal bridge as opposed to the broad nasal bridge seen in KOGS. The individual reported by Pond et al. (Asn666His) also fits this phenotype of progressive brachydactyly and osteolysis. Another individual with a variant at this amino acid, Asn666Lys reported by Zhong et al., is mosaic and it is not possible to know what the phenotype might be if the variant were germline.

OPDKD (Asn666Tyr) shares the narrow nasal bridge and keloid scarring of Penttinen syndrome, but has a much less severe phenotype. Recent work by Bredrup et al. has suggested this variability in severity is because Asn666Tyr is activating only at temperatures below 37 degrees, affecting only areas of the body with a temperature below this point, as opposed to Asn666Ser which causes continuous activation of PDGFRB at normal body temperature[230]. It has recently been proposed by Wenger et al. that *PDGFRB* activating disorders should be divided into two groups, a less severe group (*PDGFRB* activating spectrum disorder-1; (PAVS1), with individuals with IM, and a more severe group (*PDGFRB* activating spectrum disorder-2; PAVS2) with multi-systemic disease[216]. It is suggested that disorders of KOGS, Penttinen syndrome, and other novel PDGFRB variants would be subsumed into PAVS2, with PAVS2 subdivided into 'PAVS2 with consistent overgrowth' (KOGS Pro584Arg only), 'PAVS2 other' (including KOGS Trp566Arg and Pond et al. Asn666His) and 'PAVS2 with progressive osteolysis'[216]. It would seem reasonable to divide PDGFRB activating disorders

into those with and without multisystem features. However, overgrowth would not appear to be the best discriminating feature for dividing the 'PAVS2' disorders. Some individuals with Penttinen syndrome have tall stature and some individuals with KOGS do not have tall stature; and there are no apparent genotype-phenotype correlations between the two KOGS-associated phenotypes. An alternative nomenclature for activating variants in PDGFRB is proposed in Table 34.

Table 34: Proposed nomenclature for PDGFRB-activating variants

Disorder	Variants	Myofibromas	Diagnostic facial features	Diagnostic skeletal and skin features	Risk of vascular complications
Familial IM	Pro560Leu Arg561Cys (recurrent) Arg561Ser Lys567Glu Pro660Thr	++	-	-	Likely low
Penttinen syndrome	Val665Ala Asn666Ser Asn666His	Unknown	Narrow nasal bridge Progressive midface retrusion	Progressive brachydactyly Acro-osteolysis	Unknown
OPDKD	Asn666Tyr	-	Narrow nasal bridge	Digital keloid scars	Likely low
KOGS	Trp566Arg Pro584Arg Ser493Cys	+	Prominent supraorbital ridges Wide nasal bridge	Absence of brachydactly and acro-osteolysis	High
Mosaic PDGFRB-activating variants	Numerous	Phenotype dependent on variant and affected tissues			Variable; includes high risk variants e.g. Tyr562Cys

In this proposed nomenclature, distinguishing clinical features are used to delineate the disorders. In the case of mosaic variants, the phenotype may vary not only according to the variant but also the degree of mosaicism and the type of tissue affected. As further individuals with variants in *PDGFRB* are identified, this nomenclature will evolve to reflect the increasing understanding of genotype-phenotype correlations. Of note, it is possible that the

Tyr562Cys variant could represent a mosaic KOGS phenotype, given the similarities in the pattern of vascular abnormalities (cerebral and coronary aneurysms) in both of these groups.

4.1.8 *PIK3CA*: PIK3CA-related overgrowth spectrum (PROS)

4.1.8.1 Results POD 013.0

POD 013.0: mosaic *PIK3CA* c.2740G>A p.(Gly914Arg) identified on DNA extracted from fibroblasts cultured from skin biopsy. Variant not present in DNA extracted from blood lymphocytes.

POD 013.0 was born at 40 weeks weighing 3.23kg (-0.4 SD). Age seven she was 123.5cm tall (-0.6 SD) with a weight of 20.8kg (-1.4 SD) and OFC 52.3cm (-0.6 SD). She had asymmetry with non-progressive regional overgrowth of her arm and leg. She had a secondary scoliosis, cutis marmorata on her trunk, back, arms and legs, and a vascular lesion on her lower lip.

4.1.8.2 Discussion POD 013.0

Regional overgrowth and cutaneous vascular malformations are part of the known phenotype of PROS. Although p.(Gly914Arg) is commonly reported with MCAP phenotypes, it has also been described in association with a regional overgrowth phenotype without macrocephaly[241], as in POD 013.0.

4.1.8.3 Results POD 053.0

POD 053.0: mosaic *PIK3CA* c.1357G>A p.(Glu453Lys) identified in DNA extracted from buccal swab. Variant not identified in DNA extracted from blood lymphocytes.

POD 053.0 was born at 39 weeks and six days of gestation weighing 3.8kg (+0.6 SD). Antenatal scans identified a large right kidney. At birth he was noted to have a vascular birthmark on the hand, cheek and foot. Age three his height was -0.4 SD, weight +0.4 SD, and OFC 55cm (+1.9 SD). He had right sided non-progressive regional overgrowth of the right side of his face and arm, including macrodactyly of the right thumb and index finger, and left leg. He was undergoing renal USS screening every three months. He also had pectus excavatum, pes planus, strabismus, and cutis marmorata covering most of his body. He had global developmental delay and walked at 24 months and spoke his first word at 30 months.

4.1.8.4 Discussion POD 053.0

POD 053.0 has the typical PROS features of regional overgrowth and cutaneous vascular malformations. Although not strictly macrocephalic with an OFC of +1.9 SD, he has relative macrocephaly compared to his height of -0.4 SD, demonstrating the overlapping clinical features of MCAP and non-MCAP PROS. Enlarged kidney(s), strabismus, skeletal anomalies, and developmental delay have been reported in PROS[115,242,243].

The recurrent p.(Glu453Lys) variant has been reported in a number of PROS phenotypes including MCAP, regional overgrowth and macrodactyly[241].

4.1.8.5 Results POD 072.0

POD 072.0: mosaic *PIK3CA* c.1093G>A; p.(Glu365Lys) identified in DNA extracted from skin biopsy. Variant not present in DNA extracted from blood lymphocytes or buccal swab.

Germline microdeletion at 15q11.2 (22,765,637-23,217,513)x1. Parental studies are now not routinely performed in 15q11.2 microdeletions as these are frequently inherited and testing is unlikely to clarify clinical significance.

POD 072.0 was born at 40 weeks and nine days gestation weighing 4.65kg (+1.6 SD). His mother was on sertraline during the pregnancy. At birth he required facial oxygen and was admitted to NICU for five days. He had widespread bruising resulting from shoulder dystocia, a capillary vascular malformation on his trunk, and feeding difficulties. Age four, he was 110cm tall (+1.7 SD) with a weight of 22.3 kg (+2.4 SD) and OFC 53 cm (+0.5 SD). He had non-progressive regional overgrowth of the right leg. Previous medical history included asthma, febrile convulsions, seizure-like episodes with normal EEGs, and adrenal haemorrhage (incidental finding on scan). Ongoing medical problems included chronic diarrhoea, tight Achilles tendons, tibial torsion, cutis marmorata on the trunk, lymphoedema of both feet, and conductive hearing impairment with narrow ear canals. He had mild developmental delay. He sat age six months, walked age 12 months, and spoke in two word sentences at 36 months. Behavioural issues included temper tantrums, anxiety, features of autism, short attention span, and pain insensitivity. He attended a special school. His mother was 167.6 cm tall (+0.7 SD) with an OFC of 55 cm (-0.3 SD) and his father was 185.4 cm tall (+1.2 SD) with an OFC of 59.2 cm (+1.3 SD).

4.1.8.6 Discussion POD 072.0

POD 072.0's regional overgrowth (resulting in tight Achilles tendons, lower limb lymphoedema, and tibial torsion) and cutaneous vascular malformations are typical of PROS.

Developmental delay and features of autism are also described[242]. The second diagnosis of a germline 15q11.2 microdeletion may also be contributing to the developmental and behavioural phenotypes in POD 072.0. The 15q11.2 microdeletion is a neurosusceptibility locus associated with an increased risk of developmental delay, congenital malformation of the ear and palate, abnormal brain imaging, behavioural problems, autism spectrum disorder and ADHD[244]. This second diagnosis also explain his conductive hearing loss secondary to narrow ear canals.

The p.(Glu365Lys) variant is previously reported in a patient with an MCAP phenotype[241]. The identification of this variant in POD 072.0, who has a regional overgrowth phenotype without macrocephaly, illustrates the variability in phenotype with variants in *PIK3CA.*

4.1.8.7 Results POD 081.0

POD 081.0: mosaic *PIK3CA* c.2740G>A; p.(Gly914Arg) identified in DNA extracted from skin biopsy. Variant not identified in blood.

POD 081.0 was born at 40 weeks weighing 4.65kg (+2.2 SD) by emergency LSCS. Antenatal scans identified fetal overgrowth. He had a vascular birthmark. Age 29 he was 185 cm tall (+1.2 SD), weighed 82.7kg (+1.5 SD) and had an OFC of 64.2 cm (+4.0 SD). He had asymmetry with regional overgrowth of the face, arm and right leg, and had undergone a lengthening procedure on the left leg. A one-off abdominal USS was performed at the age of 17 months. He had a large right kidney, cutis marmorata, scoliosis, pes planus, crowded teeth, a high palate and myopia. He had speech and language delay, delayed social development, and features of autism. He attended a mainstream school with assistance. In adult life he lived

with family members and required assistance with activities of daily living. He was employed and had no children. His mother was 170 cm tall (+1.1 SD) with an OFC of 55 cm (-0.4 SD) and his father was 178 cm tall (+0.1 SD).

4.1.8.8 Discussion POD 081.0

POD 081.0's clinical features of macrocephaly, cutis marmorata, and regional overgrowth fall into the MCAP syndrome group of PROS[245,246]. His delayed speech and language development and intellectual disability are common in MCAP, occurring in 65% and 52% respectively of a recent series of 33 individuals with MCAP[242]. Features of autism have been described in a minority of individuals[242,247]. Enlarged kidneys and skeletal features are well described in CLOVES phenotypes but can also be seen in individuals with a clinical diagnosis of MCAP, demonstrating the overlapping features of the *PIK3CA* spectrum disorders[242]. The p.(Gly914Arg) variant present in POD 081.0 is known to be associated with MCAP and other PROS disorders[112].

4.1.8.9 Summary of participants with PROS

Four participants in the study had confirmed diagnoses of PROS. In all cases the variant was mosaic and identified on testing a tissue sample (cheek cells from buccal smears in one participant and fibroblasts from skin biopsy in three participants). All had regional overgrowth and three had associated vascular skin changes. One participant had macrocephaly and one had relative macrocephaly. Two individuals had skeletal involvement, one with pectus excavatum and pes planus and the other with scoliosis and pes planus. Two out of the three child participants had developmental delay and the adult participant had

intellectual disability. All of these features are within the known clinical spectrum of PROS. Overall the clinical features of the participants confirms that PROS is a spectrum, and not all individuals can be divided into the diagnostic categories of MCAP, CLOVES etc. The number of participants with PROS is this study is small but would corroborate the view that phenotypes are dependent on timing and location of the mosaic pathogenic variant, rather than the precise nature of the missense variant. A summary of participants is shown in Table 35.

Table 35: Summary of participants with PROS

Participant	POD 072.0	POD 053.0	POD 013.0	POD 081.0
Sex	M	M	F	M
Age	4	3	7	29
Variant	Mosaic c.1093G>A; p.(Glu365Lys)	Mosaic c.1357G>A p.(Glu453Lys)	Mosaic c.2740G>A p.(Gly914Arg)	Mosaic c.2740G>A p.(Gly914Arg)
Tissue variant identified in	skin	buccal	skin	skin
Height (SD)	+1.7	-0.4	-0.6	+1.2
OFC (SD)	+0.5	+1.9	-0.6	+4.0
Development/ID	delay	delay	normal	ID
Phenotypes	Overgrowth R lower limb Cutis marmorata Cutaneous vascular malformation Behavioural issues	Relative macrocephaly Overgrowth R face, upper limb and lower limb Cutis marmorata Pectus excavatum Pes planus	Overgrowth R upper limb and lower limb	Macrocephaly Overgrowth R face, upper limb, lower limb Cutis marmorata Scoliosis Pes planus

4.1.9 *PPP2R5D* -related neurodevelopmental disorder

4.1.9.1 Results POD 030.0

POD 030.0: *PPP2R5D* c.598G>A; p.(Glu200Lys). Parental samples not available to confirm de novo status.

POD 030.0 was born at 40 weeks and 12 days by emergency LSCS weighing 4.68 kg (+1.4 SD) with an OFC of 39.5 cm (+2.6 SD). He had neonatal hypoglycaemia. Age seven he was 127.7 cm tall (+0.8 SD) with a weight of 27.4 kg (+1.0 SD) and OFC 59.0 cm (+3.4 SD). He had mild global developmental delay. He sat at six months, walked at 12 months and spoke his first word at 24 months. MRI head scans initially showed an increase in pericerebral cerebrospinal fluid (CSF) however this normalised with age. Behavioural issues included short attention span, hyperacusis, and sensory processing difficulties. He attended mainstream school with assistance. He had a prominent forehead, widely spaced eyes, small hands and short feet. His mother's OFC was 58 cm (+1.8 SD) and his father's OFC was 61 cm (+2.4 SD).

4.1.9.2 Discussion POD 030.0

PPP2R5D-related neurodevelopmental disorder has been described in 23 individuals to date[185,248–250] with developmental delay or intellectual disability present in all reported cases. Hypotonia is common and can be pronounced. Motor milestones are delayed with age at first walking being reported between 18 months and nine years. Some individuals have an ataxic gait[185,249]. Speech delay is a universal feature with some individuals being nonverbal[250]. Behavioural issues and autism spectrum disorder have also been described in a few individuals[249,250]. POD 030.0's pattern of development is of interest because although he has mild developmental delay, he did not present with hypotonia and achieved his early motor milestones on time. His speech delay was also milder than in the previously reported individuals. This suggests that developmental delay and intellectual disability is milder in some individuals with *PPP2R5D*-related neurodevelopmental disorder than previously recognised.

Macrocephaly with OFC ranging from +2.0 to +3.8 SD has been reported in most individuals. One individual was reported to be macrocephalic at birth however detailed information on head circumference at birth has not been published. Only two individuals with height > 2 SD have been reported. POD 030.0's growth follows this pattern of pronounced macrocephaly with a SD of +3.4 but height being in the normal range.

Medical issues that have been reported include seizures, minor anomalies on MRI brain, and ophthalmic problems such as myopia, astigmatism, strabismus, nystagmus, ptosis and cataracts[185,249,250]. Less commonly, skeletal anomalies including camptodactyly, scoliosis, and hip dysplasia have been described[184,185,249]. Two individuals have been reported to have congenital heart disease. One individual as reported to have hypoglycaemia[185]. POD 030.0 had neonatal hypoglycaemia and it is possible that this may be a rare feature of *PPP2R5D*-related neurodevelopmental disorder. However, larger studies need to be performed to ascertain if this is the case.

Facial features of *PPP2R5D*-related neurodevelopmental disorder are variable and nonspecific however typical individuals have been described as having a long face, frontal bossing, widely spaced eyes and down slanted palpebral fissures. POD 030.0 shares the characteristic facial appearance of frontal bossing and wide spaced eyes. Small hands and feet have not previously been reported and again further detailed phenotyping studies of a larger number of individuals are needed to establish if this is a feature of *PPP2R5D*-related neurodevelopmental disorder.

4.1.10 *PTEN:* PTEN-hamartoma tumour syndrome (PHTS)

Two families with inherited variants in PTEN participated in the study. Participants with PHTS were identified on diagnostic testing prior to recruitment to the study.

Family 1: POD 087.0, POD 089.0, and POD 089.1: pathogenic duplication of exon 5 of *PTEN* (mosaic in POD 087.0)

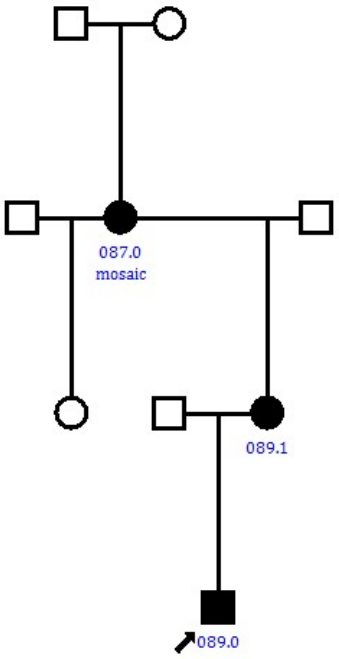


Figure 57: Pedigree POD 087.0, 089.0, and 089.1

4.1.10.1 Results POD 089.0

Participant 089.0 was born at 39 weeks of gestation weighing 3.56 kg (+0.5 SD) and had congenital anomalies including anal stenosis, anorectal malformation, chordee, hypospadias, a vesical fistula and nevus flammeus. His mother was on sodium valproate until seven weeks gestation then amitriptyline for the rest of the pregnancy. Age 13 he had an OFC of 54 cm (+2.6 SD), height 88 cm (-0.3 SD), and weight +1.3 SD. He entered puberty age 12. He was diagnosed with juvenile polyposis with >1000 bowel polyps (ganglioneuromas). He was also found to have a left frontal cystic lesion (probable glial cyst) on MRI head and thoracic hemivertebrae. He had delayed speech and language and social development and a diagnosis of autism spectrum disorder. Dysmorphic features included a prominent forehead, short ear, and preaxial polydactyly of both hands.

Investigation of bowel polyposis with targeted sequencing and dosage analysis of a panel of polyposis genes identified POD 089.0 was heterozygous for a pathogenic duplication of exon 5 of *PTEN*.

4.1.10.2 Discussion POD 089.0

POD 089.0 has several features consistent with the reported phenotype of PHTS in childhood, including macrocephaly, developmental delay and autism spectrum disorder. Macrocephaly is almost universal in PHTS[251].

Polyposis is also one of the most common features of PHTS, with 62/67 (95%) of individuals who underwent endoscopy being found to have polyps in one series[252]. Although hamartomas are the most commonly described lesion, 16/62 individuals had ganglioneuromatous polyps, with number of polyps ranging from 1 to 'carpeting'[252]. The age of diagnosis was not reported for these individuals, however there are other reports in the literature of diffuse intestinal gangliomatosis occurring in PTHS in children[253,254]. White matter cysts in the frontal lobe and other areas of the brain have also previously been reported in PHTS[255]. The juvenile polyposis and frontal cystic lesion in POD 089.0 are therefore consistent with the phenotypic spectrum of this disorder.

Polydactyly has been very rarely reported in patients with PHTS, with one other example of preaxial polydactly[256] and two patients with postaxial polydactly[257], but whether this is a true association is unknown. The other congenital anomalies in POD 089.0, anal stenosis, anorectal malformation, chordee, hypospadias, vesical fistula, and thoracic hemivertebrae, are also not known to be features of PTHS. POD 089.0 underwent array CGH and whole genome sequencing in the 100KGP and these investigations did not identify a second diagnosis. POD 089.0 could be described as having a VATER association (vertebral defects, anal atresia, tracheoesophageal fistula with oesophageal atresia, and radial and renal dysplasia)[258], as he has three features of this association (vertebral defect, anal anomaly, and radial dysplasia in the form of preaxial polydactyly). Genitourinary anomalies are also common in VATER, occurring in 30% of a recent series of 36 patients[259]. Interestingly there is a report in the literature of a child with PHTS who also had features of VATER with tracheo-oesophageal fistula and bilateral radial hand anomalies[260]. The aetiology of VATER is not yet fully understood and is likely to be extremely heterogeneous, including many different monogenic

disorders, epigenetic variants, non-genetic factors, and multifactorial explanations[261]. The maternal history of anticonvulsants in pregnancy with POD 089.0 may be highly relevant in explaining his complex phenotype, with valproate embryopathy associated with significantly increased risks of congenital anomalies including hypospadias and polydactyly[262].

4.1.10.3 Results POD 089.1

POD 089.1 was age 36 at entry to the study. She was 170.0 cm tall (+1.0 SD), weighed 215 kg (+6.4 SD) and her OFC was 64.5 cm (+6.5 SD). Medical problems included type 2 diabetes, multinodular thyroid goitre, asthma, sleep apnoea, hepatic haemangiomas, uterine fibroids, benign neoplasm of the breast and bipolar disorder. She also had skin tags and oral papillomatous papules. She had attended mainstream school and achieved GCSEs. She was undergoing annual thyroid, renal and breast screening commenced age 36 and had undergone one-off bowel screening with further screening planned for age 55.

4.1.10.4 Discussion POD 089.1

POD 089.1 has many features of PHTS, including one pathognomonic (multiple oral papillomatous lesions), one major (macrocephaly) and one minor (multinodular goitre)[263]. Her extreme macrocephaly +6.5 SD is not atypical, as the average head circumference is very large in PHTS (+5 SD in childhood) and generally remain large in adulthood[251,264]. Benign breast disease and uterine fibroids are also commonly seen in women in PHTS, but some authorities have removed these features from the minor diagnostic criteria because of frequency of these conditions in the general population[265]. Vascular anomalies including

haemangiomas are seen in PHTS, and hepatic haemangiomas are a rarely reported feature[266,267]. Bipolar disorder has described very rarely in individuals with PHTS[264,268].

The morbid obesity in POD 089.1 (BMI 74.3) is not a commonly reported feature of PHTS and the prevalence of obesity in the PHTS population compared to the general population is unknown. There may be genetic and environmental factors (such as medication for bipolar disorder) that are contributing to her obesity. However, there is a case report of another patient with morbid obesity (BMI 57.6 age 34) who was diagnosed with PHTS through gene panel testing at a bariatric surgery clinic[269]. A small study has suggested that individuals with pathogenic variants in *PTEN* may have a higher risk of obesity, yet a lower risk of type 2 diabetes (because of increased sensitivity to insulin), than individuals without PHTS[270]. It is proposed this is mediated by haploinsufficiency of PTEN increasing the activity of the PI3K/AKT pathway[270]. This is consistent with the finding that in the converse situation, germline variants in *AKT2* that reduce the activity of the PI3K/AKT pathway cause a phenotype of severe insulin resistance and partial lipodystrophy[271]. In addition, a mouse model with increased expression of PTEN has a reduced body size and reduced body fat[272].

It is possible that POD 089.1's morbid obesity may be related to the diagnosis of PHTS, although the presence of type 2 diabetes is not consistent with the proposed picture of 'obesity with increased insulin sensitivity'. Further study of obesity and insulin sensitivity in patients with PHTS is needed.

4.1.10.5 Results POD 087.0

POD 087.0 was 55 years old at recruitment. She had an OFC 59.0 cm (+2.5 SD), height 157.0 cm (-1.1 SD), and weight 98.0 kg (+3.2 SD). Dermatological features included acral keratoses and papillomatous papules. She had two renal cell carcinomas diagnosed age 52 treated with left nephrectomy, a clear cell renal cell carcinoma Fuhrman grade 3 and a chromophobe renal cell carcinoma Fuhrman grade 2. She had also undergone a subtotal thyroidectomy. She had attended mainstream school. In adult life she lived independently, was employed and had children.

4.1.10.6 Discussion POD 087.0

POD 087.0 has the pathognomonic PHTS features of acral keratoses and papillomatous lesions[263] in addition to the major feature of macrocephaly. The reason for her subtotal thyroidectomy is unknown, but both benign and malignant thyroid disease are common in PHTS[39,273]. Renal cell carcinoma (RCC) is a known association of PHTS with estimates of lifetime risks between 2-34%[39,273,274]. Papillary RCC is the most common histology type in PHTS but chromophobe and clear cell RCCs have also been described[275]. The increased risk of RCC in PTHS is thought to start in the late 40s[276], consistent with the age of diagnosis of 52 in POD 087.0. Metachronous tumours is relatively rare in PHTS. Out of 219 individuals with PTHS in one series, nine had a history of RCC and only one had metachronous tumours[275].

POD 087.0 was also obese with a BMI of 39.7.

4.1.10.7 Discussion Family 1

These participants demonstrate the range of clinical problems that can occur in PHTS even within the same family, from autism and developmental delay, to dermatological and other benign lesions, and malignancy. Some of this variability may represent age-related penetrance however the youngest member of the family has significant medical problems, notably juvenile polyposis, not present in his mother or grandmother.

Uncertainty remains about whether the phenotypic features of congenital anomalies in POD 089.0 and obesity in POD 089.1 and 087.0 could be related to the diagnosis of PTHS.

4.1.10.8 Results POD 088.0

Family 2: POD 088.0 and 088.2: *PTEN* c.469G>T; p.(Glu157Ter)

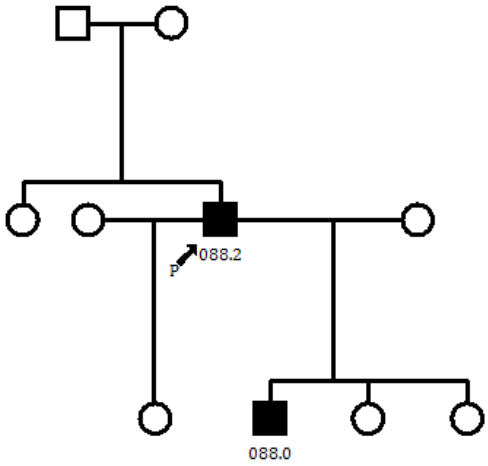


Figure 58: Pedigree POD 088.0 and 088.2

POD 088.0 was born by normal vaginal delivery at 36 weeks gestation weighing 3.57 kg (+2.2 SD) following an uneventful pregnancy. He had neonatal jaundice requiring phototherapy.

At recruitment POD 088.0 was 13 years old with a height of 166.7 cm (+1.2 SD) and weight 70.6 kg (+2.2 SD). His head circumference was previously measured at +2.5 SD.

He had bilateral hypoechoic thyroid nodules, tracheomalacia and laryngomalacia, obstructive sleep apnoea, recurrent upper respiratory tract infection, recurrent lower respiratory tract

infections, adenotonsillectomy, adenoidectomy and supraglottoplasty, hypotonia, hypermobility, juvenile idiopathic arthritis, nodular prurigo, eczema, and gynaecomastia.

POD 088.0 had global developmental delay. He walked at 18 months and said his first word age 24 months with two word sentences at 48 months. He had a diagnosis of autism spectrum disorder and behavioural features of aggression, temper tantrums, emotional lability, hyperactivity, short attention span, self-injurious behaviour, and poor sleep.

4.1.10.9 Discussion POD 088.0

POD 088.0 has several features consistent with the diagnosis of PHTS with a BRRS presentation, including macrocephaly, hypotonia, hypermobility, motor and speech delay, and autism[36,277–280]. Thyroid disease can also occur in children in PHTS, with thyroid nodules being described in children as young as five[281]. Enlargement of the tonsillar tissue and sleep apnoea responding to adenotonsillectomy has been reported[282,283]. Rarely, autoimmune conditions have been reported in association with PHTS[282], although not juvenile idiopathic arthritis as in POD 088.0.

Autism spectrum disorder is a well-recognised feature of PTHS but the behavioural features of aggression, hyperactivity and self-injurious behaviour displayed by POD 088.0 are not commonly reported. However it is increasingly recognised that PHTS can be associated with other neurodevelopmental disorders such as ADHD and there are also reports of oppositional defiant disorder and disruptive behaviour disorder[264].

Tracheomalacia and laryngomalacia requiring surgical treatment with supraglottoplasty does not appear to have been reported in PHTS previously. Gynaecomastia, nodular prurigo, and eczema are not known to be associated with PHTS.

4.1.10.10 Results POD 088.1

POD 088.2, the father of 088.0 had the familial pathogenic variant in *PTEN* c.469G>T p.(Glu157Ter) identified on a cancer panel. Parental samples were unavailable to confirm de novo status.

His height was 180cm (+0.4 SD) and head circumference 60cm (+1.6 SD). He had benign thyroid nodules managed with thyroidectomy, multiple colonic adenomas, an arteriovenous malformation (AVM) on the dorsum of the right foot, gingival hypertrophy, depression, psychosis and alcohol dependency.

4.1.10.11 Discussion POD 088.1

Unusually, POD 088.2 does not have absolute macrocephaly but his head circumference is relatively increased compared to his height.

PHTS is known to be associated with benign thyroid disease; polyposis, including adenomas;[284] and vascular malformations[267]. Gingival hypertrophy has also been reported in a small number of cases[285–287].

4.1.10.12 Summary of participants with PTHS

Psychosis has been described in one previous individual with PHTS[268]. The development of psychiatric disease in two families in this study, psychosis in POD 088.2 and bipolar disorder in POD 089.1, adds to the evidence that a range of psychiatric phenotypes are associated with PHTS[268].

4.1.11 *SUZ12:* Imagawa-Matsumoto syndrome (SUZ12-related overgrowth syndrome)

POD 103.0, POD 103.1, POD 103.3, POD 104.0, POD 104.1, POD 104.3, POD 104.4: 17q11.2 microdeletion (30,318,418-30,326,952) including exons 11-16 of *SUZ12*.

Seven individuals in this family have microdeletions of approximately 8.5 kb at 17q11.2 including exons 11-16 of *SUZ12*. To date, there are no reports of individuals with intragenic deletions of *SUZ12* in the literature and this is not a previously recognised molecular mechanism for Imagawa-Masumoto syndrome. However, there are several lines of evidence that indicate haploinsufficiency of *SUZ12* is the likely disease mechanism in this disorder.

Truncating (nonsense, frameshift and splice site) variants in *SUZ12* associated with Imagawa-Matsumoto syndrome have been shown to cause loss of PRC2 enzyme activity[82]. Further supportive evidence of *SUZ12* deletion being associated with this overgrowth disorder is provided by the phenotype of NF1 deletion patients with a recurrent ~1.4 Mb microdeletion that encompasses *SUZ12*. These individuals have a phenotype significantly different to those with NF1 point mutations and are much more likely to have overgrowth, greater dysmorphism, and intellectual disability. This is the case even for individuals whose deletion does not include another candidate overgrowth gene in this region, *RNF135*, and it is thought that haploinsufficiency of *SUZ12* is the most likely cause of this extended NF1 phenotype[288].

The intragenic deletion of SUZ12 identified in this family resulting in haploinsufficiency is therefore highly likely to cause the overgrowth disorder Imagawa-Matsumoto syndrome in these individuals.

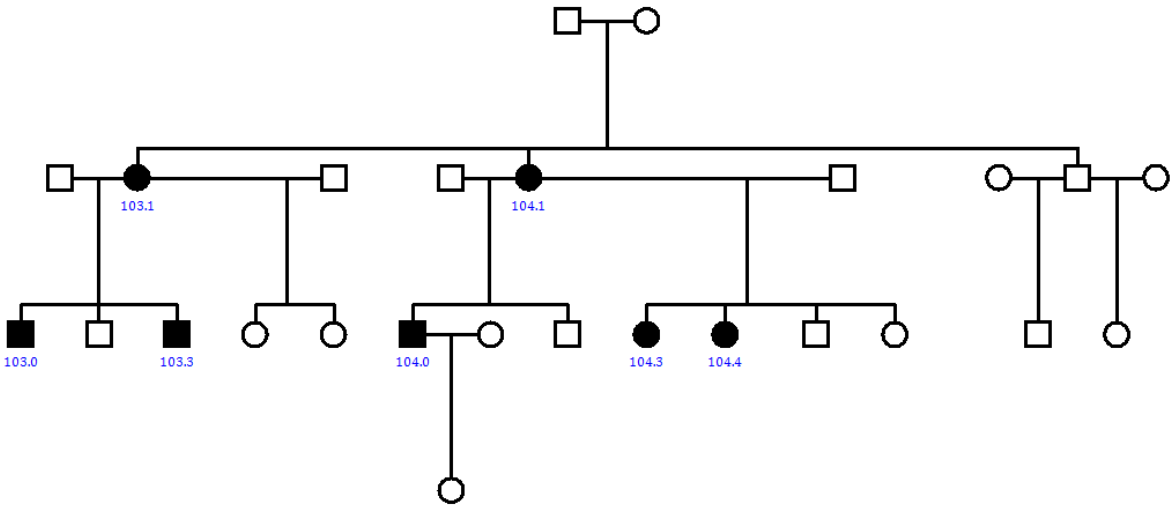


Figure 59: Pedigree of family with seven individuals with deletions of SUZ

4.1.11.1 Results POD 103.0

POD 103.0 was born at 40 weeks gestation and was noted to have an umbilical hernia. Age 11 he was 162.7 cm tall (+2.1 SD), weighed 53.6 kg (+1.7 SD) and had an OFC of 55.5 cm (+0.2 SD). Medical problems included constipation and hypermetropia. He had global developmental delay and walked at 15 months. He had behaviour issues including aggression, emotional lability, and temper tantrums. He attended a special school. Dysmorphic features included a flat occiput, epicanthic folds, almond-shaped palpebral fissures, crowded teeth, long ears, and bilateral 5th finger clinodactyly.

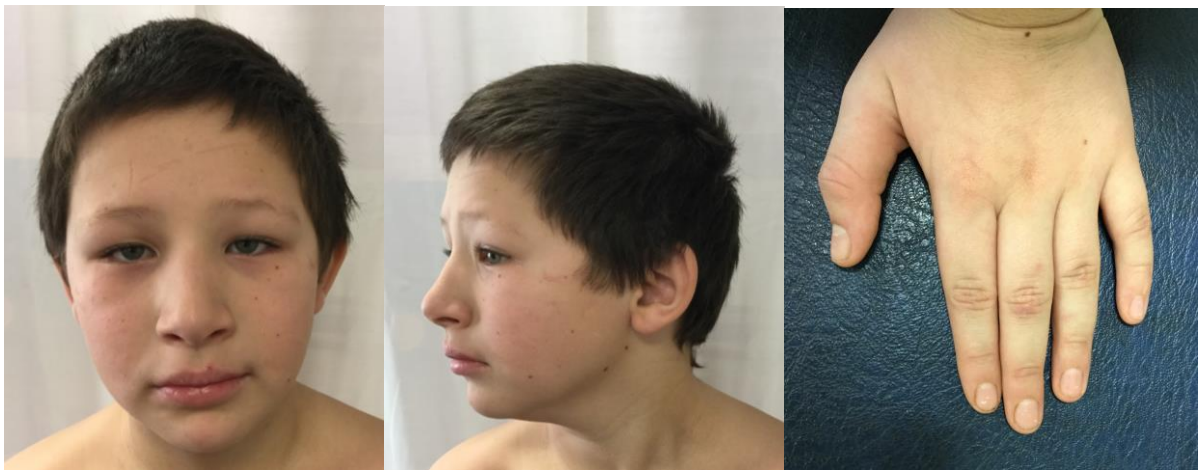


Figure 60: POD 103.0 - clinical photographs of face and hands

4.1.11.2 Results POD 103.1

POD 103.1 was 172.2 cm tall (+1.4 SD) and weighed 135.6 kg with an OFC of 55.5 cm (0 SD) at age 39. She had grommets and a tonsillectomy in childhood. Other medical problems included Gilbert's syndrome, a cholecystectomy, osteoarthritis of the knees and myopia. She had attended a mainstream school and in adult life lived independently and had three children. Dysmorphic features included a sloping forehead, long ears, and long fingers.

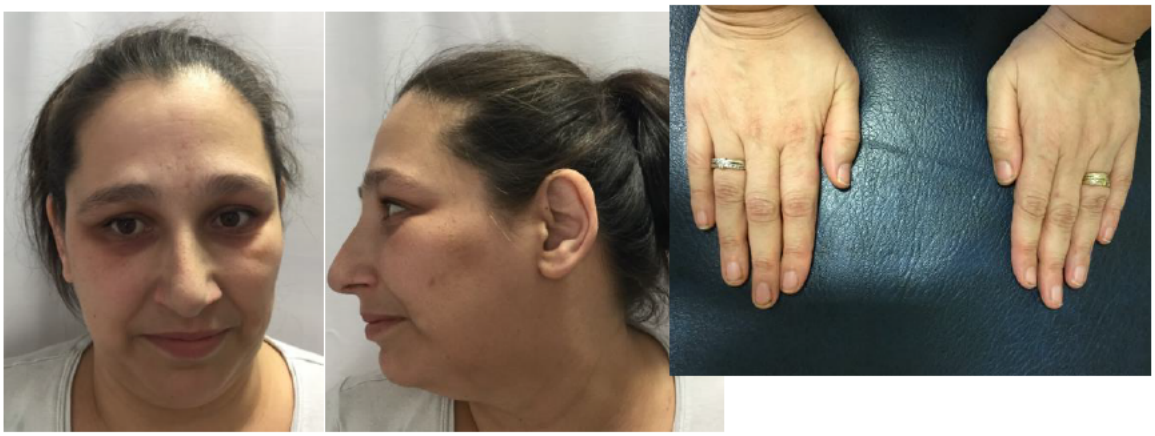
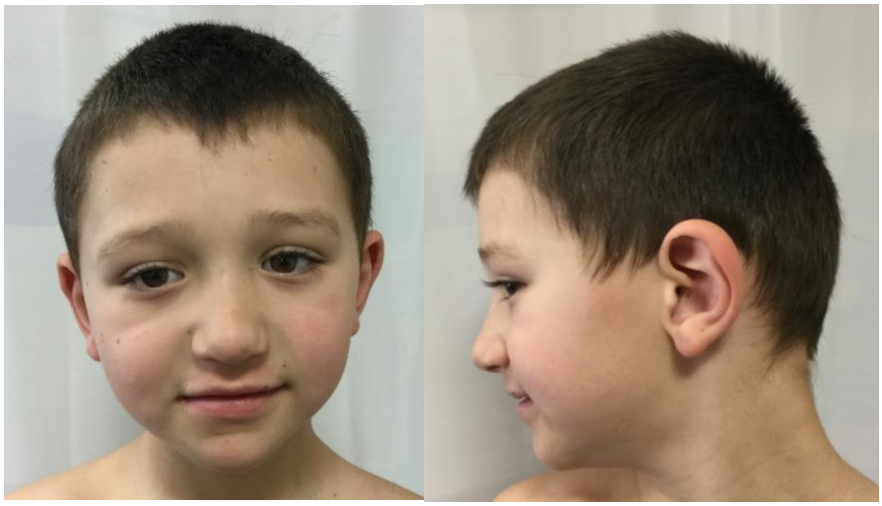
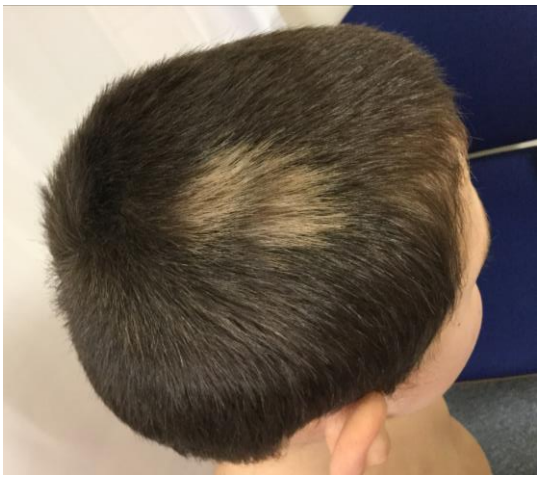


Figure 61: POD 103.1- Clinical photographs of face and hands

4.1.11.3 Results POD 103.3

POD103.3 was born at 38 weeks gestation weighing 3.5 kg (+0.9 SD) and required treatment for neonatal jaundice with phototherapy. Age six he was 131.6 cm tall (+2.4 SD) with a weight of 27.1 kg (+1.5 SD) and an OFC of 50.6 cm (-1.8 SD). He had myopia. He had global developmental delay and attended a mainstream school with a statement. He walked at 11 months and said his first word at 18 months. He was noted to have temper tantrums. He had a flat occiput, a birthmark on his scalp with hypopigmentation of the overlying hair, round face, horizontal crease in his chin, widely spaced eyes, long ears and long fingers.





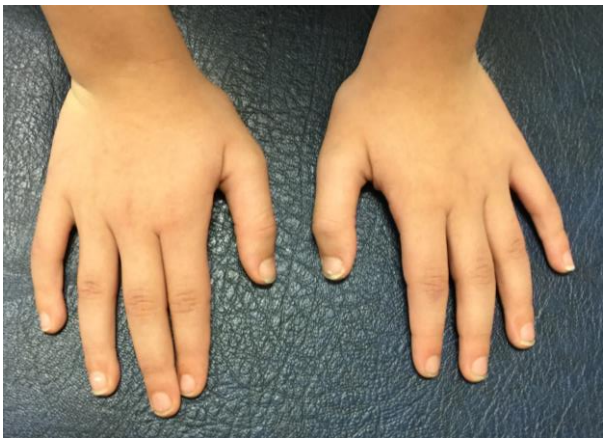


Figure 62: POD 103.3 - Clinical photographs of face, scalp hair, and hands

4.1.11.4 Results POD 104.0

POD 104.0 was born at 40 weeks and 14 days gestation with a birthweight of 3.6 kg (-0.8 SD). Age 25 he was 194.5 cm tall (+2.4 SD), his weight was 125.7 kg and his OFC 58 cm (+0.4 SD). He had neurological symptoms of pins and needles in his right hand and pes planus. He had extra help at school and achieved GCSEs at grades E and below. In adult life he lived independently, was employed and had children. Facial features included prominent nasal bridge, thick vermilion of the upper and lower lips, and dental crowding. He also had long broad palms, pes planus, and clinodactyly of the halluces.

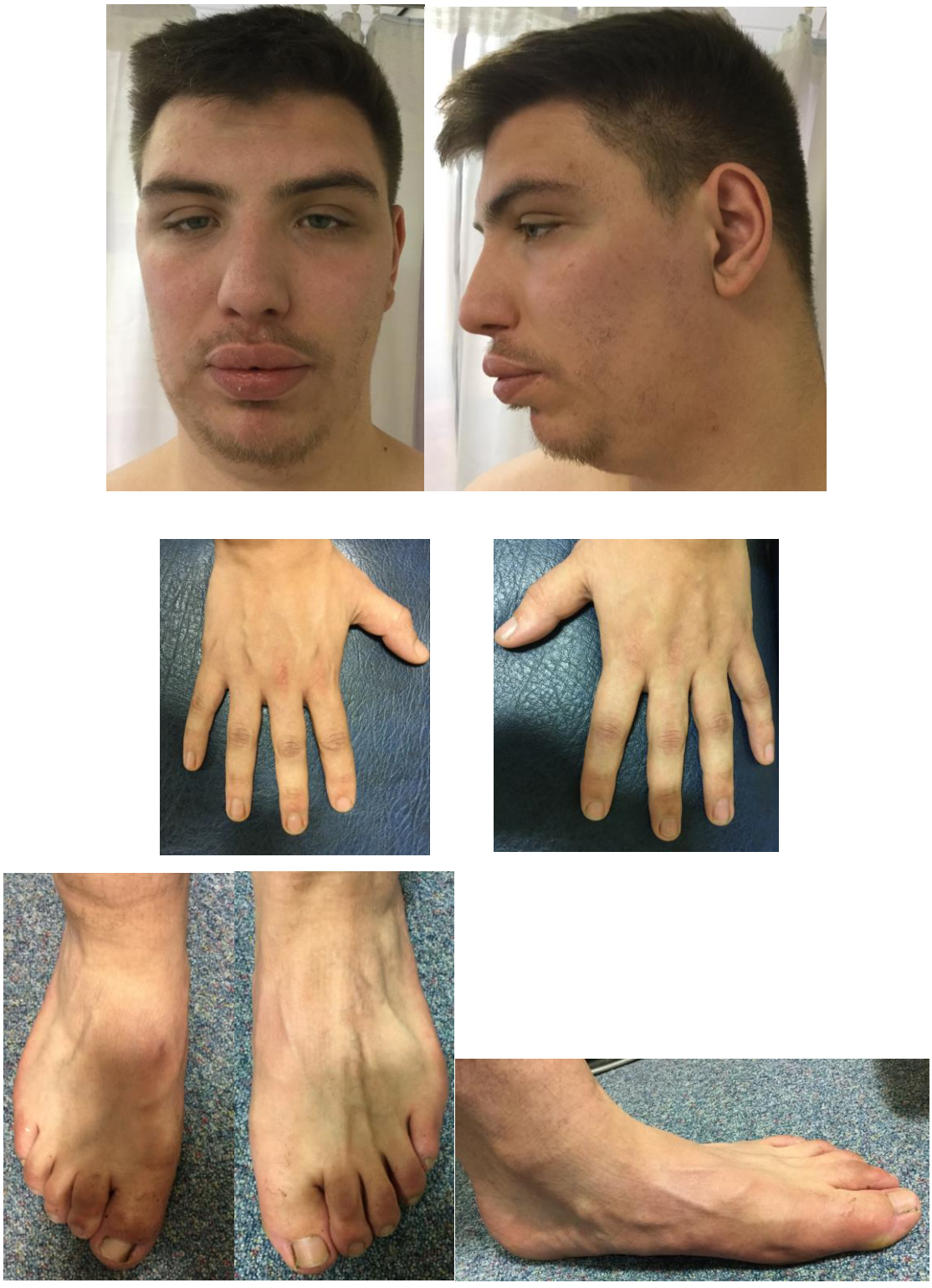
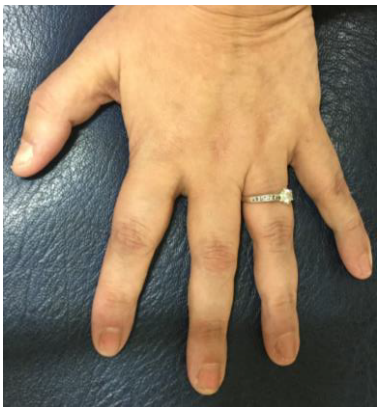


Figure 63: POD 104.0 – Clinical photographs of face, hands and feet

4.1.11.5 Results POD 104.1

POD104.1 was age 45 at recruitment and had a height of 174 cm (+1.7 cm), weight 131.1 kg (+4.5 SD) and OFC 55.1 cm (-0.3 SD). She developed hypothyroidism during pregnancy and type 2 diabetes following pregnancy. She also had hypermetropia, sciatica, and non-pitting oedema of her right leg. She had speech delay in childhood and had attended mainstream school but did not achieve any GCSEs. In adult life she was employed and had children. Dental overcrowding was noted on examination. She had a short second toe on the right foot.



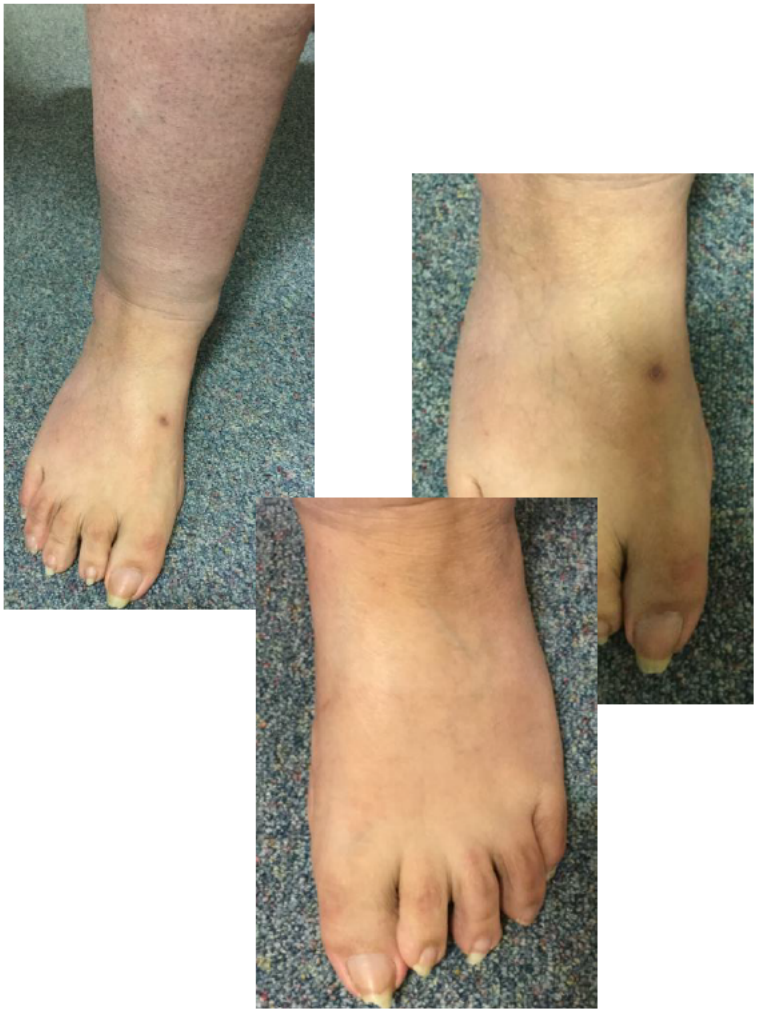


Figure 64: POD 104.1 - Clinical photographs of face, hands, right leg and feet

4.1.11.6 Results POD 104.3

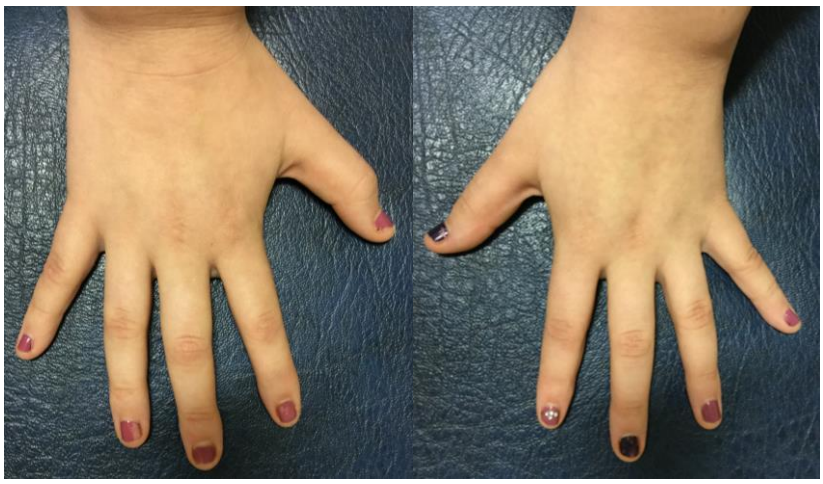
POD104.3 was born at 40 weeks gestation weighing 3.7 kg (+0.6 SD). Age 12 she was 172.8 cm tall (+2.8 SD), weighed 68.7 kg (+2.4 SD) and had an OFC of 56.4 cm (+1.5 SD). She had a medical history of two episodes of collapse of unknown cause and did not tolerate an EEG to investigate possible seizures. She had global developmental delay and attended a special school. Behavioural issues included aggression, emotional lability, temper tantrums, hyperacusis and sleep difficulties. Her facial features included widely spaced eyes, horizontal

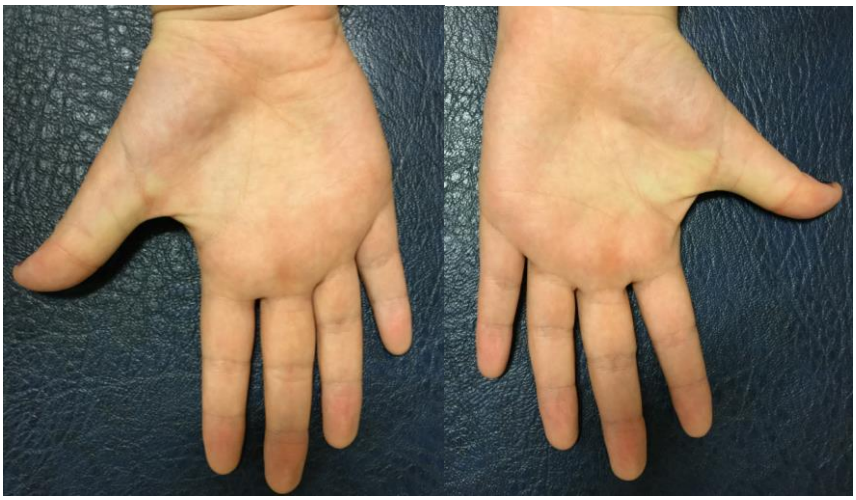
and thick eyebrows, a wide nasal base, and widely spaced teeth. She also had long palms and camptodactyly of the toes on her right foot.

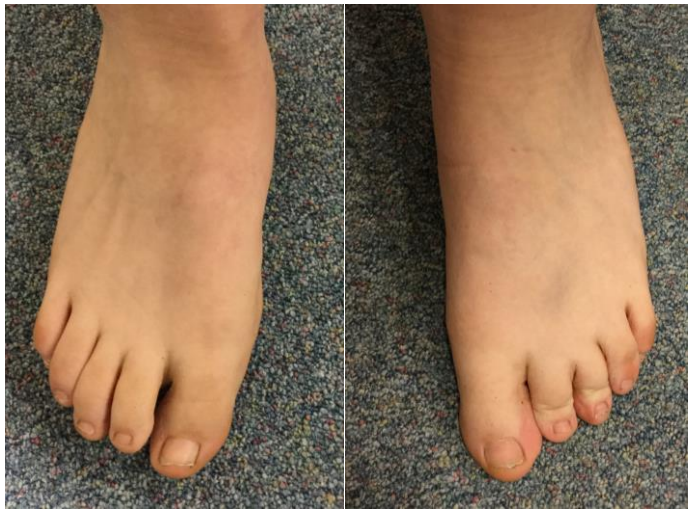
Figure 65: POD 104.3 - Clinical photographs of face, hands and feet

4.1.11.7 Results POD 104.4

POD104.4 was born at 40 weeks gestation weighing 4.9 kg (+3.0 SD) and was noted to have a brown birthmark on her left calf. Age 14 she was 181.4 cm tall (+3.3 SD), weighed 96.3 kg (+3.0 SD) and her OFC was 56.7 cm (+1.3 SD). Medical issues included vitamin D deficiency and astigmatism. She had global developmental delay and first walked at 28 months. She attended a mainstream primary school with a statement and moved to a special school for secondary education. Facial features included a round face and long ears. She also had absent palmar creases on both hands.







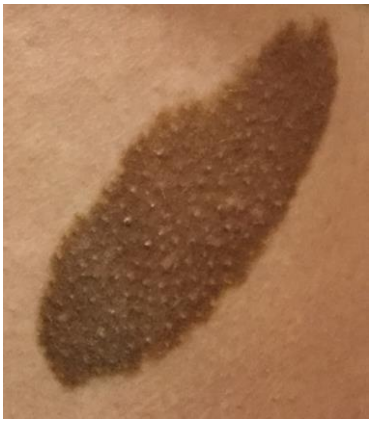


Figure 66: POD 104.4 – clinical photographs of face, hands, feet, and birthmark left calf

4.1.11.8 Expanding the phenotype of Imagawa-Matsumoto syndrome

Only 13 individuals with pathogenic variants in *SUZ12* have been described[81,289–292] in the literature. An additional two individuals on the DECIPHER database have sequence variants in SUZ12 that are classified as likely pathogenic. The identification of this family with seven affected individuals therefore substantially increases the number of individuals known to have this disorder. The identification of a group of individuals with a deletion of *SUZ12* also provides the opportunity for phenotypic comparison between missense and truncating (frameshift, nonsense, splice-site and deletion) variants.

4.1.11.8.1 Birth and neonatal history

There were no characteristic pregnancy or neonatal complications in this family, in common with the previously reported patients[290]. Birthweight was only available for three individuals. One individual was large for gestational age with a birthweight of 4.9 kg (+3.0 SD) but the other two were well within the normal range at 3.6 kg (-0.8 SD) and 3.7 kg (+0.6 SD). Birth length and OFC were not available for any family members.

4.1.11.8.2 Height

Height at recruitment to the study ranged from +1.4 SD to +3.3 SD with a mean of +2.3 SD and median of +2.4 SD. There was no relationship between age and height, indicating that final height is similarly increased in children and adults.

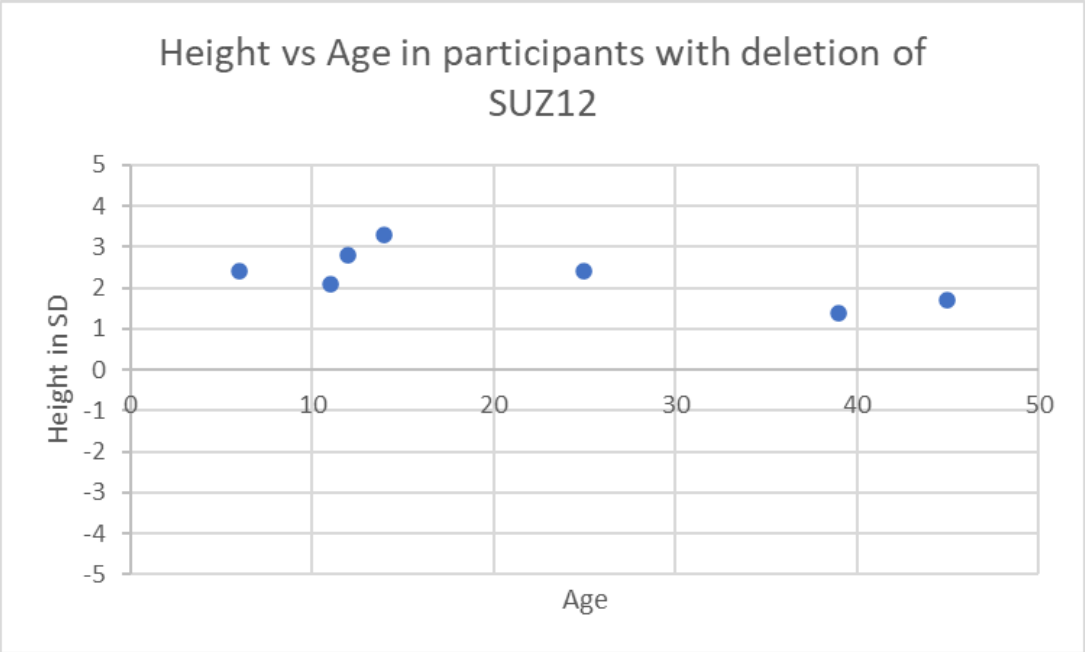


Figure 67: Height vs Age in participants with deletion of SUZ12

The heights in this family are comparable to those previously published for nine children (range -1.0 SD to +4.3 SD with a mean of +2.0 SD) and four adults (+1.3 SD to +5.9 SD with a mean of +3.8 SD)[82,293]. Reviewing data from the seven POD participants, previously reported individuals, and patients on DECIPHER, the reported range of heights is from -1.2 SD to +5.9 SD with a mean of +2.1 SD. The data from this family confirms that not all individuals have a height greater than two standard deviations above the mean and tall stature is not a universal feature of this disorder.

4.1.11.8.3 Weight

Weight ranged from +1.5 SD to +4.5 SD with a mean of +3.1 SD and median +3.0 SD.

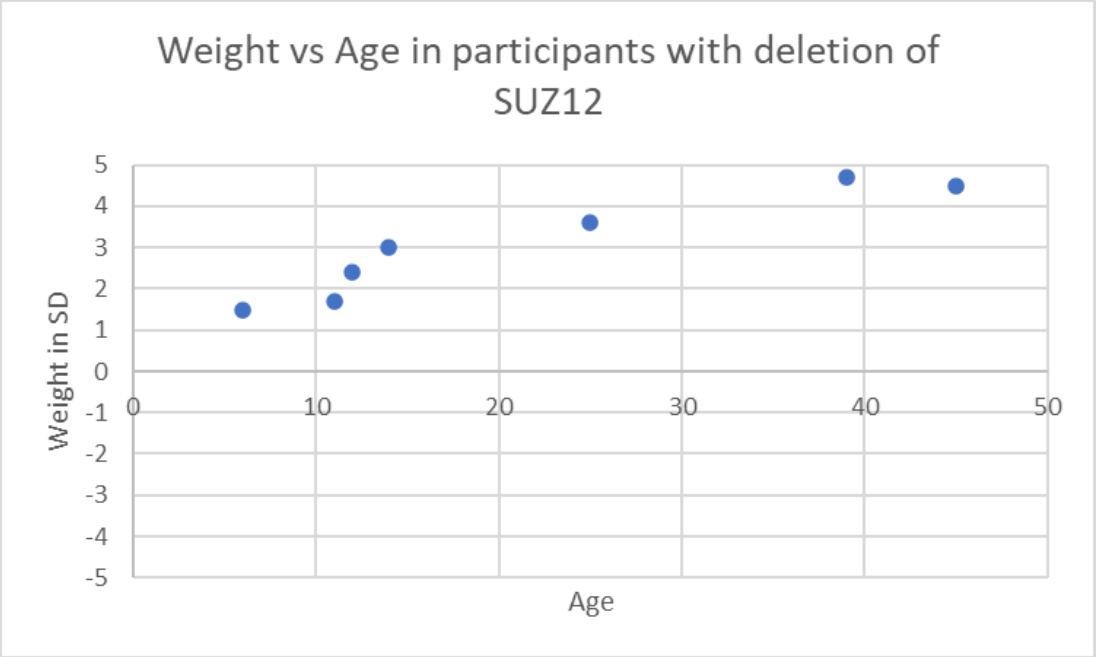


Figure 68: Weight vs Age in participants with deletion of SUZ12

Weight in all known individuals with Imagawa-Matsumoto syndrome (seven POD participants, previously reported individuals, and patients on DECIPHER) ranges from -0.6 SD to +5.7 SD with a mean of +2.3 SD.

4.1.11.8.4 Head circumference

Macrocephaly was not a feature in this family, with OFCs measuring in the normal range from -1.8 SD to +1.5 SD with a mean of +0.2 SD and median of +0.2 SD.

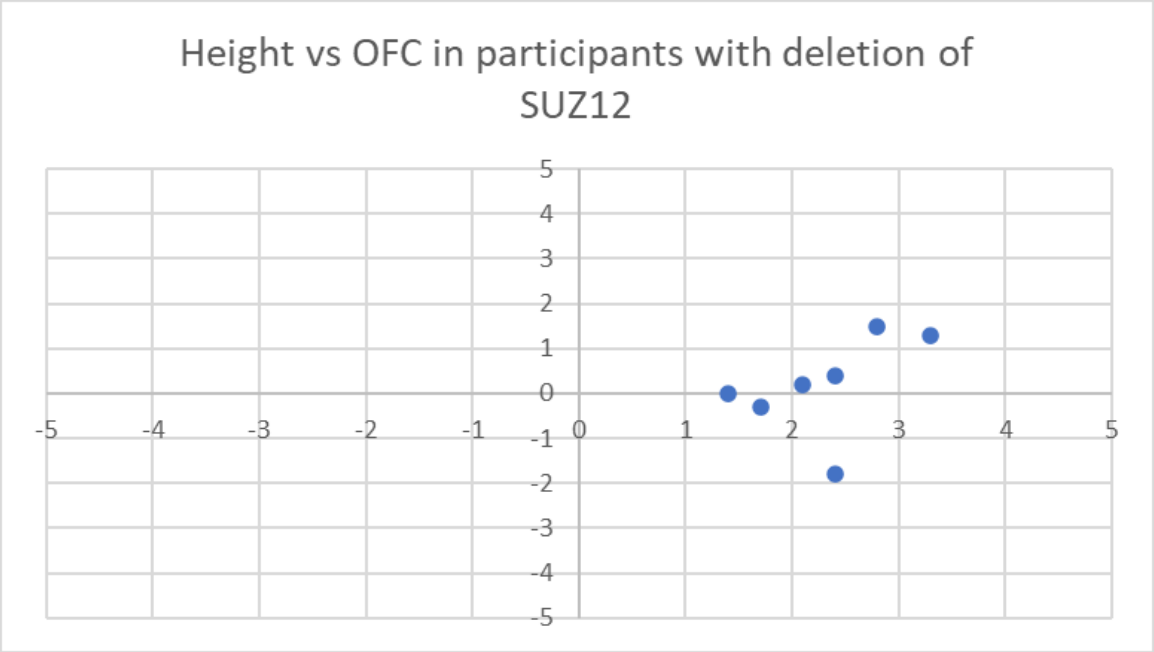


Figure 69: Height vs OFC in participants with deletion of SUZ12

Head circumference in all known individuals with Imagawa-Matsumoto syndrome (seven POD participants, previously reported individuals, and patients on DECIPHER) ranges from -1.8 SD to +6.9 SD with a mean of +1.5 SD.

4.1.11.8.5 Genotype-phenotype correlations

It has previously been noted that individuals with missense variants have a larger mean head circumference (+5.1 SD) compared to those with truncating variants (+2 SD)[293].

This finding in this family appears to confirm the observation that truncating variants and deletions in *SUZ12* are less likely to be associated with macrocephaly than missense variants (see Figures 73 and 74).

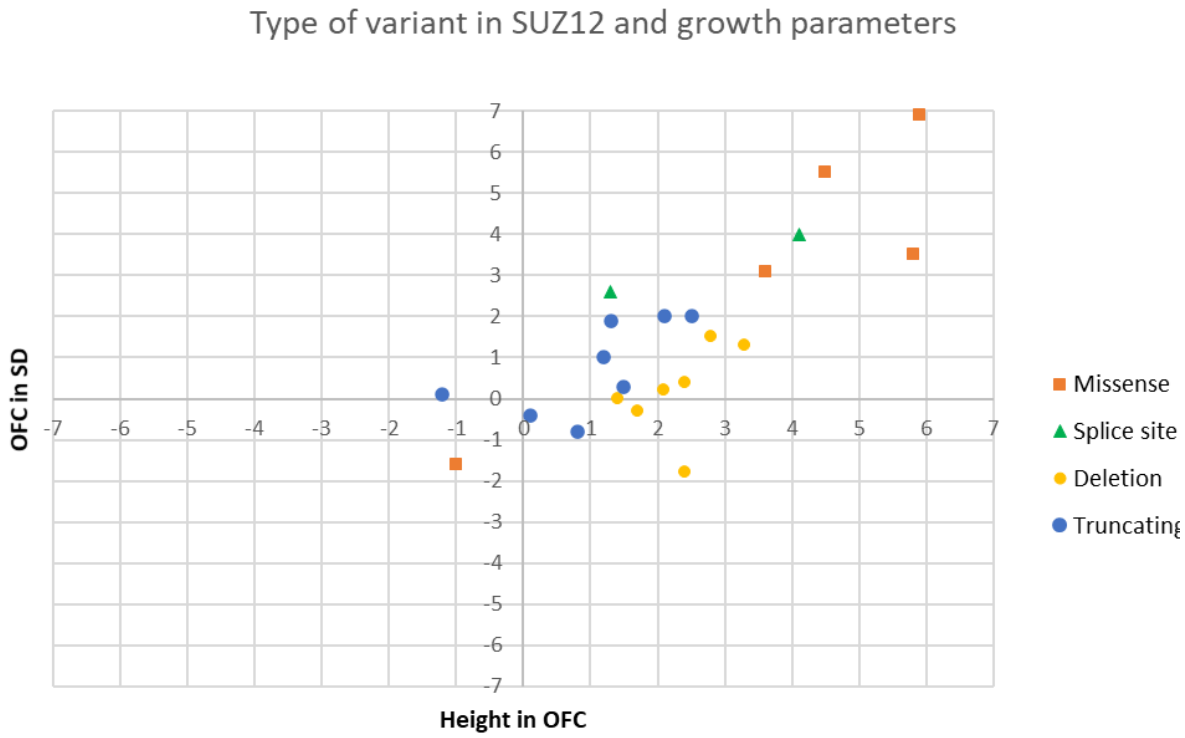


Figure 70: Scatterplot of height vs OFC according to type of SUZ12 variant in this study and in previously reported individuals

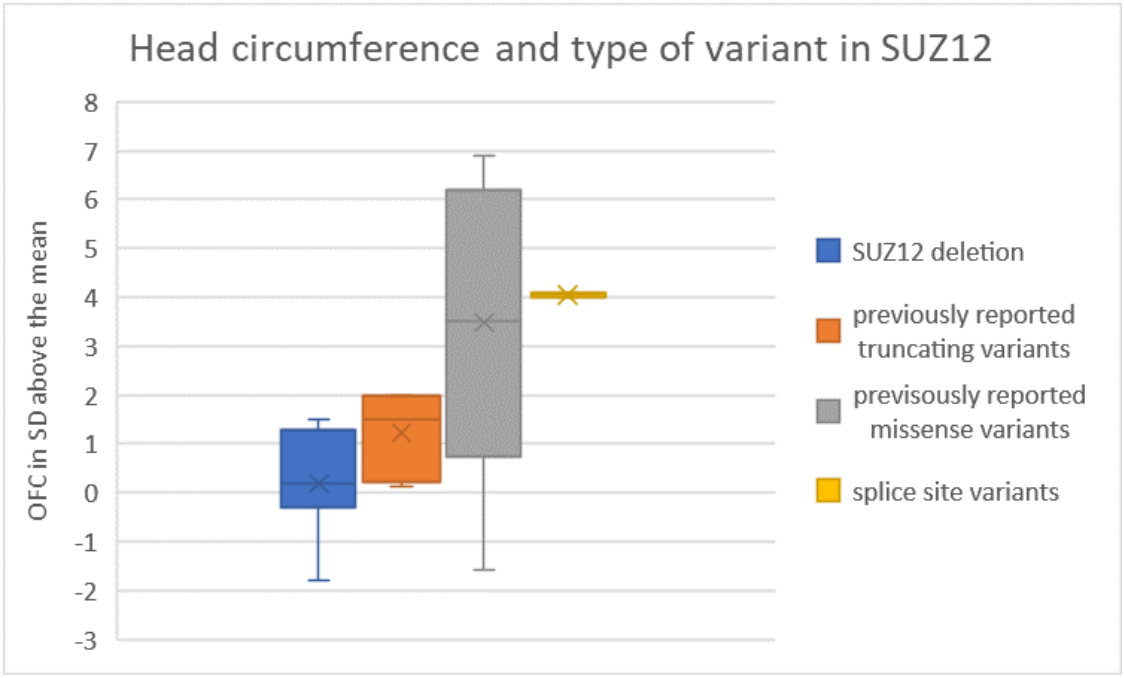


Figure 71: Head circumference and type of variant in SUZ12 in this study and in previously reported individuals

4.1.11.8.6 Clinical features

Individuals with Imagawa-Matsumoto syndrome were generally healthy with few medical issues. Minor skeletal anomalies of the hands and feet were present in six out of seven (86%) of family members. The most common findings were long palms (five individuals) followed by clinodactyly or camptodactyly of the fingers or toes (three individuals). Other features each seen in a single family member were long fingers, short second toe, pes planus and absent palmar creases. With the exception of absent palmar creases, these features have all been described in previous patients. The findings in this family confirm a high frequency of minor skeletal anomalities, including clinodactyly or camptodactyly of the digits, in Imagawa-Matsumoto syndrome.

Ophthalmic issues including myopia (two individuals), hypermetropia (two individuals), and astigmatism (one individual) were present in five out of seven (71%) of family members. These features have not been reported in other individuals with Imagawa-Matsumoto but may not have been documented. Given that refractive errors are very common in the general population it is difficult to draw the conclusion that these individuals are especially prone to ophthalmic issues.

Two individuals had pigmentary anomalies, one with a brown birthmark on the calf and the other with a pigmented scalp birthmark with overlying depigmented hair. Two other individuals have previously been reported to have pigmented naevi, suggesting that this may be a feature of Imagawa-Matsumoto syndrome[293].

Other medical issues (umbilical hernia, constipation, grommets, tonsillectomy, Gilbert's syndrome, cholecystectomy, osteoarthritis of the knees, neurological symptoms of pins and needles in the hand, hypothyroidism, type 2 diabetes, sciatica, non-pitting oedema of the leg, collapse of unknown cause and vitamin D deficiency) were each found only in one individual. Of note, umbilical hernia has previously been described in four other individuals, and therefore seems to be a relatively common finding[293]. The genitourinary anomalies described in four individuals with Imagawa-Matsumoto were not identified in our family, suggesting this is a less common feature. The findings in this family did not confirm the possible association between respiratory issues and Imagwa-Matsumoto syndrome. No members of the family had MRI brain imaging so the presence of structural brain anomalies is unknown.

Dysmorphic facial features included a round face, flat occiput, sloping forehead, widely spaced eyes, almond-shaped palpebral fissures, epicanthic folds, horizontal and thick eyebrows, long ears, horizontal crease in his chin, and thick vermilion of the upper and lower lips. The facial phenotype is reminiscent of Weaver syndrome in the younger children.

Table 36: Summary of clinical features of participants with SUZ12 deletion

	103.0	103.1	103.3	104.0	104.1	104.3	104.4	
Age	11	39	6	25	45	12	14	
Sex	M	F	M	M	F	F	F	
Tall stature	+	-	+	+	-	+	+	5/7
Macro-cephaly	-	-	-	-	-	-	-	0/7
Develop-mental delay/ ID	+	-	+	+/-	+	+	+	6/7
Skeletal	+ 5^{th} finger clinodactyly	+ long fingers	+ long fingers	+ clinodactyly halluces pes planus	+ short second toes	+ campto-dacyly toes	-	6/7
Ophthalmic	+ hyper-metropia	+ myopia	+ myopia	-	+ hyper-metropia	-	+ astigmatism	5/7
Pigmentary anomaly	-	- hypo-pigmented lesion	+	-	-	-	+ hyper-pigmented lesion	2/7
Dental	+ crowding	-	-	+ crowding	+ crowding	-	-	3/7
Other	umbilical hernia				oedema right leg			

In summary, this study increases the number of reported individuals with Imagawa-Matsumoto syndrome from 13 to 20. The phenotype is confirmed to consist of tall stature, developmental delay, and facial dysmorphism. Tall stature is not universal and a few individuals may have a height that falls within the normal range. A minority of individuals have very mild or no intellectual disability. Minor skeletal anomalies, refractory errors and dental crowding are common. This study confirms that umbilical hernia and pigmentary anomalies are part of the Imagawa-Matsumoto phenotype. The lack of macrocephaly, genitourinary anomalies and known structural brain anomalies in this family suggests a genotype-phenotype correlation, with truncating variants in *SUZ12* being associated with a less severe phenotype than missense variants. However, the number of individuals known to have this condition remains small and phenotypic analysis of a larger number of individuals is needed to be conclusive. Further work looking at the impact of the precise type of variant in

SUZ12 on the function of the PRC2 complex function is warranted to investigate this emerging relationship.

4.1.11.8.7 Inheritance

Five of the individuals in this family had maternally inherited variants. It is unknown if the variants present in the two sisters 103.1 and 104.1 were maternally or paternally inherited, or less likely the result of gonadal or somatic-gonadal mosaicism. Inherited variants appear to be not uncommon in Imagawa-Matsomoto syndrome, with two previously reported individuals having paternally inherited variants and two individuals having maternally inherited variants. Six previously reported individuals had de novo variants and in two individuals the status was unknown. One previously reported individual had somatic gonadal mosaicism for the SUZ12 variant[81].

4.2 Other single gene disorders with phenotypes overlapping overgrowth disorders.

Phenotypic descriptions of the group of participants with a molecular diagnosis of a genetic condition not usually considered to be an overgrowth disorder are described in the following section. These disorders are Marfan syndrome (three individuals), FOXP2-related speech and language disorder (one individual), Grieg syndrome (one individual), HIST1H1E syndrome (one individual), KMT5B syndrome (one individual), Gorlin syndrome (one individual), cerebro-facio-thoracic dysplasia (one individual), and Resistance to thyroid hormone alpha (one individual).

4.2.1 *FBN1:* Marfan syndrome

Marfan syndrome is an autosomal dominant connective tissue disorder caused by pathogenic variants in the *FBN1* gene[294]. The clinical features involve the cardiovascular, ocular, and skeletal systems, but can vary considerably between individuals[295]. The clinical diagnosis of Marfan syndrome is made according to the revised Ghent criteria[296] (see Appendix) and it is acknowledged that age related variability in clinical complications mean that the diagnosis can be difficult to make in children[297].

4.2.1.1 Results POD 005.0

POD 005.0: de novo splice site variant *FBN1* c.247+1G>A.

Participant POD 005.0 was born at 38 weeks and two days. During the pregnancy her mother had hypertension. Antenatal ultrasound scan identified a choroid plexus cyst at 20 weeks

gestation however this had resolved on a follow up scan. She was age ten at recruitment. Her height was 176.6 cm (+5.5 SD) and her weight 70.4 kg (+3.2 SD). Her bone age at the age of nine years one month was advanced at 12 years and four months. Her medical issues were mitral valve prolapse, mitral valve regurgitation, sacral intraspinal arachnoid cysts, severe hallux valgus, metatarsus adductus, pes planus, proximal hypermobility (Beighton score 8), striae on her back, and myopia in the left eye. She attended mainstream school. Her mother was 163.8 cm tall (+0.1 SD) and her father 189 cm (+1.7 SD) tall. Her 12 year old brother was 162.6 cm tall.

4.2.1.2 Discussion POD 005.0

POD 005.0 had several features, including mitral valve prolapse, myopia, pes planus, and striae, that give her a systemic score of at least 4 according to the revised Ghent criteria. The clinical examination proforma for the POD study does not include assessment for some systemic features of Marfan syndrome (such as wrist and thumb sign, reduced elbow extension, increased arm span/height) so it is possible that she could have a higher systemic score. However, regardless of the systemic score, in the absence of aortic dilatation, ectopia lentis, and/or a family history of this condition, she did not fulfill the diagnostic criteria for Marfan syndrome according to the revised Ghent criteria. Under the age of 20, she would be given the diagnosis of 'potential Marfan syndrome'.

Extreme tall stature as seen in POD 005.0 is known to be a feature of Marfan syndrome, with over 50% of children in a large study of 320 children having a height of >3 SD above the mean[298]. Joint hypermobility is also a common feature[298] but is not included in the Ghent

diagnostic criteria. There are case reports of sacral arachnoid cysts[299,300] and this may represent a rare feature of Marfan syndrome.

4.2.1.3 Results POD 008.0

POD 008.0: *FBN1* c.1761dupT; p.(Ile588TyrfsTer3). Parental samples not available to confirm de novo status.

POD008.0 was born at 40 weeks gestation by emergency LSCS weighing 3.8 kg (+0.5 SD) with a birth length of 58 cm (+3.5 SD). His mother had nausea and vomiting for the duration of the pregnancy. Newborn examination found he had unilateral cryptorchidism, umbilical hernia, and bilateral inguinal hernias. He also had feeding difficulties. At age 15 his height was 192.6cm (+3.9 SD), weight 65.8 kg (+1.6 SD) and OFC 55.1 cm (-0.4 SD). His bone age was not advanced and he had started puberty age 13 years. He had undergone a left orchidopexy age 15 and was on testosterone injections to limit growth. He had also undergone a right inguinal herniotomy and ligation of a left patent processus vaginalis. He had mild scoliosis, pes planus, severe planar valgus of the right foot (requiring calcaneal lengthening with an iliac crest graft), and a right hammer toe (treated with a PIP joint fusion). He attended mainstream school. His mother was 169.1 cm tall (+0.9 SD) and his father was 169.0 cm tall (-1.2 SD).

4.2.1.4 Discussion POD 008.0

POD 008.0 had a hindfoot deformity and scoliosis that give him a systemic score of 3 according to the revised Ghent criteria, but his score might be higher if an assessment for wrist and thumb sign, reduced elbow extension, and increased arm span/height were

completed. His echocardiogram and ophthalmic examination were both normal and in the absence of known family history, at under the age of 20 he would be given a diagnosis of 'potential Marfan syndrome' according to the revised Ghent criteria.

Like POD 005.0, POD 008.0 had extreme tall stature. Recurrent and/or incisional hernia are recognised as a less common feature of Marfan syndrome[301] and the presence of umbilical and inguinal hernia in POD 008.0 may therefore be related to his underlying connective tissue disorder. Cryptorchidism is often associated with congenital inguinal hernia.

4.2.1.5 Results POD 029.0

POD 029.0: *FBN1* de novo c.4444_4445delGG; p.(Gly1482Argfs*8)

Participant POD 029.0 was born at 40 weeks and six days weighing 4.2kg (+0.9 SD), a birth length of 56 cm (+2.1 SD) and OFC 35 cm (-0.6 SD). His mother has an underactive thyroid during pregnancy. On neonatal examination he had borderline hip dysplasia that was treated with bracing. At age six at entry to the study his height was 137.8 cm (+4.5 SD), weight 29.5 kg (+2.5 SD) and OFC 56.6 cm (+2.2 SD). His bone age was two years advanced. He had mitral valve prolapse and trivial mitral valve regurgitation. He had mild pectus excavatum, pes planus, a Beighton score of 4/9, and was noted to have long toes. He attended mainstream school. His mother was 170.1 cm tall (+1.1 SD) and father 182 cm tall (+0.7 SD).

4.2.1.6 Discussion POD 029.0

POD 029.0 had systemic features of mitral valve prolapse, pectus excavatum, and pes planus that would give him a score of 3 according to the revised Ghent criteria. Clinical examination

prior to recruitment to the POD study did not identify any additional systemic features according to the revised Ghent criteria. He did not have ectopia lentis, aortic dilatation and/or family history that is needed for a diagnosis of Marfan syndrome.

4.2.1.7 Summary of participants with Marfan syndrome

Table 37 gives a summary of participants with pathogenic variants in FBN1.

Table 37: Participants with a pathogenic variant in FBN1

Participant	POD 005.0	POD 008.0	POD 029.0
Variant	c.247+1G>A	c.1761dupT; p.(Ile588TyrfsTer3)	c.4444_4445delGG; p.(Gly1482Argfs*8)
Type of variant	splice site	frameshift	frameshift
Inheritance	de novo	unknown	de novo
Age	10	15	6
Birth length	nk	+3.5	+2.1
Height (SD)	+5.5	+3.9	+4.5
OFC (SD)	nk	-0.4	+2.2
Echocardiogram findings	Mitral valve prolapse and regurgitation	Normal	Mitral valve prolapse and trivial regurgitation
Ophthalmic findings	Myopia left eye	Normal	Normal
Musculoskeletal findings	Hallux valgus Metatarsus adductus Pes planus Hypermobility	Scoliosis Pes planus Planar valgus of the right foot Right hammer toe	Developmental dysplasia of the hip Pectus excavatum Pes planus Long toes
Other features	Striae Sacral intraspinal arachnoid cysts	Unilateral cryptorchidism Umbilical hernia Inguinal hernia	

4.2.2 *FOXP2*-related speech and language disorder

4.2.2.1 Results POD 038.0

POD 038.0: *FOXP2* de novo c.982C>T; p.(Arg328Ter)

POD038.0 was born by instrumental delivery at 42 weeks and 1 day and did not require resuscitation or admission to the neonatal unit. His mother was on amitriptyline and quinine during the pregnancy. Newborn examination identified a tongue tie and a birthmark on the back of his leg. He had some feeding difficulty and jaundice that did not require phototherapy. His birthweight was 4kg (50th centile) and head circumference 37 cm (+0.4 SD) at a gestation of 42 weeks and 1 day. However according to his mother her pregnancy was dated incorrectly, and he was born at 40 weeks and 1 day, which would make the birthweight +0.8 SD and OFC +1.3 SD.

At age five on recruitment to the study his height was 118 cm (+1.3 SD), weight 26 kg (+2.2 SD), and OFC 54.8 cm (+1.2 SD). His medical history included duplicated ureter, constipation, pectus excavatum, hypermobility, recurrent otitis media, glue ear with conductive hearing loss, recurrent tonsillitis and tonsillectomy, and sleep apnoea. He had mild global developmental delay with more severe delay in speech and language development. He was able to speak single words at the age of 2.5 years and join two words together at three years. He had features of autism, emotional lability, anxiety, and sleep difficulties. He had downslanted palpebral fissures and small hands and feet.

His mother's height was 165 cm (+0.2 SD) and OFC 56 cm (+0.3 SD). His father's height was 183 cm (+0.8 SD) and OFC 58 cm (+0.6 SD). His younger brother had no similar problems and there was no family history of overgrowth conditions.

4.2.2.2 Discussion POD 038.0

Molecular genetic testing in the POD study identified that POD038.0 has a de novo pathogenic variant in *FOXP2*, c.982C>T p.Arg328Ter. *FOXP2* was the first gene ever reported to be responsible for a developmental disorder in speech and language by Lai et al. in 2001[302]. A point mutation in *FOXP2* was found to segregate with affected members of a large three generation family with a severe speech and language disorder showing an autosomal dominant pattern of inheritance[302].

FOXP2-related speech and language disorders have a primary phenotype of childhood apraxia of speech (CAS)[303]. CAS is a disorder of speech motor programming causing difficulties in putting sounds together into syllables, syllables into words, and words into sentences[303]. Other speech and language difficulties including oral dyspraxia[302,304,305], dysarthria[306], moderate to severe receptive and expressive language disorder[304,307,308], and reading and spelling impairments[304] are also common. Speech development in children with FOXP2-related speech and language disorders usually begins between the ages of 18 months and seven years[304,305,309,310] and improves with age but difficulties may remain into adult life[311].

Non-speech related features of *FOXP2*-related speech and language disorders include a relatively stronger nonverbal IQ in comparison to verbal IQ[304,310]. A small number of

individuals have also been reported to have mild fine or gross motor delay, features of autism, and mild dysmorphic features of high arched palate, horizontal eyebrows and simply folded ears[310]. More significant global developmental delay and features of autism meeting the threshold of diagnosis for ASD are generally seen only in individuals where other genes are involved such as in a continuous gene deletion (*FOXP2*-plus speech and language disorder), rather than an intragenic sequence variant disrupting FOXP2 only (*FOXP2*-only speech and language disorder)[312–314].

A minority of reported affected individuals (~30%) have FOXP2-only speech and language disorder due to an intragenic sequence variant. About 70% of individuals in the literature have FOXP2-plus speech and language disorder. This is most commonly due to a non-recurrent contiguous gene deletion or less commonly due to a structural variant (such as chromosomal translocation or invertion) or maternal UPD7 that reduces FOXP2 expression[303].

Following the publication of the original family with 15 affected individuals, a further 25 individuals from 14 families have been described with intragenic sequence variants in *FOXP2.* There do not appear to be any genotype-phenotype correlations between specific intragenic variants in *FOXP2* and the clinical features although it is interesting to note the c.982C>T p.Arg328Ter variant present in POD038.0 has previously been reported in a family with three affected individuals: proband, his sibling and their mother[305]. The proband had delay in speech and language and social skills and was able to use single words at the age of four, his younger sister had motor and oropharyngeal dyspraxia, otitis media, and oesophageal

reflux, and their mother had a history of speech delay in childhood and as an adult she continued to have difficulties in communication[305].

The developmental phenotype of severe speech and language delay, mild fine and gross motor delay, and autistic features in POD038.0 is consistent with that previously described in this disorder. He does however have several other features that are not described in any other reported individuals. The double ureter, constipation, pectus excavatum, hypermobility, recurrent otitis media, glue ear with conductive hearing loss, recurrent tonsillitis and tonsillectomy, and sleep apnoea are not explained by this diagnosis. Although it is possible that one or more of these features represent an extension of the phenotype, it seems likely that several could be separate issues without a monogenic cause. Constipation, hypermobility, and recurrent otitis media and tonsillitis with associated complications are all commonly seen in the paediatric population. Mild pectus excavatum is also not uncommon and duplicated ureter is seen in approximately 1% of the population[315]. A larger cohort of patients with FOXP2-related speech and language disorder is needed to evaluate if any of these features are more common in this condition compared to the general population.

4.2.3 *GLI3*: Grieg cephalopolydactyly syndrome (GCPS)

4.2.3.1 Results POD 052.0

POD 052.0 *GLI3* c.1115C>A; p.(Ser372*). Parental samples not available to confirm possible paternal inheritance.

POD0 52.0 was born at 38 weeks and two days gestation weighing 3.2 kg (+0.3 SD) with a birth length of 49 cm (+0.1 SD). Her mother had gestational diabetes managed with insulin

and she also had a course of nitrofurantoin in pregnancy. On routine neonatal examination, POD052.0 was noted to have polydactyly with two halluces on each foot and syndactyly of the first three toes. She also had feeding difficulties, hypoglycaemia and jaundice treated with phototherapy in the newborn period.

Age five she was 106 cm tall (+0.4 SD), 29.5 kg in weight (+3.7 SD) and her OFC was 54.5 cm (+2.6 SD). She had a medical history of a pneumonia requiring treatment with oxygen and IV antibiotics. An MRI head scan showed rounded and dysplastic ventricles, hypoplastic anterior commissures, a shorter than usual corpus callosum and extra axial spaces more prominent than usual. This was reported as suspicious for Sotos syndrome. She also had gastroesophageal reflux treated with Gaviscon, pes planus, nocturnal leg pain and hypermetropia. She had delayed fine motor, social, and speech and language development, with her first word at 24 months and her first two word sentence at 31 months. Behavioural issues with polyphagia and sleeping difficulties were noted. She also had issues with temperature regulation and had increased sweating. She attended mainstream school with assistance. Dysmorphic features included a high palate, broad feet, preaxial polydactyly of the foot and osseus syndactyly of the toes. Her mother was 162.5 cm tall (-0.2 SD) with an OFC of 56.6 cm (+0.8 SD) and her father was 175 cm tall (-0.3 SD) with an OFC of 59.6 cm (+1.6 SD). Her father also had bilateral syndactyly of his first, second and third toes.

A clinical suspicion of Grieg cephalopolysyndactyly syndrome (GCPS) prompted diagnostic testing of *GLI3* and a heterozygous frameshift variant c.1115C>A; p.(Ser372*) was identified.

4.2.3.2 Discussion POD 052.0

GCPS is characterised by preaxial polydactyly, syndactyly, macrocephaly, and widely spaced eyes[316,317]. However, the craniofacial features can be absent or subtle in many individuals[316], with only 60% of patients having macrocephaly. Other less common features include anomalies of the corpus callosum, umbilical and diaphragmatic hernias, and rarely craniosynostosis[316].

Genotype-phenotype correlations are well described in *GLI3*, with deletions, duplications, missense, truncating, and splice site variants outside the middle third of *GLI3* being associated with a GCPS phenotype[318–320]. Truncating variants and a splice site variant in the middle third of *GLI3* are associated with Pallister-Hall syndrome, characterised by hypothalamic hamartoma, bifid epiglottis, and insertional (mesoaxial) polydactyly[321,322].

The location of the truncating frameshift variant in POD 052.0 in the first third of the gene is consistent with a diagnosis of GCPS. Her preaxial polydactyly, syndactyly and macrocephaly are typical of the phenotypic features of this condition. Although she does not have wide spaced eyes, this is not universal in this disorder, with only 43% of individuals having widely spaced eyes in one cohort of 55 individuals[318].

Anomalies of the corpus callosum and the ventricles, present in POD 052.0, are also reported in GCPS. In the series reported by Demurger et al. in 2015[318], 9/18 and 7/18 individuals who had undergone MRI brain imaging had hypoplasia or agenesis of the corpus callosum and

ventricular dilatation respectively. Interestingly, anomalies of the corpus callosum were generally associated with truncating variants in the C-terminus region of the protein (3' end of *GLI3)*[318]. The variant in POD 054.0 is also truncating, corroborating this type of variant is correlated with anomalies of the corpus callosum, but located towards the N-terminus of the protein. This suggests that this phenotypic feature can also be seen with variants at the 5' end of *GLI3.* Similarly, developmental delay is reported to occur in GCPS, but was previously thought to occur only in individuals with continuous gene deletions[323] and more recently also in individuals with variants towards the 3' end of the gene[318]. The developmental delay present in POD 054.0 demonstrates that developmental delay can also occur in association with variants towards the 5'end of *GLI3*. The brain anomalies and developmental delay in POD 054.0 therefore expand our knowledge of the genotype-phenotype correlations in GCPS.

4.2.4 *HIST1H1E* syndrome

4.2.4.1 Results POD 048.0

POD 048.0: *HIST1H1E* de novo c.441dup; p.(Lys148Glnfs*48)

This variant was identified through participation in the DDD study.

Participant 048.0 was born at 39 weeks gestation after a prolonged labour and face presentation with a birthweight of 3.62kg. He was hypotonic and required admission to the neonatal unit for 14 days. He was also noted to have bilateral undescended testes.

Age 12 he had macrocephaly with OFC +2.5 SD, a height of 155.1cm (-0.9 SD) and weight 48.8kg (0 SD). He also had a leg length discrepancy and facial asymmetry. In childhood he

had chronic diarrhoea for which no cause was identified but improvement was seen on a gluten free diet. An unusual pattern of recurrent periods of fever that only occurred at night with no symptoms of illness during the daytime reported. Skeletal abnormalities included fixed flexion deformities at both knees, pectus carinatum and scoliosis. He also had strabismus. Dermatological features included sparse body hair, slow growing head hair and numerous naevi. He underwent three operations for undescended testes.

POD 048.0 had severe developmental delay. Independent walking was achieved at the age of three and first word at age ten. Behavioural issues included temper tantrums and polyphagia. Facial features included a tall chin, downslanting palpebral fissures with epicanthic folds, overhanging nasal tip, long ears, tapered fingers and single palmar creases. A high narrow palate was also noted.

4.2.4.2 Discussion POD 048.0

The phenotype of 43 individuals with HIST1HE syndrome (also known as Rahman syndrome) has been described in the literature[324,325]. The neonatal hypotonia present in POD 048.0 is common, reported in 63%[325] and 73%[324] of individuals in two previous cohorts. The normal growth parameters at birth in POD 048.0 are also consistent with the literature[324,325]. Cryptorchidism in boys is also described in almost 70% of boys[325].

The growth pattern in HIST1H1E syndrome is unusual. Although initially reported as an overgrowth syndrome, it is recognised that often individuals have decreasing height percentiles with age and may be of below average height as adults[326]. Growth parameters are

very variable, with reported heights ranging from -1.8 SD to +8.3 SD and a mean of 0.4 SD[325]. Macrocephaly is more commonly seen, reported in 63% of one cohort[324] although another study found the mean OFC to be within the normal range at +1.1 SD with a range from -1.7 SD to +3.7 SD.[325] The normal height of -0.9 SD and large OFC +2.5 SD in POD 048.0 are therefore consistent with previously reported growth parameters. Lower limb asymmetry has been described in two individuals and its presence in POD 048.0 confirms this as a feature of HIST1H1E syndrome[325]. Facial asymmetry has not previously been recognised in this disorder and this finding may represent an expansion in the phenotype.

Skeletal features are commonly reported in this condition and scoliosis has been identified in four other individuals with HIST1H1E[325]. However, fixed flexion deformity at the knee and pectus carinatum are both novel skeletal features in POD 048.0. Strabismus is a common feature of HIST1H1E, reported in 53% of individuals in one cohort[324]. Ectodermal features of sparse body hair and slow growing head hair are described[325] and are suggested to be among features of premature ageing in these individuals[324]. Skin hyperpigmentation is also described as a sign of premature ageing but it is unclear if this feature is the same as the numerous naevi in POD 048.0. The problems of chronic diarrhoea in childhood and nocturnal fevers in POD 048.0, both of which remained undiagnosed, are not described in any other individuals with HIST1H1E syndrome.

Developmental delay and/or intellectual disability are universal in HIST1H1E syndrome[324,325]. Independent walking in one cohort was achieved at a mean age of 31 months (range 15-66 months), with POD 048.0's age of walking at 36 months being consistent with

this. Individuals usually have moderate ID, or less commonly mild or severe LD[324,325]. There is reported to be particularly difficulty in expressive language although there is no information on the age of first words. The age of speech development at ten years in POD 048.0 is notably delayed. The presence of developmental and behavioural issues in POD 048.0 is consistent with the literature, with anxiety, ADHD, ASD or aggression present in up to 50% of one cohort[325]. A second cohort reported fewer individuals with these behavioural issues but noted feeding issues in 60% including satiety issues in younger children[324]. The presence of polyphagia in POD 048.0 may demonstrate a persistence of satiety issues into the teenage years.

The facial features of HIST1H1E include high anterior hairline, sparse temporal hair bitemporal narrowing, frontal bossing, full cheeks, wide spaced and deep-set eyes, short and downslanted palpebral fissures, broad high nasal bridge, full nasal tip, wide spaced teeth and low set ears in later childhood[324,325]. POD 048.0 shares some of these features with downslanting palpebral fissures being particularly prominent. Single palmar creases have also been described.

The clinical features of POD 048.0 confirm the phenotypic spectrum of HIST1H1E includes lower limb asymmetry. Novel features of this condition may include facial asymmetry, fixed flexion deformity of the knee, pectus carinatum, numerous pigmented naevi, extremely delayed speech development, chronic diarrhoea in childhood, and nocturnal fevers.

4.2.5 *MAGED2*: transient antenatal Bartter syndrome

Pathogenic variants in *MAGED2* are known to cause an X-linked condition of transient antenatal Bartter syndrome. The clinical features of this condition are polyhydramnios and risk of premature birth, followed by a neonatal course of transient massive salt-wasting and polyuria, with subsequent resolution of symptoms by six weeks of age[327]. A cohort of 17 patients with pathogenic *MAGED2* variants was described by Legrand et al., with severe polyhydramnios occurring in all pregnancies and serial amnioreductions performed in each case[328]. Premature labour also occurred in all cases (16/16; one pregnancy underwent medical termination), most commonly between 26 and 33 weeks[328]. 75% of neonates had a birthweight above the 90th centile for gestational age[328].

4.2.5.1 Results POD 016.0

POD 016.0: maternally inherited hemizygous variant in *MAGED2* c.1085+1G>A

This diagnosis was identified through participation in the DDD study.

The pregnancy was complicated with severe polyhydramnios, requiring recurrent amnioreduction and treatment with the non-steroidal anti-inflammatory agent sulindac off-label to reduce production of amniotic fluid. POD 016.0 was born prematurely at 32+3 weeks gestation following a spontaneous labour. He required resuscitation at birth and received CPAP. He was large for gestational age with a birth weight of +2.3 SD however he lost a large volume of fluid through urination shortly after birth and his birth weight dropped rapidly. He underwent surgery for atypical Hirschsprung's disease.

At recruitment to the study age six, he had macrocephaly with an OFC +2.5 SD, weight of +2.4 SD, and height +1.3 SD. His medical problems included hypotonia of the lower limbs, bladder dysfunction and incontinence, gastro-oesophageal reflux, hypermobility, pes planus, and porous dentition.

He was delayed in reaching developmental milestones and walked independently at 26 months. He attended mainstream school and had a diagnosis of dyslexia but required no additional help. Behavioural issues included emotional lability.

Facial features included a square face, prominent forehead, short chin, and diastema. He also had slender fingers and bilateral 5th finger clinodactyly. He had a café-au-lait patch on his left thigh and a capillary haemangioma on his left upper arm.

His mother was 175 cm tall (+1.9 SD) with an OFC of 55.4 cm (-0.1 SD) and his father was 172 cm tall (-0.8 SD). Of note in the family history, his maternal grandmother had a pregnancy complicated by polyhydramnios and the male offspring died at six weeks of age.

4.2.5.2 Discussion POD 016.0

The pregnancy and neonatal history of severe polyhydramnios, prematurity, and high birth weight for gestational age in POD 016.0 is therefore consistent with the phenotype described in the literature. Although the cause of death of his maternal uncle is not known, the X-linked

pattern of inheritance and history of polyhydramnios suggests that he may also have had transient antenatal Bartter syndrome. Mortality related to premature birth has been described in several cases of this condition[327].

The 1085+1G>A variant has been reported in one individual with antenatal Bartter syndrome in the literature[328]. The identification of this variant partially explains POD 016.0's phenotype; but does not account for his clinical features of macrocephaly, atypical Hirschsprung's disease, bladder dysfunction, lower limb hypotonia, hypermobility, and dysmorphic features. Interestingly, dysmorphic facies have been reported in three individuals with pathogenic variants in *MAGED2*[328], although the specific facial features were not described so it is not possible to assess if these were similar to POD 016.0. One individual has been reported to have hydrocephalus and thus will have had macrocephaly[328], but the macrocephaly in POD 016.0 is not due to hydrocephalus and it is not possible to conclude this is a shared feature.

4.2.6 KMT5B syndrome

KMT5B (SUV420H1) was first described as a neurodevelopmental gene in 2017 by McRae et al. and Stessman et al. in two large scale exome sequencing projects[329,330]. In total, 46 individuals have been described in the literature with copy number or single nucleotide variants in *KMT5B*[331,332].

Phenotypic data on the 27 individuals with de novo likely pathogenic and pathogenic intragenic (missense, frameshift, stop gain) variants and one pathogenic splice donor variant

show that global developmental delay, speech and language delay, motor delay, intellectual disability, autism spectrum disorder, hypotonia, febrile convulsions, seizures, brain anomalies, hypermobility, sleep problems and dysmorphic features [329,331,332] are found in this cohort. MRI brain anomalies include macrocephaly, hydrocephalus, hypoplasia of the corpus callosum, enlarged ventricles, enlarged perivascular spaces, brain atrophy and Chiari type 1 malformation. [329,331]. Information on facial dysmorphism is available for a small number of individuals and features described include triangular face, high forehead, broad forehead, horizontal palpebral fissures, sparse lateral eyebrows, small ears, low set ears, cupped ears, posteriorly rotated ears, smooth philtrum, wide nasal bridge, wide nasal base, and low columella[329,331,332]. Additional features reported in single individuals include unilateral cryptorchidism[329], pes equinovarus[329], pes planus[329], scoliosis[331], chronic variable immune deficiency[329], high palate[332], hypermetropia[331], strabismus[331], inverted nipples[331], and tics[331].

Growth parameters have been reported available for relatively few individuals however macrocephaly has been described in two individuals, one with an OFC +4 SD, and five have been reported to have overgrowth, tall stature and/or height over 2 SD. *KMT5B* was initially identified through sequencing of cohorts of individuals with intellectual disability or autism and in these studies the growth phenotype was not fully characterised. Subsequently this condition has been described as an OGID disorder[333].

4.2.6.1 Results POD 085.0

POD 085.0: *KMT5B* c.2347C>T; p.(Arg783Ter) and microduplications at 1q21.1 (145415156_145899418)x3 and 16p13.11 (15048732+16194575)x3. These variants were not maternally inherited. A paternal sample was not available to confirm de novo status.

POD085.0 was three years old on recruitment to the study. He was known to have microduplications at 1q21.1 and 16p13.11 however these microarray abnormalities were not thought to fully account for his phenotype.

He was born following an uneventful pregnancy at 39 weeks and six days of gestation weighing 3.402 kg (-0.3 SD). He required resuscitation with inflation breaths and was admitted to the neonatal unit for seven days for respiratory distress. He was also noted to have hypotonia and developed neonatal jaundice.

At age three he was 103 cm tall (+2.5 SD), weighed 17.3 kg (+1.6 SD) and his OFC measured 52.4 cm (+0.8 SD). He was hypotonic. An MRI brain scan showed small patchy areas of high signal around the trigones of the lateral ventricles suggestive of previous white matter injury and a mild degree of colpocephaly. He also had eczema. He had global developmental delay and first walked at 18 months. At the age of 36 months he did not yet have any words and temper tantrums were an issue. He had a number of dysmorphic facial features including a high anterior hairline, broad forehead, full cheeks, widely spaced eyes, upslanted palpebral fissures, a depressed nasal bridge, broad nasal tip, tented upper lip vermilion, overfolded helices and increased posterior angulation of the ears.

His parents and half-sister were well and there were no family members with similar problems. His mother's birthweight was 3.09 kg. Her current height was 167 cm (+0.6 SD) and OFC 56.3 cm (-0.4 SD) and his father's height 190 cm (+1.8 SD).

4.2.6.2 Discussion POD 085.0

Singleton whole exome sequencing in the POD study identified a likely pathogenic stop gain variant c.2347C>T p.Arg783Ter in *KMT5B.*

POD 038.0's features fit the described phenotype of global developmental delay, speech delay, hypotonia, and brain anomalies. His height of +2.5 SD would be consistent with the emerging evidence that *KMT5B*-related disorders could be classified as an OGID disorder. However, growth data on larger cohorts of individuals with pathogenic variants in *KMT5B* is required to further establish the overgrowth phenotype of this condition.

4.2.7 *PTCH1*: Nevoid basal cell carcinoma syndrome (BCNS; Gorlin syndrome)

Nevoid basal cell carcinoma syndrome (NBCCS) is characterised by basal cell carcinomas and odontogenic keratocysts, although these features are often not present in children and usually develop from the teenage years onwards[334]. Other features include palmar and plantar pits, ectopic calcification of the falx cerebri, and skeletal anomalies include bifid rib and wedge shaped vertebrae[334,335]. Macrocephaly is a recognised feature of NBCCS, present in at least 50% of individuals[336].

4.2.7.1 Results POD 035.0

POD 035.0: *PTCH1* de novo c.2611_2624del; p.(Asn871Trpfs*20) identified by DDD study. Maternally inherited deletion 15q11.2.

POD 035.0 was born at 39 weeks gestation by instrumental delivery weighing 4.12 kg (+1.6 SD) and required resuscitation at birth. His mother had gestational diabetes and was treated with metformin. In the neonatal period he became jaundiced and was treated with phototherapy. Age ten he was 156.0 cm (+2.7 SD), his weight was 70.6 kg (+3.3 kg), and his OFC was 59.7 cm (+3.2 SD). His medical history included undescended testis managed with orchidopexy, febrile convulsions, recurrent tonsillitis, glue ear, hypermetropia, and joint pains. He also had global developmental delay and several behavioural issues including aggression, emotional lability, temper tantrums, anxiety, features of autism, short attention span and diagnosis of ADHD, self-injury, and sleep difficulties. He was also noted to have hypersensitivity to light and excessive sweating. His mother was 162.5 cm tall (-0.2 SD) with an OFC of 58 cm (+1.8 SD) and his father was 172.7 cm tall (-0.7 SD) with an OFC of 56.5 cm (-0.5 SD).

Following diagnosis of Gorlin syndrome (naevoid basal cell carcinoma syndrome), he underwent surgical removal of multiple jaw keratocysts at the age of 14.

4.2.7.2 Discussion POD 035.0

The degree of macrocephaly in POD 035.0, +3.2 SD, is greater than the mean OFC of +2 SD in children and teenagers with Gorlin syndrome described by Kimonis et al.[336]. The tall stature of POD 035.0 (height +2.7 SD) is not reported as a feature of NBCCS. However, the same study noted an average height of 20 affected individuals to be +0.75 SD compared to +0.1 SD in 18 unaffected siblings. This was not statistically significant ($p = 0.18$) but suggests a trend towards increased height in individuals with NBCCS[336].

Although gross motor delay is sometimes present, there is no evidence for global delay and development is said to be normal by the age of five years[337]. The developmental delay and behavioural issues in POD 035.0 are likely to be due to his second diagnosis of a microdeletion at 15q11.2, Burnside-Butler syndrome. This susceptibility locus is associated with issues including an increased risk of developmental delay, ADHD, autism spectrum disorder, and general behavioural problems[244].

The other features present in POD 035.0 (undescended testis, febrile convulsions, recurrent tonsillitis, glue ear, hypermetropia, and joint pains) are not known to be associated with NBCSS or 15q11.2 microdeletion but are not uncommon in the general paediatric population.

POD 035.0's phenotype of macrocephaly is consistent with the diagnosis of NBCCS but his tall stature, developmental delay and behavioural phenotype prompted further investigation on the overgrowth panel. No further variants were identified and the two diagnoses of NBCCS

and 15q11.2 microdeletion are likely to fully explain his phenotype. The tall stature in POD 035.0 suggests this may be an under recognised feature of NBCCS and further study of the height of individuals with NBCCS would be needed to confirm if this is the case.

4.2.8 *TMCO1:* Cerebro-facio-thoracic dysplasia (CFTD)

TMCO1 (transmembrane coiled-coil domains 1) defect syndrome was first described in 2010 by Xin et al. as an autosomal recessive condition affecting 11 individuals from the Old Order Amish in Northeastern Ohio with global developmental delay, skeletal anomalies and facial dysmorphism[338]. It was subsequently recognised by that four Turkish families with cerebrofaciothoracic dysplasia (CFTD), a condition first described by Pascual-Castroviejo in 1975, also had biallelic variants in TMCO1[339]. In total, 27 patients have now been reported to have CFTD resulting from TMCO1 deficiency[340].

4.2.8.1 Results POD 054.0

POD 054.0: *TMCO* homozygous variants c.233G>A p.(Trp78*). Parents confirmed to be heterozygous carriers.

This participant's diagnosis was identified through participation in the 100,000 Genomes Project.

The pregnancy was complicated by Group B Strep infection. Antenatal ultrasound scans identified polyhydramnios and positional talipes. She was born by emergency Caesarian section because of fetal distress at 40 weeks and 4 days weighing 2.6kg (-2.3 SD) and

required resuscitation with inflation breaths. On neonatal examination she was hypotonic and had an umbilical hernia and heart murmur that spontaneously resolved.

Her hypotonia and stiffness and weakness of the right arm were investigated and she was found to have agenesis of the corpus callosum. Other medical problems included bilateral divergent strabismus, constipation and gastro-oesophageal reflux. On review age two, her height was +3.5 SD, weight +2.7 SD and OFC +1.2 SD. Previous growth measurements at the age of ten months were height +3.4 SD, weight +1.9 SD and OFC +2 SD.

She had global developmental delay and had no independent steps or words at the age of two. She had self-injurious behaviours of hitting her head and pulling her hair and poor sleep. She was also noted to have increased sweating and reduced sensitivity to pain. Dysmorphic facial features included facial asymmetry, coarse facies, frontal bossing, broad eyebrows, synophrys, a wide nasal bridge, a tented vermilion of the upper lip, and low set ears. She also had a hypopigmented line down the front of her chest.

Following diagnosis, AP and lateral x-ray images of the chest and cervical spine identified a mild modelling deformity of the right upper ribs and a small osteochondroma arising from the fourth rib.

Her parents were consanguineous. Her mother was 167cm tall (+0.6 SD) with an OFC of 56.7 cm (+0.9 SD) and her father was 188cm tall (+1.5 SD) with an OFC of 60 cm (+1.8 SD).

4.2.8.2 Discussion POD 054.0

The antenatal scan finding of polyhydramnios in the pregnancy with POD 054.0 is previously reported as a feature of TMCO1 deficiency, complicating 7/22 (32%) of pregnancies[338,339,341–343]. Talipes has been reported in 7/25 (28%)[338,339,342–344] and hypotonia in 17/21 (81%)[338,341–345] of individuals. Umbilical hernia has not previously been reported and is a novel feature of TMCO1 deficiency in this participant.

Although not defined as an overgrowth disorder, the reported TMCO1 deficiency phenotype includes high birth weight in 7/12 (58%)[339,341,343,344] and macrocephaly in 13/23 (57%)[338,339,341,343–345]. The low birth weight in POD 054.0 is unusual for this condition. Reported height is variable, with tall stature in 5/16 (31%)[338,343,344] and short stature in 5/13 (38%)[338,341,344,345]. There is limited information available about the growth parameters of the previously reported individuals with tall stature. This participant's height demonstrates that tall stature in this condition can be as much as +3.5 SD above the mean.

Brain anomalies are common in TMCO1 deficiency, with 8/14 (57%) sharing POD 054.0's finding of hypoplastic or absent corpus callosum[339,341–345]. The finding of stiffness and weakness of a limb has not previously been reported however. Her other medical problems of strabismus and constipation have been reported in 5/11 (35%)[338,339,343] and 7/11 (64%)[338] of individuals respectively. Gastro-oesophageal reflux has not been identified as an issue but it is likely that this problem has been classified under the broader category of poor feeding,

affecting the majority of patients (18/22; 82%)[338,339,343–345]. Rib anomalies are common in CFTD (20/26; 77%)[338,339,341–345] but osteochondromas have not previously been reported.

Developmental delay appears to be universal in this disorder (23/23 reported individuals)[338,339,342–345]. 11/22 (50%) of reported individuals are non-verbal[338,339,342,344,345], as is the case for POD 054.0, but as she is only two years old she may acquire spoken language in future. The self-injurious behaviour in this participant is shared by three other reported patients (3/7; 43%)[339,342,345]. Reduced sensitivity to pain and increased sweating have not been described in any other patients.

A characteristic facial appearance has been described in CFTD[340]. POD 054.0 shares many of the typical facial features, including frontal bossing, synophrys, wide nasal bridge and low set ears. Facial asymmetry and pigmentary anomalies have not been described in other individuals in the literature.

In summary, POD 054.0 has many of the known phenotypic features of CFTD, including developmental delay, hypotonia, feeding problems, MRI brain anomaly, characteristic facial features, and skeletal features of rib anomalies and talipes. Novel features in this participant include umbilical hernia, weakness and stiffness of a single limb, osteochondroma, reduced sensitivity to pain, increased sweating, facial asymmetry and pigmentary anomaly, potentially widening the phenotypic spectrum of CFTD/TMCO1 defect disorder.

4.2.9 *THRA:* Resistance to thyroid hormone alpha (RTH-alpha)

4.2.9.1 Results POD 067.0

POD 067.0: *THRA* de novo c.1195C>T; p.(Pro399Ser)

POD067.0 was born at 42 weeks gestation with a birth weight of 3.26 kg (-1.4 SD). His mother was on citalopram and smoked during the pregnancy. Age four his height was 89 cm (-3.4 SD), weight 15.3 kg (-0.7 SD) and OFC 54 cm (+1.2 SD). He was recruited to the study based on previous growth parameters at 18 months of age with a length of 73.5 cm (-2.8 SD), weight of 9.5 kg (-1.7 SD) and macrocephaly with an OFC of 52cm (+2.3 SD) with markedly increased head size relative to his short stature. He had hypotonia with a normal MRI head scan. Other medical problems were renal stones, constipation, and hypermobility (Beighton score 6). He had global developmental delay with independent sitting at 12 months and walking at 30 months. His first word was at 12 months and two word sentences at 24 months. His mother and father were both 170 cm tall (+1.1 SD and -1.1 SD respectively).

4.2.9.2 Discussion POD 067.0

THRA encodes the thyroid hormone receptor alpha, one of the two receptors for thyroid hormone (TH)[346]. Loss of function variants leading to lack of T3 binding were first described in 2012[347] and to date over 30 individuals have been reported[348]. Their phenotype is similar to that of congenital hypothyroidism, with short stature, developmental delay, and constipation; despite low to normal serum T4, high –normal to high T3, low rT3, and normal TSH levels[346].

The clinical features of short stature, macrocephaly, global developmental delay, constipation, hypotonia, and hypermobility have been described in this disorder[347–349]. Renal stones have not previously been reported in Resistance to Thyroid Hormone alpha and may represent a novel clinical feature in this disorder.

The Pro399Ser variant identified in this participant is located in the ligand (T3) binding domain and is adjacent to the previously reported pathogenic P398R variant[350].

4.3 Microdeletion and microduplication syndromes

4.3.1 16p13.11 microduplication syndrome

Three participants, POD 055.0, POD 073.0, and POD 085.0, had microduplications at 16p13.11.

The 16p13.11 region is rich in low copy repeats (LCRs) and therefore prone to non-allelic homologous recombination (NAHR), leading to recurrent microdeletions and microduplications[351]. The ~1.15Mb microduplication identified in these three participants encompasses the critical region (interval II) of the 16p13.11 neurodevelopmental susceptibility locus[352]. Although 16p13.11 microduplications are usually inherited from an apparently healthy parent, they are often pathogenic when detected in the investigation of developmental delay or autism spectrum disorder[353]. The phenotype includes an increased chance of speech delay, intellectual disability, autism spectrum disorder, ADHD, seizures, abnormal brain MRI and cardiac malformations[352–354]. It has also been suggested that microduplications of this region predispose to thoracic aortic aneurysms and dissections (TAAD)[355]. Candidates for the neurocognitive phenotype are the NDE1 (nudE nuclear distribution gene E homolog 1) gene, which has a role in microtubule organisation and neuronal proliferation and migration; and the miRNA miR-484, which appears to be associated with a hyperactivity phenotype in mouse[353,356]. It has been suggested that duplication of MYH11, the smooth muscle cell (SMC)-specific beta myosin heavy chain isoform, could account for a possible increased risk of TAAD[355].

4.3.1.1 Results POD 055.0

POD 055.0: paternally inherited microduplication 16p13.11 (15048732_16194575)x3.

RB1 pathogenic variant c.1333C>T p.(Arg445X) and LOH identified in tumour; not present in the germline.

POD 055.0 was born at 38 weeks gestation by elective LSCS weighing 3.36 kg (+0.5 SD). Age 18 he was 190 cm tall (+1.8 SD), weighed 147.5 kg (+4.5 SD) and his OFC was 63.5 cm (+3.6 SD). He entered puberty at 15 years and six months of age. He had a unilateral non-heritable retinoblastoma diagnosed age three. His renal function was borderline abnormal secondary to chemotherapy. He also complained of headaches and clumsiness and an MRI head scan was reported as normal. Other medical problems included intolerance to milk, mandibular prognathism requiring surgical correction, and myopia. Psychiatric problems of anxiety and depression were being treated with fluoxetine, melatonin and aripiprazole. He had developmental delay and first walked age 18 months. Developmental and behavioural issues included aggression, emotional lability, temper tantrums, autism spectrum disorder, short attention span, obsessive compulsive behaviour, hyperacusis, pain insensitivity and eating and sleeping issues. Autonomic features of abnormality of temperature regulation and increased sweating were also present. He had attended a mainstream primary school with a Statement of Special Educational Needs and a moved to a special school for secondary education. Age 18 he was completely dependent on family members for activities of daily living. Facial features included prognathism, thick vermilion of the upper and lower lip, and a high palate. His mother was 165 cm tall (+0.2 SD) with an OFC of 56.6 cm (+0.7 SD) and his father was 175 cm tall (-0.3 SD) with an OFC of 59 cm (+1.2 SD).

4.3.1.2 Discussion POD 055.0

This participant's unilateral retinoblastoma is explained by the identification of two hits in *RB1* in DNA extracted from the tumour. His developmental delay and behavioural issues could be the result of the 16p13.11 microduplication. However, obesity and macrocephaly have not previously been described in association with this microduplication. His mild dysmorphic features are also unexplained.

4.3.1.3 Results POD 073.0

POD073.0: microduplication 16p13.11 (15048732_16194575)x3. This variant was not maternally inherited. A paternal sample was unavailable to confirm de novo status.

POD 073.0 was born at 39 weeks gestation weighing 3.2 kg (-0.3 SD). His mother smoked five cigarettes per day in the pregnancy. He had a naevus flammeus on neonatal examination. Age 11 he was 169 cm tall (+3.3 SD), his weight was 71.5 cm (+2.9 SD) and his OFC was 57.5 cm (+1.6 SD). His bone age was advanced to 13.6 years at an actual age of ten years and two months. He entered puberty age 11. Medical problems included mild asthma, constipation, and eczema. Behavioural difficulties included aggression, emotional lability, temper tantrums, features of autism, and polyphagia. He also had autonomic features of abnormal temperature regulation and increased sweating. He attended a mainstream school with assistance. His mother was 173 cm tall (+1.6 SD) with an OFC of 55.8 cm (+0.2 SD) and his father was 175 cm tall (-0.3 SD). He had anterior creases on both earlobes, 5th finger camptodactyly of his right hand, and pes planus.

4.3.1.4 Discussion POD 073.0

POD 073.0's behavioural issues could be explained by the 16p13.11 microduplication. Tall stature has not previously been described in association with this microduplication however.

4.3.1.5 Results POD 085.0

POD 085.0: *KMT5B* c.2347C>T; p.(Arg783Ter) and microduplications at 1q21.1 (145415156_145899418)x3 and 16p13.11 (15048732_16194575)x3. These variants were not maternally inherited. A paternal sample was not available to confirm de novo status.

POD085.0 was born following an uneventful pregnancy at 39 weeks and six days of gestation weighing 3.402 kg (-0.3 SD). He required resuscitation with inflation breaths and was admitted to the neonatal unit for seven days for respiratory distress. He was also noted to have hypotonia and developed neonatal jaundice.

At age three he was 103 cm tall (+2.5 SD), weighed 17.3 kg (+1.6 SD) and his OFC measured 52.4 cm (+0.8 SD). He was hypotonic. An MRI brain scan showed small patchy areas of high signal around the trigones of the lateral ventricles suggestive of previous white matter injury and a mild degree of colpocephaly. He also had eczema. He had global developmental delay and first walked at 18 months. At the age of 36 months he did not yet have any words and temper tantrums were an issue. He had a number of dysmorphic facial features including a high anterior hairline, broad forehead, full cheeks, widely spaced eyes, upslanted palpebral

fissures, a depressed nasal bridge, broad nasal tip, tented upper lip vermilion, overfolded helices and increased posterior angulation of the ears.

His mother's birthweight was 3.09 kg, her current height was 167 cm (+0.6 SD) and OFC 56.3 cm (+0.6 SD) and his father's height was 190 cm (+1.8 SD).

4.3.1.6 Discussion POD 085.0

POD 085.0 was also diagnosed with KMT5B syndrome on exome sequencing, which would account for his tall stature, developmental delay, dysmorphic features and abnormal MRI brain imaging. It is possible that the 16p13.11 microduplication is contributing to his neurobehavioural phenotype.

4.3.1.7 Discussion 16p13.11 microduplication syndrome

It is notable that three out of 100 participants in a cohort of individuals with overgrowth disorders have precisely the same 16p13.11 microduplication, and raises the possibility that this duplication could be associated with overgrowth. The reciprocal 16p13.11 microdeletion has been associated with microcephaly[352], and homozygous loss of function variants in the candidate gene NDE1 have been identified as a cause of extreme microcephaly[357]. Given that there are many examples of reciprocal duplications and deletions causing opposite phenotypes, duplication of 16p13.11 and consequently an additional copy of NDE1 could be a plausible cause of macrocephaly. This would not obviously explain a phenotype of tall stature however.

To date the literature on 16p13.11 does not indicate a predisposition towards tall stature or macrocephaly. Enquiry with Unique, the rare chromosome patient support group, identified that out of almost 200 families with 16p13.11 duplications, only two families mentioned their child had tall stature, one had a child with obesity and one had a child with mild macrocephaly. In the POD cohort, only POD 055.0 had macrocephaly and the other two participants had tall stature, and of these had a concurrent diagnosis of KMT5B syndrome that would explain their tall stature.

It therefore seems unlikely that 16p13.11 microduplications cause an overgrowth phenotype. An alternative explanation is that the identification of these participants in an overgrowth cohort is likely to reflect their shared phenotypic features of developmental disorders and intellectual disability that are part of the eligibility criteria for participation in the study.

4.3.2 15q11.2 BP1-BP2 microdeletion syndrome (Burnside-Butler syndrome)

Three participants, POD 035.0, POD 036.0, and POD 072.0 had 15q11.2 microdeletion syndrome (Burnside-Butler syndrome).

The recurrent deletion of ~452kb at 15q11.2 (BP1-BP2) is in a region of commonly recorded copy number variation (CNV) in the Database of Genomic Variants (DGV). Literature evidence suggests that deletions of this region may increase susceptibility to a broad spectrum of neurodevelopmental disorders[244,358]. However, they have also been observed in normal-

random-controls (not phenotypically characterized) and in unaffected relatives, with the deletion frequently inherited from an unaffected or very mildly affected parent[244,358].

4.3.2.1 Results POD 035.0

POD 035.0: *PTCH1* de novo c.2611_2624del; p.(Asn871Trpfs*20) and maternally inherited deletion 15q11.2

POD 035.0 was born at 39 weeks gestation by instrumental delivery weighing 4.12 kg (+1.6 SD) and required resuscitation at birth. His mother had gestational diabetes and was treated with metformin. In the neonatal period he became jaundiced and was treated with phototherapy. Age ten he was 156.0 cm (+2.7 SD), his weight was 70.6 kg (+3.3 kg) and his OFC was 59.7 cm (+3.2 SD). His medical history included febrile convulsions investigated with MRI head scan which was reported as normal, hypermetropia, and joint pains. He also had global developmental delay and several behavioural issues including aggression, emotional lability, temper tantrums, anxiety, features of autism, short attention span and diagnosis of ADHD, self-injury, and sleep difficulties. He was also noted to have hypersensitivity to light and excessive sweating. His mother was 162.5 cm tall (-0.2 SD) with an OFC of 58 cm (+1.8 SD) and his father was 172.7 cm tall (-0.7 SD) with an OFC of 56.5 cm (-0.3 SD).

4.3.2.2 Results POD 036.0

POD 036.0: paternally inherited microdeletion 15q11.2 (22,765,637-23,217,513)x1

POD 036.0 was born at 29 weeks gestation weighing 1.3 SD (0 SD). He required resuscitation and was on NICU for 70 days. He had feeding difficulties and was treated with phototherapy for neonatal jaundice. Age eight his height was 149 cm (+2.8 SD), weight 53.6 kg (+3.0 SD) and OFC 57.5 cm (+2.1 SD). His bone age was advanced to 11 years at an actual age of seven years and ten months. Medical problems included a heart murmur, post-nasal drip and nocturnal cough, constipation, leg pain, and recurrent chest infections. He had two seizures age seven and an MRI head scan was reported as normal. He had developmental delay and walked independently at 16 months and started speaking age three and a half. Behavioural difficulties included aggression, emotional lability, temper tantrums, anxiety, autism spectrum disorder, hyperactivity, pain insensitivity, polyphagia and sleep difficulties. He had autonomic symptoms of abnormal temperature regulation and increased sweating. He attended a mainstream school with assistance. Bilateral 2,3 toe syndactyly was noted. His mother was 165 cm tall (+0.2 SD) with an OFC of 55.2 cm (-0.2 SD) and his father was 182 cm tall (+0.7 SD).

4.3.2.3 Results POD 072.0

POD 072.0: mosaic *PIK3CA* c.1093G>A; p.(Glu365Lys) identified in DNA extracted from skin biopsy. Variant not present in DNA extracted from blood lymphocytes or buccal swab. Germline microdeletion at 15q11.2 (22,765,637-23,217,513)x1. Parental studies are now not routinely performed in 15q11.2 microdeletions as they are frequently inherited and testing is unlikely to clarify clinical significance.

POD 072.0 was born at 40 weeks and nine days gestation weighing 4.65kg (+1.6 SD). His mother was on sertraline during the pregnancy. At birth he required facial oxygen and was

admitted to NICU for five days. He had widespread bruising resulting from shoulder dystocia, a capillary vascular malformation on her trunk and feeding difficulties. Age four he was 110 cm tall (+1.7 SD) with a weight of 22.3 kg +2.4 SD) and OFC 53 cm (+0.5 SD). He had non-progressive regional overgrowth of the right leg. Medical problems include asthma, febrile convulsions, absence seizures, chronic diarrhoea, an adrenal haemorrhage, tight Achilles tendons, tibial torsion, cutis marmorata on the trunk, leg lymphoedema, and conductive hearing impairment with narrow ear canals. He had mild developmental delay. He sat age six months, walked age 12 months, and spoke in two word sentences at 36 months. Behavioural issues included temper tantrums, anxiety, features of autism, short attention span, pain insensitivity and hyperacusis. He attended a special school. His mother was 167.6 cm tall (+0.6 SD) with an OFC of 55 cm (-0.4 SD) and her father was 185.4 cm tall (+1.2 SD) with an OFC of 59.2 cm (+1.3 SD).

The recurrent deletion of ~452kb at 15q11.2 (BP1-BP2) is in a region of commonly recorded copy number variation (CNV) in the Database of Genomic Variants (DGV). Literature evidence suggests that deletions of this region may increase susceptibility to a broad spectrum of neurodevelopmental disorders. However, they have also been observed in normal-random-controls (not phenotypically characterized) and in unaffected relatives[358,359]. Deletion of this region is frequently an inherited finding.

4.3.2.4 Discussion 15q11.2 microdeletion syndrome

The 15q11.2 microdeletion syndrome is emerging as the most common copy number variant neurosusceptibility locus.[360] As with the 16p13.11 microduplication, the identification of this

finding in three individuals in an overgrowth cohort is likely to reflect shared phenotypic features, including developmental delay/ID, rather than an association with overgrowth. Tall stature and macrocephaly have not been reported in series of over 200 individuals with 15q11.2 microdeletions[244].

4.3.3 Participants with other findings on microarray

The details of participants with other pathogenic and likely pathogenic microarray findings are described in this section.

4.3.3.1 Results POD 101.0

POD 101.0: Deletion of 1.155 Mb at 5q35.2q35.3(175,329,033-177,484,097)x1 with 26 OMIM referenced genes including *NSD1*. Parental samples not available to confirm de novo status.

POD 101.0 was born at 42 weeks gestation weighing 4.18 kg (+0.18 SD) with an OFC of 36.8 cm (+0.9 SD). She required facial oxygen at birth. She developed neonatal jaundice that did not reach the threshold for treatment. At age five she was 120 cm tall (+2.1 SD), weighed 21.8 cm (+1.1 SD) and had an OFC of 53.0 cm (+1.0 SD). Medical problems included hypermobility, bilateral convergent squint, and conductive hearing impairment due to glue ear. She had global development delay. She sat at seven months, walked at 19 months and spoke her first word at 48 months. She attended a mainstream school with a statement of educational needs. Facial features included midface prominence and deeply set eyes. She had

broad toes and halluces. Her mother was 165 cm tall (+0.2 SD) and her father was 168 cm tall (-1.3 SD) with an OFC of 60 cm (+1.8 SD).

4.3.3.2 Discussion POD 101.0

POD 101.0 does not have the typical facial features of Sotos syndrome and this is likely to be because she has a larger deletion at 5q35.2q35.3 encompassing 25 other OMIM genes in addition to *NSD1*. The absence of macrocephaly may also be due to the effect of other genes in the deletion.

4.3.3.3 Results POD 069.0

POD 069.0: 17q24.2q24.3 arr[hg19](64,222,212-69,107,492)x3 4.93 Mb microduplication encompassing 62 HGNC mapped genes. Parental samples unavailable to confirm de novo status.

POD 069.0 was born at 42 weeks gestation weighing 4.0 kg (+0.4 SD) and required resuscitation at birth. She required admission to the neonatal unit for nine days and had an episode of cyanosis requiring oxygen. She was also reported to have had hypoglycaemia. Further details of the pregnancy and birth history were unknown because she was adopted.

At age 12 she was 160 cm tall (+0.9 D), weighed 85.0 kg (+3.3 SD) and had an OFC of 56.5 cm (+1.5 SD). She was eligible for the study based on previous measurements age three years and nine months when her height was 107.8 cm (+2.0 SD), weight 23.9 kg (+3.1 SD) and

OFC 52.4cm (+1.2 SD). At age three years and eleven months, her bone age was advanced to six years eight months. Age 12 she had not yet reached menarche. Medical problems included mild scoliosis, worn dental enamel, and a visual processing disorder.

She had developmental delay. Assessment at a Child Development Centre age 34 months identified eye and hand coordination equivalent to 24 months (<1st centile), visual matching equivalent to 24 months (<1st centile) and verbal and expressive language in the normal range (20th centile). She attended mainstream school. Behavioural difficulties included aggressive, emotional lability, temper tantrums, anxiety, features of autism spectrum disorder, and obsessive-compulsive behaviour. She also had abnormality of temperature regulation (always hot), hypersensitivity to light and sound, and sleep problems (difficulty falling asleep).

Both biological parents were known to have intellectual disability. Her biological mother's height was reported to be 175 cm (+1.8 SD) and her biological father's height 167 cm (-1.5 SD).

On examination she had mild scoliosis, hyperextensible elbow joints, and pes planus. Dysmorphic features included tapering fingers and toes, broad feet, widely spaced eyes, and a short nose with anteverted nares.

4.3.3.4 Discussion POD 069.0

Duplications of this region are not known to be associated with an overgrowth phenotype. However, the size and gene content of the duplication mean it is likely to be pathogenic. Interestingly, 17q24.2q24.3 microdeletions have been reported to cause growth retardation and microcephaly in addition to developmental delay, specific dysmorphic features and congenital anomalies[361], suggesting an association between the reciprocal duplication and overgrowth is plausible.

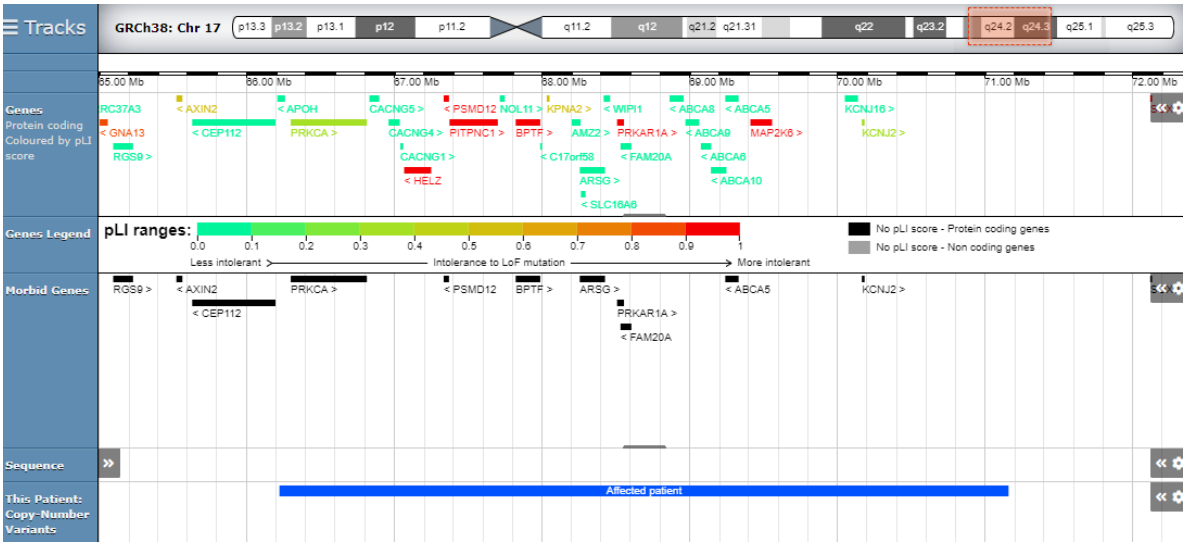


Figure 72: 17q24.2q24.3 duplication in DECIPHER database browser

Table 38: OMIM disease genes in the 17q24.224.3 duplication

Gene	Gene name	Phenotype	Mechanism
APOH	Apolipoprotein H		
PRKCA	Protein kinase C alpha	Candidate gene for autism spectrum disorder[362]	AD
CACNG5	Calcium voltage-gated channel auxiliary subunit gamma 5		
CACNG4	Calcium voltage-gated channel auxiliary subunit gamma 4		
CACNG1	Calcium voltage-gated channel auxiliary subunit gamma 1		
HELZ	Helicase with zinc finger	Candidate gene for intellectual disability[363]	AD
PSMD12	Proteasome 26S subunit, non-ATPase 12	Syndromic neurodevelopmental disorder[364]	AD
PITPNC1	Phosphatidylinositol transfer protein, cytoplasmic 1		
NOL11	Nucleolar protein 11		
BPTF	Bromodomain PHD finger transcription factor	Syndromic neurodevelopmental disorder including microcephaly[365]	AD
KPNA2	Karyopherin subunit alpha 2		
AMZ2	Archaelysin family metallopeptidase	Candidate gene for intellectual disability[366]	AR
SLC16A6	Solute carrier family 16 member 16		
PRKAR1A	Protein kinase cAMP-dependent type 1 regulatory subunit alpha	Carney complex, type 1	AD
WIPI1	WD repeat domain, phospoinositide interacting 1		
FAM20A	Golgi associated secretory pathway pseudokinase	Amelogenesis imperfecta, type 1G (enamel-renal syndrome)	AR
ABCA8	ATP binding cassette subfamily A member 6		
ABCA9	ATP binding cassette subfamily A member 6		
ABCA6	ATP binding cassette subfamily A member 6		
ABCA10	ATP binding cassette subfamily A member 6		
ABCA5	ATP binding cassette subfamily A member 6		
MAP2K6	Mitogen-activated protein kinase 6		
KCNJ16	Potassium voltage-gated channel subfamily J member 16	Anderson-Tawil syndrome	AD activating mutations
KCNJ2	Potassium voltage-gated channel subfamily J member 2		

Further work is needed to identify the gene(s) responsible for the phenotype in POD 069.0.

4.3.3.5 Results POD 060.0

POD 060.0: paternally inherited microduplication 17q11.2(29,311,620-29,316,562)x3~4

CGH microarray in this individual identified a paternally inherited duplication at 17q11.2. The mean log ratio of the imbalance was reported as higher than expected for a duplication (three copies of the region instead of the usual two) and might represent a multiplication with four copies of the region. This multiplication involves and potentially disrupts part of *RNF135*.

RNF135 was reported as an overgrowth gene in 2007 based on six families with variable intellectual disability, tall stature, macrocephaly and autism[177]. The stop gain, frameshift and missense variants described were all inherited from a parent[177]. The five parents for whom data was available were described as mildly affected, with only one having mild intellectual disability and two having only macrocephaly[177].

POD 060.0 has a phenotype of macrocephaly +2.6 SD, obesity, developmental delay and autism that would be consistent with the clinical features reported to be associated with variants in *RNF135*. The multiplication was inherited from the participant's father who has macrocephaly (OFC 62 cm; +3.0 SD), in keeping with the intrafamilial variability described.

4.4.3.6 Discussion POD 060.0

However, *RNF135* was reclassified in 2019 by Wright et al[367] based on three lines of evidence. Firstly, the lack of association between frameshift or stop gain variants in *RNF135*

and developmental traits in the UK BioBank data. Secondly, the ExAC browser gives this gene a pLI 'probability of being loss of function intolerant' score of 0. The pLI is determined from the observed and expected variant counts for a given gene in the ExAC dataset. A pLI value > 0.9 is consistent with a gene that is intolerant of loss of function variants. Finally, there is a lack of de novo *RNF135* variants in the DDD study[367]. Based on this, it was suggested that *RNF135* is not associated with a developmental disorder. *RNF135* is on the 'Red List' on the 100KGP PanelApp as a gene for which there is low evidence of pathogenicity.

The microarray finding in this participant therefore remains of uncertain significance. Testing on the 44 gene panel in this study and by whole genome sequencing in the 100KGP (Intellectual disability panel v.2.654) did not identify an alternative diagnosis.

Chapter 5. RESULTS: Genomic analysis of overgrowth disorders

Molecular genetic diagnoses in POD study participants were established through one of three routes: first line diagnostic molecular genetic testing in a regional genetics laboratory prior to or after recruitment; participation in other studies; or through NGS panel and/or whole exome sequencing in the POD study.

5.1 Molecular data generated outside the study

5.1.1 Clinical diagnostic testing

First line diagnostic testing in regional genetics laboratories included CGH microarray and single gene testing.

5.1.1.1 CGH microarray

CGH microarray is a first line diagnostic molecular genetic test for individuals with generalised overgrowth and developmental delay or intellectual disability.

16 participants had findings on CGH microarray that fully or partially explained their phenotype (see Table 39). Seven individuals from one family had a 17q11.2 (30,318,418-30,326,952) microdeletion. This was initially classified as a variant of uncertain significance, however given it encompasses the overgrowth gene *SUZ12,* it is likely to be pathogenic. A deletion of 5q35.2q35.3 including *NSD1* (POD 101.0) is also pathogenic. Three unrelated

participants had microdeletions at 15q11.2, classified as pathogenic of mild effect size, and three unrelated participants had microduplications at 16p13.11, classified as pathogenic.

A further two participants had microarray findings of interest. POD 60.0 has had microduplication involving the proposed overgrowth gene *RNF135* that is classified as a variant of uncertain significance. POD 069.0 has a large deletion of 17q24.2q24.3 that is classified as likely pathogenic based on its size.

The phenotypes of these participants are described in Chapter 4 section 4.3.

Table 39: Participants with pathogenic variants and likely pathogenic variants on CGH microarray

POD	Variant	Duplication/ Deletion	Relevant genes	Inheritance	Classification
101.0	5q35.2q35.3 (175,329,033-177,484,097)x1	del	*NSD1*	de novo	pathogenic
036.0	15q11.2 (22,765,637-23,217,513)x1	del		paternal	pathogenic of mild effect size
035.0	15q11.2 (22,765,658-23,146,103)x1	del		maternal	pathogenic of mild effect size
072.0	15q11.2 (22,765,637-23,217,513)x1	del		unknown	pathogenic of mild effect size
055.0	16p13.11 (15048732-16194575)x3	dup		paternal	likely pathogenic
073.0	16p13.11 (15,048,732-16,194,575)x3	dup		unknown	likely pathogenic
085.0	16p13.11 (15,048,732-16,194,575)x3	dup		unknown	likely pathogenic
060.0	17q11.2 (29,311,620-29,316,562)x3~4	dup (multiplication)	*RNF135*	paternally	uncertain significance
103.0	17q11.2 (30,318,418-30,326,952)x1	del	*SUZ12*	maternal	uncertain significance
103.1	17q11.2 (30,318,418-30,326,952)x1	del	*SUZ12*	unknown	uncertain significance
103.3	17q11.2 (30,318,418-30,326,952)x1	del	*SUZ12*	maternal	uncertain significance
104.0	17q11.2 (30,318,418-30,326,952)x1	del	*SUZ12*	maternal	uncertain significance
104.1	17q11.2 (30,318,418-30,326,952)x1	del	*SUZ12*	unknown	uncertain significance
104.3	17q11.2 (30,318,418-30,326,952)x1	del	*SUZ12*	maternal	uncertain significance
104.4	17q11.2 (30,318,418-30,326,952)x1	del	*SUZ12*	maternal	uncertain significance
069.0	17q24.2q24.3 (64,222,212-69,107,492)x3	dup		unknown	likely pathogenic

5.1.1.2 Other clinical diagnostic testing

Other molecular diagnoses on clinical testing:

Ten participants in the POD study had molecularly confirmed diagnoses resulting from targeted single gene testing in a regional genetics laboratory prior to recruitment. Following recruitment to the study, a further seven participants underwent targeted single gene testing in a regional genetics laboratory that identified a molecular diagnosis (see table 40).

Table 40: Participants with molecular genetic diagnoses made on diagnostic testing

POD	Gene	Variant	Diagnosis	Prior to/after study entry
017.0	*EZH2*	c.1876 G>A; p.(Val626Met)	Weaver syndrome	prior
046.0	*EZH2*	c.1299C>T	Weaver syndrome	prior
074.0	*NSD1*	c.5279_5282delTCTC; p.(Val1760Glyfs*2)	Sotos syndrome	prior
080.0	*NSD1*	c.1187delC; p.(Pro396Leufs*23)	Sotos syndrome	prior
095.0	*NSD1*	c.4833T>G; p.(Cys1611Trp)	Sotos syndrome	prior
087.0	*PTEN*	Mosaic pathogenic duplication of exon 5	PTHTS	prior
088.0	*PTEN*	c.469G>T; p.(Glu157Ter)	PHTS	prior
088.2	*PTEN*	c.469G>T; p.(Glu157Ter)	PHTS	prior
089.0	*PTEN*	pathogenic duplication of exon 5	PHTS	prior
089.1	*PTEN*	pathogenic duplication of exon 5	PHTS	prior
029.0	*FBN1*	c.4444_4445delGG; p.(Gly1482Argfs*8)	Marfan syndrome	after
052.0	*GLI3*	c.1115C>A; p.(Ser372*)	Grieg syndrome	after
065.0	*NSD1*	intragenic duplication exons 11-22	Sotos syndrome	after
061.0		hypermethylation at H19/IGF2:IG-DMR within 11p15.5 on buccal sample*	Beckwith-Wiedemann syndrome	after
053.0	*PIK3CA*	c.1357G>A p.(Glu453Lys) in tissue	PROS	after
072.0	*PIK3CA*	c.1093G>A; p.(Glu365Lys) in tissue	PROS	after
081.0	*PIK3CA*	c.2740G>A; p.(Gly914Arg) in tissue	PROS	after

*initial testing on DNA extracted from blood lymphocytes did not detect any changes and testing was repeated twice on buccal samples

The most common monogenic disorders diagnosed in participants prior to recruitment were PHTS (five participants from two families), Sotos syndrome (three unrelated individuals), and Weaver syndrome (two unrelated participants).

Seven participants had clinical features consistent with a specific disorder for which clinical diagnostic testing was available. These participants underwent testing and were diagnosed with PROS (three individuals), Marfan syndrome, Grieg syndrome, BWS, and Sotos syndrome.

5.1.2 Molecular diagnoses made through other studies

Some participants were recruited to both POD and to other projects. Six participants had molecular diagnoses in other studies, one in the Segmental Overgrowth Study (SOS) study, two in the Deciphering Developmental Disorders (DDD) study, and two in the 100,000 Genomes Project (100KGP). POD 013.0 was identified to have a somatic pathogenic variant in *PIK3CA* by ISO and this was confirmed on NGS panel testing in the POD study. POD 035.0 was found to have a pathogenic variant in *PTCH1* by DDD and this was confirmed on NGS panel testing in the POD study. POD 016.0 was found to have a pathogenic variant in *MAGED2* by DDD and this was confirmed on review of whole exome sequencing data in the POD study.

Table 41: Participants with molecular diagnoses made through other studies

POD	Study	Gene	Variant	Diagnosis
013.0	ISO	*PIK3CA*	c.2740G>A p.(Gly914Arg) in tissue	PROS
035.0	DDD	*PTCH1*	c.2611_2624del; p.(Asn871Trpfs*20)	Gorlin syndrome
048.0	DDD	*HIST1H1E*	p.(Lys148Glnfs*48)	HIST1H1E syndrome
016.0	DDD	*MAGED2*	c.1085+1G>A	transient neonatal Bartter syndrome
054.0	100KGP	*TMCO1*	Homozygous c.233G>A; p.(Trp78*)	Cerebrofaciothoracic dysplasia
067.0	100KGP	*THRA*	c.1195C>T; p.(Pro399Ser)	Resistance to thyroid hormone alpha

5.2 NGS panel of overgrowth genes

Molecular diagnoses were made in POD study participants through testing on an NGS panel of overgrowth genes and/or whole exome sequencing. The initial 20 gene panel (panel v.1) was subsequently redesigned to include 44 overgrowth genes (panel v.2) as novel genes were published in the literature.

60 samples from 57 participants underwent testing on an overgrowth panel. Two participants with regional overgrowth (POD 012.0 and POD 013.0) had multiple samples tested. POD 012.0 had testing on DNA extracted from blood, saliva and fibroblasts generated from skin biopsy and POD 013.0 had testing on DNA extracted from blood and fibroblasts generated from skin biopsy. All other participants had testing on DNA samples extracted from blood.

22 samples from 19 participants were tested on panel v.1. 42 samples from 42 participants were tested on panel v.2. Four participants (POD 006.0, POD 020.0, POD 050.0, and POD 064.0) were initially tested on panel v.1 and proceeded to have testing on panel v.2.

Seven different molecular diagnoses were made in seven individuals (see tables 42 and 43). The diagnostic rate for panel v.1 was three diagnoses in 19 participants tested (16%) and the diagnostic rate for panel v.2 was four diagnoses out of 42 participants tested (10%). The overall diagnostic rate for panel testing was seven diagnoses out of 58 participants tested (12%).

Table 42: Molecular diagnoses made on 20 gene panel (panel v.1)

Participant	Gene	Variant	Classification	Diagnosis
068.0	*DNMT3A*	c.499C>T p.(Gln167*)	likely pathogenic	TBRS
002.0	*NFIX*	c.248T>G p.(Ile83Ser)	likely pathogenic	Malan syndrome
013.0	*PIK3CA*	c.2740G>A p.(Gly914Arg) in tissue	pathogenic	PROS

Table 43: Molecular diagnoses made on 44 gene panel (panel v.2)

Participant	Gene	Variant	Classification	Diagnosis
028.0	*NSD1*	c.5791T>C; p.(Cys1931Arg)	likely pathogenic	Sotos syndrome
064.0	*PDGFRB*	c.1751C>G; p.(Pro584Arg)	pathogenic	Kosaki overgrowth syndrome
030.0	*PPP2R5D*	c.598G>A; p.(Glu200Lys)	pathogenic	*PPP2R5D*-related neurodelelopmental disorder
035.0	*PTCH1*	c.2611_2624del; p.(Asn871Trpfs*20)		Gorlin syndrome

5.2.1 POD 068.0 *DNMT3A*

POD068.0: *DNMT3A* c.499C>T; p.(Gln167*)

A stop gain variant c.499C>T predicted to result in a premature stop codon p.(Gln167*) was identified in exon 5 of *DNMT3A*. This variant is not reported in population databases (dbSNP, ESP, ExAC, gnomAD). Pathogenic variants (missense, truncating and frameshift) in *DNMT3A* cause Tatton-Brown-Rahman syndrome (TBRS). This truncating variant is located in the functional ATRX-Dnmt3-Dnmt3L (ADD) domain where other nonsense variants have been reported in TBRS[89]. The participant's phenotype of overgrowth and intellectual disability are consistent with a diagnosis of TBRS. Parental samples were not available to confirm de novo status. The variant was classed as likely pathogenic according to ACMG guidelines.

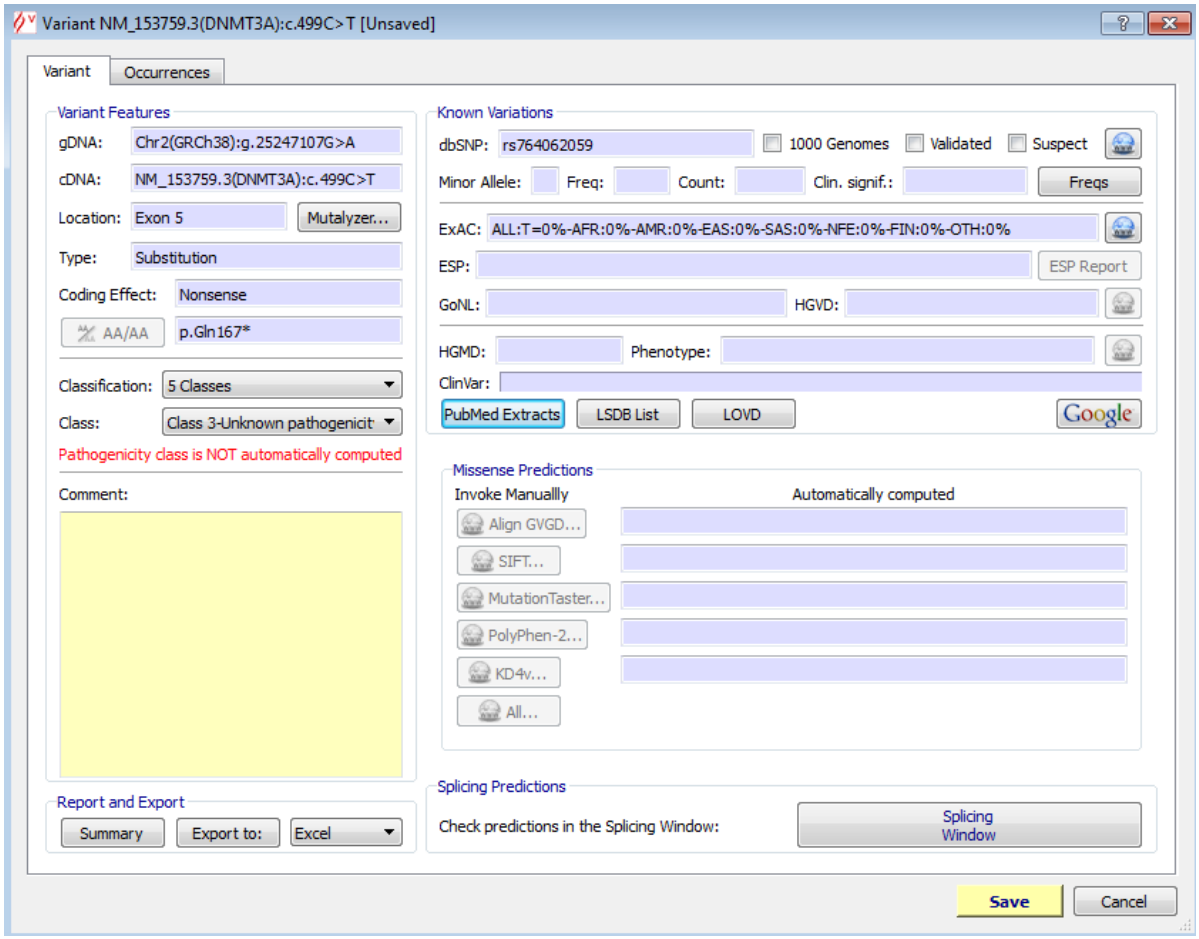


Figure 73: Screenshot of DNMT3A c.499C>T in Alamut Visual

5.2.2 POD 002.0 *NFIX*

POD 002.0: *NFIX* c.248T>G; p.(Ile83Ser)

A missense variant c.248T>G p.(Ile83Ser) was identified in exon 2 of *NFIX*. This variant is not previously reported in the literature and is absent from the population databases dbSNP, 1000Genomes, ExAC, gnomAD and ESP. It is predicted to be deleterious by *in silico* tools SIFT, MutationTaster and PolyPhen2 and is located in the DNA binding domain where other pathogenic variants have been reported [9,134,368]. Sanger sequencing of the proband and both parents by the WMRGL confirmed the variant to be de novo. The participant's phenotype of generalised overgrowth and intellectual disability is consistent with a diagnosis of Malan syndrome. Classification according to the ACMG guidelines is Class 4, likely pathogenic.

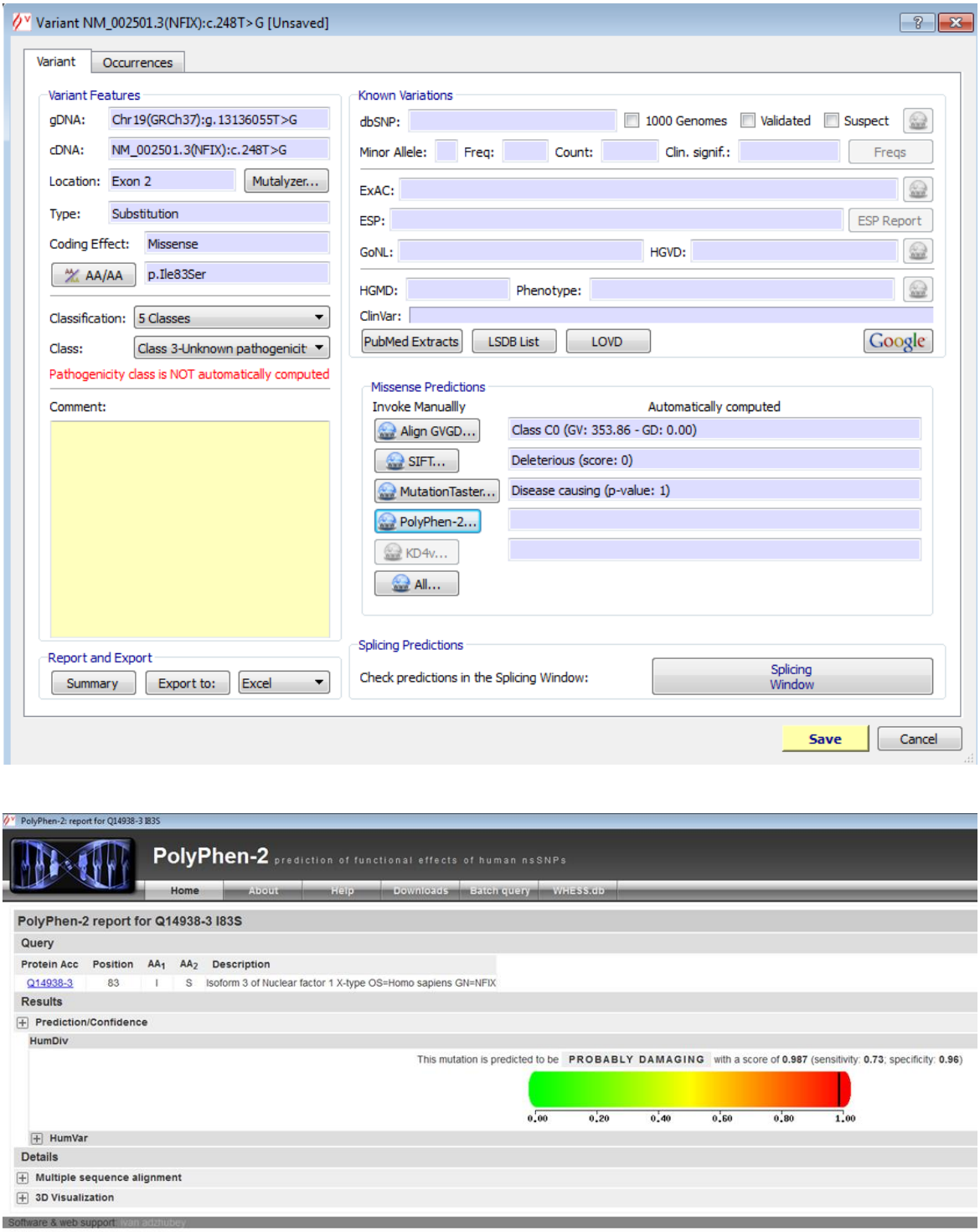


Figure 74: Screenshot of NFIX c.248T>G in Alamut Visual and PolyPhen-2

5.2.3 POD 013.0 *PIK3CA*

POD 013.0: *PIK3CA* somatic c.2740G>A p.(Gly914Arg) in tissue

Panel testing of DNA extracted from blood did not identify any pathogenic variants. However, panel testing of DNA extracted from tissue (fibroblasts from skin biopsy) identified a missense variant c.2740G>A in exon 19 of PIK3CA. This variant is absent from population databases, is predicted to be pathogenic by the *in silico* tools SIFT, MutationTaster and PolyPhen-2, and has been previously reported in several individuals with PROS (MCAP phenotype)[103]. This variant is classified as pathogenic.

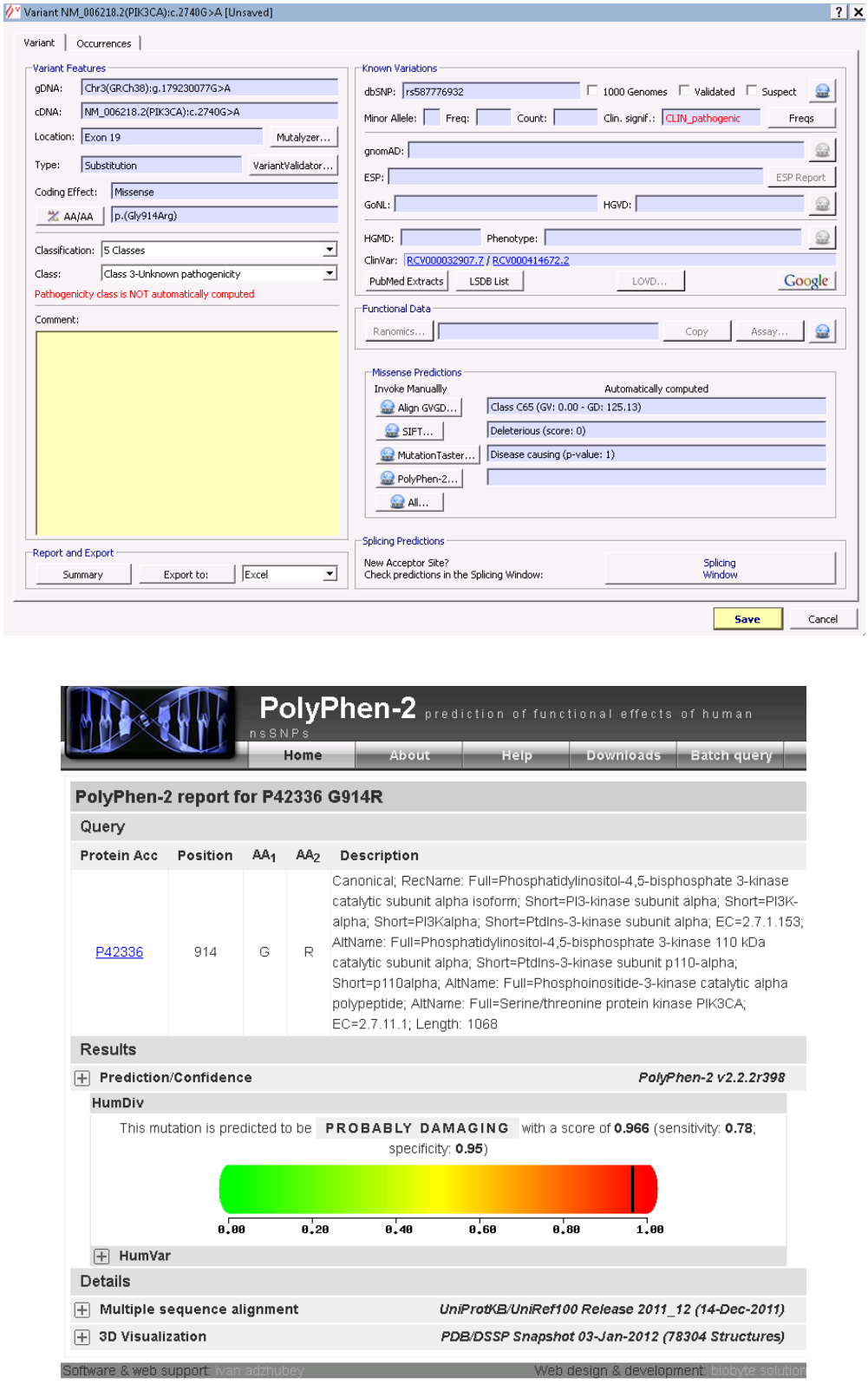


Figure 75: Screenshots of PIK3CA c.2740G>A in Alamut and PolyPhen-2

5.2.4 POD 028.0 *NSD1*

POD 028.0: *NSD1* c.5791T>C; p.(Cys1931Arg)

A missense variant c.5791T>C p.(Cys1931Arg) was identified in exon 19 of *NSD1*. This is absent from population databases dbSNP, 1000G, gnomAD and ESP. This is predicted to result in the substitution of a highly conserved cysteine residue with an arginine and predicted to be deleterious by *in silico* tools SIFT, Mutation Taster, and PolyPhen2. Confirmatory Sanger sequencing by the WMRGL of the proband and both parents confirmed the variant is de novo. The participant's phenotype of tall stature, developmental delay, seizures and congenital kidney anomaly is consistent with a diagnosis of Sotos syndrome. According to the ACMG guidelines, this variant is Class 4 likely pathogenic.

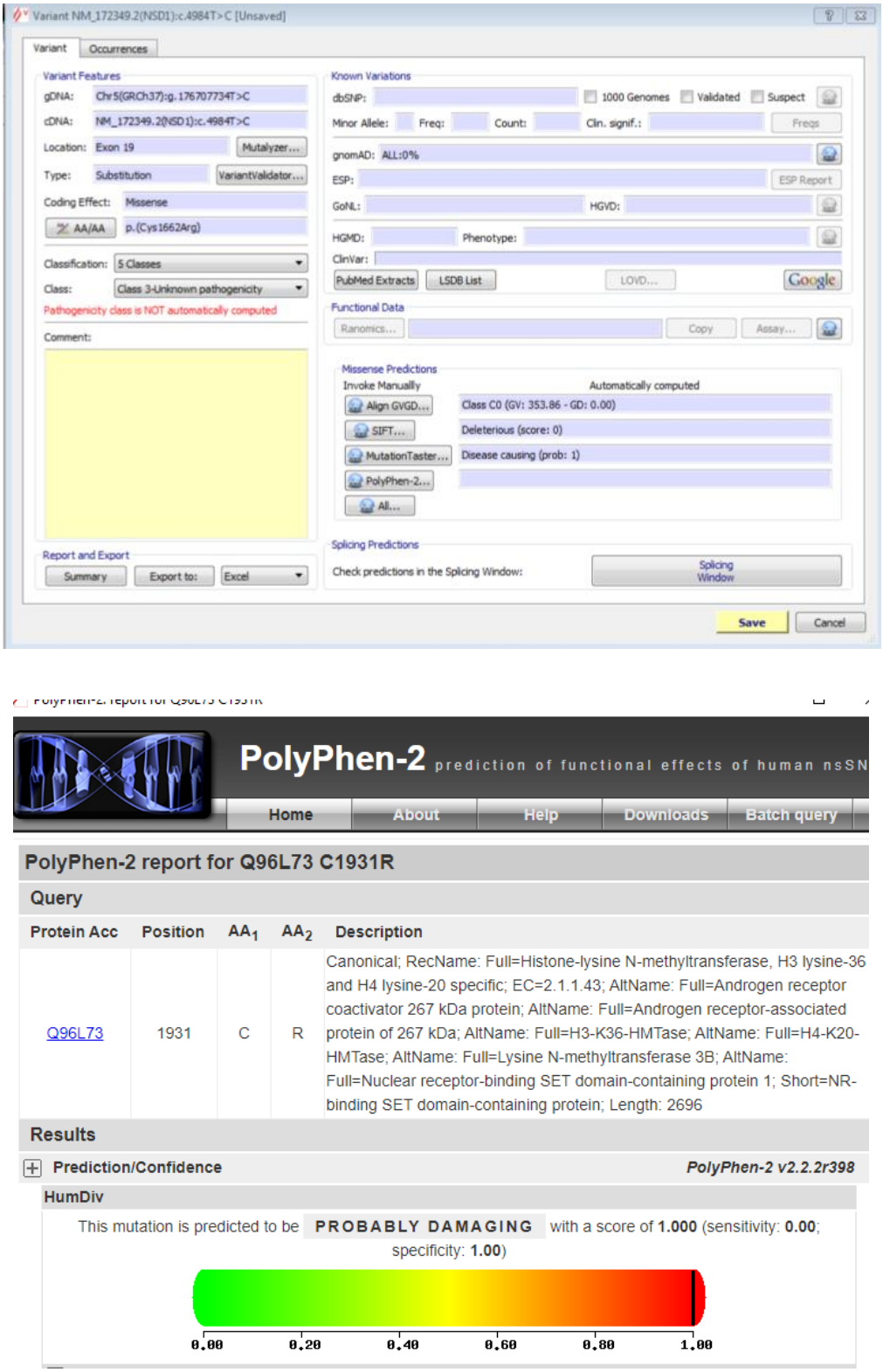


Figure 76: Screenshots of NSD1 c.5791T>C in Alamut and Polyphen2

5.2.5 POD 064.0 *PDGFRB*

POD 064.0: *PDGFRB* c.1751C>G; p.(Pro584Arg)

A heterozygous missense variant c.1751C>G was identified in exon 12 of *PDGFRB*. This variant is absent from population databases dbSNP, 1000G, gnomAD and ESP. *In silico* tools give conflicting predictions of pathogenicity, with SIFT giving a score of 0.08 (tolerated) and both MutationTaster and PolyPhen-2 giving scores of 1 (disease causing/probably damaging). However this variant has been reported in three unrelated individuals in the literature with Kosaki overgrowth syndrome[104,132] and functional studies indicate it has a damaging effect on protein function[369]. Parental studies performed by the WMRGL confirmed this variant to be de novo.

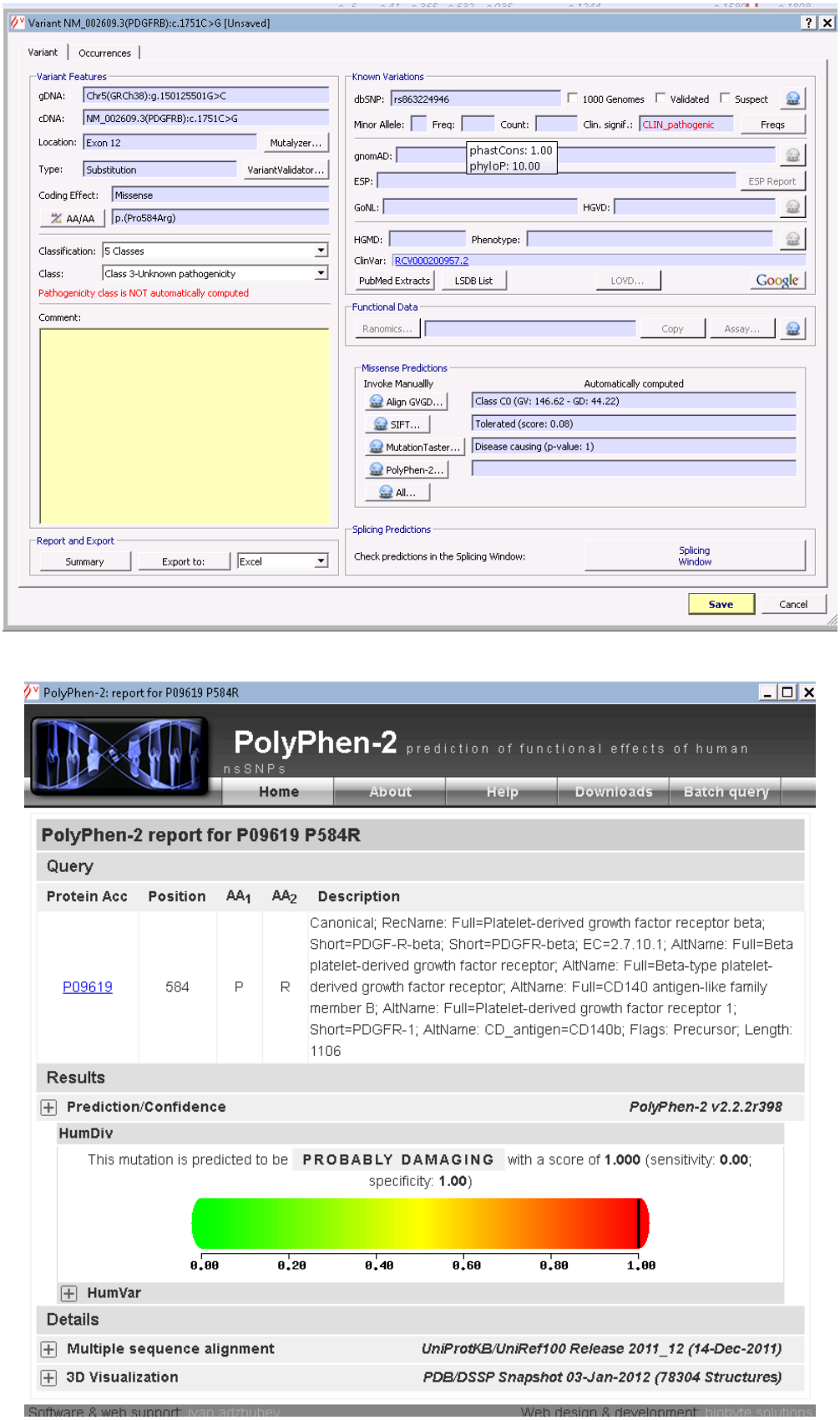
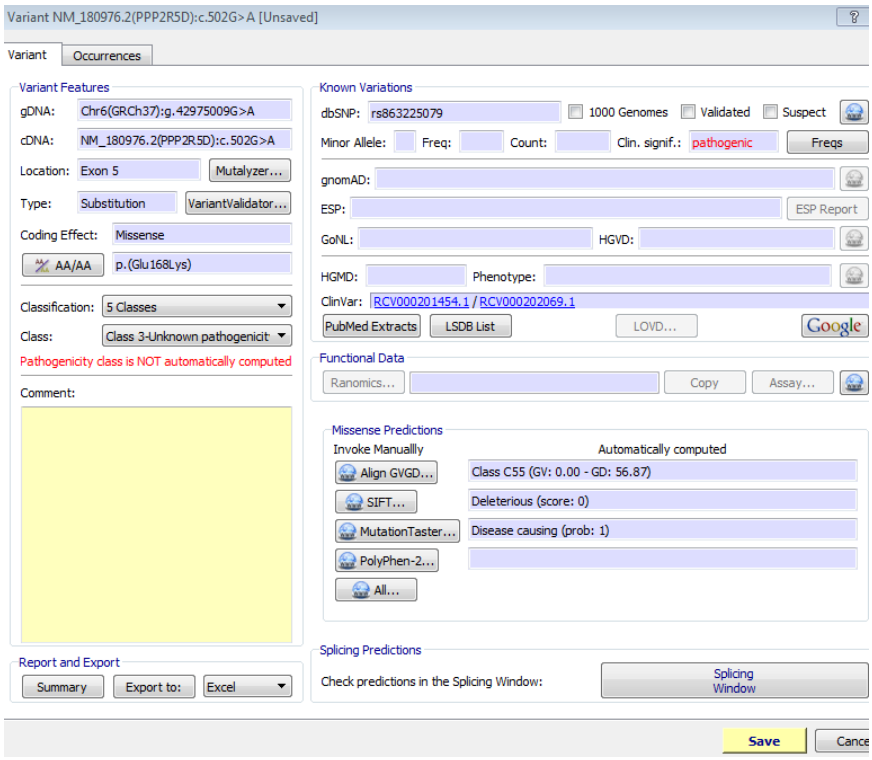


Figure 77: Screenshots of PDGFRB c.1751C>G in Alamut and PolyPhen-2

5.2.6 POD 030.0 *PPP2R5D*

POD 030.0: *PPP2R5D* c.598G>A; p.(Glu200Lys)

A missense variant c.982G>A p.(Glu200Lys) was identified in exon 5 of *PPP2R5D.* This variant is absent from the population databases dbSNP, 1000G, gnomAD and ESP, is predicted to be pathogenic by *in silico* tools SIFT and MutationTaster, and has been previously reported as pathogenic in five individuals in the literature[184,185,249]. The participant's phenotype of macrocephaly and global developmental delay is consistent with PPP2R5D-related neurodevelopmental disorder. This variant is classified as pathogenic.



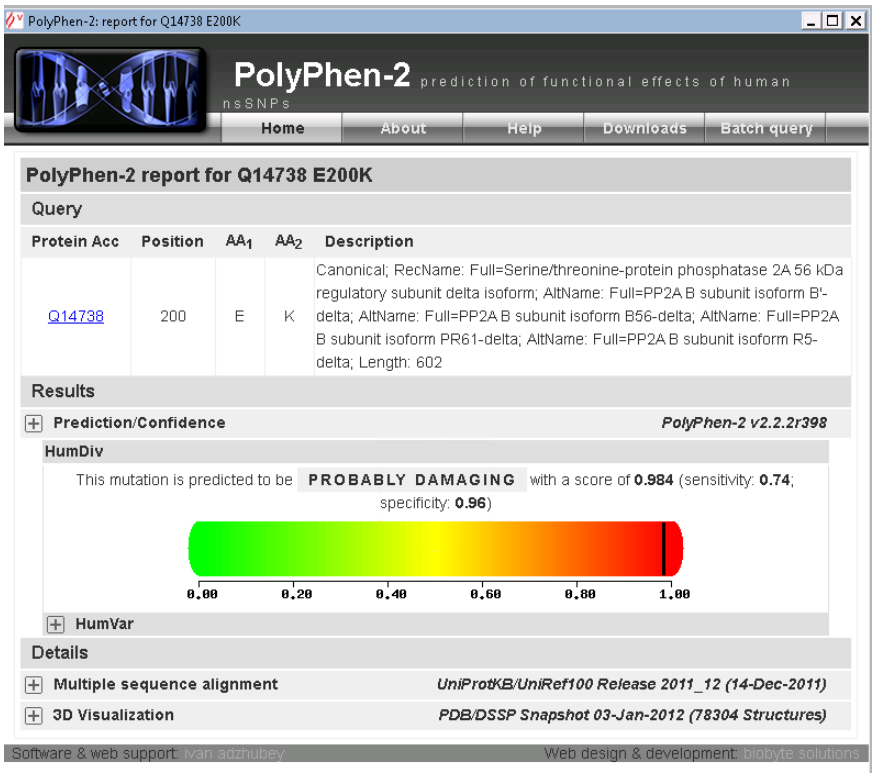


Figure 78: Screenshots of PPP2R5D c.598G>A in Alamut and PolyPhen-2

5.2.7 POD 035.0 *PTCH1*

POD 035.0: *PTCH1* c.2611_2624del; p.(Asn871Trpfs*20)

A heterozygous frameshift variant was identified in *PTCH1*. This variant is absent from population databases dbSNP, 1000G, gnomAD and ESP. Pathogenic mutations in *PTCH1* cause Gorlin syndrome. This participant has clinical features that would be consistent with this diagnosis. Parental studies performed by the WMRGL confirmed this variant to be de novo. This variant is classified as pathogenic.

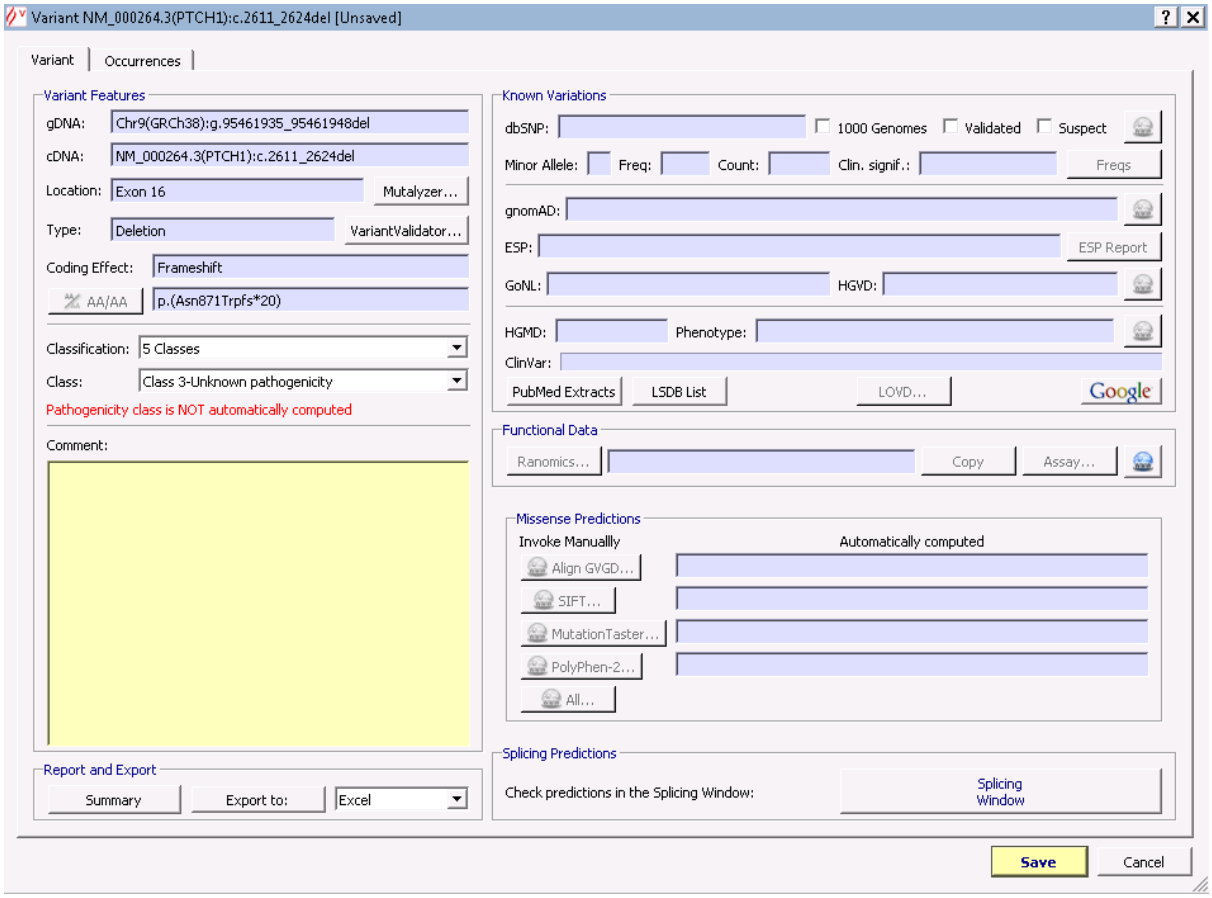


Figure 79: Screenshot of PTCH1 c.2611_2624del in Alamut Visual

5.3 Whole Exome Sequencing

16 participants proceeded to have whole exome sequencing following a panel-negative result. 14 participants had whole exome sequencing as a first line NGS investigation (see Table: Summary of NGS investigations in the POD study). Ten participants (005.0, 090.1, 010.0, 014.0, 022.0, 038.0, 039.0, 041.0, 051.0, and 070.0) underwent whole exome sequencing as trios (ten participants and 20 parent participants) and 20 participants underwent whole exome sequencing as singletons. All participants had generalised overgrowth except 094.0 who had regional overgrowth. DNA tested was extracted from blood samples in all cases.

Pathogenic variants were identified in six genes, *CHD8*, *DNMT3A*, *FBN1*, *FOXP2*, *KTM5B* and *MAGED2*, in seven individuals (see Table 44). Only two of these genes, *CHD8* and *DNMT3A*, are overgrowth genes on the v.2 overgrowth panel. *FBN1*, *FOXP2*, and *KTM5B* are on the DDDG2P list of genes. *MAGED2* is not on the DDG2P list.

Molecular diagnoses were made in three out of ten trios (30%) and four out of 20 singletons (20%). The overall diagnostic rate was seven diagnoses in 30 participants (23%).

Table 44: Molecular diagnoses made on exome sequencing

POD	Trio/ Singleton	Gene	Variant	Type of variant	Inherited/de novo	Classification	Diagnosis
077.0	singleton	*DNMT3A*	c.993delC; p.(Phe331LeufsTer14)	frameshift	unknown	pathogenic	TBRS
008.0	singleton	*FBN1*	c.1761dupT; p.(Ile588TyrfsTer3)	frameshift	unknown	pathogenic	Marfan syndrome
085.0	singleton	*KMT5B*	c.2347C>T; p.(Arg783Ter)	stop gain	unknown	likely pathogenic	KMT5B syndrome
016.0	singleton	*MAGED2*	c.1085+1G>A	splice site	unknown	likely pathogenic	Bartter syndrome type 5 (antenatal, transient)
009.0	trio	*CHD8*	c.716delA; p.(Lys239ArgfsTer22)	frameshift	de novo	pathogenic	CHD8 overgrowth syndrome
005.0	trio	*FBN1*	c.247+1G>A	splice site	de novo	pathogenic	Marfan syndrome
038.0	trio	*FOXP2*	c.982C>T; p.(Arg328Ter)	stop gain	de novo	pathogenic	FOXP2-related speech and language disorder

Variant in *MAGED2* initially identified by DDD

5.3.1 POD 077.0 *DNMT3A*

POD 077.0: c.993delC; p.(Phe331LeufsTer14) *DNMT3A*

This participant underwent whole exome sequencing as a singleton. Filtering for heterozygous variants on the virtual 44 gene overgrowth panel identified one variant. This frameshift variant c.993delC; p.(Phe331LeufsTer14) in *DNMT3A* results in premature truncation of the protein and is absent from population databases including ExAC, UK10K and 1000G. Loss of function variants in *DNMT3A* are associated with Tatton-Brown-Rahman syndrome (TBRS)[88]. This participant's phenotype of overgrowth and intellectual disability is consistent with this disorder.

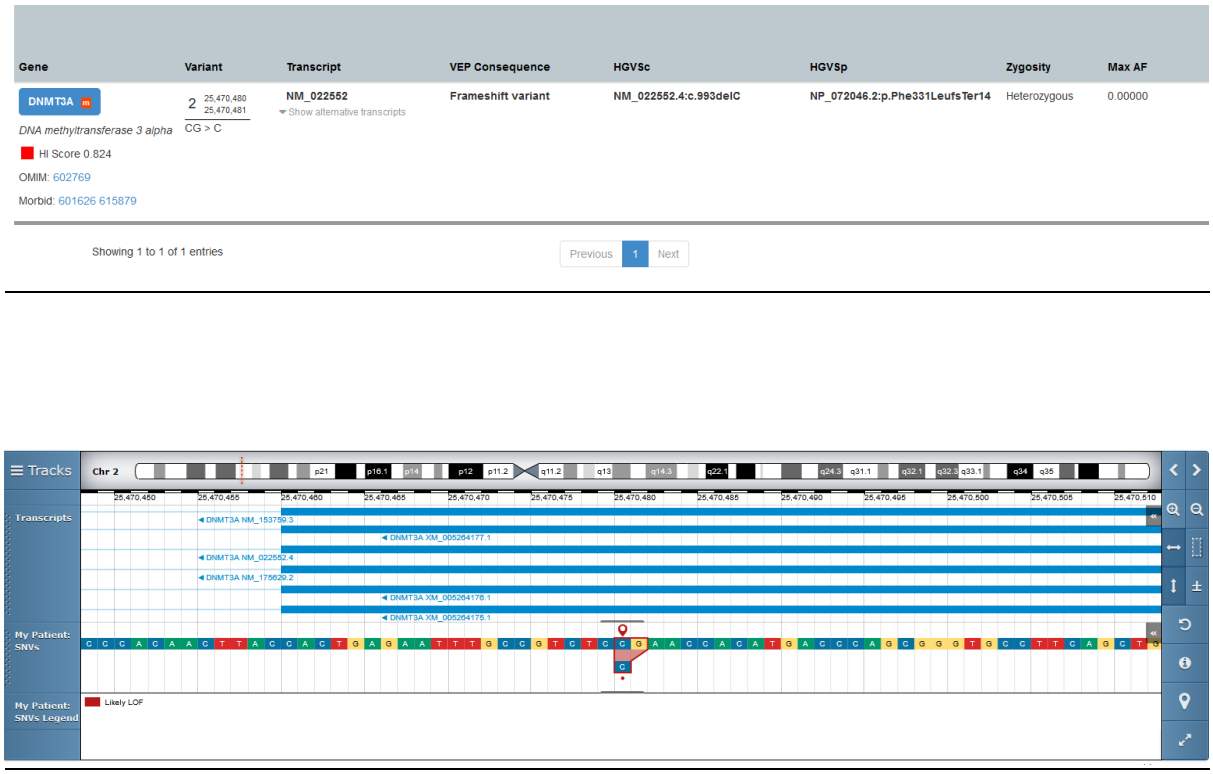


Figure 80: Screenshots of DNMT3A variant in Congenica

5.3.2 POD 008.0 *FBN1*

POD 008.0: *FBN1* c.1761dupT; p.(Ile588TyrfsTer3)

This participant underwent whole exome sequencing as a singleton. Filtering for heterozygous variants on the virtual 44 gene overgrowth panel identified one variant. This was an in-frame deletion in *CDKN1C* reported by ClinVar as likely benign and therefore unlikely to be responsible for this participant's phenotype. Filtering for homozygous or hemizygous variants on the virtual overgrowth panel identified no variants. A third analysis filtering for heterozygous frameshift or stop gain variants in the DDG2P gene list identified two variants. The first was a deep intronic variant in *CRYAA*, a gene associated with autosomal recessive cataracts. Heterozygous variants are not thought to have a phenotype and this variant is not consistent with the participant's phenotype. The second was a frameshift c.1761dupT p.(Ile588TyrfsTer3) variant in *FBN1*. Loss of function variants in FBN1 are associated with Marfan syndrome and this variant is classified as pathogenic.

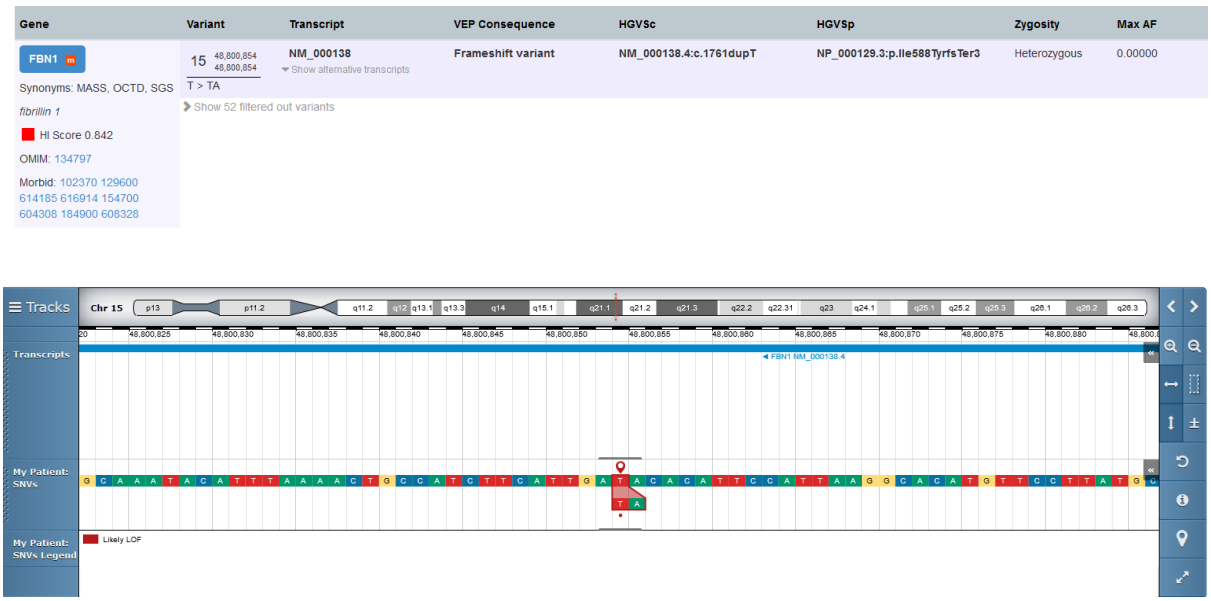


Figure 81: Screenshots of FBN1 variant in Congenica

5.3.3 POD 085.0 *KMT5B*

POD 085.0: *KMT5B* c.2347C>T; p.(Arg783Ter)

This participant underwent whole exome sequencing as a singleton. Initial filtering for heterozygous, homozygous or hemizygous variants on the virtual overgrowth panel identified no variants. A second analysis for heterozygous frameshift or stop gain variants in the DDG2P gene list identified two variants. The first was a frameshift 5 prime UTR variant in *ALDH7A1*, a gene associated with autosomal recessive pyridoxine-deficient epilepsy. The second was a stop gain variant in *KMT5B* c.2347C>T; p.(Arg783Ter). This variant is absent from the population databases ExAC, UK10K and 1000G and is reported by DECIPHER as a likely pathogenic variant. *KMT5B* has recently been identified as a novel OGID gene[370]. This variant is classified as likely pathogenic.

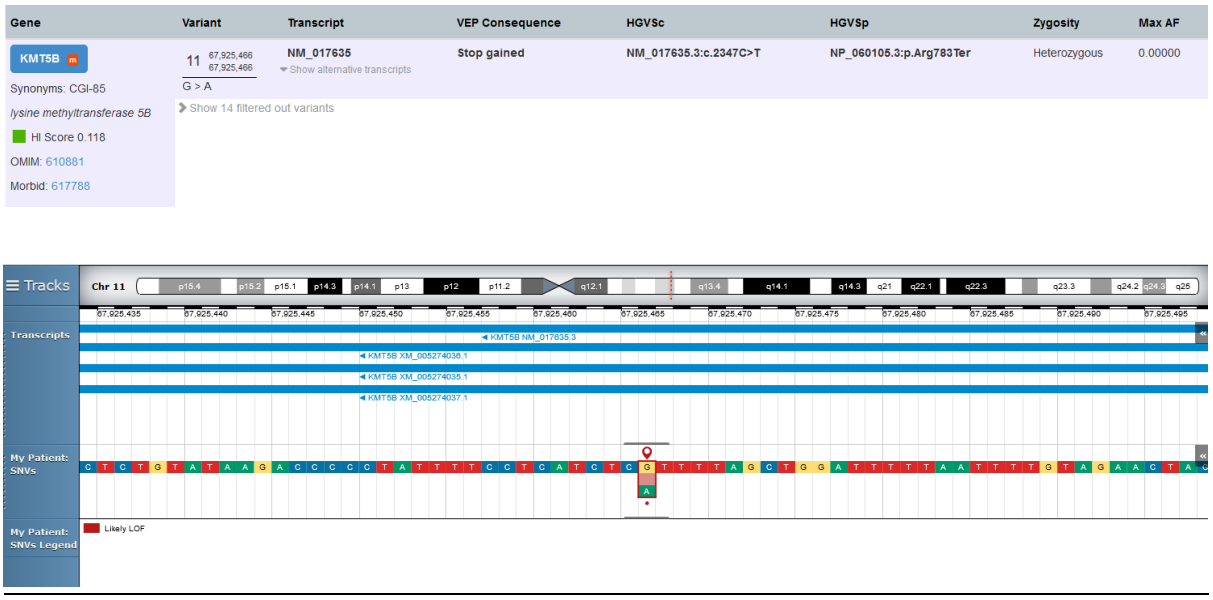


Figure 82: Screenshot of KMT5B variant in Congenica

5.3.4 POD 016.0 *MAGED2*

POD 016.0 *MAGED2* c.1085+1G>A

This participant underwent exome sequencing as a singleton. Initial filtering for heterozygous variants in the virtual 44 gene overgrowth panel did not identify any variants. Filtering for homozygous and hemizygous variants on the virtual overgrowth panel identified homozygous deep intronic variants in *RNF135*. Filtering for heterozygous, homozygous, and hemizygous stop gain, frameshift, or splice acceptor/donor variants in the DDG2P list of genes identified one variant, a heterozygous splice acceptor variant in *GJC2*. Homozygous variants in this gene are associated with recessive hypomyelinating leukodystrophy and spastic paraplegia. Heterozygous 143C-T variants are associated with autosomal dominant hereditary lymphoedema.

Following identification of the c.1085+1G>A variant in *MAGED2* by DDD, filtering was performed for hemizygous splice donor variants in whole exome data. One variant, a splice donor variant c.1085+1G>A in *MAGED2,* was identified. This variant is absent from the population databases ExAC, UK10K and 1000G. Pathogenic variants in this gene are associated with Bartter syndrome type 5, with clinical features of polyhydramnios, prematurity, and salt-wasting and polyuria in the neonatal period. This is consistent with this participant's clinical history of antenatal polyhydramnios and polyuria with presumed diabetes insipidus as a neonate. It does not explain his macrocephaly, developmental issues or bowel problems. This variant therefore partially accounts for his phenotype.

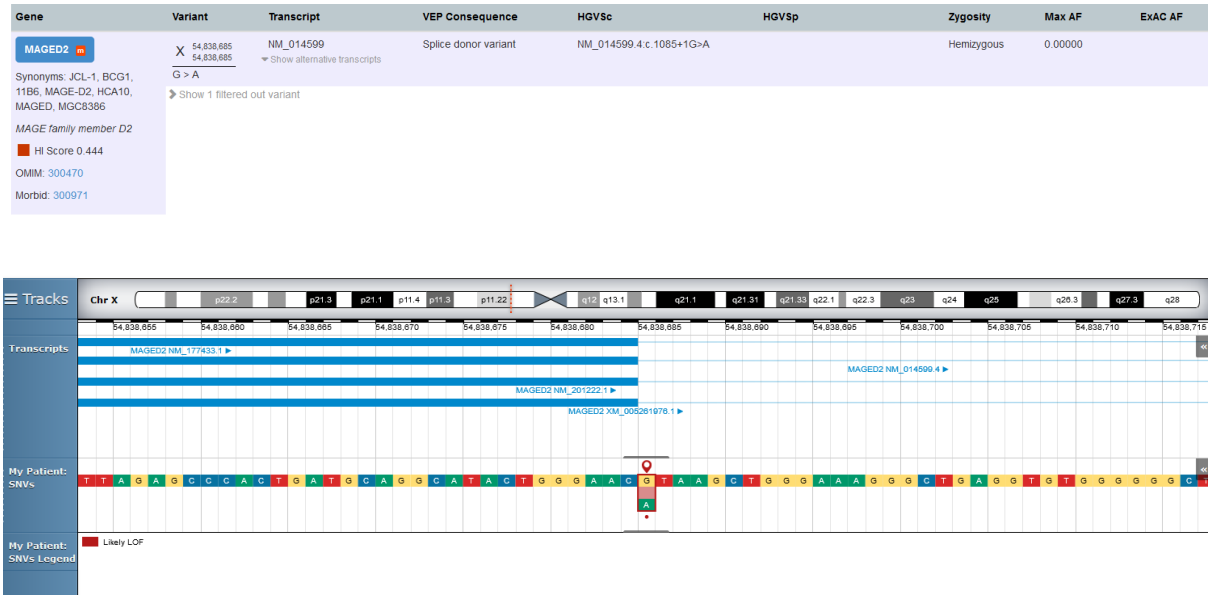


Figure 83: Screenshot of MAGED2 variant in Congenica

5.3.5 POD 009.0 *CHD8*

POD 009.0 *CHD8* c.716delA; p.(Lys239ArgfsTer22)

This participant was initially tested on the 20 gene panel before undergoing trio whole exome sequencing. Filtering for de novo heterozygous variants on the virtual 44 gene overgrowth gene panel identified one variant, a de novo frameshift c.716delA p.(Lys239ArgfsTer22) variant in exon 4 of *CHD8.* This variant is absent from the population databases ExAC, UK10K and 1000G. Pathogenic loss of function variants in *CHD8* cause CHD8 overgrowth disorder (also called CHD8-related developmental disorder) associated with macrocephaly (80%), tall stature (85%), intellectual disability (60%) and GI problems (80%) (OMIM #615032). This participant's clinical features of tall stature and developmental delay are consistent with this diagnosis. Parental studies performed by the WMRGL confirmed this variant was de novo. This variant is classified as pathogenic.

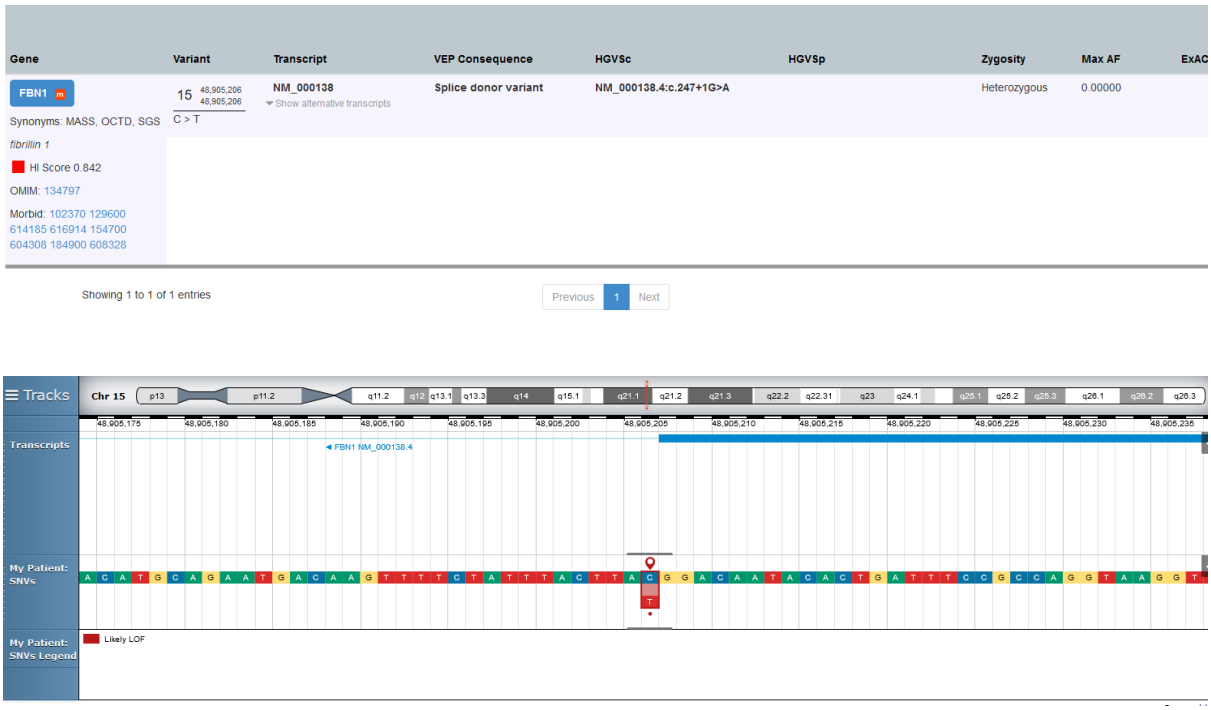
Depth	**76**
Reads split	**46/30**
Quality score	**859.55**
QC status	PASS

Figure 84: Screenshots of POD 009.0 CHD8 c.716delA; p.(Lys239ArgfsTer22) in Congenica

5.3.6 POD 005.0 *FBN1*

POD 005.0: *FBN1* c.247+1G>A

POD 005.0 underwent trio exome analysis. Filtering for de novo heterozygous, homozygous and hemizygous variants on the 44 gene overgrowth panel did not identify any variants. Filtering for de novo heterozygous variants in the DDG2P list of genes identified one variant, a splice donor variant c.247+1G>A in exon 3 of *FBN1*. This variant is absent from population databases ExAC, UK10K and 1000G. It is reported as pathogenic by ClinVar. This variant lies within the conserved splice donor site and has previously been reported in the literature in a patient with Marfan syndrome[371]. Analysis of this variant by Guo et al. demonstrated splicing out of exon 3, resulting in the creation of a frameshift and premature termination codon[371]. Further functional work with in vitro analysis showed this variant caused reduced synthesis and deposition of fibrillin[371]. This variant is classified as pathogenic.



Depth	**112**
Reads split	**58/54**
Quality score	**1698.48**
QC status	PASS

Figure 85: Screenshots of POD 005.0: FBN1 c.247+1G>A in Congenica

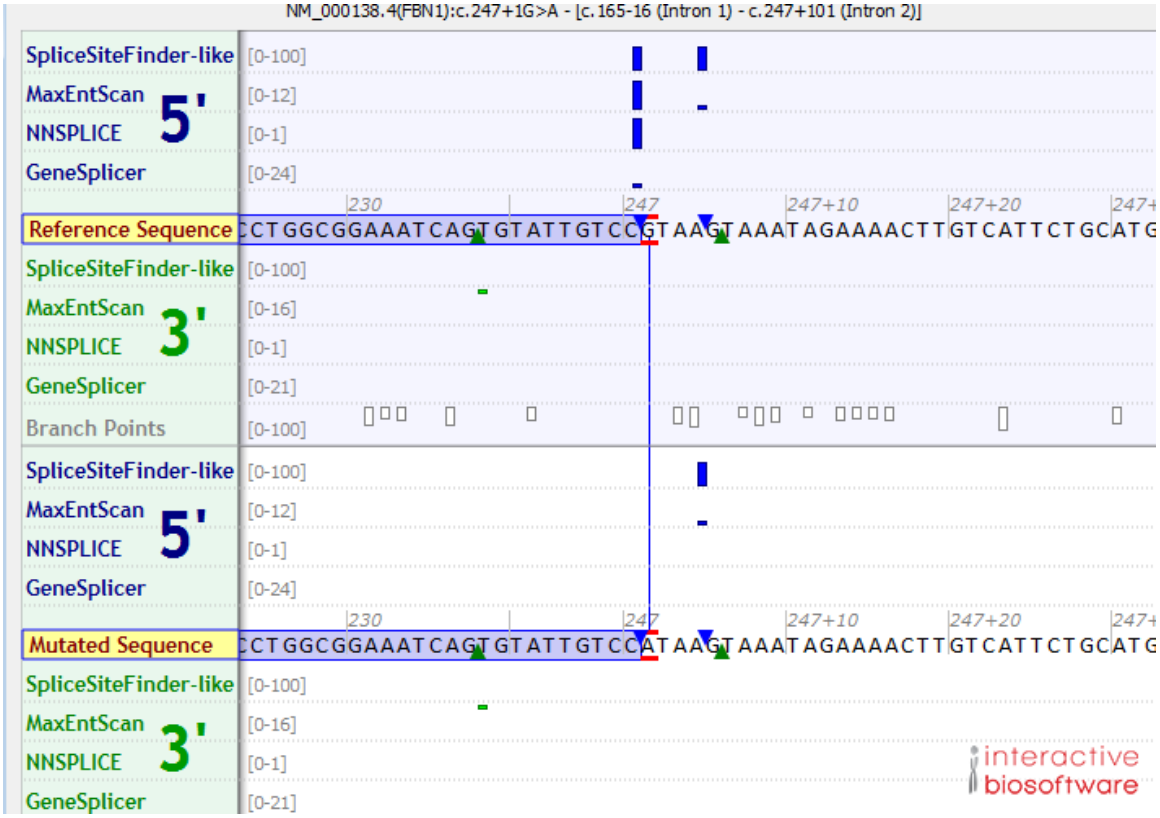


Figure 86: Screenshot from Alamut Visual showing effect of the FBN1 c.247+1G>A variant on the splice site

5.3.7 POD 038.0 *FOXP2*

POD 038.0 underwent trio exome analysis. Filtering for variants on the virtual 44 gene overgrowth panel did not identify any variants. Filtering for de novo heterozygous variants on the DDG2P panel identified a stop gain variant c.982C>T p.Arg328Ter in *FOXP2*. This variant is absent from population databases ExAC, 1000G and UK 10K, is listed as pathogenic in ClinVar, and has been previously reported in the literature as causing *FOXP2*-related speech and language disorder[305]. This variant is classified as pathogenic.

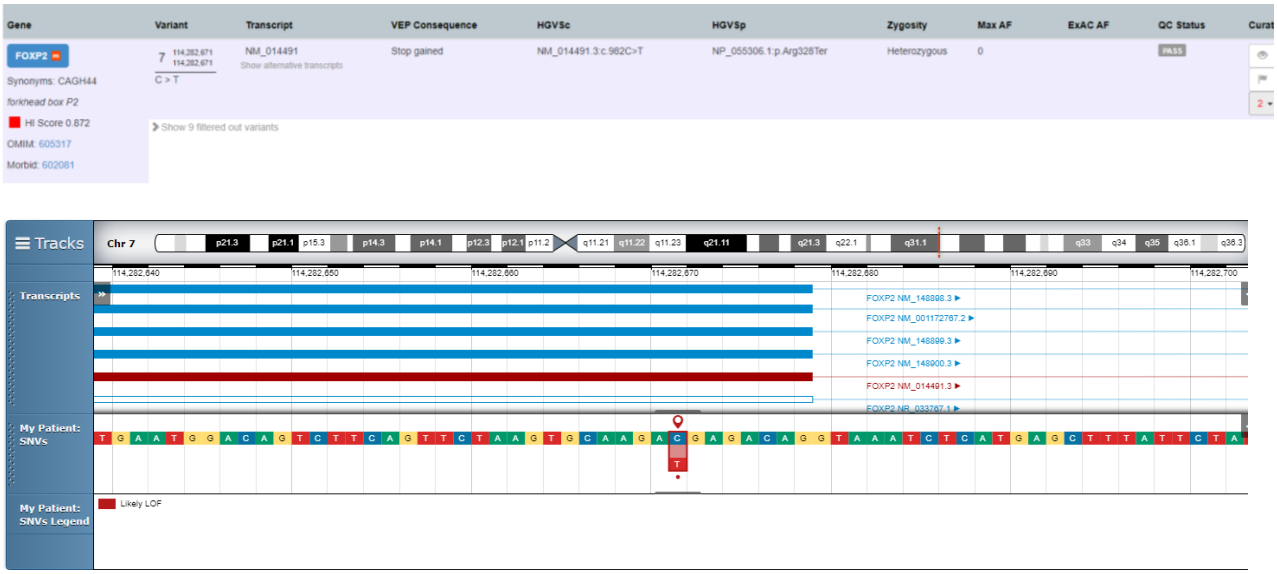


Figure 87: Screenshots of POD 038.0: FOXP2 c.982C>T in Congenica

5.4 Summary of molecular diagnoses

5.4.1 Diagnoses made on testing in the POD study

In total, 71 individuals underwent NGS investigations in the POD study through either panel testing and/or whole exome sequencing. NGS identified a molecular diagnosis in 14 individuals (20%) that fully (13 individuals, 18%) or partially (one individual, POD 016.0) explained their clinical features (see Table 45).

Table 45: Summary of NGS performed in the POD study

POD	Panel v.1	Panelv.2	Singleton WES	Trio WES	Gene	Diagnosis
6	-	-				
20	-	-				
50	-	-				
64	-	+			*PDGFRB*	Kosaki overgrowth syndrome
3	-		-			
4	-		-			
11	-		-			
15	-		-			
19	-		-			
8	-		+		*FBN1*	Marfan syndrome
16	-		+		*MAGED2*	Transient antenatal Bartter syndrome
10	-			-		
14	-			-		
5	-			+	*FBN1*	Marfan syndrome
9	-			+	*CHD8*	CHD8 overgrowth syndrome
12	-					
2	+				*NFIX*	Malan syndrome
13	+				*PIK3CA*	PROS
68	+				*DNMT3A*	Tatton-Brown-Rahman syndrome
39		-		-		
41		-		-		
51		-		-		
70		-		-		
38		-		+	*FOXP2*	*FOXP2*-related speech and language disorder
1		-				
18		-				
21		-				
29		-				
31		-				
32		-				
33		-				
34		-				
36		-				
37		-				
40		-				
42		-				
43		-				
45		-				
47		-				

48		-				
49		-				
52		-				
53		-				
54		-				
55		-				
56		-				
57		-				
60		-				
61		-				
62		-				
63		-				
67		-				
69		-				
82		-				
28		+			*NSD1*	Sotos syndrome
30		+			*PPP2R5D*	PPP2R5D-related neurodevelopmental disorder
35		+			*PTCH1*	Gorlin syndrome
71			-			
75			-			
76			-			
83			-			
84			-			
90			-			
92			-			
93			-			
94			-			
77			+		*DNMT3A*	Tatton-Brown-Rahman syndrome
78			+			
79			+			
85			+		*KMT5B*	KMT5B syndrome
22				-		

- Test performed, no diagnosis + Test performed, diagnosis made

5.4.2 Diagnostic rate according to testing strategy

The number of NGS tests performed and number of diagnoses made are shown in Figure 91.

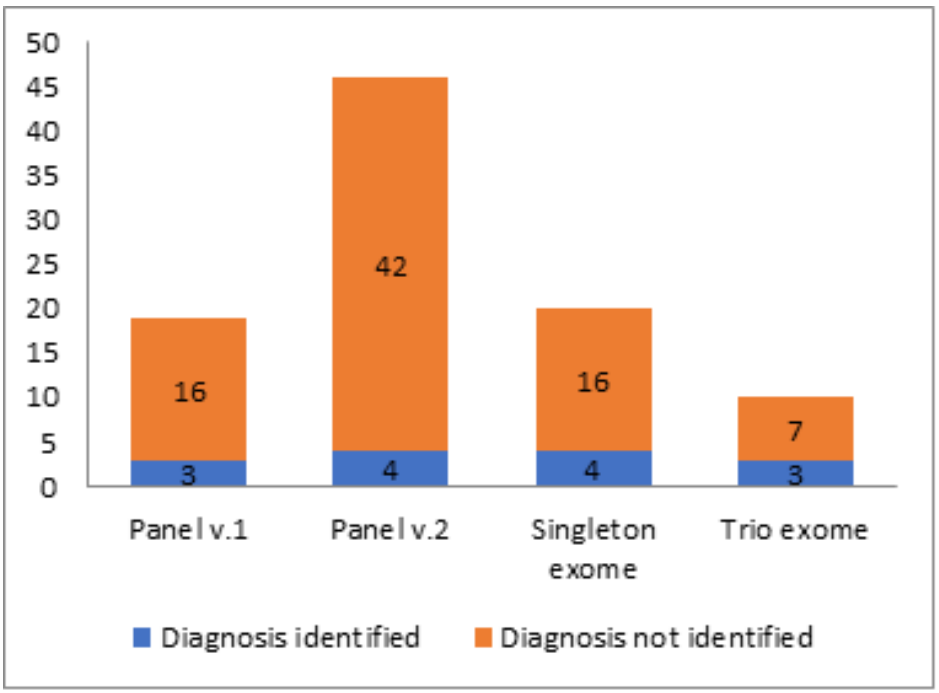


Figure 88: Number of tests performed and number of diagnoses made according to testing strategy

Trio WES was the most successful approach of identifying molecular diagnoses. This gave a diagnostic rate of 30%, compared to 20% for singleton exome (analysis of large virtual panel of genes) and 8% for a targeted panel of overgrowth genes (combined data for panel v.1 and v.2). The diagnostic rates are shown in Figure 92.

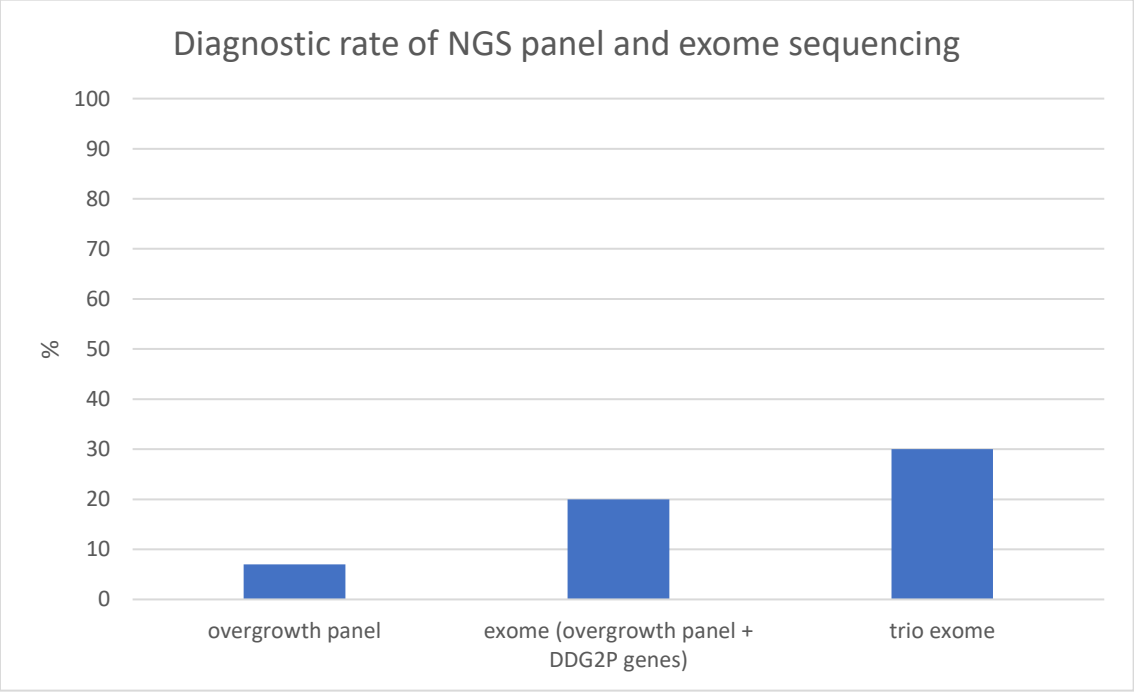


Figure 89: Diagnostic rate of different testing strategies

5.4.3 All molecular diagnoses

42 out of 100 participants in the POD study had a molecular diagnosis of a single gene disorder or imprinting disorder identified either through clinical testing in a regional genetics laboratory, NGS panel or whole exome sequencing in the POD study, or through other studies (see Table 46). One of these diagnoses only partially explained their phenotype (POD 016.0 *MAGED2*; transient neonatal Bartter syndrome). Three participants had two diagnoses, with a single gene disorder and microarray anomaly (POD 035.0 had Gorlin syndrome and a 15q11.2 microdeletion, POD 072.0 had PROS and a 15q11.2 microdeletion, and POD 085.0 had KMT5B syndrome and 16p11.3 microduplication). A further three participants had microarray anomalies that partially explained their phenotype.

Table 46: POD participants with a molecular diagnosis

POD	type of variant	gene	diagnosis	contribution to phenotype
9	monogenic	*CHD8*	CHD8 overgrowth syndrome	full
68	monogenic	*DNMT3A*	Tatton-Brown-Rahman syndrome	full
77	monogenic	*DNMT3A*	Tatton-Brown-Rahman syndrome	full
17	monogenic	*EZH2*	Weaver syndrome	full
46	monogenic	*EZH2*	Weaver syndrome	full
5	monogenic	*FBN1*	Marfan syndrome	full
8	monogenic	*FBN1*	Marfan syndrome	full
29	monogenic	*FBN1*	Marfan syndrome	full
38	monogenic	*FOXP2*	*FOXP2*-related speech and language disorder	full
52	monogenic	*GLI3*	Grieg syndrome	full
48	monogenic	*HIST1H1E*	Rahman syndrome	full
85	monogenic	*KMT5B*	*KMT5B* syndrome	full
2	monogenic	*NFIX*	Malan syndrome	full
28	monogenic	*NSD1*	Sotos syndrome	full
65	monogenic	*NSD1*	Sotos syndrome	full
74	monogenic	*NSD1*	Sotos syndrome	full
80	monogenic	*NSD1*	Sotos syndrome	full
95	monogenic	*NSD1*	Sotos syndrome	full
64	monogenic	*PDGFRB*	Kosaki overgrowth syndrome	full

30	monogenic	*PPP2R5D*	*PPP2R5D*-related neurodevelopmental disorder	full
88	monogenic	*PTEN*	*PTEN* hamartoma tumour syndrome	full
88.2	monogenic	*PTEN*	*PTEN* hamartoma tumour syndrome	full
89	monogenic	*PTEN*	*PTEN* hamartoma tumour syndrome	full
89.1	monogenic	*PTEN*	*PTEN* hamartoma tumour syndrome	full
103	monogenic	*SUZ12*	Imagawa-Matsumoto syndrome	full
103.1	monogenic	*SUZ12*	Imagawa-Matsumoto syndrome	full
103.3	monogenic	*SUZ12*	Imagawa-Matsumoto syndrome	full
104	monogenic	*SUZ12*	Imagawa-Matsumoto syndrome	full
104.1	monogenic	*SUZ12*	Imagawa-Matsumoto syndrome	full
104.3	monogenic	*SUZ12*	Imagawa-Matsumoto syndrome	full
104.4	monogenic	*SUZ12*	Imagawa-Matsumoto syndrome	full
67	monogenic	*THRA*	Resistance to thyroid hormone alpha syndrome	full
54	monogenic	*TMCO1*	Cerebrofaciothoracic dysplasia	full
35	monogenic	*PTCH1*	Gorlin syndrome	partial
16	monogenic	*MAGED2*	Transient antenatal Bartter syndrome	partial
13	mosaic	*PIK3CA*	*PIK3CA*-related overgrowth spectrum	full
53	mosaic	*PIK3CA*	*PIK3CA*-related overgrowth spectrum	full
81	mosaic	*PIK3CA*	*PIK3CA*-related overgrowth spectrum	full
72	mosaic	*PIK3CA*	*PIK3CA*-related overgrowth spectrum	partial
87	mosaic	*PTEN*	PTEN hamartoma tumour syndrome	full
61	imprinting		Beckwith-Wiedemann syndrome	full
101	chromosomal	*NSD1*	5q35.2q35.3 microdeletion; Sotos syndrome	full
35	chromosomal		15q11.2 BP1-BP2 microdeletion syndrome	partial
36	chromosomal		15q11.2 BP1-BP2 microdeletion syndrome	partial
72	chromosomal		15q11.2 BP1-BP2 microdeletion syndrome	partial
85	chromosomal		16p13.11 microduplication syndrome	partial
55	chromosomal		16p13.11 microduplication syndrome	partial
73	chromosomal		16p13.11 microduplication syndrome	partial

Note POD 035.0, 072.0 and 085.0 have two contributory diagnoses

Chapter 6. DISCUSSION:

Phenotypic and molecular analysis of a cohort of individuals with overgrowth has confirmed that this is both a clinically and genetically heterogenous group of disorders, with over 20 different genetic disorders identified in participants.

6.1 Limitations of the study

The study has a number of limitations to consider in relation to the phenotypic and genotypic results.

Although recruiting from clinical genetics centres across the UK was important for developing a large cohort with rare disease, it also meant that clinical phenotyping was performed by the local clinician in some cases. This may have led to inter-observer variability in phenotyping. In some cases, participants were recruited by the genetics research team at BWH. The phenotyping data for these participants was obtained from the clinical record made by another clinician in the clinical genetics team, causing variability in the method of data collection and again potentially affecting the consistency of phenotypic data.

With regard to genomic analysis, parental samples were not available for all participants. This prevented WES trio analysis from being performed in some cases. It is possible that a greater number of participants could have achieved a diagnosis with this investigation.

Lack of parental samples could have also led to failure to identify inheritance of a variant from a mildly affected parent in some cases. Although it would be unexpected to identify a variant in an apparently unaffected parent, it would be an important expansion of the spectrum in clinical phenotype of a disorder if this were identified. An unidentified parental diagnosis would also have important implications for recurrence risk in the family.

6.2 Overgrowth disorders are a heterogenous group of conditions

Because of the large number of overgrowth disorders and relatively small numbers of participants in each disorder group, it is difficult to analyse the phenotypic differences between the conditions. The cohort was divided into participants with a molecular diagnosis and participants without a molecular diagnosis to investigate if there were any phenotypic differences between these two groups. Very few differences were identified, with the exception that those without a molecular diagnosis were more likely to have autism spectrum disorder than those with a molecular diagnosis. This lack of difference between the groups is likely to reflect the heterogeneity in the molecular genetic aetiology in the group with a diagnosis, and the variability in clinical features between participants with the same disorder. It is reasonable to speculate that similar heterogeneity exists in the group without a diagnosis. It can be concluded that it is difficult to suggest features that are predictive of a molecular diagnosis being reached for an individual with a suspected overgrowth disorder.

The variability in phenotype of individuals with overgrowth disorders has implications for the identification and diagnosis of individuals with these conditions. The results of this study show that none of the features of overgrowth disorders are universal. The absence of any of

the features of tall stature, macrocephaly, or intellectual disability, do not exclude the possibility of a pathogenic variant in an overgrowth gene in an individual. This has important implications for selecting individuals for genetic testing. The eligibility criteria for testing needs to be broad and subject to the discretion of the clinician, so that an opportunity for diagnosis is not missed.

6.3 Definition of overgrowth and overgrowth disorders

The identification of several diagnoses in the study that are not usually considered to be overgrowth disorders suggests that the concept of what constitutes an 'overgrowth disorder' may need to be reviewed. For example, looking at the example of the participant with Greig syndrome (GCPS), the phenotypic features that warranted inclusion in the study were macrocephaly and congenital anomalies. From a molecular perspective, *GLI3* encodes a transcription factor that functions in the Hedgehog signal transduction pathway[372]. Another participant with macrocephaly was found to have Gorlin syndrome, caused by variants in *PTCH1* which encodes the PTCH receptor in the Hedgehog pathway[373]. There are parallels to be drawn with SGB, a condition long considered to be an overgrowth syndrome, with *GPC3* encoding a negative regulator of the Hedgehog signalling pathway. Further phenotypic similarity between GCPS and SGB includes polydactyly and umbilical and diaphragmatic hernia. It could be suggested that Gorlin syndrome and GCPS should be grouped with SGB based on their shared phenotypes and molecular mechanisms through the Hedgehog pathway.

The Hedgehog signalling pathway interacts with other signalling pathways, notably the mTOR pathway and the TGF-beta signalling pathway[374]. Another condition identified in several participants that is not usually considered an overgrowth disorder, Marfan syndrome, results from an abnormal fibrillin-1 protein causing overactivity of the TGF-beta signalling pathway[375]. Many individuals with Marfan syndrome will meet the proposed definition of an overgrowth disorder based on their tall stature and associated anomalies. It is clear from the number of individuals recruited to the study that individuals with Marfan syndrome do not always have distinctive features that readily exclude them from suspected diagnosis of an 'overgrowth disorder'.

A solution to the issue of what conditions to define as overgrowth disorders would be to dispense with the use of this term 'overgrowth disorders' entirely. Classification of groups of disorders according to the genomic mechanism (imprinting, epigenetic, single gene etc.), and molecular pathways involved (PI3K/AKT/mTOR, Hedgehog, TGF-beta etc.), would be a more biologically accurate nomenclature.

From a clinical perspective, genomic testing is now widely available to achieve a molecular diagnosis in a patient, and the proportion of patients without a molecular diagnosis is likely to decrease further in future with further development of testing technologies. With the advent of molecularly targeted therapies, stratification according to the genomic aetiology will also be relevant for appropriate medical management. For patients who do not have a molecular diagnosis, a clinical diagnostic label is needed. The term 'overgrowth disorder' refers to a wide range of growth phenotypes. More detailed clinical diagnoses of 'regional overgrowth',

‘generalised overgrowth’, ‘tall stature’, or ‘macrocephaly’, listed with any associated features (such as intellectual disability or congenital anomaly) would be more useful as alternative clinical labels.

6.4 Expanding the phenotypes of overgrowth disorders

This work has described deep phenotypic information for each individual with a molecular diagnosis of a genetic disorder and discussed how it expands our knowledge of the known phenotype of each condition. In these rare disorders, where only a single individual or very small number of individuals with each condition are identified, it is challenging to draw genotype-phenotype correlations with confidence. The POD study has been extended for three years to recruit larger numbers of participants to further expand our knowledge of the phenotypes, natural history and genotype-phenotype correlations.

However, even a small number of individuals can provide important information about the clinical features and genotype-phenotype correlations in a rare disorder. This study has increased the number of individuals known to have Imagawa-Matsumoto syndrome from 13 to 20. Seven individuals in this family with intragenic deletions of *SUZ12*, who do not have macrocephaly or major congenital anomalies, have a milder phenotype than previously reported individuals with missense variants. A genotype-phenotype correlation appears to be emerging, with truncating variants in *SUZ12* causing a less severe phenotype than missense variants. This could be explained by haploinsufficiency causing a less damaging loss of function effect on the protein, and missense variants causing a more damaging gain of

function effect. Further work, including phenotyping of a larger cohort of individuals and functional analysis, is needed to establish if this is a true genotype-phenotype association.

The importance of phenotyping individuals with rare disorders is exemplified by the three individuals described in this work diagnosed with KOGS. The identification of the oldest known individual with this disorder has provided valuable information about the natural history of this disorder, indicating normal intellectual ability is part of the phenotypic spectrum in later adult life. The premature death of the third individual from a ruptured basilar artery aneurysm, precisely the same pathology responsible for a stroke in the second individual, led to the discovery that vascular complications are common in this disorder and associated with significant morbidity and mortality. The novel *PDGFRB* variant in this individual prompted review of the wider spectrum of PDGFRB activating disorders, identifying that vascular complications may be a feature in common. Detailed genotypic and phenotypic analysis of this group of disorders has culminated in a new proposed nomenclature for their classification. This has clinically relevant implications for patient management, with patients in high or uncertain vascular risk groups warranting vascular imaging.

6.4.1 Future work: Targeted molecular treatment

The identification of a risk of vascular complications with outcomes including sudden premature death makes developing a therapeutic treatment of great importance to these patients and their families and clinicians. Targeted molecular treatment is now possible by

integrating knowledge of the phenotype with understanding of the genetic aetiology and molecular pathways. PDGFRB is a receptor tyrosine kinase that stimulates the PI3K-AKT pathway[133]. Myeloproliferative disorders associated with PDGFRB fusion genes have been successfully treated for many years with imatinib mesylate, a targeted agent that inhibits tyrosine kinases including PDGFRB[131]. It has been established that the specific variants in PDGFRB that are associated with KOGS are activating and are responsive to imatinib[369]. A small number of patients in the literature with germline variants in *PDGFRB* are reported to have been treated with imatinib. These include a patient with infantile myofibromatosis whose multiple tumours responded to this therapy[376], and the patient with the Asn666His variant reported by Pond et al., where treatment was well tolerated and resulted in improvement in contractures of the hands[234]. An ERN-ITHACA consortium of six international teams has now been established to explore the efficacy and tolerance of imatinib treatment in KOGS.

6.5 Molecular diagnosis in overgrowth disorders

In this study, molecular genetic analysis enabled a diagnosis to be made in 42% of individuals with a clinical diagnosis of overgrowth, a diagnostic rate comparable to previous studies of cohorts with overgrowth[6]. There are several possible reasons why a genetic variant was not identified in a participant.

Firstly, it may be that the causative gene was not included in a panel. If testing is limited to a set panel of overgrowth genes, diagnoses will be missed in some individuals. This study showed a gene agnostic trio exome sequencing approach gave a higher diagnostic yield than a large virtual panel of developmental disorder genes, which in turn gave a higher diagnostic yield than a targeted panel including only overgrowth genes.

The genetic diagnoses made in this cohort included conditions not traditionally considered to fall in the overgrowth category, such as Marfan syndrome and Gorlin syndrome, as well as known overgrowth disorders. It is likely that these diagnoses were not considered on a clinical basis because the childhood phenotype is not as distinctive as in adult life, with features such as aortic dilatation and basal cell carcinoma respectively developing with increasing age. Considering the clinical consequences of these complications, and the known significant impact of appropriate management on reducing morbidity and mortality, including these disorders in the differential diagnosis in a child with a suspected overgrowth is an important learning point from this work. This has relevance to the genetic testing now available for patients in the NHS as specified by the National Genomic Test Directory. Testing of

‘individuals with syndromic overgrowth or overgrowth in combination with intellectual disability or developmental delay’ falls under R27, ‘Congenital malformation and dysmorphism syndromes’. Including generalised overgrowth disorders in this large paediatric super-panel, rather than a separate ‘overgrowth’ panel, is a testing strategy supported by the findings of this study.

Another reason why a molecular diagnosis was not achieved in a participant may be that the causative genetic variant is mosaic. This is particularly relevant in individuals with regional overgrowth as these disorders are very often mosaic, as illustrated by the participants with PIK3CA-related overgrowth spectrum in this study. Pathogenic variants in this disorder are usually absent from blood and may be present at very low levels in other tissues. Only sequencing of DNA extracted from affected tissue, to sufficient depth, is likely to yield a diagnostic result. This highlights a potential issue for testing regional overgrowth disorders on the National Genomic Test Directory. The testing for R110 ‘Segmental overgrowth disorders’ acknowledges that ‘many of these disorders are anticipated to be mosaic and sample type and test technology need to take account of this e.g. in planning coverage of NGS assay’. However, on recent request of this testing, the laboratory designated to perform this testing does not have a platform able to sequence at the necessary depth to detect mosaic variants. The laboratory performing R327 ‘Mosaic skin disorders – deep sequencing’ has a suitable platform and this panel covers the relevant genes. However the testing criteria is ‘Dermatological abnormality’, excluding patients with regional overgrowth without a dermatological phenotype. This issue needs to be addressed to ensure these patients are able to access appropriate testing.

Mosaicism is also very relevant in testing for BWS, as seen in POD 061.0, where a diagnosis was not identified on testing DNA extracted from lymphocytes and was only achieved on a second buccal sampling. The example of BWS also illustrates another potential reason why a participant might not receive a molecular diagnosis. Methylation abnormalities will not be picked up by straightforward NGS and the appropriate test must be applied to reach a diagnosis. This is also true for other mechanisms of genetic disease, for example Fragile X syndrome caused by triplet repeat expansion.

Clinical phenotyping is key in determining the likely differential diagnosis for an individual. This highlights the value of clinical phenotyping in informing selection of the testing method required to detect the associated genomic aberration and ultimately achieving a molecular diagnosis.

Other possible causes of not achieving a molecular diagnosis include that sequencing of the causative gene was not of sufficient quality or depth, or that the bioinformatic strategy did not call the variant. Alternatively the disorder may be caused by a variant not detected on exome sequencing, such as an alteration in a regulatory region. In addition, methylation analysis of participants in this study was limited to clinical testing for Beckwith-Wiedemann syndrome in the diagnostic laboratory. Some participants in the undiagnosed cohort may have other abnormalities of methylation that could be identified on more extensive methylation analysis.

It is also possible that an individual could have an overgrowth disorder phenotype without a single identifiable genetic cause. Phenotypic features such as height and head size are familial and polygenic, and an individual may have a growth parameter >2 SD in parallel with a congenital anomaly or learning disability, without a single identifiable genetic cause. The number of participants in the study who had more than one diagnosis, for example KMT5B syndrome and 16p13.11 duplication in POD 085.0, and participants with a diagnosis that only partially explained their phenotype, for example Bartter syndrome in POD 016.0, illustrates that there is not always a simple relationship of a single genotype explaining a phenotype.

Further strategies to increase the diagnostic rate in this cohort could include 1) in participants who have undergone testing on panel or virtual panel, make further attempts to obtain parental samples for trio WES to enable gene agnostic analysis; 2) utilise a whole genome sequencing instead of panel or WES based approach; 3) repeat tissue sampling from participants with regional overgrowth, 4) employ alternative sequencing technology such as Oxford Nanopore for sequencing long fragments of DNA and covering regions of the genome that are not well sequenced by the Illumina technology, 5) develop diagnostic testing based on DNA methylation signatures for individuals with suspected Mendelian disorder of the epigenetic machinery (MDEMs), and 6) undertake genome wide methylation testing.

6.5.1 Future work: DNA methylation signature analysis

It has recently been established that in disorders associated with variants in epigenes (MDEMs), the normal DNA methylation pattern is altered in a disorder-specific pattern[377]. In addition to facilitating classification of variants of uncertain significance, this provides an

opportunity for classifying known MDEMs based on their impact on the genome and for identification of novel MDEMs.

6.5.1.1 Functional classification of PRC2 complex genes

There is considerable phenotypic overlap between Weaver syndrome, Cohen-Gibson syndrome, and Imagawa-Matsumoto syndrome, the three conditions caused by variants in genes in the PRC2 complex. Given the relatively small numbers of individuals in the literature with the latter two conditions, it is possible the reported differences in phenotype may represent variability within the same condition, and variants in *EZH2*, *EED* and *SUZ12* could be grouped together into the disorder known as 'Weaver syndrome'. However, this study has demonstrated that there may be genotype-phenotype correlations even within Imagawa-Matsumoto syndrome, with truncating variants in *SUZ12* potentially causing a less severe phenotype than missense variants. On a practical level, the decision to 'lump' or 'split' is best guided by the likely clinical consequences for an individual with a particular variant. At the time of diagnosis of a young child, or identification of a variant in the prenatal setting, information regarding likelihood and spectrum of specific clinical features is critical for clinicians and families. Assessment of genotype-phenotype correlations in a larger number of individuals is needed to provide this information. The POD study has been extended for a further three years to increase the cohort of individuals with rare genetic overgrowth disorders, including the PRC2 complex disorders.

A strategy to help elucidate the functional impact of a particular variant, and potentially help predict the likely severity of clinical consequence, is the use of DNA methylation signatures.

In 2020, Choufani et al. identified a genome-wide DNA methylation 'signature' caused by pathogenic variants in *EZH2*[378]. A small number of individuals with variants in *EED* and *SUZ12* also underwent methylome analysis, with resulting DNA methylation signatures that grouped them with the Weaver syndrome individuals[378]. DNA methylation analysis of a larger number of individuals with variants in these genes may identify if the type of variant in each of these genes has an impact on the resulting signature. If this is the case, it may be possible to draw correlations between specific methylation signatures and associated features. If the converse is true and there is no difference between the methylation signatures caused by any type of variant in any of the three genes, and in the absence of large numbers of individuals for phenotypic analysis, it would seem reasonable to group these conditions together as a single entity of Weaver syndrome. This would indicate considerable and unpredictable clinical variability within this disorder, an important point in genetic counselling for a family.

6.5.1.2 Further investigation of overgrowth disorders without a known molecular aetiology

The use of trio exome sequencing in a selected group of participants allowed for a gene agnostic approach to analysis and thus identification of potential candidate genes for novel overgrowth disorders. One gene of interest was identified in POD 051.0. This participant had a long standing clinical diagnosis of Sotos syndrome. His phenotype included high birth weight, hypotonia, tall stature, macrocephaly, severe developmental delay, features of autism spectrum disorder, advanced bone age, constipation, increased sweating, uncontrolled epilepsy, cortical visual impairment, recurrent otitis media and upper respiratory tract infections, and dysmorphic facial features suggestive of Sotos syndrome. However, no

variants in *NSD1* were identified on single gene testing in the diagnostic lab, or through molecular investigations in the study.

Trio whole exome sequencing identified a de novo missense variant in a gene encoding a histone. The variant is absent from population databases and *in silico* tools predict it is likely to be pathogenic. Germline variants in this gene have not previously been reported in association with human disease, although there is a single report of a somatic variant in a paediatric tumour. Examining the DNA methylation signature of this participant would be of great interest. A Sotos-specific methylation signature has been reported in the literature[377]; and if this participant's signature were typical of Sotos syndrome, this would provide confirmation of his clinical diagnosis. If the signature showed a normal methylation pattern, this would suggest a variant in an MDEM is not responsible for his medical problems. Finally, if an abnormal methylation pattern distinct from any previously described signature were found, this would suggest the discovery of a novel overgrowth disorder.

Chapter 7. CONCLUSION:

In conclusion, this study has expanded the known phenotypes of over ten different rare genetic overgrowth disorders, including the recently identified disorders Imagawa-Matsumoto syndrome and Kosaki overgrowth syndrome.

Molecular investigations have confirmed the genetic heterogeneity of this group of conditions, with over twenty different molecular aetiologies identified in the cohort. Clinical phenotyping is essential to guide the approach to diagnostic genetic testing in overgrowth disorders; for example regional overgrowth suggests mosaicism and the need for testing a tissue sample. The identification of several non-overgrowth disorders in the cohort demonstrates a large panel or trio WES/WGS approach, instead of a small panel of overgrowth genes, is optimal for investigating individuals with generalised overgrowth. The overall diagnostic rate in this cohort of individuals with overgrowth disorders is 42%, indicating the development of additional diagnostic strategies is required to fully understand the molecular aetiology of overgrowth.

Importantly, this work has shown how integrating genomic testing and deep phenotyping can impact on the management of patients with rare genetic overgrowth disorders. The identification of serious vascular complications in Kosaki overgrowth syndrome has led to the recommendation for vascular investigations in this group of patients. A new nomenclature for stratification according to genotype and vascular risk is proposed based on detailed genotype-

phenotype correlations in the wider PDGFRB activating disorder spectrum. Finally, an international consortium has been established for targeted molecular therapy in these patients.

List of References

1. Tatton-Brown, K. & Weksberg, R. Molecular mechanisms of childhood overgrowth. *Am. J. Med. Genet. C. Semin. Med. Genet.* **163C**, 71–5 (2013).
2. Kalish, J. M. *et al.* Nomenclature and definition in asymmetric regional body overgrowth. *Am. J. Med. Genet. Part A* 1–4 (2017) doi:10.1002/ajmg.a.38266.
3. Türkmen, S. *et al.* Mutations in NSD1 are responsible for Sotos syndrome, but are not a frequent finding in other overgrowth phenotypes. *Eur. J. Hum. Genet.* **11**, 858–865 (2003).
4. Firth, H. & Hurst, J. *Oxford Desk Reference Clinical Genetics*. (Oxford University Press, 2005).
5. Freeman, J. V *et al.* Cross sectional stature and weight reference curves for the UK, 1990. *Arch. Dis. Child.* **73**, 17–24 (1995).
6. Tatton-Brown, K. *et al.* Mutations in Epigenetic Regulation Genes Are a Major Cause of Overgrowth with Intellectual Disability. *Am. J. Hum. Genet.* **100**, 725–736 (2017).
7. Grange, D. *et al.* Cantú syndrome: Findings from 74 patients in the International Cantú Syndrome Registry. *Am. J. Med. Genet. C. Semin. Med. Genet.* **181**, 658–681 (2019).
8. Kapoor, R. R. *et al.* Clinical and molecular characterisation of 300 patients with congenital hyperinsulinism. *Eur. J. Endocrinol.* **168**, 557 (2013).
9. Malan, V. *et al.* Distinct effects of allelic NFIX mutations on nonsense-mediated mRNA decay engender either a Sotos-like or a Marshall-Smith syndrome. *Am. J. Hum. Genet.* **87**, 189–98 (2010).
10. Gripp, K. W. *et al.* Costello syndrome: Clinical phenotype, genotype, and management guidelines. *Am. J. Med. Genet. A* **179**, 1725 (2019).
11. Wilkens, A. *et al.* Novel clinical manifestations in Pallister-Killian syndrome: comprehensive evaluation of 59 affected individuals and review of previously reported cases. *Am. J. Med. Genet. A* **158A**, 3002–3017 (2012).
12. Sotos, J. F., Dodge, P. R., Muirhead, D., Crawford, J. D. & Talbot, N. B. Cerebral gigantism in childhood - a syndrome of excessively rapid growth with acromegalic features and a nonprogressive neurologic disorder. *N. Engl. J. Med.* **271**, 109–116 (1964).
13. Cole, T. R. & Hughes, H. E. Sotos syndrome: a study of the diagnostic criteria and natural history. *J. Med. Genet.* **31**, 20–32 (1994).
14. Tatton-Brown, K. *et al.* Genotype-phenotype associations in Sotos syndrome: an analysis of 266 individuals with NSD1 aberrations. *Am. J. Hum. Genet.* **77**, 193–204 (2005).
15. Allanson, J. E. & Cole, T. R. P. Sotos syndrome: Evolution of facial phenotype

subjective and objective assessment. *Am. J. Med. Genet.* **65**, 13–20 (1996).

16. Horikoshi, H. *et al.* Neuroradiologic Findings in Sotos Syndrome. *J. Child Neurol.* **21**, 614–618 (2006).

17. Schaefer, G. B., Bodensteiner, J. B., Buehler, B. A., Lin, A. & Cole, T. R. P. The neuroimaging findings in Sotos syndrome. *Am. J. Med. Genet.* **68**, 462–465 (1997).

18. Sheth, K. *et al.* The behavioral characteristics of Sotos syndrome. *Am. J. Med. Genet. Part A* **167**, 2945–2956 (2015).

19. Lane, C., Milne, E. & Freeth, M. Characteristics of Autism Spectrum Disorder in Sotos Syndrome. *J. Autism Dev. Disord.* **47**, 1–9 (2016).

20. Lane, C., Milne, E. & Freeth, M. The cognitive profile of Sotos syndrome. *J. Neuropsychol.* (2018).

21. Kurotaki, N. *et al.* Haploinsufficiency of NSD1 causes Sotos syndrome. *Nat. Genet.* **30**, 365–366 (2002).

22. Weaver, D. D., Graham, C. B., Thomas, I. T. & Smith, D. W. A new overgrowth syndrome with accelerated skeletal maturation, unusual facies, and camptodactyly. *J. Pediatr.* **84**, 547–52 (1974).

23. Cole, T. R., Dennis, N. R. & Hughes, H. E. Weaver syndrome. *J. Med. Genet.* **29**, 332–7 (1992).

24. Tatton-Brown, K. *et al.* Weaver syndrome and EZH2 mutations: Clarifying the clinical phenotype. *Am. J. Med. Genet. A* **161A**, 2972–80 (2013).

25. Douglas, J. *et al.* NSD1 mutations are the major cause of Sotos syndrome and occur in some cases of Weaver syndrome but are rare in other overgrowth phenotypes. *Am. J. Hum. Genet.* **72**, 132–143 (2003).

26. Rio, M. *et al.* Spectrum of NSD1 mutations in Sotos and Weaver syndromes. *J Med Genet* **40**, 436–440 (2003).

27. Gibson, W. T. *et al.* Mutations in EZH2 cause Weaver syndrome. *Am. J. Hum. Genet.* **90**, 110–8 (2012).

28. Tatton-Brown, K. *et al.* Germline mutations in the oncogene EZH2 cause Weaver syndrome and increased human height. *Oncotarget* **2**, 1127–1133 (2011).

29. Lui, J. C. *et al.* Ezh2 Mutations Found in the Weaver Overgrowth Syndrome Cause a Partial Loss of H3K27 Histone Methyltransferase Activity. *J. Clin. Endocrinol. Metab.* **103**, 1470–1478 (2018).

30. Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P. & Reinberg, D. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. *Genes Dev.* **16**, 2893–905 (2002).

31. Lloyd, K. & Dennis, M. Cowden's Disease: A Possible New Symptom Complex With Multiple System Involvement. *Ann. Intern. Med.* (1963).

32. Padberg, G. W., Schot, J. D. L., Vielvoye, G. J., Bots, G. T. A. M. & De Beer, F. C.

Lhermitte-duclos disease and cowden disease: A single phakomatosis. *Ann. Neurol.* **29**, 517–523 (1991).

33. Riley, H. & Smith, W. Macrocephaly, pseudopapilledema and multiple hemangiomata. *Pediatrics* **26**, 293–300 (1960).

34. Zonana, J., Rimoin, D. & Davis, D. Macrocephaly with multiple lipomas and hemangiomas. *J. Pediatr.* **89**, 600–603 (1976).

35. Ruvalcaba, R., Myrhe, S. & Smith, D. Sotos syndrome with intestinal polyposis and pigmentary changes of the genitalia. *Clin. Genet.* **18**, 413–416 (1980).

36. Gorlin, R., Cohen, M., Condon, L. & Burke, B. Bannayan-Riley-Ruvalcaba syndrome. *Am. J. Med. Genet.* **44**, 307–14 (1992).

37. Mester, J. & Eng, C. When overgrowth bumps into cancer: the PTEN-opathies. *Am. J. Med. Genet. C. Semin. Med. Genet.* **163C**, 114–21 (2013).

38. Stambolic, V. *et al.* Negative regulation of PKB/Akt-dependent cell survival by the tumor suppressor PTEN. *Cell* **95**, 29–39 (1998).

39. Bubien, V. *et al.* High cumulative risks of cancer in patients with PTEN hamartoma tumour syndrome. *J. Med. Genet.* **50**, 255–63 (2013).

40. Zhou, X.-P. *et al.* Germline and germline mosaic PTEN mutations associated with a Proteus-like syndrome of hemihypertrophy, lower limb asymmetry, arteriovenous malformations and lipomatosis. *Hum. Mol. Genet.* **9**, 765–768 (2000).

41. Caux, F. *et al. Segmental overgrowth, lipomatosis, arteriovenous malformation and epidermal nevus (SOLAMEN) syndrome is related to mosaic PTEN nullizygosity. EJHG* 767–773 (2007). doi:10.1038/sj.ejhg.5201823.

42. Elliott, M. *et al.* Clinical features and natural history of Beckwith-Wiedemann syndrome : presentation of 74 new cases. 168–174 (1994).

43. Brioude, F. *et al.* Expert consensus document: Clinical and molecular diagnosis, screening and management of Beckwith-Wiedemann syndrome: an international consensus statement. *Nat Rev Endocrinol.* **14**, 229–249 (2018).

44. Cooper, W. N. *et al.* Molecular subtypes and phenotypic expression of Beckwith–Wiedemann syndrome. *Eur. J. Hum. Genet.* **13**, 1025–1032 (2005).

45. Eggermann, T. *et al.* Clinical utility gene card for: Beckwith–Wiedemann Syndrome. *Eur. J. Hum. Genet.* **22**, 7–10 (2013).

46. Choufani, S., Shuman, C. & Weksberg, R. Beckwith-Wiedemann syndrome. *Am. J. Med. Genet. Part C Semin. Med. Genet.* **154C**, 343–354 (2010).

47. Maas, S. M. *et al.* Phenotype, cancer risk, and surveillance in Beckwith–Wiedemann syndrome depending on molecular genetic subgroups. *Am. J. Med. Genet. Part A* **170**, 2248–2260 (2016).

48. Mussa, A. *et al.* Cancer Risk in Beckwith-Wiedemann Syndrome: A Systematic Review and Meta-Analysis Outlining a Novel (Epi)Genotype Specific Histotype Targeted Screening Protocol. *J. Pediatr.* **176**, 142-149.e1 (2016).

49. Lam, W. W. *et al.* Analysis of germline CDKN1C (p57KIP2) mutations in familial and sporadic Beckwith-Wiedemann syndrome (BWS) provides a novel genotype-phenotype correlation. *J. Med. Genet.* **36**, 518–523 (1999).

50. Baskin, B. *et al.* High frequency of copy number variations (CNVs) in the chromosome 11p15 region in patients with Beckwith-Wiedemann syndrome. *Hum. Genet.* **133**, 321–30 (2014).

51. Bliek, J. *et al.* Phenotypic discordance upon paternal or maternal transmission of duplications of the 11p15 imprinted regions. *Eur. J. Med. Genet.* **52**, 404–408 (2009).

52. Simpson JL, Landey S, New M, G. J. A previously unrecognized X-linked syndrome of dysmorphia. *Birth Defects Orig Artic Ser* **11**, 18–24 (1975).

53. Cottereau, E. *et al.* Phenotypic spectrum of Simpson-Golabi-Behmel syndrome in a series of 42 cases with a mutation in GPC3 and review of the literature. *Am. J. Med. Genet. C. Semin. Med. Genet.* **163C**, 92–105 (2013).

54. Pilia, G. *et al.* Mutations in GPC3, a glypican gene, cause the Simpson-Golabi-Behmel overgrowth syndrome. *Nat. Genet.* **12**, 241–247 (1996).

55. Filmus, J. & Capurro, M. The role of glypican-3 in the regulation of body size and cancer. *Cell Cycle* **7**, 2787–2790 (2008).

56. Morris, M. R., Astuti, D. & Maher, E. R. Perlman syndrome: overgrowth, Wilms tumor predisposition and DIS3L2. *Am. J. Med. Genet. C. Semin. Med. Genet.* **163C**, 106–13 (2013).

57. Astuti, D. *et al.* Germline mutations in DIS3L2 cause the Perlman syndrome of overgrowth and Wilms tumor susceptibility. *Nature Genetics* vol. 44 277–284 at https://doi.org/10.1038/ng.1071 (2012).

58. Lindhurst, M. J. *et al.* A mosaic activating mutation in AKT1 associated with the Proteus syndrome. *N. Engl. J. Med.* **365**, 611–9 (2011).

59. Cohen, M. & Hayden, P. A newly recognized hamartomatous syndrome. *Birth Defects Orig. Artic. Ser.* **15**, 291–296 (1979).

60. Wiedemann, H. *et al.* The Proteus syndrome. Partial gigantism of the hands and/or feet, nevi, hemihypertrophy, subcutaneous tumors, macrocephaly or other skull anomalies and possible accelerated growth and visceral affections. *Eur. J. Pediatr.* **140**, 5–12 (1983).

61. Murray, P. G. & Clayton, P. E. Endocrine control of growth. *Am. J. Med. Genet. C. Semin. Med. Genet.* **163C**, 76–85 (2013).

62. Koukoura, O., Sifakis, S. & Spandidos, D. A. DNA methylation in the human placenta and fetal growth (Review). *Mol. Med. Rep.* **5**, 883 (2012).

63. Jee, Y. H. & Baron, J. The Biology of Stature. *J. Pediatr.* **173**, 32 (2016).

64. Lui, J. C. K. *et al.* Spatial and Temporal Regulation of Gene Expression in the Mammalian Growth Plate. *Bone* **46**, 1380 (2010).

65. Hakuno, F. & Takahashi, S. I. IGF1 receptor signaling pathways. *J. Mol. Endocrinol.*

61, T69–T86 (2018).

66. Gupta, M. B. & Jansson, T. Novel roles of mechanistic target of rapamycin signaling in regulating fetal growth. *Biol. Reprod.* **100**, 872 (2019).

67. Watanabe, H. *et al.* DNA methylation analysis of multiple imprinted DMRs in Sotos syndrome reveals IGF2-DMR0 as a DNA methylation-dependent, P0 promoter-specific enhancer. *FASEB J.* **34**, 960–973 (2020).

68. Quintero-Rivera, F. *et al.* 5q35 duplication presents with psychiatric and undergrowth phenotypes mediated by NSD1 overexpression and mTOR signaling downregulation. *Hum. Genet.* **140**, 681 (2021).

69. Wei, F. Z. *et al.* Epigenetic regulation of autophagy by the methyltransferase EZH2 through an MTOR-dependent pathway. *Autophagy* **11**, 2309 (2015).

70. Brennan, K. *et al.* NSD1 mutations deregulate transcription and DNA methylation of bivalent developmental genes in Sotos syndrome. *Hum. Mol. Genet.* **31**, 2164–2184 (2022).

71. Cytrynbaum, C., Choufani, S. & Weksberg, R. Epigenetic signatures in overgrowth syndromes: Translational opportunities. *Am. J. Med. Genet. Part C Semin. Med. Genet.* **181**, 491–501 (2019).

72. Luger, K. & Hansen, J. C. Nucleosome and chromatin fiber dynamics. *Current Opinion in Structural Biology* vol. 15 188–196 at https://doi.org/10.1016/j.sbi.2005.03.006 (2005).

73. Berger, S. L., Kouzarides, T., Shiekhattar, R. & Shilatifard, A. An operational definition of epigenetics. *Genes Dev.* **23**, 781 (2009).

74. Bjornsson, H. T. The Mendelian disorders of the epigenetic machinery. *Genome Res.* **25**, 1473 (2015).

75. Hyun, K., Jeon, J., Park, K. & Kim, J. Writing, erasing and reading histone lysine methylations. *Experimental and Molecular Medicine* vol. 49 e324 at https://doi.org/10.1038/emm.2017.11 (2017).

76. Fahrner, J. A. & Bjornsson, H. T. Mendelian Disorders of the Epigenetic Machinery: Tipping the Balance of Chromatin States. *Annu. Rev. Genomics Hum. Genet.* **15**, 269 (2014).

77. Cao, R. *et al.* Role of histone H3 lysine 27 methylation in polycomb-group silencing. *Science (80-.).* **298**, 1039–1043 (2002).

78. Cooney, E., Bi, W., Schlesinger, A. E., Vinson, S. & Potocki, L. Novel EED mutation in patient with Weaver syndrome. *Am. J. Med. Genet. Part A* **173**, 541–545 (2017).

79. Cohen, A. S. A. *et al.* A novel mutation in EED associated with overgrowth. *J. Hum. Genet.* **60**, 339–342 (2015).

80. Cohen, A. & Gibson, W. EED-associated overgrowth in a second male patient. *J. Hum. Genet.* **61**, 831–834 (2016).

81. Imagawa, E. *et al.* Mutations in genes encoding polycomb repressive complex 2

subunits cause Weaver syndrome. *Hum. Mutat.* **38**, 637–648 (2017).

82. Imagawa, E. *et al.* Novel *SUZ12* mutations in Weaver-like syndrome. *Clin. Genet.* **94**, 461–466 (2018).

83. Griffiths, S. *et al.* EED and EZH2 constitutive variants: A study to expand the Cohen-Gibson syndrome phenotype and contrast it with Weaver syndrome. *Am. J. Med. Genet. Part A* **179**, 588–594 (2019).

84. Spellicy, C. J. *et al.* Three additional patients with EED-associated overgrowth: potential mutation hotspots identified? *J. Hum. Genet.* **64**, 561–572 (2019).

85. Cyrus, S., Burkardt, D., Weaver, D. D. & Gibson, W. T. PRC2-complex related dysfunction in overgrowth syndromes: A review of EZH2, EED, and SUZ12 and their syndromic phenotypes. *Am. J. Med. Genet. Part C Semin. Med. Genet.* (2019) doi:10.1002/ajmg.c.31754.

86. Cooney, E., Bi, W., Schlesinger, A. E., Vinson, S. & Potocki, L. Novel *EED* mutation in patient with Weaver syndrome. *Am. J. Med. Genet. Part A* **173**, 541–545 (2017).

87. Smigiel, R. *et al.* Novel de novo mutation affecting two adjacent aminoacids in the EED gene in a patient with Weaver syndrome. *J. Hum. Genet.* **63**, 517–520 (2018).

88. Tatton-Brown, K. *et al.* Mutations in the DNA methyltransferase gene DNMT3A cause an overgrowth syndrome with intellectual disability. *Nat. Genet.* **46**, 385–8 (2014).

89. Tatton-Brown, K. *et al.* The Tatton-Brown-Rahman Syndrome: A clinical study of 55 individuals with de novo constitutive DNMT3A variants. *Wellcome Open Res.* **3**, 1–16 (2018).

90. Kosaki, R., Terashima, H., Kubota, M. & Kosaki, K. Acute myeloid leukemia-associated *DNMT3A* p.Arg882His mutation in a patient with Tatton-Brown-Rahman overgrowth syndrome as a constitutional mutation. *Am. J. Med. Genet. Part A* **173**, 250–253 (2017).

91. Hollink, I. H. I. M. *et al.* Acute myeloid leukaemia in a case with Tatton-Brown-Rahman syndrome: The peculiar DNMT3A R882 mutation. *J. Med. Genet.* **54**, 805–808 (2017).

92. Sweeney, K. J. *et al.* The first case report of medulloblastoma associated with Tatton-Brown–Rahman syndrome. *Am. J. Med. Genet. Part A* **179**, ajmg.a.61180 (2019).

93. Zahir, F. *et al.* Novel deletions of 14q11.2 associated with developmental delay, cognitive impairment and similar minor anomalies in three children. *J. Med. Genet.* **44**, 556–561 (2007).

94. O'Roak, B. J. *et al.* Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. *Nature* **485**, 246–250 (2012).

95. O'Roak, B. J. *et al.* Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science (80-.).* **338**, 1619–1622 (2012).

96. Bernier, R. *et al.* Disruptive CHD8 mutations define a subtype of autism early in

development. *Cell. July* **17**, 263–276 (2014).

97. Yasin, H. *et al.* A distinct neurodevelopmental syndrome with intellectual disability, autism spectrum disorder, characteristic facies, and macrocephaly is caused by defects in CHD8. *J. Hum. Genet.* **64**, 271–280 (2019).

98. Douzgou, S. *et al.* The clinical presentation caused by truncating *CHD8* variants. *Clin. Genet.* **96**, 72–84 (2019).

99. Ostrowski, P. J. *et al.* The *CHD8* overgrowth syndrome: A detailed evaluation of an emerging overgrowth phenotype in 27 patients. *Am. J. Med. Genet. Part C Semin. Med. Genet.* ajmg.c.31749 (2019) doi:10.1002/ajmg.c.31749.

100. Thompson, B. A., Tremblay, V., Lin, G. & Bochar, D. A. CHD8 Is an ATP-Dependent Chromatin Remodeling Factor That Regulates -Catenin Target Genes. *Mol. Cell. Biol.* **28**, 3894–3904 (2008).

101. Nishiyama, M., Skoultchi, A. I. & Nakayama, K. I. Histone H1 Recruitment by CHD8 Is Essential for Suppression of the Wnt- -Catenin Signaling Pathway. *Mol. Cell. Biol.* **32**, 501–512 (2012).

102. Arya, V. B. *et al.* Activating AKT2 mutation: Hypoinsulinemic hypoketotic hypoglycemia. *J. Clin. Endocrinol. Metab.* **99**, 391–394 (2014).

103. Rivière, J. B. *et al.* De novo germline and postzygotic mutations in AKT3, PIK3R2 and PIK3CA cause a spectrum of related megalencephaly syndromes. *Nat. Genet.* **44**, 934–940 (2012).

104. Takenouchi, T. *et al.* Novel Overgrowth Syndrome Phenotype Due to Recurrent De Novo PDGFRB Mutation. *J. Pediatr.* **166**, 483–486 (2015).

105. Kurek, K. C. *et al.* Somatic mosaic activating mutations in PIK3CA cause CLOVES syndrome. *Am. J. Hum. Genet.* **90**, 1108–1115 (2012).

106. Lindhurst, M. J. *et al.* Mosaic overgrowth with fibroadipose hyperplasia is caused by somatic activating mutations in PIK3CA. *Nat. Genet.* **44**, 928–933 (2012).

107. Baynam, G. *et al.* A germline MTOR mutation in Aboriginal Australian siblings with intellectual disability, dysmorphism, macrocephaly, and small thoraces. *Am. J. Med. Genet. Part A* **9999**, 1–9 (2015).

108. Hussain, K. *et al.* An Activating Mutation of AKT2 and Human Hypoglycemia. *Science* **334**, 474 (2011).

109. George, S. *et al.* A Family with Severe Insulin Resistance and Diabetes Due to a Mutation in AKT2. *Science (80-.).* **304**, 1325–1328 (2004).

110. Dobyns, W. B. & Mirzaa, G. M. Megalencephaly syndromes associated with mutations of core components of the PI3K-AKT–MTOR pathway: PIK3CA, PIK3R2, AKT3, and MTOR. *Am. J. Med. Genet. Part C Semin. Med. Genet.* 582–590 (2019) doi:10.1002/ajmg.c.31736.

111. Ciaccio, C., Cellini, E., Guerrini, R., Pantaleoni, C. & Masson, R. Mirror syndromes regarding AKT3 mutations: Loss of function variant leading to microcephaly. *Am. J.*

Med. Genet. A **182**, 2800–2802 (2020).

112. Keppler-Noreuil, K. M. *et al.* PIK3CA-Related Overgrowth Spectrum (PROS): Diagnostic and Testing Eligibility Criteria, Differential Diagnosis, and Evaluation. *Am. J. Med. Genet. A* **0**, 287 (2015).

113. Thauvin-Robinet, C. *et al.* PIK3R1 Mutations Cause Syndromic Insulin Resistance with Lipoatrophy. *Am. J. Hum. Genet.* **93**, 141 (2013).

114. Mirzaa, G. *et al.* Characterization of mutations of the phosphoinositide-3-kinaseregulatory subunit, PIK3R2, in perisylvian polymicrogyria: anext generation sequencing study. *Lancet. Neurol.* **14**, 1182 (2015).

115. Keppler-Noreuil, K. M. *et al.* Clinical Delineation and Natural History of the PIK3CA-Related Overgrowth Spectrum. *Am. J. Med. Genet. A* **164**, 1713 (2014).

116. Mirzaa, G. M., Rivière, J.-B. & Dobyns, W. B. Megalencephaly syndromes and activating mutations in the PI3K-AKT pathway: MPPH and MCAP. *Am. J. Med. Genet. C. Semin. Med. Genet.* **163C**, 122–30 (2013).

117. Couto, J. A. *et al.* Somatic PIK3CA Mutations are Present in Multiple Tissues of Facial Infiltrating Lipomatosis. *Pediatr. Res.* **82**, 850 (2017).

118. Vahidnezhad, H., Youssefian, L. & Uitto, J. Klippel–Trenaunay syndrome belongs to the PIK3CA-related overgrowth spectrum (PROS). *Exp. Dermatol.* **25**, 17–19 (2016).

119. Rios, J. J. *et al.* Somatic gain-of-function mutations in PIK3CA in patients with macrodactyly. *Hum. Mol. Genet.* **22**, 444 (2013).

120. Douzgou, S. *et al.* A standard of care for individuals with PIK3CA-related disorders: An international expert consensus statement. *Clin. Genet.* cge.14027 (2021) doi:10.1111/CGE.14027.

121. Mirzaa, G. *et al.* PIK3CA-associated developmental disorders exhibit distinct classes of mutations with variable expression and tissue distribution. *JCI Insight* **1**, (2016).

122. Kandoth, C. *et al.* Mutational landscape and significance across 12 major cancer types. *Nat. 2013 5027471* **502**, 333–339 (2013).

123. Kuentz, P. *et al.* Molecular diagnosis of PIK3CA-related overgrowth spectrum (PROS) in 162 patients and recommendations for genetic testing. *Genet. Med. 2017 199* **19**, 989–997 (2017).

124. Cheung, Y. H. *et al.* A recurrent PDGFRB mutation causes familial infantile myofibromatosis. *Am. J. Hum. Genet.* **92**, 996–1000 (2013).

125. Martignetti, J. A. *et al.* Mutations in PDGFRB cause autosomal-dominant infantile myofibromatosis. *Am. J. Hum. Genet.* **92**, 1001–7 (2013).

126. Murray, N. *et al.* The spectrum of infantile myofibromatosis includes both non-penetrance and adult recurrence. *European Journal of Medical Genetics* vol. 60 353–358 at https://doi.org/10.1016/j.ejmg.2017.02.005 (2017).

127. Lepelletier, C. Heterozygous PDGFRB mutation in a Three-generation Family with Autosomal Dominant Infantile Myofibromatosis. *Acta Derm. Venereol.* **97**, 858–859

(2017).

128. Johnston, J. J. *et al.* A Point Mutation in PDGFRB Causes Autosomal-Dominant Penttinen Syndrome. *Am. J. Hum. Genet.* **97**, 465 (2015).

129. Minatogawa, M. *et al.* Expansion of the phenotype of Kosaki overgrowth syndrome. *Am. J. Med. Genet. Part A* 2422–2427 (2017) doi:10.1002/ajmg.a.38310.

130. Nicolas, G., Pottier, C., Maltete, D., Coutant, S. & Rovelet-Lecrux, A. Mutation of the PDGFRB gene as a cause of idiopathic basal ganglia calcification. *Neurology* **80**, 181–187 (2013).

131. Apperley, J. F. *et al.* Response to imatinib mesylate in patients with chronic myeloproliferative diseases with rearrangements of the platelet-derived growth factor receptor beta. *N. Engl. J. Med.* **347**, 481–487 (2002).

132. Gawliński, P. *et al.* Phenotype expansion and development in Kosaki overgrowth syndrome. *Clin. Genet.* **93**, 919–924 (2018).

133. Zarate, Y. A. *et al.* Constitutive activation of the PI3K-AKT pathway and cardiovascular abnormalities in an individual with Kosaki overgrowth syndrome. *Am. J. Med. Genet. Part A* ajmg.a.61145 (2019) doi:10.1002/ajmg.a.61145.

134. Klaassens, M. *et al.* Malan syndrome: Sotos-like overgrowth with de novo NFIX sequence variants and deletions in six new patients and a review of the literature. *Eur. J. Hum. Genet.* 1–6 (2014) doi:10.1038/ejhg.2014.162.

135. Moncla, A. *et al.* A cluster of translocation breakpoints in 2q37 is associated with overexpression of NPPC in patients with a similar overgrowth phenotype. *Hum. Mutat.* **28**, 1183–1188 (2007).

136. Ko, J. M. *et al.* Skeletal overgrowth syndrome caused by overexpression of C-type natriuretic peptide in a girl with balanced chromosomal translocation, t(1;2)(q41;q37.1). *Am. J. Med. Genet. Part A* **167**, 1033–1038 (2015).

137. Molin, A.-M. *et al.* Original article: A novel microdeletion syndrome at 3q13.31 characterised by developmental delay, postnatal overgrowth, hypoplastic male genitals, and characteristic facial features. *J. Med. Genet.* **49**, 104 (2012).

138. Cordeddu, V. *et al.* Mutations in ZBTB20 cause Primrose syndrome. *Nat. Genet.* **46**, 815–7 (2014).

139. Partington, M. W., Fagan, K., Soubjaki, V. & Turner, G. Translocations involving 4p16.3 in three families: deletion causing the Pitt-Rogers-Danks syndrome and duplication resulting in a new overgrowth syndrome. *J. Med. Genet.* **34**, 719 (1997).

140. Palumbo, O. *et al.* Report of a patient and further clinical and molecular characterization of interstitial 4p16.3 microduplication. *Mol. Cytogenet.* **8**, (2015).

141. Toydemir, R. M. *et al.* A Novel Mutation in FGFR3 Causes Camptodactyly, Tall Stature, and Hearing Loss (CATSHL) Syndrome. *Am. J. Hum. Genet.* **79**, 935 (2006).

142. Escobar, L., Tucker, M. & Bamshad, M. A second family with CATSHL syndrome: Confirmatory report of another unique FGFR3 syndrome. *Am. J. Med. Genet. A* **170**,

1908–1911 (2016).

143. Kamien, B. *et al.* Narrowing the critical region for overgrowth within 13q14.2-q14.3 microdeletions. *Eur. J. Med. Genet.* **58**, 629–633 (2015).

144. Kannu, P. *et al.* Post-axial polydactyly type A2, overgrowth and autistic traits associated with a chromosome 13q31.3 microduplication encompassing miR-17-92 and GPC5. *Eur. J. Med. Genet.* **56**, 452–457 (2013).

145. Siavrienė, E. *et al.* A de novo 13q31.3 microduplication encompassing the miR-17 ~ 92 cluster results in features mirroring those associated with Feingold syndrome 2. *Gene* **753**, 144816 (2020).

146. Tatton-Brown, K. *et al.* 15q overgrowth syndrome: A newly recognized phenotype associated with overgrowth, learning difficulties, characteristic facial appearance, renal anomalies and increased dosage of distal chromosome 15q. *Am. J. Med. Genet. Part A* **149**, 147–154 (2009).

147. Leffler, M. *et al.* Two familial microduplications of 15q26.3 causing overgrowth and variable intellectual disability with normal copy number of IGF1R. *Eur. J. Med. Genet.* **59**, 257–262 (2016).

148. Dolan, M. *et al.* A novel microdeletion/microduplication syndrome of 19p13.13. *Genet. Med. 2010 128* **12**, 503–511 (2010).

149. Phelan, M. *et al.* 22q13 deletion syndrome. *Am. J. Med. Genet.* **101**, 91–99 (2001).

150. Disciglio, V. *et al.* Interstitial 22q13 deletions not involving SHANK3 gene: A new contiguous gene syndrome. *Am. J. Med. Genet. Part A* **164**, 1666–1676 (2014).

151. Sanger, F., Nicklen, S. & Coulson, a R. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A.* **74**, 5463–5467 (1977).

152. Anderson, M. W. & Schrijver, I. Next generation DNA sequencing and the future of genomic medicine. *Genes (Basel).* **1**, 38–69 (2010).

153. Lander, E. S. *et al.* Initial sequencing and analysis of the human genome. *Nature* **409**, 860–921 (2001).

154. Venter, J. C. *et al.* The Sequence of the Human Genome. *Science (80-.).* **291**, 1304–1351 (2001).

155. Collins, F. S. Positional cloning moves from perditional to traditional. *Nat. Genet. 1995 94* **9**, 347–350 (1995).

156. Ng, S. B. *et al.* Exome sequencing identifies the cause of a Mendelian disorder. *Nat. Genet.* **42**, 30 (2010).

157. Head, S. R. *et al.* Library construction for next-generation sequencing: overviews and challenges. *Biotechniques* **56**, 61–4, 66, 68, passim (2014).

158. van Dijk, E. L., Auger, H., Jaszczyszyn, Y. & Thermes, C. Ten years of next-generation sequencing technology. *Trends Genet.* **30**, (2014).

159. van Dijk, E. L., Jaszczyszyn, Y. & Thermes, C. Library preparation methods for next-

generation sequencing: tone down the bias. *Exp. Cell Res.* **322**, 12–20 (2014).

160. Linnarsson, S. Recent advances in DNA sequencing methods - general principles of sample preparation. *Exp. Cell Res.* **316**, 1339–1343 (2010).

161. Consortium, T. 1000 G. P. A global reference for human genetic variation. *Nature* **526**, 68 (2015).

162. Fu, W. *et al.* Analysis of 6,515 exomes reveals the recent origin of most human protein-coding variants. *Nature* **493**, 216–220 (2013).

163. Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285–291 (2016).

164. Sherry, S. *et al.* dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res.* **29**, 308–311 (2001).

165. Wright, C. *et al.* Genetic diagnosis of developmental disorders in the DDD study: a scalable analysis of genome-wide research data. *Lancet (London, England)* **385**, 1305–1314 (2015).

166. Collod-Béroud, G. *et al.* Update of the UMD-FBN1 mutation database and creation of an FBN1 polymorphism database. *Hum. Mutat.* **22**, 199–208 (2003).

167. Ng, P. & Henikoff, S. SIFT: Predicting amino acid changes that affect protein function. *Nucleic Acids Res.* **31**, 3812–3814 (2003).

168. Adzhubei, I. A. *et al.* A method and server for predicting damaging missense mutations. *Nat. Methods* **7**, 248 (2010).

169. Schwarz, J., Rödelsperger, C., Schuelke, M. & Seelow, D. MutationTaster evaluates disease-causing potential of sequence alterations. *Nat. Methods* **7**, 575–576 (2010).

170. Leman, R. *et al.* Novel diagnostic tool for prediction of variant spliceogenicity derived from a set of 395 combined in silico/in vitro studies: an international collaborative effort. *Nucleic Acids Res.* **46**, 7913 (2018).

171. Richards, S. *et al.* 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. *Genet. Med.* **17**, 405–423 (2015).

172. Robinson, P. N. Deep phenotyping for precision medicine. *Hum. Mutat.* **33**, 777–80 (2012).

173. Köhler, S. *et al.* The Human Phenotype Ontology project: Linking molecular biology and disease through phenotype data. *Nucleic Acids Res.* **42**, 966–974 (2014).

174. Cole, T. The LMS method for constructing normalized growth standards. *undefined* (1990).

175. Marsh, D. J. *et al.* Germline mutations in PTEN are present in Bannayan-Zonana syndrome. *Nat. Genet.* **16**, 333–334 (1997).

176. Liaw, D. *et al.* Germline mutations of the PTEN gene in Cowden disease, an inherited

breast and thyroid cancer syndrome. *Nat. Genet.* **16**, 64–67 (1997).

177. Douglas, J. *et al.* Mutations in RNF135, a gene within the NF1 microdeletion region, cause phenotypic abnormalities including overgrowth. *Nat. Genet.* **39**, 963–5 (2007).

178. Miura, K. *et al.* Overgrowth syndrome associated with a gain-of-function mutation of the natriuretic peptide receptor 2 (NPR2) gene. *Am. J. Med. Genet. A* **164A**, 156–63 (2014).

179. Klein, S. *et al.* Expanding the phenotype of mutations in DICER1: mosaic missense mutations in the RNase IIIb domain of DICER1 causes GLOW syndrome. *J. Med. Genet.* **51**, 294–302 (2014).

180. Luscan, A. *et al.* Mutations in SETD2 cause a novel overgrowth condition. doi:10.1136/jmedgenet-2014-102402.

181. Tenorio, J. *et al.* A new overgrowth syndrome is due to mutations in RNF125. 1–20 (2014) doi:10.1002/humu.22689.This.

182. Mirzaa, G. M. *et al.* De novo CCND2 mutations leading to stabilization of cyclin D2 cause megalencephaly-polymicrogyria-polydactyly-hydrocephalus syndrome. *Nat. Genet.* **46**, 510–5 (2014).

183. Almuriekhi, M. *et al.* Loss-of-Function Mutation in APC2 Causes Sotos Syndrome Features. *CellReports* **10**, 1585–1598 (2015).

184. Loveday, Chey, Tatton-Brown, Katrina, Rahman, N. Mutations on the PP2A regulatory subunit B family genes PPP2R5B, PPP2R5C and PPP2R5D cause human overgrowth. *Hum. Mol. Genet.* (2015).

185. Houge, G. *et al.* B56delta-related protein phosphatase 2A dysfunction identified in patients with intellectual disability. *J. Clin. Invest.* **125**, (2015).

186. Aggarwal, S., Bhowmik, A. Das, Ramprasad, V. L., Murugan, S. & Dalal, A. A splice site mutation in *HERC1* leads to syndromic intellectual disability with macrocephaly and facial dysmorphism: Further delineation of the phenotypic spectrum. *Am. J. Med. Genet. Part A* **170**, 1868–1873 (2016).

187. Fauth, C. *et al.* A recurrent germline mutation in the PIGA gene causes Simpson-Golabi-Behmel syndrome type 2. *Am. J. Med. Genet. A* 1–11 (2015) doi:10.1002/ajmg.a.37452.

188. Prontera, P. *et al.* Recurrent ~100 Kb microdeletion in the chromosomal region 14q11.2, involving *CHD8* gene, is associated with autism and macrocephaly. *Am. J. Med. Genet. Part A* **164**, 3137–3141 (2014).

189. Grotto, S. *et al.* Clinical assessment of five patients with BRWD3 mutation at Xq21.1 gives further evidence for mild to moderate intellectual disability and macrocephaly. *Eur. J. Med. Genet.* **57**, 200–206 (2014).

190. Schäfgen, J. *et al.* De novo nonsense and frameshift variants of TCF20 in individuals with intellectual disability and postnatal overgrowth. *Eur. J. Hum. Genet.* **24**, 1739 (2016).

191. Thauvin-Robinet, C. *et al.* Homozygous FIBP nonsense variant responsible of syndromic overgrowth, with overgrowth, macrocephaly, retinal coloboma and learning disabilities. *Clin. Genet.* **89**, e1–e4 (2016).

192. Parente, D. J. *et al.* Neuroligin 2 nonsense variant associated with anxiety, autism, intellectual disability, hyperphagia, and obesity. *Am. J. Med. Genet. Part A* **173**, 213–216 (2017).

193. Weiss, K. *et al.* De Novo Mutations in CHD4, an ATP-Dependent Chromatin Remodeler Gene, Cause an Intellectual Disability Syndrome with Distinctive Dysmorphisms. *Am. J. Hum. Genet.* **99**, 934–941 (2016).

194. Labonne, J. D. J. *et al.* Comparative deletion mapping at 1p31.3-p32.2 implies NFIA responsible for intellectual disability coupled with macrocephaly and the presence of several other genes for syndromic intellectual disability. *Mol. Cytogenet.* **9**, (2016).

195. Baple, E. L. *et al.* Mutations in KPTN Cause Macrocephaly, Neurodevelopmental Delay, and Seizures. *Am. J. Hum. Genet.* **94**, 87–94 (2014).

196. Hahn, H. *et al.* Mutations of the Human Homolog of Drosophila patched in the Nevoid Basal Cell Carcinoma Syndrome. *Cell* **85**, 841–851 (1996).

197. Pastorino, L. *et al.* Identification of a SUFU germline mutation in a family with Gorlin syndrome. *Am. J. Med. Genet. Part A* **149A**, 1539–1543 (2009).

198. Risheg, H. *et al.* A recurrent mutation in MED12 leading to R961W causes Opitz-Kaveggia syndrome. doi:10.1038/ng1992.

199. Capo-Chichi, J.-M. *et al.* Disruption of TBC1D7, a subunit of the TSC1-TSC2 protein complex, in intellectual disability and megalencephaly. *J. Med. Genet.* **50**, 740–744 (2013).

200. Richards, S. *et al.* 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. *Genet. Med.* **17**, 405–424 (2015).

201. Foster, A., Titheradge, H. & Morton, J. Genetics of learning disability. *Paediatr. Child Health (Oxford).* **25**, 450–457 (2015).

202. Brioude, F. Clinical and molecular diagnosis of Beckwith-Wiedemann syndrome: an international consensus statement. *Nat. Rev. Endocrinol.* (2018).

203. Ostrowski, P. J. *et al.* The CHD8 overgrowth syndrome: A detailed evaluation of an emerging overgrowth phenotype in 27 patients. *Am. J. Med. Genet. Part C Semin. Med. Genet.* 1–8 (2019) doi:10.1002/ajmg.c.31749.

204. Nakamura, H. & Puri, P. Concurrent Hirschsprung's disease and anorectal malformation: a systematic review. *Pediatr. Surg. Int.* **36**, 21–24 (2020).

205. Balci, T. B. *et al.* Tatton-Brown-Rahman syndrome: Six individuals with novel features. *Am. J. Med. Genet. Part A* **182**, 673–680 (2020).

206. Xin, B. *et al.* Novel *DNMT3A* germline mutations are associated with inherited Tatton-

Brown-Rahman syndrome. *Clin. Genet.* **91**, 623–628 (2017).

207. Sweeney, K. J. *et al.* The first case report of medulloblastoma associated with Tatton-Brown–Rahman syndrome. *Am. J. Med. Genet. Part A* **179**, 1357–1361 (2019).

208. Griffiths, S. *et al.* EED and EZH2 constitutive variants: A study to expand the Cohen-Gibson syndrome phenotype and contrast it with Weaver syndrome. *Am. J. Med. Genet. A* **179**, 588–594 (2019).

209. Priolo, M. *et al.* Further delineation of Malan syndrome. *Hum Mutat* 1226–1237 (2018) doi:10.1002/humu.23563.

210. Grand, K. *et al.* Hyperinsulinemic hypoglycemia in seven patients with de novo NSD1 mutations. *Am. J. Med. Genet. A* **179**, 542–551 (2019).

211. Schaefer, G., Bodensteiner, J., Buehler, B. & Cole, T. The neuroimaging findings in Sotos syndrome. *Am J Med Genet A* **68**, 462–5 (1997).

212. Foster, A. *et al.* The phenotype of Sotos syndrome in adulthood: A review of 44 individuals. *Am. J. Med. Genet. Part C Semin. Med. Genet.* ajmg.c.31738 (2019) doi:10.1002/ajmg.c.31738.

213. Foster, A. *et al.* Kosaki overgrowth syndrome: A novel pathogenic variant in PDGFRB and expansion of the phenotype including cerebrovascular complications. *Clin. Genet.* 19–31 (2020) doi:10.1111/cge.13752.

214. Stoll, C., Juif, J. G., Grosshans, E. & Brini, A. [Ehlers-Danlos syndrome associated with gigantism, craniosynostosis and melanoderma]. *Pediatrie* **29**, 81–9 (1974).

215. Takenouchi, T., Kodo, K., Yamazaki, F., Nakatomi, H. & Kosaki, K. Progressive cerebral and coronary aneurysms in the original two patients with Kosaki overgrowth syndrome. *Am. J. Med. Genet. Part A* **185**, 999–1003 (2021).

216. Wenger, T. L. *et al.* Activating variants in <scp> *PDGFRB* </scp> result in a spectrum of disorders responsive to imatinib monotherapy. *Am. J. Med. Genet. Part A* **182**, 1576–1591 (2020).

217. Zufferey, F. *et al.* Acro-osteolysis, keloid like-lesions, distinctive facial features, and overgrowth: Two newly recognized patients with premature aging syndrome, penttinen type. *Am. J. Med. Genet. Part A* **161**, 1786–1791 (2013).

218. Karasozen, Y. *et al.* Somatic PDGFRB Activating Variants in Fusiform Cerebral Aneurysms. *Am. J. Hum. Genet.* **104**, 968–976 (2019).

219. Chenbhanich, J. *et al.* Segmental overgrowth and aneurysms due to mosaic PDGFRB p.(Tyr562Cys). *Am. J. Med. Genet. Part A* (2021) doi:10.1002/ajmg.a.62126.

220. Johnston, J. J. *et al.* A Point Mutation in PDGFRB Causes Autosomal-Dominant Penttinen Syndrome. *Am. J. Hum. Genet.* **97**, 465–474 (2015).

221. Zhang, Z. *et al.* Acta Dermato-Venereologica Premature Aging Syndrome, Penttinen Type: Report of a Chinese Case with a PDGFRB Mutation. (2018) doi:10.2340/00015555-2993.

222. Wright, C., Corbally, M. T., Hayes, R. & McDermott, M. B. Multifocal Infantile

Myofibromatosis and Generalized Fibromuscular Dysplasia in a Child: Evidence for a Common Pathologic Process?: *https://doi.org/10.1007/s10024-003-0107-4* **7**, 385–390 (2004).

223. Frezin, J., Fusaro, F., Reding, R. & Godefroid, N. Kidney transplantation in infantile myofibromatosis and fibromuscular dysplasia: A case report. *J. Med. Case Rep.* **9**, (2015).

224. Brasseur, B. *et al.* Development of renal and iliac aneurysms in a child with generalized infantile myofibromatosis. *Pediatr. Nephrol. 2009 255* **25**, 983–986 (2009).

225. Rustad, C. F. *et al.* Positive response to imatinib in PDGFRB-related Kosaki overgrowth syndrome. *American Journal of Medical Genetics, Part A* at https://doi.org/10.1002/ajmg.a.62264 (2021).

226. Penttinen, M., Niemi, K. M., Vinkka-Puhakka, H., Johansson, R. & Aula, P. New progeroid disorder. *Am. J. Med. Genet.* **69**, 182–7 (1997).

227. Johnston, J. J. *et al.* A Point Mutation in PDGFRB Causes Autosomal-Dominant Penttinen Syndrome. *Am. J. Hum. Genet.* **97**, 465–74 (2015).

228. Bredrup, C. *et al.* A tyrosine kinase-activating variant Asn666Ser in PDGFRB causes a progeria-like condition in the severe end of Penttinen syndrome. *Eur. J. Hum. Genet.* **27**, 574–581 (2019).

229. Abarca, H. *et al.* Ocular pterygium—Digital keloid dysplasia. *Am. J. Med. Genet. Part A* **164**, 2901–2907 (2014).

230. Bredrup, C. *et al.* Temperature-dependent autoactivation associated with clinical variability of PDGFRB Asn666 substitutions. *Hum. Mol. Genet.* **30**, 72 (2021).

231. Hettmer, S. *et al.* Genetic testing and surveillance in infantile myofibromatosis: a report from the SIOPE Host Genome Working Group. *Fam. Cancer 2020* **1**, 1–10 (2020).

232. Arts, F. A. *et al.* PDGFRB gain-of-function mutations in sporadic infantile myofibromatosis. *Hum. Mol. Genet.* **26**, 1801–1810 (2017).

233. Qawahmed, R. Al *et al.* Infantile Myofibromatosis With Intracranial Extradural Involvement and PDGFRB Mutation: A Case Report and Review of the Literature: *https://doi.org/10.1177/1093526618787736* **22**, 258–264 (2018).

234. Pond, D. *et al.* A patient with germ-line gain-of-function PDGFRB p.N666H mutation and marked clinical response to imatinib. *Genet. Med.* **20**, 142–150 (2018).

235. Guimier, A. *et al.* A novel de novo *PDGFRB* variant in a child with severe cerebral malformations, intracerebral calcifications, and infantile myofibromatosis. *Am. J. Med. Genet. Part A* ajmg.a.61151 (2019) doi:10.1002/ajmg.a.61151.

236. Zhong, C. S. *et al.* Myofibromatosis presenting as reticulated vascular changes and subcutaneous atrophy in a patient with somatic mosaicism of PDGFRB mutation. *Br. J. Dermatol.* **179**, 1408–1409 (2018).

237. Li, Z. *et al.* Recurrent PDGFRB mutations in unicentric Castleman disease. *Leukemia* vol. 33 1035–1038 at https://doi.org/10.1038/s41375-018-0323-6 (2019).

238. Mathorne, S. W., Sørensen, K., Fagerberg, C., Bode, M. & Hertz, J. M. A novel PDGFRB sequence variant in a family with a mild form of primary familial brain calcification: a case report and a review of the literature. *BMC Neurol.* **19**, 60 (2019).

239. Golub, T., Barker, G., Lovett, M. & Gilliland, D. Fusion of PDGF receptor beta to a novel ets-like gene, tel, in chronic myelomonocytic leukemia with t(5;12) chromosomal translocation. *Cell* **77**, 307–316 (1994).

240. Savage, N., George, T. & Gotlib, J. Myeloid neoplasms associated with eosinophilia and rearrangement of PDGFRA, PDGFRB, and FGFR1: a review. *Int. J. Lab. Hematol.* **35**, 491–500 (2013).

241. Mirzaa, G. *et al.* PIK3CA-associated developmental disorders exhibit distinct classes of mutations with variable expression and tissue distribution. *JCI Insight* **1**, (2016).

242. Garde, A. *et al.* Clinical and neuroimaging findings in 33 patients with MCAP syndrome: A survey to evaluate relevant endpoints for future clinical trials. *Clin. Genet.* **99**, 650–661 (2021).

243. Sapp, J. C. *et al.* Newly delineated syndrome of congenital lipomatous overgrowth, vascular malformations, and epidermal nevi (CLOVE syndrome) in seven patients. in *American Journal of Medical Genetics, Part A* vol. 143 2944–2958 (2007).

244. Butler, M. G. Clinical and genetic aspects of the 15q11.2 BP1–BP2 microdeletion disorder. *J. Intellect. Disabil. Res.* **61**, 568–579 (2017).

245. Clayton-Smith, J. *et al.* Macrocephaly with cutis marmorata, haemangioma and syndactyly--a distinctive overgrowth syndrome. *Clin. Dysmorphol.* **6**, 291–302 (1997).

246. Kamien, B. *et al.* A Clinical Review of Generalized Overgrowth Syndromes in the Era of Massively Parallel Sequencing. *Mol. Syndromol.* **9**, 70–82 (2018).

247. Mirzaa, G. *et al.* Megalencephaly-capillary malformation (MCAP) and megalencephaly-polydactyly-polymicrogyria-hydrocephalus (MPPH) syndromes: two closely related disorders of brain overgrowth and abnormal brain and body morphogenesis. *Am. J. Med. Genet. A* **158A**, 269–291 (2012).

248. Loveday, C. *et al.* Mutations in the PP2A regulatory subunit B family genes PPP2R5B, PPP2R5C and PPP2R5D cause human overgrowth. 1–39 (2015).

249. Shang, L. *et al.* De Novo Missense Variants in PPP2R5D Are Associated with Intellectual Disability, Macrocephaly, Hypotonia, and Autism HHS Public Access. *Neurogenetics* **17**, 43–49 (2016).

250. Yeung, K. S. *et al.* Identification of mutations in the PI3K-AKT-mTOR signalling pathway in patients with macrocephaly and developmental delay and/or autism. *Mol. Autism* **8**, (2017).

251. Macken, W. L., Tischkowitz, M. & Lachlan, K. L. *PTEN* Hamartoma tumor syndrome in childhood: A review of the clinical literature. *Am. J. Med. Genet. Part C Semin. Med. Genet.* ajmg.c.31743 (2019) doi:10.1002/ajmg.c.31743.

252. Heald, B. *et al.* Frequent gastrointestinal polyps and colorectal adenocarcinomas in prospective series of PTEN mutation carriers. *Gastroenterology* **139**, 1927–1933 (2010).

253. Vinitsky, A., Zaleski, C. A., Sajjad, S. M. & McPherson, E. W. Intestinal Ganglioneuromatosis: Unusual Presentation of Cowden Syndrome Resulting in Delayed Diagnosis. *Am. J. Med. Genet. Part A* **161**, 1085–1090 (2013).

254. Rosenfeld, E. H., Chumpitazi, B. P., Castro, E. & Naik-Mathuria, B. Diffuse Intestinal Ganglioneuromatosis Causing Severe Intestinal Dysmotility in a Child With a PTEN Mutation. *J. Pediatr. Gastroenterol. Nutr.* **68**, e35–e37 (2019).

255. Bhargava, R., Yong, K. J. A. & Leonard, N. Bannayan-Riley-Ruvalcaba Syndrome: MRI Neuroimaging Features in a Series of 7 Patients. *AJNR Am. J. Neuroradiol.* **35**, 402 (2014).

256. Delatycki, M. B., Danks, A., Churchyard, A. & Zhou, X.-P. De novo germline PTEN mutation in a man with Lhermitte-Duclos disease which arose on the paternal chromosome and was transmitted to his child with polydactyly and Wormian bones. *J Med Genet* **40**, 92 (2003).

257. Buxbaum, J. D. *et al.* Mutation Screening of the PTEN Gene in Patients With Autism Spectrum Disorders and Macrocephaly. *Am. J. Med. Genet.* **144B**, 484 (2007).

258. Quan, L. & Smith, D. The VATER association. Vertebral defects, Anal atresia, T-E fistula with esophageal atresia, Radial and Renal dysplasia: a spectrum of associated defects. *J. Pediatr.* **82**, 231–240 (1973).

259. Husain, M. *et al.* Phenotypic diversity of patients diagnosed with VACTERL association. *Am. J. Med. Genet. Part A* **176**, 1830–1837 (2018).

260. Reardon, W., Zhou, X. & Eng, C. A novel germline mutation of the PTEN gene in a patient with macrocephaly, ventricular dilatation, and features of VATER association. *J. Med. Genet.* **38**, 820 (2001).

261. Solomon, B. The etiology of VACTERL association: Current knowledge and hypotheses. *Am. J. Med. Genet. C. Semin. Med. Genet.* **178**, 440–446 (2018).

262. Jentink, J. *et al.* Valproic Acid Monotherapy in Pregnancy and Major Congenital Malformations. *http://dx.doi.org/10.1056/NEJMoa0907328* **362**, 2185–2193 (2010).

263. Pilarski, R. & Eng, C. Will the real Cowden syndrome please stand up (again)? Expanding mutational and clinical spectra of the PTEN hamartoma tumour syndrome. *J. Med. Genet.* **41**, 323 (2004).

264. Hansen-Kiss, E. *et al.* A retrospective chart review of the features of PTEN hamartoma tumour syndrome in children. *J. Med. Genet.* **54**, 471–478 (2017).

265. Pilarski, R. PTEN hamartoma tumor syndrome: A clinical overview. *Cancers* vol. 11 at https://doi.org/10.3390/cancers11060844 (2019).

266. Kato, N. *et al.* Germline mutation of the PTEN gene in a Japanese patient with Cowden's disease. *Int. J. Oncol.* **18**, 1017–1022 (2001).

267. Tan, W. *et al.* The spectrum of vascular anomalies in patients with PTEN mutations: implications for diagnosis and management. *J. Med. Genet.* **44**, 594 (2007).

268. Balci, T. B. *et al.* Broad spectrum of neuropsychiatric phenotypes associated with white matter disease in PTEN hamartoma tumor syndrome. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* **177**, 101–109 (2018).

269. Cooiman, M. I. *et al.* Genetic analysis in the bariatric clinic; impact of a PTEN gene mutation. *Mol. Genet. Genomic Med.* **7**, (2019).

270. Pal, A. *et al.* PTEN Mutations as a Cause of Constitutive Insulin Sensitivity and Obesity. *http://dx.doi.org/10.1056/NEJMoa1113966* **367**, 1002–1011 (2012).

271. George, S. *et al.* A Family with Severe Insulin Resistance and Diabetes Mellitus due to a Missense Mutation in AKT2. *Science* **304**, 1325 (2004).

272. Garcia-Cao, I. *et al.* Systemic elevation of PTEN induces a tumor suppressive metabolic state. *Cell* **149**, 49 (2012).

273. Tan, M.-H. *et al.* Lifetime cancer risks in individuals with germline PTEN mutations. *Clin. Cancer Res.* **18**, 400–7 (2012).

274. Riegert-Johnson, D. L. *et al.* Cancer and Lhermitte-Duclos disease are common in Cowden syndrome patients. *Hered. Cancer Clin. Pract.* **8**, 6 (2010).

275. Mester, J. L., Zhou, M., Prescott, N. & Eng, C. Papillary Renal Cell Carcinoma is Associated with PTEN Hamartoma Tumor Syndrome. *Urology* **79**, 1187.e1 (2012).

276. Tischkowitz, M. *et al.* Cancer Surveillance Guideline for individuals with PTEN hamartoma tumour syndrome. *Eur. J. Hum. Genet.* 1–7 (2020) doi:10.1038/s41431-020-0651-7.

277. Frazier, T. W. *et al.* Molecular and Phenotypic Abnormalities in Individuals with Germline Heterozygous PTEN Mutations and Autism. *Mol. Psychiatry* **20**, 1132 (2015).

278. Herman, G. E. *et al.* Increasing knowledge ofPTEN germline mutations: Two additional patients with autism and macrocephaly. *Am. J. Med. Genet. Part A* **143A**, 589–593 (2007).

279. Butler, M. G. *et al.* Subset of individuals with autism spectrum disorders and extreme macrocephaly associated with germline PTEN tumour suppressor gene mutations. *J. Med. Genet.* **42**, 318–321 (2005).

280. Parisi, M. *et al.* The spectrum and evolution of phenotypic findings in PTEN mutation positive cases of Bannayan-Riley-Ruvalcaba syndrome. *J. Med. Genet.* **38**, 52 (2001).

281. Smpokou, P., Fox, V. L. & Tan, W.-H. PTEN hamartoma tumour syndrome: early tumour development in children. *Arch. Dis. Child.* 1–4 (2014) doi:10.1136/archdischild-2014-305997.

282. Heindl, M. *et al.* Autoimmunity, Intestinal Lymphoid Hyperplasia, and Defects in Mucosal B-Cell Homeostasis in Patients With PTEN Hamartoma Tumor Syndrome. *Gastroenterology* **142**, 1093-1096.e6 (2012).

283. Sharma, M. R., Petty, E. M. & Lesperance, M. M. Airway Obstruction Caused by PTEN Hamartoma (Bannayan-Riley-Ruvalcaba) Syndrome. *Arch. Otolaryngol. Neck Surg.* **133**, 1157–1160 (2007).

284. Stanich, P. P. *et al.* Colonic manifestations of PTEN hamartoma tumor syndrome: Case series and systematic review. *World J. Gastroenterol.* **20**, 1833 (2014).

285. Sabir, A., Parry, G., Heaton, T. & Ong, K. R. Cowden syndrome: new clinical features in a large family; joint hyperextensibility, dental abnormalities and gingival enlargement. *BMJ Case Reports CP* **14**, e236768 (2021).

286. Villeneuve, H. *et al.* Acinic cell carcinoma of the retromolar trigone region: expanding the tumor phenotype in Cowden syndrome? *Fam. Cancer* **10**, 691–694 (2011).

287. Masmoudi, A. *et al.* Cowden syndrome. *J. Dermatol. Case Rep.* **5**, 8 (2011).

288. Kehrer-Sawatzki, H. & Cooper, D. N. Classification of NF1 microdeletions and its importance for establishing genotype/phenotype correlations in patients with NF1 microdeletions. *Hum. Genet. 2021* 1–15 (2021) doi:10.1007/S00439-021-02363-3.

289. Imagawa, E. *et al.* Novel SUZ12 mutations in Weaver-like syndrome. *Clin. Genet.* (2018).

290. Cyrus, S. S. *et al.* Rare *SUZ12* variants commonly cause an overgrowth phenotype. *Am. J. Med. Genet. Part C Semin. Med. Genet.* ajmg.c.31748 (2019) doi:10.1002/ajmg.c.31748.

291. de Albuquerque Albuquerque, E. V. *et al.* Genetic investigation of patients with tall stature. *Eur. J. Endocrinol.* **182**, 139–147 (2020).

292. Choufani, S. *et al.* DNA Methylation Signature for EZH2 Functionally Classifies Sequence Variants in Three PRC2 Complex Genes. *Am. J. Hum. Genet.* 1–15 (2020) doi:10.1016/j.ajhg.2020.03.008.

293. Cyrus, S., Burkardt, D., Weaver, D. D. & Gibson, W. T. PRC2-complex related dysfunction in overgrowth syndromes: A review of EZH2, EED, and SUZ12 and their syndromic phenotypes. *Am. J. Med. Genet. Part C Semin. Med. Genet.* 1–13 (2019) doi:10.1002/ajmg.c.31754.

294. Dietz, H. *et al.* Marfan syndrome caused by a recurrent de novo missense mutation in the fibrillin gene. *Nature* **352**, 337–339 (1991).

295. Vanem, T. T. *et al.* Marfan syndrome: Evolving organ manifestations—A 10-year follow-up study. *Am. J. Med. Genet. Part A* **182**, 397–408 (2020).

296. Loeys, B. L. *et al.* The revised Ghent nosology for the Marfan syndrome. *J. Med. Genet.* **47**, 476–485 (2010).

297. Faivre, G. *et al.* The new Ghent criteria for Marfan syndrome: what do they change? *Clin. Genet.* **81**, 433–442 (2012).

298. Faivre, L. *et al.* Clinical and molecular study of 320 children with Marfan syndrome and related type I fibrillinopathies in a series of 1009 probands with pathogenic FBN1 mutations. *Pediatrics* **123**, 391–8 (2009).

299. Hochino, Y. *et al.* Sacral Arachnoid Cyst Associated with Marfan Syndrome. *Intern. Med.* **44**, 271–273 (2005).

300. Newman, P. K. & Tilley, P. J. B. Myelopathy in Marfan's syndrome. *Neurosurgery, and Psychiatry* **42**, 176–178 (1979).

301. De Paepe, A., Devereux, R., Dietz, H., Hennekam, R. & Pyeritz, R. Revised diagnostic criteria for the Marfan syndrome. *Am. J. Med. Genet.* **62**, 417–426 (1996).

302. Lai, C. S. L., Fisher, S. E., Hurst, J. A., Vargha-Khadem, F. & Monaco, A. P. A forkhead-domain gene is mutated in a severe speech and language disorder. *Nature* **413**, 519–523 (2001).

303. Morgan, A., Fisher, S., Scheffer, I. & Hildebrand, M. FOXP2-related speech and language disorder. *In: Adam MP, Ardinger HH, Pagon RA, et al., editors. GeneReviews® [Internet]. Seattle (WA): University of Washington, Seattle; 1993-2020.* (2017).

304. Vargha-Khadem, F., Gadian, D. G., Copp, A. & Mishkin, M. FOXP2 and the neuroanatomy of speech and language. *Nature Reviews Neuroscience* vol. 6 131–138 at https://doi.org/10.1038/nrn1605 (2005).

305. MacDermot, K. D. *et al.* Identification of FOXP2 truncation as a novel cause of developmental speech and language deficits. *Am. J. Hum. Genet.* **76**, 1074–1080 (2005).

306. Turner, S. J. *et al.* Small intragenic deletion in FOXP2 associated with childhood apraxia of speech and dysarthria. *Am. J. Med. Genet. Part A* **161**, 2321–2326 (2013).

307. Zeesman, S. *et al.* Speech and language impairment and oromotor dyspraxia due to deletion of 7q31 that involves FOXP2. *Am. J. Med. Genet.* **140 A**, 509–514 (2006).

308. Tomblin, J. *et al.* Language Features in a Mother and Daughter of a Chromosome 7;13 Translocation Involving FOXP2. *J. Speech. Lang. Hear. Res.* **52**, (2009).

309. Laffin, J. J. S. *et al.* Novel candidate genes and regions for childhood apraxia of speech identified by array comparative genomic hybridization. *Genet. Med.* **14**, 928–936 (2012).

310. Reuter, M. S. *et al.* FOXP2 variants in 14 individuals with developmental speech and language disorders broaden the mutational and clinical spectrum. *J. Med. Genet.* **54**, 64–72 (2017).

311. Fee, E. J. The phonological system of a specifically language-impaired population. *Clin. Linguist. Phon.* **9**, 189–209 (1995).

312. Feuk, L. *et al.* Absence of a paternally inherited FOXP2 gene in developmental verbal dyspraxia. *Am. J. Hum. Genet.* **79**, 965–972 (2006).

313. Rice, G. M. *et al.* Phenotype of FOXP2 haploinsufficiency in a mother and son. *Am. J. Med. Genet. Part A* **158 A**, 174–181 (2012).

314. Žilina, O. *et al.* Maternally and paternally inherited deletion of 7q31 involving the FOXP2 gene in two families. *American Journal of Medical Genetics, Part A* vol. 158

A 254–256 at https://doi.org/10.1002/ajmg.a.34378 (2012).

315. Cylke, R., Karpeta, E., Bieniasz, M. & Kosieradzki, M. Urologic Complications After Transplantation of Kidneys With Duplicated Ureter: A Retrospective Study. *Transplant. Proc.* **51**, 779–782 (2019).

316. Biesecker, L. G. The Greig cephalopolysyndactyly syndrome. *Orphanet J. Rare Dis.* **3**, 10 (2008).

317. Baraitser, M., Winter, R. M. & Brett, E. M. Greig cephalopolysyndactyly: report of 13 affected individuals in three families. *Clin. Genet.* **24**, 257–265 (1983).

318. Démurger, F. *et al.* New insights into genotype–phenotype correlation for GLI3 mutations. *Eur. J. Hum. Genet.* **23**, 92 (2015).

319. Johnston, J. J. *et al.* Molecular and Clinical Analyses of Greig Cephalopolysyndactyly and Pallister-Hall Syndromes: Robust Phenotype Prediction from the Type and Position of GLI3 Mutations. *Am. J. Hum. Genet.* **76**, 609 (2005).

320. Vortkamp, A. & Grzeschik, K.-H. Deletion of GLI3 supports the homology of the human Greig cephalopolysyndactyly syndrome (GCPS) and the mouse mutant extra toes (Xt). *Mamm. Genome* **3**, 461–463 (1992).

321. Hall, J. *et al.* Congenital hypothalamic hamartoblastoma, hypopituitarism, imperforate anus and postaxial polydactyly--a new syndrome? Part I: clinical, causal, and pathogenetic considerations. *Am. J. Med. Genet.* **7**, 47–74 (1980).

322. Johnston, J. J. *et al.* Molecular analysis expands the spectrum of phenotypes associated with GLI3 mutations. *Hum. Mutat.* **31**, 1142 (2010).

323. Johnston, J. J. *et al.* Zoom-in comparative genomic hybridisation arrays for the characterisation of variable breakpoint contiguous gene syndromes. *J. Med. Genet.* **44**, e59 (2007).

324. Flex, E. *et al.* Aberrant Function of the C-Terminal Tail of HIST1H1E Accelerates Cellular Senescence and Causes Premature Aging. *Am. J. Hum. Genet.* **105**, 493–508 (2019).

325. Burkardt, D. D. C. *et al.* HIST1H1E heterozygous protein-truncating variants cause a recognizable syndrome with intellectual disability and distinctive facial gestalt: A study to clarify the HIST1H1E syndrome phenotype in 30 individuals. *Am. J. Med. Genet. Part A* **179**, 2049–2055 (2019).

326. Takenouchi, T., Uehara, T., Kosaki, K. & Mizuno, S. Growth pattern of Rahman syndrome. *Am. J. Med. Genet. Part A* **176**, 712–714 (2018).

327. K, L. *et al.* Polyhydramnios, Transient Antenatal Bartter's Syndrome, and MAGED2 Mutations. *N. Engl. J. Med.* **374**, 1853–1863 (2016).

328. Legrand, A. *et al.* Prevalence of Novel MAGED2 Mutations in Antenatal Bartter Syndrome. *Clin. J. Am. Soc. Nephrol.* **13**, 242 (2018).

329. Stessman, H. A. F. *et al.* Targeted sequencing identifies 91 neurodevelopmental-disorder risk genes with autism and developmental-disability biases. *Nat. Genet.* **49**,

515–526 (2017).

330. McRae, J. F. *et al.* Prevalence and architecture of de novo mutations in developmental disorders. *Nature* **542**, 433–438 (2017).

331. Wickramasekara, R. N. & Stessman, H. A. F. Histone 4 lysine 20 methylation: A case for neurodevelopmental disease. *Biology (Basel).* **8**, (2019).

332. Trinh, J. *et al.* Novel pathogenic variants and multiple molecular diagnoses in neurodevelopmental disorders. *J. Neurodev. Disord.* **11**, 11 (2019).

333. Faundes, V. *et al.* Histone Lysine Methylases and Demethylases in the Landscape of Human Developmental Disorders. *Am. J. Hum. Genet.* **102**, 175–187 (2018).

334. Gorlin, R. J. & Goltz, R. W. Multiple Nevoid Basal-Cell Epithelioma, Jaw Cysts and Bifid Rib. *http://dx.doi.org/10.1056/NEJM196005052621803* **262**, 908–912 (2010).

335. Evans, D. G. R. *et al.* Complications of the naevoid basal cell carcinoma syndrome: results of a population based study. *4J Med Genet* **30**, 460–464 (2018).

336. Kimonis, V. E. *et al.* Clinical and radiological features in young individuals with nevoid basal cell carcinoma syndrome. *Genet. Med. 2013 151* **15**, 79–83 (2012).

337. Evans, G. & Farndon, P. A. Nevoid Basal Cell Carcinoma Syndrome Synonyms: Basal Cell Nevus Syndrome (BCNS), Gorlin Syndrome, NBCCS. (1993).

338. Xin, B. *et al.* Homozygous frameshift mutation in TMCO1 causes a syndrome with craniofacial dysmorphism, skeletal anomalies, and mental retardation. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 258 (2010).

339. Alanay, Y. *et al.* TMCO1 deficiency causes autosomal recessive cerebrofaciothoracic dysplasia. *Am. J. Med. Genet. Part A* **164**, 291–304 (2014).

340. Batchelor-Regan, H., Xin, B., Zhou, A. & Wang, H. From Disease Description and Gene Discovery to Functional Cell Pathway: A Decade-Long Journey for TMCO1. *Front. Genet.* **12**, (2021).

341. Pehlivan, D. *et al.* Whole-exome sequencing links TMCO1 defect syndrome with cerebro-facio-thoracic dysplasia. *Eur. J. Hum. Genet.* **22**, 1145 (2014).

342. Tender, J. A. F. & Ferreira, C. R. Cerebro-facio-thoracic dysplasia (Pascual-Castroviejo syndrome): Identification of a novel mutation, use of facial recognition analysis, and review of the literature. *Transl. Sci. Rare Dis.* **3**, 37 (2018).

343. Sharkia, R. *et al.* A novel biallelic loss-of-function mutation in TMCO1 gene confirming and expanding the phenotype spectrum of cerebro-facio-thoracic dysplasia. *Am. J. Med. Genet. Part A* **179**, 1338–1345 (2019).

344. Yates, T. M. *et al.* Cerebrofaciothoracic dysplasia: Four new patients with a recurrent TMCO1 pathogenic variant. *Am. J. Med. Genet. Part A* **179**, 43–49 (2019).

345. Caglayan, A. *et al.* Whole-exome sequencing identified a patient with TMCO1 defect syndrome and expands the phenotic spectrum. *Clin. Genet.* **84**, 394 (2013).

346. van Mullem, A. A., Visser, T. J. & Peeters, R. P. Clinical Consequences of Mutations

in Thyroid Hormone Receptor-α1. *Eur. Thyroid J.* **3**, 17–24 (2014).

347. Bochukova, E. *et al.* A Mutation in the Thyroid Hormone Receptor Alpha Gene. *N. Engl. J. Med.* **366**, 243–249 (2012).

348. Erbaş, İ. M. & Demir, K. The Clinical Spectrum of Resistance to Thyroid Hormone Alpha in Children and Adults. *J. Clin. Res. Pediatr. Endocrinol.* **0**, 0–0 (2020).

349. Moran, C. & Chatterjee, K. Resistance to thyroid hormone due to defective thyroid receptor alpha. *Best Practice and Research: Clinical Endocrinology and Metabolism* vol. 29 647–657 at https://doi.org/10.1016/j.beem.2015.07.007 (2015).

350. Tylki-Szymańska, A. *et al.* Thyroid hormone resistance syndrome due to mutations in the thyroid hormone receptor α gene (THRA). *J. Med. Genet.* **52**, 312–6 (2015).

351. Tropeano, M., Andrieux, J. & Collier, D. A. Clinical utility gene card for: 16p13.11 microdeletion syndrome. *Eur. J. Hum. Genet.* 22 (2014) doi:10.1038/ejhg.2013.230.

352. Nagamani, S. C. S. *et al.* Phenotypic manifestations of copy number variation in chromosome 16p13.11. *Eur. J. Hum. Genet.* **19**, 280–286 (2011).

353. El Khattabi, A. 16p13.11 microduplication in 45 new patients: refined clinical significance and genotype-phenotype correlations. *J Med Genet* **0**, 1–7 (2018).

354. Ramalingam, A. *et al.* 16p13.11 duplication is a risk factor for a wide spectrum of neuropsychiatric disorders. *J. Hum. Genet.* **56**, 541–544 (2011).

355. Kuang, S. Q. *et al.* Recurrent chromosome 16p13.1 duplications are a risk factor for aortic dissections. *PLoS Genet.* **7**, (2011).

356. Bakircioglu, M. *et al.* The Essential Role of Centrosomal NDE1 in Human Cerebral Cortex Neurogenesis. (2011) doi:10.1016/j.ajhg.2011.03.019.

357. Alkuraya, F. S. *et al.* Human Mutations in NDE1 Cause Extreme Microcephaly with Lissencephaly. (2011) doi:10.1016/j.ajhg.2011.04.003.

358. Burnside, R. D. *et al.* Microdeletion/microduplication of proximal 15q11.2 between BP1 and BP2: a susceptibility region for neurological dysfunction including developmental and language delay. *Hum. Genet.* **130**, 517 (2011).

359. Cafferkey, M., Ahn, J. W., Flinter, F. & Ogilvie, C. Phenotypic features in patients with 15q11.2(BP1-BP2) deletion: Further delineation of an emerging syndrome. *Am. J. Med. Genet. Part A* **164**, 1916–1922 (2014).

360. Rafi, S. K. & Butler, M. G. The 15q11.2 bp1-bp2 microdeletion (burnside–butler) syndrome: In silico analyses of the four coding genes reveal functional associations with neurodevelopmental phenotypes. *Int. J. Mol. Sci.* **21**, (2020).

361. Hancarova, M. *et al.* Association of 17q24.2-q24.3 deletions with recognizable phenotype and short telomeres. *Am. J. Med. Genet. Part A* **176**, 1438–1442 (2018).

362. Turner, T. N. *et al.* Genome Sequencing of Autism-Affected Families Reveals Disruption of Putative Noncoding Regulatory DNA. *Am. J. Hum. Genet.* **98**, 58–74 (2016).

363. Anazi, S. *et al.* Clinical genomics expands the morbid genome of intellectual disability and offers a high diagnostic yield. (2016) doi:10.1038/mp.2016.113.

364. Küry, S. *et al.* De Novo Disruption of the Proteasome Regulatory Subunit PSMD12 Causes a Syndromic Neurodevelopmental Disorder. *Am. J. Hum. Genet.* **100**, 352–363 (2017).

365. Stankiewicz, P. *et al.* Haploinsufficiency of the Chromatin Remodeler BPTF Causes Syndromic Developmental and Speech Delay, Postnatal Microcephaly, and Dysmorphic Features. *Am. J. Hum. Genet.* **101**, 503–515 (2017).

366. Reuter, M. S. *et al.* Diagnostic yield and novel candidate genes by exome sequencing in 152 consanguineous families with neurodevelopmental disorders. *JAMA Psychiatry* **74**, 293–299 (2017).

367. Wright, C. F. *et al.* Assessing the Pathogenicity, Penetrance, and Expressivity of Putative Disease-Causing Variants in a Population Setting. doi:10.1016/j.ajhg.2018.12.015.

368. Gurrieri, F. *et al.* NFIX mutations affecting the DNA-binding domain cause a peculiar overgrowth syndrome (Malan syndrome): A new patients series. *Eur. J. Med. Genet.* 4–7 (2015) doi:10.1016/j.ejmg.2015.06.009.

369. Arts, F. A. *et al.* PDGFRB mutants found in patients with familial infantile myofibromatosis or overgrowth syndrome are oncogenic and sensitive to imatinib. *Oncogene* 1–10 (2015) doi:10.1038/onc.2015.383.

370. Faundes, V. *et al.* Histone Lysine Methylases and Demethylases in the Landscape of Human Developmental Disorders. *Am. J. Hum. Genet.* **102**, 175–187 (2018).

371. Guo, D., Tan, F. K., Cantu, A., Plon, S. E. & Milewicz, D. M. FBN1 exon 2 splicing error in a patient with Marfan syndrome. *Am. J. Med. Genet.* **101**, 130–4 (2001).

372. Matissek, S. J. & Elsawa, S. F. GLI3: a mediator of genetic diseases, development and cancer. *Cell Commun. Signal.* **18**, (2020).

373. Reinders, M. G. *et al.* New mutations and an updated database for the patched-1 (PTCH1) gene. *Mol. Genet. Genomic Med.* **6**, 409 (2018).

374. Larsen, L. J. & Møller, L. B. Crosstalk of Hedgehog and mTORC1 Pathways. *Cells* **9**, (2020).

375. Takeda, N. *et al.* TGF-β Signaling-Related Genes and Thoracic Aortic Aneurysms and Dissections. *Int. J. Mol. Sci.* **19**, (2018).

376. Mudry, P. *et al.* Case report: Rapid and durable response to PDGFR targeted therapy in a child with refractory multiple infantile myofibromatosis and a heterozygous germline mutation of the PDGFRB gene. *BMC Cancer* **17**, (2017).

377. Choufani, S. *et al.* NSD1 mutations generate a genome-wide DNA methylation signature. *Nat. Commun.* **6**, 10207 (2015).

378. Choufani, S. *et al.* DNA Methylation Signature for EZH2 Functionally Classifies Sequence Variants in Three PRC2 Complex Genes. *Am. J. Hum. Genet.* (2020)

doi:10.1016/j.ajhg.2020.03.008.

Appendix

A. ACMG and AMP standards and guidelines for the interpretation of sequence variants

ACMG STANDARDS AND GUIDELINES

Table 3 Criteria for classifying pathogenic variants

Evidence of pathogenicity	Category
Very strong	PVS1 null variant (nonsense, frameshift, canonical ±1 or 2 splice sites, initiation codon, single or multiexon deletion) in a gene where 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	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, and so on, can contribute to nonmaternity. 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 with the prevalence in controls Note 1: Relative risk or OR, as obtained from case–control studies, is >5.0, and the confidence interval around the estimate of relative risk or OR does not include 1.0. See the article 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	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) (Table 6) in Exome Sequencing Project, 1000 Genomes Project, or Exome Aggregation Consortium Caveat: Population data for insertions/deletions 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 as a result of in-frame deletions/insertions in a nonrepeat 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 acid/protein level. PM6 Assumed de novo, but without confirmation of paternity and maternity
Supporting	PP1 Cosegregation 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 in which 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: Because 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

LOF, loss of function; OR, odds ratio.

Taken from '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' by Richards et al.[171]

B. Documentation for the POD study

B.1 IRAS application form

Full Set of Project Data IRAS Version 5.21

Welcome to the Integrated Research Application System

IRAS Project Filter

The integrated dataset required for your project will be created from the answers you give to the following questions. The system will generate only those questions and sections which (a) apply to your study type and (b) are required by the bodies reviewing your study. Please ensure you answer all the questions before proceeding with your applications.

Please complete the questions in order. If you change the response to a question, please select 'Save' and review all the questions as your change may have affected subsequent questions.

Please enter a short title for this project (maximum 70 characters)
POD study

1. Is your project research?

(●) Yes ( ) No

2. Select one category from the list below:

( ) Clinical trial of an investigational medicinal product
( ) Clinical investigation or other study of a medical device
( ) Combined trial of an investigational medicinal product and an investigational medical device
( ) Other clinical trial to study a novel intervention or randomised clinical trial to compare interventions in clinical practice
( ) Basic science study involving procedures with human participants
( ) Study administering questionnaires/interviews for quantitative analysis, or using mixed quantitative/qualitative methodology
( ) Study involving qualitative methods only
(●) Study limited to working with human tissue samples (or other human biological samples) and data (specific project only)
( ) Study limited to working with data (specific project only)
( ) Research tissue bank
( ) Research database

If your work does not fit any of these categories, select the option below:

( ) Other study

2a. Please answer the following question(s):

a) Will you be taking new samples primarily for research purposes (i.e. not surplus or existing stored samples), including any removal of organs or tissue from the deceased?	(●) Yes	( ) No
b) Will you be using surplus tissue or existing stored samples identifiable to the researcher?	(●) Yes	( ) No
c) Will you be using only surplus tissue or existing stored samples not identifiable to the researcher?	( ) Yes	(●) No
d) Will you be processing identifiable data at any stage of the research (including in the identification of participants)?	(●) Yes	( ) No

3. In which countries of the UK will the research sites be located? *(Tick all that apply)*

- [x] England
- [ ] Scotland
- [x] Wales
- [ ] Northern Ireland

3a. In which country of the UK will the lead NHS R&D office be located:

- (•) England
- ( ) Scotland
- ( ) Wales
- ( ) Northern Ireland
- ( ) This study does not involve the NHS

4. Which applications do you require?

- [ ] IRAS Form
- [x] NHS/HSC Research and Development offices
- [ ] Social Care Research Ethics Committee
- [x] Research Ethics Committee
- [ ] Confidentiality Advisory Group (CAG)
- [ ] Her Majesty's Prison and Probation Service (HMPPS)

5. Will any research sites in this study be NHS organisations?

(•) Yes ( ) No

5a. Are all the research costs and infrastructure costs (funding for the support and facilities needed to carry out the research e.g. NHS support costs) for this study provided by a NIHR Biomedical Research Centre (BRC), NIHR Applied Research Collaboration (ARC), NIHR Patient Safety Translational Research Centre (PSTRC), or an NIHR Medtech and In Vitro Diagnostic Co-operative (MIC) in all study sites?

Please see information button for further details.

( ) Yes (•) No

Please see information button for further details.

5b. Do you wish to make an application for the study to be considered for NIHR Clinical Research Network (CRN) Support and inclusion in the NIHR Clinical Research Network Portfolio?

Please see information button for further details.

(•) Yes ( ) No

The NIHR Clinical Research Network (CRN) provides researchers with the practical support they need to make clinical studies happen in the NHS in England e.g. by providing access to the people and facilities needed to carry out research "on the ground".

If you select yes to this question, information from your IRAS submission will automatically be shared with the NIHR CRN. ***Submission of a Portfolio Application Form (PAF) is no longer required.***

6. Do you plan to include any participants who are children?

◉ Yes ○ No

7. Do you plan at any stage of the project to undertake intrusive research involving adults lacking capacity to consent for themselves?

◉ Yes ○ No

Answer Yes if you plan to recruit living participants aged 16 or over who lack capacity, or to retain them in the study following loss of capacity. Intrusive research means any research with the living requiring consent in law. This includes use of identifiable tissue samples or personal information, except where application is being made to the Confidentiality Advisory Group to set aside the common law duty of confidentiality in England and Wales. Please consult the guidance notes for further information on the legal frameworks for research involving adults lacking capacity in the UK.

8. Do you plan to include any participants who are prisoners or young offenders in the custody of HM Prison Service or who are offenders supervised by the probation service in England or Wales?

○ Yes ◉ No

9. Is the study or any part of it being undertaken as an educational project?

◉ Yes ○ No

Please describe briefly the involvement of the student(s):
This project is being undertaken as part of a PhD, with the student being the CI.

9a. Is the project being undertaken in part fulfilment of a PhD or other doctorate?

◉ Yes ○ No

10. Will this research be financially supported by the United States Department of Health and Human Services or any of its divisions, agencies or programs?

○ Yes ◉ No

11. Will identifiable patient data be accessed outside the care team without prior consent at any stage of the project (including identification of potential participants)?

○ Yes ◉ No

Integrated Research Application System
Application Form for Research limited to working with human tissue samples and/or data

The Chief Investigator should complete this form. Guidance on the questions is available wherever you see this symbol displayed. We recommend reading the guidance first. The complete guidance and a glossary are available by selecting Help.

Please define any terms or acronyms that might not be familar to lay reviewers of the application.

Short title and version number: (maximum 70 characters - this will be inserted as header on all forms)
POD study

PART A: Core study information

1. ADMINISTRATIVE DETAILS

A1. Full title of the research:

Phenotyping of rare genetic overgrowth disorders

A2-1. Educational projects

Name and contact details of student(s):

Name and contact details of academic supervisor(s):

Please state which academic supervisor(s) has responsibility for which student(s):
Please click "Save now" before completing this table. This will ensure that all of the student and academic supervisor details are shown correctly.

Student(s)	Academic supervisor(s)

A copy of a current CV for the student and the academic supervisor (maximum 2 pages of A4) must be submitted with the application.

A2-2. Who will act as Chief Investigator for this study?

- ○ Student
- ○ Academic supervisor
- ◉ Other

A3-1. Chief Investigator:

	Title	Forename/Initials	Surname
	Dr	Alison	Foster
Post	Research Fellow		
Qualifications	BSc MBBS MRCPCH		
ORCID ID			
Employer	Birmingham Women's Hospital NHS Foundation Trust		

Work Address	Mindelsohn Way Edgbaston Birmingham
Post Code	B15 2TG
Work E-mail	
* Personal E-mail	
Work Telephone	01214721377
* Personal Telephone/Mobile	
Fax	

** This information is optional. It will not be placed in the public domain or disclosed to any other third party without prior consent.*
A copy of a current CV (maximum 2 pages of A4) for the Chief Investigator must be submitted with the application.

A4. Who is the contact on behalf of the sponsor for all correspondence relating to applications for this project?
This contact will receive copies of all correspondence from REC and HRA/R&D reviewers that is sent to the CI.

Title	Forename/Initials	Surname
Dr	Sean	Jennings

Address	Research Governance and Ethics Manager, Research Support Group Room 119, Aston Webb Building University of Birmingham, Edgbaston
Post Code	B15 2TT
E-mail	researchgovernance@contacts.bham.ac.uk
Telephone	01214158011
Fax	

A5-1. Research reference numbers. *Please give any relevant references for your study:*

Applicant's/organisation's own reference number, e.g. R & D (if available):	
Sponsor's/protocol number:	RG_14-249
Protocol Version:	1.0
Protocol Date:	17/04/2015
Funder's reference number (enter the reference number or state not applicable):	
Project website:	

Additional reference number(s):

Ref.Number Description	Reference Number

Registration of research studies is encouraged wherever possible. You may be able to register your study through your NHS organisation or a register run by a medical research charity, or publish your protocol through an open access publisher. If you have registered your study please give details in the "Additional reference number(s)" section.

A5-2. Is this application linked to a previous study or another current application?

○ Yes ◉ No

Please give brief details and reference numbers.

2. OVERVIEW OF THE RESEARCH

To provide all the information required by review bodies and research information systems, we ask a number of specific questions. This section invites you to give an overview using language comprehensible to lay reviewers and members of the public. Please read the guidance notes for advice on this section.

A6-1. Summary of the study. *Please provide a brief summary of the research (maximum 300 words) using language easily understood by lay reviewers and members of the public. Where the research is reviewed by a REC within the UK Health Departments' Research Ethics Service, this summary will be published on the Health Research Authority (HRA) website following the ethical review. Please refer to the question specific guidance for this question.*

Overgrowth disorders are a group of rare genetic conditions that cause children to be larger than others of the same age. They are associated with a wide spectrum of medical problems, including learning disability, congenital abnormalities, and in some cases an increased risk of developing tumours. Overgrowth disorders are genetic, either inherited from a parent or occurring for the first time in a child, and lifelong.

The medical complications, prognosis and recurrence risks for an individual with overgrowth are determined by the underlying cause, and achieving a diagnosis enables optimal care to be provided. In recent years a number of novel genes have been identified, but the clinical course of these conditions is not yet known and access to genetic testing is limited. Even in individuals with a diagnosis of a relatively well known condition, the clinical features can differ from expected, suggesting the existence of genetic modifying factors. There are also many individuals who do not have a clinical or molecular diagnosis, indicating that there are other novel causes of overgrowth yet to be discovered.

This project will study many individuals with overgrowth disorders. Data including history and examination, complications, laboratory investigations, imaging, clinical photography and molecular genetic data will be recorded in detail and held on an access controlled secure public database managed by the NIHR Rare Diseases Translational Research Collaboration.

The aim of this study is to
a) understand the clinical course of overgrowth disorders
b) investigate the underlying genetic causes of overgrowth
c) study the associations between genetic causes and clinical features in individuals with overgrowth

Understanding the genetic basis of these disorders may also help identify molecular genetic targets for potential future therapeutic interventions.

A6-2. Summary of main issues. *Please summarise the main ethical, legal, or management issues arising from your study and say how you have addressed them.*

Not all studies raise significant issues. Some studies may have straightforward ethical or other issues that can be identified and managed routinely. Others may present significant issues requiring further consideration by a REC, R&D office or other review body (as appropriate to the issue). Studies that present a minimal risk to participants may raise complex organisational or legal issues. You should try to consider all the types of issues that the different reviewers may need to consider.

This project aims to increase our understanding of the causes and impacts of a rare disease, genetic overgrowth disorders. Recruitment will be from those specialties most likely to see patients with overgrowth disorders, paediatric endocrinology and clinical genetics. These disorders are life-long conditions and in some cases are associated with variable degrees of intellectual disability therefore children, adults, and adults with learning disabilities will be asked to participate.

Children under the age of 16 will be asked to assent to the study and their parents will give consent. Individuals who turn 16 during the course of the study will be contacted and re-consented as adults.

It is important that adults with learning difficulties who may not be able to give consent themselves are not denied the opportunity to participate in research that may be beneficial in the future to other individuals with these conditions. Therefore an appointed consultee (friend, relative or other appropriate individual) will be asked to sign a declaration form to allow participation in the study.

Participants will be requested to provide a blood or saliva sample for potential genetic investigations and storage.

The results of genetic investigations are best interpreted in the context of the results of parental investigations (trio

analysis). Therefore informed consent will also need to be taken from parents who participate in this aspect of the research. Individuals will not be excluded from the study if their parents are unavailable or do not wish to participate.

Venepuncture may cause discomfort and/or bruising and there is a minimal risk of infection. The blood sample may be taken at a time when the individual is having venepuncture for a separate reason and therefore will not always require an additional venepuncture. Saliva sampling will be offered as an alternative if venepuncture is not possible or declined.

A small number of individuals may have overgrowth affecting a single region of the body (eg a single limb) instead of constitutional overgrowth affecting the whole body. In these individuals a blood or saliva sample is unlikely to contain DNA with the genetic variant causing overgrowth. In this case a skin biopsy of the area overlying the region of overgrowth is an appropriate sample for analysis. Skin biopsies will be performed under local anaesthetic by a health professional trained in this technique. Alternatively, if a participant is undergoing a procedure or operation under general anaesthetic as part of their clinical care, they may opt to have the skin biopsy taken at this time. Skin biopsy is a low risk procedure. It may cause a small scar and there is a small risk of bleeding or infection.

Participants who have previously had a operation or biopsy on an area of overgrowth or tumour will be asked to consent for analysis of stored tissue.

Participants will be able to opt out of any or all of these sampling methods and subsequent analysis. Participants who choose not provide a sample will still be able to participate in the clinical phenotyping aspects of the study.

The ethical issues around incidental findings as a result of new genomic technologies have been widely debated. As our ability to perform more detailed genetic tests improves, with the aim of finding the cause of an individual's condition, we are also more likely to identify other variants that are unrelated to the presenting condition. In some cases the clinical significance of the variant may be unknown or uncertain. In others the clinical significance may be more certain, but the condition may be one without treatment. In children there is the additional issue that testing for adult onset disorders in childhood is not thought to be in the child's best interest. However, it is also possible that a finding of no clinical significance in a child may be relevant to other family members.

There are currently no best practice guidelines as to whether incidental findings identified through research (as opposed to clinical practice) should be disclosed to participants. The arguments in favour of non-disclosure are that feedback may be psychologically harmful and there is no evidence of any benefit in doing so. In addition, any findings generated through genomic testing are not visible to the researcher and must be actively sought through the data analysis process, validated and interpreted by a clinical laboratory, and fed back by a professional with expertise in genetic counselling. Each of these steps requires additional resources and adversely affects the feasibility of performing research in this area. Finally, participation in research is an altruistic activity and while the potential for harm must be minimised as much as possible, participants do not usually expect any direct benefit from their participation. The disadvantages of not disclosing incidental findings are the potential reduction in autonomy and missed opportunity for detection of possible preventable disease.

At present several large UK research studies, including the NIHR BioResource for Rare Disease, and the Deciphering Developmental Disorders study (DDD), do not disclose any incidental findings. In line with this current practice, this study will disclose results that are directly relevant to the individuals' presenting condition to their recruiting clinician to help inform their clinical management. These results should be confirmed in a clinical laboratory. However individuals will be able to opt out of disclosure of research findings if they choose. No incidental findings will be disclosed.

There is currently considerable debate on a national scale regarding whether incidental findings should be disclosed in a research setting. Should the consensus on best practice change in the future, an amendment will be sought and new consent forms put in place for prospective use.

Patient data will be submitted by recruiting clinicians. All data will be held securely on the NIHR RD-TRC OpenClinica database. The Rare Diseases Clinical Infrastructure Team (RDCIT) are responsible for implementing and maintaining the OpenClinica database. The database is hosted in the UK and information governance accredited by the Health and Social Care Information Centre, the national provider of information, data and IT systems for health and social care in the UK.

3. PURPOSE AND DESIGN OF THE RESEARCH

A7. Select the appropriate methodology description for this research. *Please tick all that apply:*

- ☐ Case series/ case note review
- ☐ Case control

- [x] Cohort observation
- [ ] Controlled trial without randomisation
- [ ] Cross-sectional study
- [ ] Database analysis
- [ ] Epidemiology
- [ ] Feasibility/ pilot study
- [x] Laboratory study
- [ ] Metanalysis
- [ ] Qualitative research
- [ ] Questionnaire, interview or observation study
- [ ] Randomised controlled trial
- [ ] Other (please specify)

A10. What is the principal research question/objective? *Please put this in language comprehensible to a lay person.*

The objective of this study is to characterise a cohort of individuals with overgrowth disorders. The primary research questions are:
What is the natural history of overgrowth disorders in terms of growth pattern, associated medical problems, and tumour risk?
What are the genetic causes of overgrowth?
How do different genetic causes of overgrowth correlate with the different medical problems that occur?

A11. What are the secondary research questions/objectives if applicable? *Please put this in language comprehensible to a lay person.*

•to provide a cohort with which to undertake future academic studies into the pathophysiology of overgrowth conditions
•to support recruitment to future multicentre studies of therapeutic interventions in overgrowth disorders

A12. What is the scientific justification for the research? *Please put this in language comprehensible to a lay person.*

The study is part of the NIHR Rare Diseases Translational Research Collaboration (NIHR RD-TRC), a national initiative to support research on rare disease. Rare diseases are individually rare, but there are several thousand rare diseases and collectively they affect 1 in 17 of the UK population. Historically rare diseases have been difficult to study because of their low prevalence, and this has been a barrier to increasing our understanding of these conditions and therefore improving clinical management. This detailed cohort study will increase our understanding of the clinical features and genetic aetiology of these rare diseases.

Our understanding of the molecular genetic causes of overgrowth disorders is likely to increase rapidly in the near future with the advent of new genomic technologies and the introduction of the 100,000 Genomes Project. Large quantities of genomic data can now be generated quickly and at low cost. However, detailed phenotypic (clinical) data is essential in order to interpret this information and translate it into patient benefit. This study will record detailed phenotypic data and correlate this with genotypic data.

Establishing a cohort of patients with detailed phenotypic information and comprehensive genomic analysis will facilitate future research on overgrowth disorders including trials of therapeutic intervention.

This study may also help increase our understanding of growth in a wider context. Several known genes that cause overgrowth syndromes are also implicated in the development of cancers, and therefore this research may increase our understanding of the molecular aetiology of cancer and potentially identify new targets for therapy of overgrowth and malignancy.

A13. Please summarise your design and methodology. *It should be clear exactly what will happen to the research participant, how many times and in what order. Please complete this section in language comprehensible to the lay person.*

Do not simply reproduce or refer to the protocol. Further guidance is available in the guidance notes.

Potential participants will be identified by their clinician (clinical geneticist or paediatric endocrinologist) and invited to participate in the study during a routine outpatient appointment. If a clinician identifies an individual as a potential participant prior to clinic, a leaflet introducing the study may be sent out prior to the appointment. Alternatively, individuals may choose to participate in the study at the Child Growth Foundation (CGF) national conference.

Participants will fall into one of the following groups: competent adults, children age under 16, or adults unable to give consent due to a learning disability. In each case, the parents of a participant will also be asked to participate in the study. Appropriate patient information leaflets and consent forms will be provided in each case. In cases where an adult is unable to give consent, the opinion of a consultee will be sought. Children age under 16 will be asked to give assent to participate in the study, and a parent will give consent on their behalf. Individuals will be given at least 24 hours to decide whether they would like to participate in the study. Given the non-invasive nature of the research, individuals may choose to consent after a shorter time period.

The clinical data collection will consist of comprehensive history taking and clinical examination in conjunction with review of the medical notes. Clinicians will record this data on a proforma or enter it directly into the database. Investigation results such as imaging (and clinical photography if consent given for this) will also be uploaded either at the time of the appointment or subsequently by a member of the research team.

If a participant has not already had genetic diagnostic testing as part of their clinical care this should be offered, for example CGH microarray testing, single gene testing if appropriate, or through a gene panel of overgrowth genes. If a participant does not have a molecular genetic diagnosis, their sample may undergo genetic analysis through exome sequencing and/or methylation analysis as part of the study. Alternatively sequencing may be available through the 100,000 Genomes Project as part of a participant's clinical care.

Participants will be asked to provide a blood sample. This sample could be taken at the same time as other bloods are being taken for routine clinical testing. If the individual does not wish to provide a blood sample, a saliva sample may be given instead.

A small number of participants who have overgrowth of a single region of the body will be be invited to provide a skin sample. This will be taken by a trained professional under local anaesthetic. Alternatively if the individual is having a general anaesthestic for clinical reasons as part of their routine health care, they may choose to have the skin biopsy taken at this time.

Individuals who have previously had tissue from an area of overgrowth or tumour stored as a result of prior biopsy or surgery will be asked to give consent for analysis of this tissue.

Blood, saliva, skin biopsy and tissue samples will be divided into two samples. The first will be sent to the West Midlands Regional Genetics Laboratory for DNA extraction and storage. The second will be sent to a licensed biorepository for storage.

Participants may choose to opt out of providing any or all samples.

Individuals' parents will be asked to participate in the study, as trio analysis facilitates interpretation of the genetic data. However if one or both parents decline this this will not exclude an individual from participating in the study.

Participants will only need to be seen once, at their routine outpatient appointment. If they have multiple clinical appointments during the time period of the study the clinician will be asked to update the clinical data collection. Following the end point of the study, participants may be contacted about other ethically approved research studies they may wish to participate in.

If the study identifies the genetic cause of a participant's overgrowth, and the participant has consented to this information being fed back, the results will be sent to the recruiting clinician. These results will need to be confirmed in a clinical laboratory and fed back to the individual by their health care team.

If a child turns 16 during the course of the study, they will be contacted and re-consented as an adult if they wish to continue participating in the study.

A14-1. In which aspects of the research process have you actively involved, or will you involve, patients, service users, and/or their carers, or members of the public?

☑ Design of the research

- [ ] Management of the research
- [x] Undertaking the research
- [ ] Analysis of results
- [x] Dissemination of findings
- [ ] None of the above

Give details of involvement, or if none please justify the absence of involvement.
The study proposal has been presented at the Child Growth Foundation (CGF) annual conference and publicised in the CGF newsletter. Members of this leading charity for children's growth and endocrine conditions have been invited to contribute towards the design of the study. The research findings will be disseminated via the CGF annual conference, website, and newsletter. Members of the CGF will be able to participate in the study at the annual conference.

4. RISKS AND ETHICAL ISSUES

RESEARCH PARTICIPANTS

A15. What is the sample group or cohort to be studied in this research?

Select all that apply:

- [ ] Blood
- [ ] Cancer
- [ ] Cardiovascular
- [x] Congenital Disorders
- [ ] Dementias and Neurodegenerative Diseases
- [ ] Diabetes
- [ ] Ear
- [ ] Eye
- [ ] Generic Health Relevance
- [ ] Infection
- [ ] Inflammatory and Immune System
- [ ] Injuries and Accidents
- [ ] Mental Health
- [ ] Metabolic and Endocrine
- [ ] Musculoskeletal
- [ ] Neurological
- [ ] Oral and Gastrointestinal
- [x] Paediatrics
- [ ] Renal and Urogenital
- [ ] Reproductive Health and Childbirth
- [ ] Respiratory
- [ ] Skin
- [ ] Stroke

Gender: Male and female participants

Lower age limit: 0	Days
Upper age limit:	No upper age limit

A17-1. Please list the principal inclusion criteria (list the most important, max 5000 characters).

1.Participants may be of any age
2.Participants must have a height and or head circumference greater than or equal to two standard deviations above the mean in association with either learning difficulties, congenital abnormality, childhood tumour, or characteristic facial features OR height and/or head circumference greater than or equal to three standard deviations above the mean OR regional overgrowth/hemihypertrophy OR a known genomic variant associated with overgrowth OR be a parent of an individual meeting the above criteria.

A17-2. Please list the principal exclusion criteria (list the most important, max 5000 characters).

1.Individuals with a clinical diagnosis of a genetic condition causing tall stature or increased head circumference that is not considered to be a primary overgrowth disorder (eg connective tissue disorder such as Marfan syndrome).
2.Individuals with tall stature or increased head circumference due to a secondary cause, for example an acquired hormonal condition such as acromegaly.
3.Individuals who do not give consent for participation in the study.

RESEARCH PROCEDURES, RISKS AND BENEFITS

A18. Give details of all non-clinical intervention(s) or procedure(s) that will be received by participants as part of the research protocol. *These include seeking consent, interviews, non-clinical observations and use of questionnaires.*

Please complete the columns for each intervention/procedure as follows:

1. Total number of interventions/procedures to be received by each participant as part of the research protocol.
2. If this intervention/procedure would be routinely given to participants as part of their care outside the research, how many of the total would be routine?
3. Average time taken per intervention/procedure (minutes, hours or days)
4. Details of who will conduct the intervention/procedure, and where it will take place.

Intervention or procedure	1	2	3	4
Providing information about the project	1	0	15	A letter of introduction to the study and relevant participant information leaflet(s) (PILs) may be sent to a potential participant with the appointment letter for a routine outpatient clinic. Alternatively the PILs may be given to a potential participant in the clinic.
Seeking informed consent	1	0	15	A trained member of staff such as clinician, nurse or study coordinator will discuss the project with the potential participant (and their family if applicable) in clinic and answer any questions prior to consent being given to join the study.

A19. Give details of any clinical intervention(s) or procedure(s) to be received by participants as part of the research protocol. *These include uses of medicinal products or devices, other medical treatments or assessments, mental health interventions, imaging investigations and taking samples of human biological material. Include procedures which might be received as routine clinical care outside of the research.*

Please complete the columns for each intervention/procedure as follows:

1. Total number of interventions/procedures to be received by each participant as part of the research protocol.
2. If this intervention/procedure would be routinely given to participants as part of their care outside the research, how many of the total would be routine?
3. Average time taken per intervention/procedure (minutes, hours or days).
4. Details of who will conduct the intervention/procedure, and where it will take place.

Intervention or procedure	1	2	3	4
Participants will be asked to provide a blood sample (20ml or smaller volume depending on the age of the individual)	1	1	10	A health care professional trained in venepuncture will take the sample in the routine clinic setting.
Participants may be asked to provide a saliva sample if they do not wish to provide a blood sample. Samples will be collected in a specialised kit eg Oragene.	1	0	10	The sample may be provided in clinic, or the participant may take the sample kit home and post the sample in the provided envelope.
Participants may be asked to provide a skin sample if indicated. Samples will be obtained using a skin biopsy kit.	1	0	20	A health care professional trained in this technique will take the sample in an appropriate clinic room or ward area. Alternatively, if the participant is having a procedure or surgery under general anesthetic as part of their routine care, the skin biopsy may be taken at this time if the participant prefers.

A21. How long do you expect each participant to be in the study in total?

Participants will take part in the study from when they first give informed consent until the end of September 2017. Following this, participants may be contacted in the future with information about other relevant ethically approved studies, in line with the aims of the NIHR RD-TRC. Individuals will be contacted a maximum of four times per year.

A22. What are the potential risks and burdens for research participants and how will you minimise them?

For all studies, describe any potential adverse effects, pain, discomfort, distress, intrusion, inconvenience or changes to lifestyle. Only describe risks or burdens that could occur as a result of participation in the research. Say what steps would be taken to minimise risks and burdens as far as possible.

1. Blood sampling - discomfort, chance of minor bruising. Samples will be taken by trained health professionals. Topical anaesthetic cream and play therapists may be offered if children are anxious. If blood sampling is required as part of the individuals clinical care, the samples for research may be taken at the same time. Individuals may choose to opt out of providing a blood sample and can provide a saliva sample instead.

2. Skin biopsy - small risk of bleeding and/or infection. May leave a small scar. Samples will be taken by a trained health professional. Local anaesthetic will be used to minimise any discomfort. Play therapy may be offered if children are anxious. If an individual is having a procedure or operation under general anaesthetic as part of their routine care, the skin biopsy may be taken at the same time. Individuals may choose to opt out of providing a skin biopsy.

3. Confidentiality - Every effort will be made to maintain the confidentiality and security of patient data. The NIHR Rare Disease Translational Research Collaboration will maintain the OpenClinica database system.

4. Potential psychological impact of receiving molecular genetic test results - Only results directly relevant to the condition will be disclosed, so there will be no unexpected findings or results of uncertain clinical significance. Participants may opt out of receiving any results.

A24. What is the potential for benefit to research participants?

A participant may benefit directly from taking part in the research project if it results in a molecular diagnosis in individual with an undiagnosed overgrowth disorder. This will inform their medical care, for example in deciding whether surveillance for childhood tumour risk is indicated. Identification of a molecular diagnosis will also enable accurate assessment of the recurrence risk for the individual and family, and facilitate reproductive options if this is desired.
Participants may benefit indirectly from access to detailed clinical information about their condition and increased awareness of the availability of diagnostic clinical testing.
Participants who do not directly benefit will be making a contribution to science and future improvements in the care of individuals with overgrowth conditions.

RECRUITMENT AND INFORMED CONSENT

In this section we ask you to describe the recruitment procedures for the study. Please give separate details for different study groups where appropriate.

A27-1. How will potential participants, records or samples be identified? Who will carry this out and what resources will be used? *For example, identification may involve a disease register, computerised search of social care or GP records, or review of medical records. Indicate whether this will be done by the direct care team or by researchers acting under arrangements with the responsible care organisation(s).*

Participants will be identified by their existing clinical care team. This can be done prior or during a clinic appointment and will take the recruiting clinician a minimal amount of time. As overgrowth disorders are rare, each individual clinician is likely to see only a small number of eligible patients so the burden on resources will be minimal.

A27-2. Will the identification of potential participants involve reviewing or screening the identifiable personal information of patients, service users or any other person?

○ Yes ◉ No

Please give details below:
Potential participants will be identified by their clinician and only members of the patient's existing clinical care team will access the patient records.

A28. Will any participants be recruited by publicity through posters, leaflets, adverts or websites?

◉ Yes ○ No

If Yes, please give details of how and where publicity will be conducted, and enclose copy of all advertising material (with version numbers and dates).

The study will be publicised through leaflets and posters at the CGF conference and the CGF newsletter and website.

A29. How and by whom will potential participants first be approached?

Potential participants will be identified by their health care team in paediatric endocrinology or clinical genetics. A letter of introduction and participant information leaflet may be sent out with the clinic appointment letter, or the potential participant may be approached at the clinic appointment.

A30-1. Will you obtain informed consent from or on behalf of research participants?

◉ Yes ○ No

If you will be obtaining consent from adult participants, please give details of who will take consent and how it will be done, with details of any steps to provide information (a written information sheet, videos, or interactive material). Arrangements for adults unable to consent for themselves should be described separately in Part B Section 6, and for children in Part B Section 7.

If you plan to seek informed consent from vulnerable groups, say how you will ensure that consent is voluntary and fully informed.

The potential participant will be approached by their clinician and provided with the appropriate patient information leaflet. They will be given the opportunity to ask any questions they may have. Written consent will be taken in all cases by a trained member of the clinical team.
Parents will be asked to consent on behalf of children aged under 16 and assent will be sought from children who are able to assent to the study. If a child does not wish to assent to the study, they will not be recruited. If a child turns 16 during the study, they will be contacted and re-consented as an adult if they wish to continue participating in the study.
Adults who are unable to consent because of a learning disability will have a consultee (relative, friend, or other appropriate individual) whose opinion will be sought. If an individual shows any sign they do not want to participate they will not be recruited into the study. Participants will be given at least 24 hours to decide whether they would like to participate in the research, but they may choose to decide within a shorter timeframe.

If you are not obtaining consent, please explain why not.

Please enclose a copy of the information sheet(s) and consent form(s).

A30-2. Will you record informed consent (or advice from consultees) in writing?

◉ Yes ○ No

A31. How long will you allow potential participants to decide whether or not to take part?

Potential participants will have at least 24 hours to decide whether to take part in the study.
If a potential participant feels that a shorter time period is sufficient for full consideration of the study they may decide to given consent within the 24 hour time frame.

A33-1. What arrangements have been made for persons who might not adequately understand verbal explanations or written information given in English, or who have special communication needs? *(e.g. translation, use of interpreters)*

Hospital interpreting services will be used as needed to communicate effectively with potential participants with additional communication needs.

A33-2. What arrangements will you make to comply with the principles of the Welsh Language Act in the provision of information to participants in Wales?

Hospital interpreting services will be used as needed to communicate effectively with Welsh-speaking potential participants.

A35. What steps would you take if a participant, who has given informed consent, loses capacity to consent during the study? *Tick one option only.*

○ The participant and all identifiable data or tissue collected would be withdrawn from the study. Data or tissue which is not identifiable to the research team may be retained.

◉ The participant would be withdrawn from the study. Identifiable data or tissue already collected with consent would be retained and used in the study. No further data or tissue would be collected or any other research procedures carried out on or in relation to the participant.

○ The participant would continue to be included in the study.

○ Not applicable – informed consent will not be sought from any participants in this research.

○ Not applicable – it is not practicable for the research team to monitor capacity and continued capacity will be assumed.

Further details:

In the unlikely event of a loss of capacity, samples and personal data would be retained and continued to be used confidentially in this study and potentially in other future ethically approved research. No further data or tissue would be collected or any other research procedure carried out.

If you plan to retain and make further use of identifiable data/tissue following loss of capacity, you should inform participants about this when seeking their consent initially.

CONFIDENTIALITY

In this section, personal data means any data relating to a participant who could potentially be identified. It includes pseudonymised data capable of being linked to a participant through a unique code number.

Storage and use of personal data during the study

A36. Will you be undertaking any of the following activities at any stage (including in the identification of potential participants)? *(Tick as appropriate)*

- [x] Access to medical records by those outside the direct healthcare team
- [ ] Access to social care records by those outside the direct social care team
- [x] Electronic transfer by magnetic or optical media, email or computer networks
- [ ] Sharing of personal data with other organisations
- [ ] Export of personal data outside the EEA
- [x] Use of personal addresses, postcodes, faxes, emails or telephone numbers
- [ ] Publication of direct quotations from respondents
- [ ] Publication of data that might allow identification of individuals
- [x] Use of audio/visual recording devices
- [x] Storage of personal data on any of the following:
  - [x] Manual files (includes paper or film)
  - [x] NHS computers
  - [ ] Social Care Service computers
  - [ ] Home or other personal computers
  - [x] University computers
  - [ ] Private company computers
  - [ ] Laptop computers

Further details:
The NIHR Rare Diseases Clinical Infrastructure Team (RDCIT) is responsible for implementing and maintaining the OpenClinica database, an access controlled secure public database. The database is hosted in the UK and information governance accredited by the Health and Social Care Information Centre, the national provider of information, data and IT systems for health and social care in the UK.

Participants' clinical data, including clinical photography, will be held on the NIHR RD-TRC OpenClinica database. Data will only be accessible via password by named study staff and authorised personnel at the RD-TRC.

Each participant will be given a unique study identifier. Personal identifiable data will be held separately and securely by the RD-TRC. The unique study identifier will be linked to the personal data to allow individuals to be contacted about ethically approved studies in the future, as per the aims of the RD-TRC. This will include sharing data and samples with commercial companies. Commercial companies will not have access to any personal data.

A37. Please describe the physical security arrangements for storage of personal data during the study?

Paper files will be held securely in a locked filing cabinet in a lockable room within NHS offices only accessible by keypad.

A38. How will you ensure the confidentiality of personal data? *Please provide a general statement of the policy and procedures for ensuring confidentiality, e.g. anonymisation or pseudonymisation of data.*

The Rare Diseases Clinical Infrastructure Team (RDCIT) is responsible for implementing and maintaining the OpenClinica database. The database is hosted in the UK (Liverpool)and information governance accredited by the Health and Social Care Information Centre, the national provider of information, data and IT systems for health and social care in the UK.

Clinical information held on the database will be identifiable only by a unique study identifier. Access to all data will be controlled by password and limited to named individuals. Patient identifiable information will be held by the NIHR RD-TRC securely and separately from clinical information and will be linked only by the study identifier.

A40. Who will have access to participants' personal data during the study? *Where access is by individuals outside the direct care team, please justify and say whether consent will be sought.*

Information may be accessed by individuals from the RD-TRC, regulatory authorities or the NHS Trust where it is relevant to participation in the research. Written consent will be sought from participants for this access.

Storage and use of data after the end of the study

A41. Where will the data generated by the study be analysed and by whom?

Data will be analysed by members of the research team. Data will be held on the NIHR RD-TRC database hosted in Liverpool and will not be exported out of the UK.

A42. Who will have control of and act as the custodian for the data generated by the study?

	Title	Forename/Initials	Surname
	Dr	Alison	Foster
Post	Academic Clinical Research Fellow		
Qualifications	BSc MBBS MRCPCH		
Work Address	Birmingham Women's NHS Foundation Trust		
	Mindelsohn Way		
	Birmingham		
Post Code	B179JJ		
Work Email	[redacted]		
Work Telephone	[redacted]		
Fax	[redacted]		

A43. How long will personal data be stored or accessed after the study has ended?

- ○ Less than 3 months
- ○ 3 – 6 months
- ○ 6 – 12 months
- ○ 12 months – 3 years
- ◉ Over 3 years

If longer than 12 months, please justify:
Personal anonymised data will be held indefinitely for future studies in line with the remit of the NIHR RD-TRC. As overgrowth disorders are rare diseases, this will be an important resource for future research for the benefit of patients with these conditions.

A44. For how long will you store research data generated by the study?

Years: 30
Months:

A45. Please give details of the long term arrangements for storage of research data after the study has ended. *Say where data will be stored, who will have access and the arrangements to ensure security.*

Data will be held on a secure database managed by the NIHR RD-TRC and will be held as long as the NIHR RD-TRC runs. Paper files will be archived in accordance with local arrangements for as long as the RD-TRC requests.

INCENTIVES AND PAYMENTS

A46. Will research participants receive any payments, reimbursement of expenses or any other benefits or incentives for taking part in this research?

○ Yes ◉ No

A47. Will individual researchers receive any personal payment over and above normal salary, or any other benefits or incentives, for taking part in this research?

○ Yes ◉ No

A48. Does the Chief Investigator or any other investigator/collaborator have any direct personal involvement (e.g. financial, share holding, personal relationship etc.) in the organisations sponsoring or funding the research that may give rise to a possible conflict of interest?

○ Yes ◉ No

NOTIFICATION OF OTHER PROFESSIONALS

A49-1. Will you inform the participants' General Practitioners (and/or any other health or care professional responsible for their care) that they are taking part in the study?

◉ Yes ○ No

If Yes, please enclose a copy of the information sheet/letter for the GP/health professional with a version number and date.

A49-2. Will you seek permission from the research participants to inform their GP or other health/ care professional?

◉ Yes ○ No

It should be made clear in the participant's information sheet if the GP/health professional will be informed.

PUBLICATION AND DISSEMINATION

A50-1. Will the research be registered on a public database?

◉ Yes ○ No

Please give details, or justify if not registering the research.
The project will be submitted for inclusion on the NIHR portfolio.

Registration of research studies is encouraged wherever possible.
You may be able to register your study through your NHS organisation or a register run by a medical research charity, or publish your protocol through an open access publisher. If you are aware of a suitable register or other method of publication, please give details. If not, you may indicate that no suitable register exists. Please ensure that you have entered registry reference number(s) in question A5-1.

A51. How do you intend to report and disseminate the results of the study? *Tick as appropriate:*

- [x] Peer reviewed scientific journals
- [x] Internal report
- [x] Conference presentation
- [x] Publication on website
- [x] Other publication
- [ ] Submission to regulatory authorities
- [ ] Access to raw data and right to publish freely by all investigators in study or by Independent Steering Committee on behalf of all investigators
- [ ] No plans to report or disseminate the results
- [ ] Other (please specify)

A52. If you will be using identifiable personal data, how will you ensure that anonymity will be maintained when publishing the results?

As these conditions are rare, it is possible that an individual could be identifiable despite the lack of patient identifiable information. Efforts will be made to reduce the chances of this occurring by including a patient cohort as large as possible when publishing the results of this study.

A53. How and when will you inform participants of the study results?

If there will be no arrangements in place to inform participants please justify this.
Results that are relevant to a patient's clinical presentation will be confirmed by testing in a clinical laboratory and sent to their local clinical care team to be fed back to the participant and inform their clinical management. Alternatively, participants may choose to opt out of receiving any results identified through the research.
The results of the study will be published in the Child Growth Foundation Newsletter and presented at the CGF annual conference.

5. Scientific and Statistical Review

A54-1. How has the scientific quality of the research been assessed? *Tick as appropriate:*

- [x] Independent external review
- [ ] Review within a company
- [ ] Review within a multi-centre research group
- [ ] Review within the Chief Investigator's institution or host organisation
- [ ] Review within the research team
- [ ] Review by educational supervisor
- [ ] Other

Justify and describe the review process and outcome. If the review has been undertaken but not seen by the researcher, give details of the body which has undertaken the review:
This project has been independently peer reviewed in a transparent iterative process by a panel of the NIHR Rare Disease Translational Research Collaboration Theme Leads. The RD-TRC Operations Team can provide assurance of the peer review process if required.

For all studies except non-doctoral student research, please enclose a copy of any available scientific critique reports, together with any related correspondence.

For non-doctoral student research, please enclose a copy of the assessment from your educational supervisor/ institution.

A56. How have the statistical aspects of the research been reviewed? *Tick as appropriate:*

☐ Other review by independent statistician

☐ Review by company statistician

☐ Review by a statistician within the Chief Investigator's institution

☐ Review by a statistician within the research team or multi−centre group

☑ Review by educational supervisor

☐ Other review by individual with relevant statistical expertise

☐ No review necessary as only frequencies and associations will be assessed – details of statistical input not required

In all cases please give details below of the individual responsible for reviewing the statistical aspects. If advice has been provided in confidence, give details of the department and institution concerned.

	Title Forename/Initials	Surname
	Prof T	Barrett
Department	School of Clinical and Experimental Medicine	
Institution	University of Birmingham	
Work Address	c/o Diabetes Unit, Birmingham Children's Hospital	
	Steelhouse Lane	
	Birmingham	
Post Code	B4 6NH	
Telephone	01213339267	
Fax		
Mobile		
E-mail	[redacted]	

Please enclose a copy of any available comments or reports from a statistician.

A57. What is the primary outcome measure for the study?

A well characterised cohort study with comprehensive genotype-phenotype correlations established for rare overgrowth disorders.

A58. What are the secondary outcome measures?*(if any)*

To provide a cohort with which to undertake future academic studies into the pathophysiology of overgrowth conditions

To support recruitment to future multicentre studies of therapeutic interventions in overgrowth disorders

A59. What is the sample size for the research? *How many participants/samples/data records do you plan to study in total? If there is more than one group, please give further details below.*

Total UK sample size: 100

Total international sample size (including UK):

Total in European Economic Area:

Further details:
There are more than ten known overgrowth syndromes. The most common of these, Sotos syndrome and Beckwith-Wiedemann syndrome, each have a prevalance of 1 in 15,000. The prevalence of other overgrowth syndromes is unknown but likely to be significantly lower.

A60. How was the sample size decided upon? *If a formal sample size calculation was used, indicate how this was done,*

giving sufficient information to justify and reproduce the calculation.

Overgrowth disorders are rare diseases, with the two most common disorders, Sotos syndrome and Beckwith Wiedemann syndrome, each thought to have a prevalence of 1 in 15,000. The prevalence of other overgrowth disorders is unknown and likely to be much lower. Given the population of England (53.9 million in 2013 according to the Office for National Statistics), this would suggest about 7000 individuals with an overgrowth disorder. However most adults with overgrowth disorders will not be under routine clinical follow up and it is likely that the majority of participants will be recruited from the paediatric population. We are aiming to recruit a minimum of 100 participants.

A61-1. Will participants be allocated to groups at random?

○ Yes ◉ No

A62. Please describe the methods of analysis (statistical or other appropriate methods, e.g. for qualitative research) by which the data will be evaluated to meet the study objectives.

Over 100 phenotypic features will be entered in the database, including demographic data, family history, birth and neonatal history, developmental milestones, behavioural characteristics, growth measurements, medical history, and investigation results. Study participants will be categorised into clinical conditions and then by genotype.

The phenotypic data can be divided into continuous, categorical, or time to event. Appropriate statistical tests will be used to analyse the data. For continuous data, for example height and head circumference, ANOVA or linear regression models will be used to compare groups. For binary data (Y/N), for example presence or absence of seizures, Chi squared or Fisher's exact test will be used. For time to event data, such as time to diagnosis of tumour, a Kaplan Meier curve will be generated.

6. MANAGEMENT OF THE RESEARCH

A63. Other key investigators/collaborators. *Please include all grant co-applicants, protocol co-authors and other key members of the Chief Investigator's team, including non-doctoral student researchers.*

	Title	Forename/Initials	Surname
	Professor	Timothy	Barrett
Post	Professor of Paediatrics		
Qualifications	MBBS DCH PhD FHEA FRCPCH FRCP		
Employer	University of Birmingham		
Work Address	School of Clinical and Experimental Medicine, c/o Diabetes Unit, Birmingham Children's Hospital Steelhouse Lane		
Post Code	B4 6NH		
Telephone	01213339267		
Fax			
Mobile			
Work Email			

	Title	Forename/Initials	Surname
	Dr	Trevor	Cole
Post	Consultant Clinical Geneticist		
Qualifications	MBChB FRCP		
Employer	Birmingham Women's Hospital NHS Foundation Trust		
Work Address	Mindelsohn Way, Birmingham, W Midlands		

Post Code	B15 2TG
Telephone	01214721377
Fax	
Mobile	
Work Email	

Title	Forename/Initials	Surname
Dr	Derek	Lim

Post	Consultant Clinical Geneticist
Qualifications	MBChB MRCPCH DCH
Employer	Birmingham Women's Hospital NHS Foundation Trust
Work Address	Mindelsohn Way, Birmingham, W Midlands

Post Code	B15 2TG
Telephone	
Fax	
Mobile	
Work Email	

A64. Details of research sponsor(s)

A64-1. Sponsor

Lead Sponsor

Status:
- ◯ NHS or HSC care organisation
- ◉ Academic
- ◯ Pharmaceutical industry
- ◯ Medical device industry
- ◯ Local Authority
- ◯ Other social care provider (including voluntary sector or private organisation)
- ◯ Other

Commercial status:

If Other, please specify:

Contact person

Name of organisation	University of Birmingham
Given name	Sean
Family name	Jennings
Address	Room 119, Aston Webb Building, Edgbaston
Town/city	Birmingham
Post code	B15 2TT

Country	United Kingdom
Telephone	01214147618
Fax	
E-mail	researchgovernance@contacts.bham.ac.uk

Legal representative for clinical investigation of medical device (studies involving Northern Ireland only)
Clinical Investigations of Medical Devices that take place in Northern Ireland must have a legal representative of the sponsor that is based in Northern Ireland or the EU

Contact person

Name of organisation	
Given name	
Family name	
Address	
Town/city	
Post code	
Country	
Telephone	
Fax	
E-mail	

A65. Has external funding for the research been secured?

Please tick at least one check box.

- [x] Funding secured from one or more funders
- [ ] External funding application to one or more funders in progress
- [ ] No application for external funding will be made

What type of research project is this?

- ( ) Standalone project
- ( ) Project that is part of a programme grant
- ( ) Project that is part of a Centre grant
- (•) Project that is part of a fellowship/ personal award/ research training award
- ( ) Other

Other – please state:

Please give details of funding applications.

Organisation	NIHR Rare Diseases Translational Research Collaboration
Address	Barton House, Level 5, Box 406 Cambridge Biomedical Campus Hills Road, Cambridge

Post Code	CB2 0QQ
Telephone	01223254601
Fax	
Mobile	
Email	

Funding Application Status: (•) Secured ( ) In progress

Amount: £210,162

Duration
Years: 2
Months: 6

If applicable, please specify the programme/ funding stream:

What is the funding stream/ programme for this research project?
NIHR Rare Disease TRC Fellowship Scheme

A66. Has responsibility for any specific research activities or procedures been delegated to a subcontractor (other than a co-sponsor listed in A64-1) ? *Please give details of subcontractors if applicable.*

( ) Yes (•) No

A67. Has this or a similar application been previously rejected by a Research Ethics Committee in the UK or another country?

( ) Yes (•) No

Please provide a copy of the unfavourable opinion letter(s). You should explain in your answer to question A6-2 how the reasons for the unfavourable opinion have been addressed in this application.

A68-1. Give details of the lead NHS R&D contact for this research:

	Title	Forename/Initials	Surname
	Ms	Kelly	Hard
Organisation	Birmingham Women's Hospital		
Address	R&D Department		
	Birmingham Women's Hospital		
Post Code	B15 2TG		
Work Email			
Telephone			
Fax			
Mobile			

Details can be obtained from the NHS R&D Forum website: http://www.rdforum.nhs.uk

A68-2. Select Local Clinical Research Network for NHS Organisation identified in A68-1:

-- Not Selected --

For more information, please refer to the question specific guidance.

A69-1. How long do you expect the study to last in the UK?

Planned start date: 01/06/2015

Planned end date: 28/02/2019

Total duration:

Years: 3 Months: 8 Days: 28

A70.

Definition of the end of trial, and justification in the case where it is not the last visit of the last subject undergoing the trial

The end of the data collection for the study will be the last visit of the last subject. Data analysis will continue beyond this date. End of study will be formally declared once all data and samples have been analysed and the study findings can be written up. Participants may also be re-contacted in the future (up to 4 times per year) to be invited to participate in other ethically approved studies.

A71-1. Is this study?

- ○ Single centre
- ◉ Multicentre

A71-2. Where will the research take place? *(Tick as appropriate)*

- ☑ England
- ☐ Scotland
- ☑ Wales
- ☐ Northern Ireland
- ☐ Other countries in European Economic Area

Total UK sites in study 10

Does this trial involve countries outside the EU?

○ Yes ◉ No

A72. Which organisations in the UK will host the research? *Please indicate the type of organisation by ticking the box and give approximate numbers if known:*

☑ NHS organisations in England	9
☑ NHS organisations in Wales	1
☐ NHS organisations in Scotland	
☐ HSC organisations in Northern Ireland	
☐ GP practices in England	
☐ GP practices in Wales	
☐ GP practices in Scotland	
☐ GP practices in Northern Ireland	

☐ Joint health and social care agencies (eg community mental health teams)
☐ Local authorities
☐ Phase 1 trial units
☐ Prison establishments
☐ Probation areas
☐ Independent (private or voluntary sector) organisations
☐ Educational establishments
☐ Independent research units
☐ Other (give details)

Total UK sites in study: 10

A73-1. Will potential participants be identified through any organisations other than the research sites listed above?

◉ Yes ○ No

A73-2. If yes, will any of these organisations be NHS organisations?

○ Yes ◉ No

If yes, details should be given in Part C.

A74. What arrangements are in place for monitoring and auditing the conduct of the research?

The CI will monitor and audit the conduct of the research at Birmingham Children's Hospital and Birmingham Women's Hospital.

A76. Insurance/ indemnity to meet potential legal liabilities

Note: in this question to NHS indemnity schemes include equivalent schemes provided by Health and Social Care (HSC) in Northern Ireland

A76-1. What arrangements will be made for insurance and/or indemnity to meet the potential legal liability of the sponsor(s) for harm to participants arising from the management of the research? *Please tick box(es) as applicable.*

Note: Where a NHS organisation has agreed to act as sponsor or co-sponsor, indemnity is provided through NHS schemes. Indicate if this applies (there is no need to provide documentary evidence). For all other sponsors, please describe the arrangements and provide evidence.

☐ NHS indemnity scheme will apply (NHS sponsors only)
☑ Other insurance or indemnity arrangements will apply (give details below)

The University of Birmingham has in force a Public Liability Policy and/or Clinical Trials policy which provides cover for claims for "negligent harm" and the activities here are included within that coverage.

Please enclose a copy of relevant documents.

A76-2. What arrangements will be made for insurance and/ or indemnity to meet the potential legal liability of the sponsor(s) or employer(s) for harm to participants arising from the design of the research? *Please tick box(es) as applicable.*

☐ NHS indemnity scheme will apply (protocol authors with NHS contracts only)

☑ Other insurance or indemnity arrangements will apply (give details below)

The University of Birmingham has in force a Public Liability Policy and/or Clinical Trials policy which provides cover for claims for "negligent harm" and the activities here are included within that coverage.

Please enclose a copy of relevant documents.

A76-3. What arrangements will be made for insurance and/ or indemnity to meet the potential legal liability of investigators/collaborators arising from harm to participants in the conduct of the research?

Note: Where the participants are NHS patients, indemnity is provided through the NHS schemes or through professional indemnity. Indicate if this applies to the whole study (there is no need to provide documentary evidence). Where non-NHS sites are to be included in the research, including private practices, please describe the arrangements which will be made at these sites and provide evidence.

☑ NHS indemnity scheme or professional indemnity will apply (participants recruited at NHS sites only)

☐ Research includes non-NHS sites (give details of insurance/ indemnity arrangements for these sites below)

Please enclose a copy of relevant documents.

A78. Could the research lead to the development of a new product/process or the generation of intellectual property?

○ Yes ◉ No ○ Not sure

A79. Please select the level of commercial participation in this project.

◉ None

○ Industry funding, but not industry sponsored

○ Industry funding and industry sponsored

○ Industry sponsored, but not industry funded

A80. Please select the main subject area of research. Additional sub-topics may be selected, if required

☐ Age and Ageing

☐ Anaesthetics

☐ Cancer (includes malignant haematology

☐ Cardiovascular

☐ Clinical

☐ Critical Care

☐ Dementias and Neurodegenerative Diseases

☐ Dermatology

☐ Diabetes

☐ Ear, Nose and Throat

☐ Gastrointestinal

- [x] Genetics
- [ ] Health Services Research
- [ ] Hepatology
- [ ] Immunology and Inflammation
- [ ] Infectious Disease and Microbiology
- [ ] Injuries and Accidents
- [ ] Medicines for Children (does not include Paediatrics)
- [ ] Mental Health
- [ ] Metabolic and Endocrine
- [ ] Musculoskeletal (Rheumatoid Arthritis is a separate category)
- [ ] Nervous System Disorders
- [ ] Non-malignant Haematology
- [ ] Ophthalmology
- [ ] Oral and Dental
- [x] Paediatrics (does not include Medicines for Children)
- [ ] Primary Care
- [ ] Public Health Research
- [ ] Renal
- [ ] Reproductive Health and Childbirth
- [ ] Respiratory
- [ ] Rheumatoid Arthritis
- [ ] Stroke
- [ ] Surgery
- [ ] Urogenital

9. Has the study been the subject of a scientific review/opinion (Expert Panel)?

○ Yes ○ No

If yes, please provide a copy of the review as part of your application.

Part B: Section 4 – Use of residual or existing stored human tissue(or other human biological materials)

1. What types of human tissue or other biological material will be included in the study?

A small number of individuals in the study may have previously had a biopsy or surgery of an area of overgrowth or tumour and there may be stored tissue samples. These participants will be asked to consent for genetic analysis of this stored tissue.

2. Will the samples be released to the researcher:

In fully anonymised form? *(link to stored tissue and data is broken)*

○ Yes ◉ No

In linked anonymised form? *(linked to stored tissue but donor not identifiable to researchers)*
◉ Yes ○ No

In a form in which the donor could be identifiable to researchers?
○ Yes ◉ No

3. Has consent been obtained previously to use the samples for research

○ Consent has been given for all samples
○ Consent has been given for some of the samples
◉ No consent has been given

5. Is it proposed to seek further consent to use the samples in this research?

◉ Yes ○ No

6. Will any tissues or cells be used for human application or to carry out testing for human application in this research?

○ Yes ◉ No

8. What types of test or analysis will be carried out on the samples?

Existing stored tissue samples may be used for DNA analysis, including exome sequencing and/or methylation studies.
It is possible that this may identify the genetic cause of overgrowth in a participant, which may have implications for their clinical care.
Participants will be asked to give consent for results that are clinically relevant to the cause of their overgrowth to be fed back to them via their health care team. Participants will be able to opt out of receiving these results.

9. Will the research involve the analysis or use of human DNA in the samples?

◉ Yes ○ No

10. Is it possible that the research could produce findings of clinical significance for donors or their relatives?

◉ Yes ○ No

11. If so, will arrangements be made to notify the individuals concerned?

◉ Yes
○ No
○ Not applicable

If No, please justify. If Yes, say what arrangements will be made and give details of the support or counselling service.

Results that are relevant to the patient's clinical presentation will be confirmed by testing in a clinical laboratory and sent to their local clinical care team to be fed back to the participant and inform their clinical management.
Alternatively, participants may choose to opt out of receiving any results identified through the research.

12. Who is the holder of the samples?

Please tick either/both boxes as applicable.

☑ NHS pathology department(s) / diagnostic archive(s)
Specific details of each department/archive are not required

☐ Other research tissue bank(s) or sample collection(s)
Please provide further details of each bank/collection below

13. Will any of the samples be imported from outside the UK?

○ Yes ◉ No

14. Please give details of where the samples will be stored, who will have access and the custodial arrangements.

Samples will be stored in the West Midlands Regional Genetics Laboratory, a CPA accredited NHS laboratory, and at the Human Biomaterials Resource Centre, a Human Tissue Authority licensed human tissue biorepository at the University of Birmingham.

15. What will happen to the samples at the end of the research? *Please tick all that apply and give further details.*

☐ Return to current holder of the samples

☑ Transfer to another tissue bank

(If the bank is in England, Wales or Northern Ireland a licence from the Human Tissue Authority will be required to store relevant material for possible further research.)

☐ Storage by research team pending ethical approval for use in another project

(Unless the researcher's institution holds a storage licence from the Human Tissue Authority, or the tissue is stored in Scotland, or it is not relevant material, a further application for ethical review should be submitted before the end of this project.)

☐ Storage by research team as part of a new research tissue bank

(The institution will require a storage licence for research from the Human Tissue Authority if the bank will be storing relevant material in England, Wales or Northern Ireland. A separate application for ethical review of the tissue bank may also be submitted.)

☐ Storage by research team of biological material which is not "relevant material" for the purposes of the Human Tissue Act

☐ Disposal in accordance with the Human Tissue Authority Code of Practice

☐ Other

☐ Not yet known

Please give further details of the proposed arrangements:

Tissue samples will be stored at the Human Biomaterials Resource Centre, a Human Tissue Authority licensed human tissue biorepository at the University of Birmingham.

Part B: Section 5 – Use of newly obtained human tissue(or other human biological materials) for research purposes

1. What types of human tissue or other biological material will be included in the study?

Participants in the study will be asked to provide a blood sample.
Saliva samples may be taken instead of blood in some cases if blood sampling is declined or not possible.
A small number of participants may have regional overgrowth affecting one part of the body instead of constitutional overgrowth. These individuals will be asked to provide a skin sample.
Part of each sample will undergo DNA extraction. The other part of the sample will be stored.

2. Who will collect the samples?

Samples will be collected by a suitable qualified member of the health care team at the local centre. In the case of blood samples, this is likely to be the research nurse. Skin biopsies will be taken by a recruiting clinician who is trained in this technique. Saliva samples will be collected by participants.

3. Who will the samples be removed from?

- [x] Living donors
- [ ] The deceased

4. Will informed consent be obtained from living donors for use of the samples? *Please tick as appropriate*

In this research?
(•) Yes ( ) No

In future research?
(•) Yes ( ) No ( ) Not applicable

6. Will any tissues or cells be used for human application or to carry out testing for human application in this research?

( ) Yes (•) No

8. Will the samples be stored: *[Tick as appropriate]*

In fully anonymised form? *(link to donor broken)*
( ) Yes (•) No

In linked anonymised form? *(linked to stored tissue but donor not identifiable to researchers)*
(•) Yes ( ) No

If Yes, say who will have access to the code and personal information about the donor.
Only authorised members of the NIHR RD-TRC will have access to the code and personal information about the donor.

In a form in which the donor could be identifiable to researchers?
( ) Yes (•) No

9. What types of test or analysis will be carried out on the samples?

Genetic tests including whole exome sequencing and methylation studies will be carried out on samples.

These tests may identify the cause of an individual's overgrowth. Participants will be asked to consent to the feedback of findings related to their overgrowth. They will be able to opt out of receiving any findings if they choose. Any findings will be confirmed in a clinical laboratory and fed back to the participant's recruiting clinician to feedback to the participant.

With detailed genetic tests such as whole exome sequencing, it is possible that other variants unrelated to the presenting condition may be identified. In some cases the clinical significance of the variant may be unknown or uncertain. In others the clinical significance may be more certain, but the condition may be one without treatment. In children there is the additional issue that testing for adult onset disorders in childhood is not thought to be in the child's best interest. However, it is also possible that a finding of no clinical significance in a child may be relevant to other family members.

The ethical issues around incidental findings as a result of new genomic technologies have been widely debated.There are currently no best practice guidelines as to whether incidental findings identified through research (as opposed to clinical practice) should be disclosed to participants. The arguments in favour of non-disclosure are that feedback may be psychologically harmful and there is no evidence of any benefit in doing so. In addition, any findings generated through genomic testing are not visible to the researcher and must be actively sought through the data analysis process, validated and interpreted by a clinical laboratory, and fed back by a professional with expertise in genetic counselling. Each of these steps requires additional resources and adversely affects the feasibility of performing research in this area. Finally, participation in research is an altruistic activity and while the potential for harm must be minimised as much as possible, participants do not usually expect any direct benefit from their participation. The disadvantages of not disclosing incidental findings are the potential reduction in autonomy and missed opportunity for detection of possible preventable disease.

At present several large UK research studies, including the NIHR BioResource for Rare Disease, and the Deciphering Developmental Disorders study (DDD), do not disclose any incidental findings. In line with this current practice, this study will disclose results that are directly relevant to the individuals' presenting condition to their recruiting clinician to help inform their clinical management. These results should be confirmed in a clinical laboratory. However individuals will be able to opt out of disclosure of research findings if they choose. No incidental findings will be disclosed.

There is currently considerable debate on a national scale regarding whether incidental findings should be disclosed in a research setting. Should the consensus on best practice change in the future, an amendment will be sought and new consent forms put in place for prospective use.

10. Will the research involve the analysis or use of human DNA in the samples?

◉ Yes ○ No

11. Is it possible that the research could produce findings of clinical significance for donors or their relatives?

◉ Yes ○ No

12. If so, will arrangements be made to notify the individuals concerned?

◉ Yes ○ No ○ Not applicable

If No, please justify. If Yes, say what arrangements will be made and give details of the support or counselling service.

Results that are relevant to the patient's clinical presentation will be confirmed on by testing in a clinical laboratory and sent to their local clinical care team to be fed back to the participant and inform their clinical management. Alternatively, participants may choose to opt out of receiving any results identified through the research.

13. Give details of where the samples will be stored, who will have access and the custodial arrangements.

Samples will be stored in the West Midlands Regional Genetics Laboratory, a CPA accredited NHS laboratory, and at

the Human Biomaterials Resource Centre, a Human Tissue Authority licensed human tissue biorepository at the University of Birmingham.

14. What will happen to the samples at the end of the research? *Please tick all that apply and give further details.*

☑ Transfer to research tissue bank

(If the bank is in England, Wales or Northern Ireland the institution will require a licence from the Human Tissue Authority to store relevant material for possible further research.)

☐ Storage by research team pending ethical approval for use in another project

(Unless the researcher's institution holds a storage licence from the Human Tissue Authority, or the tissue is stored in Scotland, or it is not relevant material, a further application for ethical review should be submitted before the end of this project.)

☐ Storage by research team as part of a new research tissue bank

(The institution will require a licence from the Human Tissue Authority if the bank will be storing relevant material in England, Wales or Northern Ireland. A separate application for ethical review of the tissue bank may also be submitted.)

☐ Storage by research team of biological material which is not "relevant material" for the purposes of the Human Tissue Act

☐ Disposal in accordance with the Human Tissue Authority's Code of Practice

☐ Other

☐ Not yet known

Please give further details of the proposed arrangements:

At the end of the study, samples will be stored at the Human Biomaterials Resource Centre, a Human Tissue Authority licensed human tissue biorepository at the University of Birmingham.

B. All research other than CTIMPs

In this sub-section, an adult means a person aged 16 or over.

B1. What impairing condition(s) will the participants have?

The study must be connected to this condition or its treatment.

Overgrowth disorders are often associated with learning disabilities which can be mild, moderate or severe.

B2. Justify the inclusion of adults unable to consent for themselves. It should be clear why the research could not be carried out as effectively if confined to adults capable of giving consent.

The majority of individuals with overgrowth disorders have a degree of learning difficulty and a proportion of these will have a learning disability severe enough that they will not be able to consent for themselves. Excluding these individuals would prevent a full study of these disorders.

B3. Who in the research team will decide whether or not the participants have the capacity to give consent? What training/experience will they have to enable them to reach this decision?

The recruiting clinician will decide whether or not a participant has the capacity to give consent.

B4. Does the research have the potential to benefit participants who are unable to consent for themselves?

(●) Yes ( ) No

If Yes, please indicate the nature of this benefit. You may refer back to your answer to Question A24.

A participant may benefit directly from taking part in the research project as it may identify a molecular diagnosis in an individual with an undiagnosed overgrowth disorder. This will inform their medical care, for example in deciding whether surveillance for childhood tumours is indicated.
Participants may benefit indirectly from access to detailed clinical information about their condition and increased awareness of the availability of diagnostic clinical testing.
Participants who do not benefit directly will be making a contribution to science and future improvements in the care of individuals with overgrowth conditions.

B5. Will the research contribute to knowledge of the causes or the treatment or care of persons with the same impairing condition (or a similar condition)?

(●) Yes ( ) No

If Yes, please explain how the research will achieve this:

Increasing our understanding of the molecular aetiology, genotype-phenotype correlations, phenotypic modifiers, and natural history of overgrowth disorders will improve our ability to accurately assess recurrence risks, give prognostic information, and institute appropriate management for individual patients (e.g. surveillance for childhood cancer). In addition, understanding the genetic aetiology and resulting phenotype will give insight into the molecular mechanisms underlying overgrowth and therefore identify potential targets for treatment.

B6. Will the research involve any foreseeable risk or burden for these participants, or interfere in any way with their freedom of action or privacy?

( ) Yes (●) No

Questions B7 and B8 apply to any participants recruited in England and Wales.

B7. What arrangements will be made to identify and consult persons able to advise on the presumed wishes and

feelings of participants unable to consent for themselves and on their inclusion in the research?

A personal consultee (family member, carer or friend) will be consulted in regard to whether the individual would want to take part. If such an individual does not exist, a nominated consultee will be appointed. The consultee will be given an information sheet and asked to sign a declaration form. If the consultee feels that the individual would not want to participate they will not be recruited into the study. If the individual indicates or shows any signs that they do not want to participate in the study they will not be recruited. If the consultee feels or the individual shows any signs that they do not wish to continue in the study they will be withdrawn from the study.

Please enclose a copy of the written information to be provided to consultees. This should describe their role under section 32 of the Mental Capacity Act and provide information about the research similar to that which might be given to participants able to consent for themselves.

B8. Is it possible that a participant requiring urgent treatment might need to be recruited into research before it is possible to identify and consult a person under B7?

○ Yes ◉ No

If Yes, say whether arrangements will be made instead to seek agreement from a registered medical practitioner and outline these arrangements. Or, if this is also not feasible, outline how decisions will be made on the inclusion of participants and what arrangements will be made to seek consent from the participant (if capacity has been recovered) or advice from a consultee as soon as practicable thereafter.

B9. What arrangements will be made to continue to consult such persons during the course of the research where necessary?

Not applicable

B10. What steps will you take, if appropriate, to provide participants who are unable to consent for themselves with information about the research, and to consider their wishes and feelings?

If the consultee feels that the individual would not want to participate they will not be recruited into the study. If the individual indicates or shows any signs that they do not want to participate in the study they will not be recruited. If the consultee feels or the individual shows any signs that they do not wish to continue in the study they will be withdrawn from the study.

B11. Is it possible that the capacity of participants could fluctuate during the research? How would this be handled?

It is unlikely that the capacity of participants would fluctuate during the research.

B12-1. What will be the criteria for withdrawal of participants?

If the consultee feels or the individual shows any signs that they do not wish to continue in the study they will be withdrawn from the study.

B13. Describe what steps will be taken to ensure that nothing is done to which participants appear to object (unless it is to protect them from harm or minimise pain or discomfort).

If the consultee feels that the individual would not want to participate they will not be recruited into the study. If the individual indicates or shows any signs that they do not want to participate in the study they will not be recruited. If the consultee feels or the individual shows any signs that they do not wish to continue in the study they will be withdrawn from the study.

B14. Describe what steps will be taken to ensure that nothing is done which is contrary to any advance decision or statement by the participant?

Participants will be provided with written information and given sufficient time to consider. Written consent will be taken from participants who are able to give consent. In the case of adults who are not able to give consent, a consultee (friend, relative or other appropriate individual) will be consulted as to the wishes of the individual.

PART B: Section 7 - Children

1. Please specify the potential age range of children under 16 who will be included and give reasons for carrying out the research in this age group.

Children up to the age of 16 will be included, as primary overgrowth disorders are genetic conditions that affect individuals from birth. Features of these disorders, for example congenital abnormalities, will present in the neonatal period. Other complications such as learning difficulties will present during childhood. One of the most important complications, the risk of tumours, is highest in early childhood in a number of these conditions.

2. Indicate whether any children under 16 will be recruited as controls and give further details.

No controls will need to be recruited.

3-2. Please describe the arrangements for seeking informed consent from a person with parental responsibility and/or from children able to give consent for themselves.

Informed consent will be taken from a person with parental responsibility. Children under the age of 16 will be asked to assent to the study.

4. If you intend to provide children under 16 with information about the research and seek their consent or agreement, please outline how this process will vary according to their age and level of understanding.

Appropriate information leaflets will be provided depending on the age of the child and level of understanding. Assent forms for children will be provided.

Copies of written information sheet(s) for parents and children, consent/assent form(s) and any other explanatory material should be enclosed with the application.

PART C: Overview of research sites

Please enter details of the host organisations (Local Authority, NHS or other) in the UK that will be responsible for the research sites. *For further information please refer to guidance.*

Investigator identifier	Research site	Investigator Name
IN1 ☐	◉ NHS/HSC Site ○ Non-NHS/HSC Site Country: England	Forename: Timothy Middle name: G Family name: Barrett Email: [redacted] Qualification (MD...): MBBS DCH PhD FHEA FRCPCH FRCP Country: United Kingdom
	Organisation name: BIRMINGHAM CHILDREN'S HOSPITAL NHS FOUNDATION TRUST Address: STEELHOUSE LANE BIRMINGHAM WEST MIDLANDS Post Code: B4 6NH	
	Institution name: BIRMINGHAM CHILDREN'S HOSPITAL NHS FOUNDATION TRUST Department name: Street address: STEELHOUSE LANE Town/city: Post Code: B4 6NH	
IN3 ☐	◉ NHS/HSC Site ○ Non-NHS/HSC Site Country: England	Forename: Trevor Middle name: R Family name: Cole Email: [redacted] Qualification (MD...): Mb ChB FRCP Country: United Kingdom
	Organisation name: BIRMINGHAM WOMEN'S NHS FOUNDATION TRUST Address: BIRMINGHAM WOMENS HOSPITAL METCHLEY PARK ROAD BIRMINGHAM WEST MIDLANDS Post Code: B15 2TG	
	Institution name: BIRMINGHAM WOMEN'S NHS FOUNDATION TRUST Department name:	

Street address BIRMINGHAM WOMENS HOSPITAL
Town/city METCHLEY PARK ROAD
Post Code B15 2TG

IN4 ☐

◯ NHS/HSC Site
◉ Non-NHS/HSC Site

Forename
Middle name
Family name
Email
Qualification (MD...)
Country

Institution name
Department name
Street address
Town/city
Post Code

Institution name
Department name
Street address
Town/city
Post Code
Country

B.2 Participant information

B.2.1 Competent adult

INFORMATION LEAFLET FOR ADULT PARTICIPANTS

PHENOTYPING OF OVERGROWTH DISORDERS (POD)

Version 1.1 16/6/15

You are invited to take part in a research project. Before you decide whether or not you wish to take part it is important for you to understand why the project is being done and what it will involve. Please take the time to read the following information carefully. Feel free to discuss it with your family or close friends and ask us if there is anything that is not clear, or if you would like more information. You will be given as much time as you like to make a decision.

What is the purpose of this study?

We would like to increase our understanding of the clinical and genetic features of rare genetic overgrowth disorders. At present we do not fully understand the genetic causes of these conditions and the medical problems that are associated with each condition. Studying the clinical features (the 'phenotype') of individuals with overgrowth will increase our knowledge of these conditions.

Improving our understanding of these disorders will enable health care professionals to provide more accurate information and the best possible care to individuals with overgrowth conditions. Identifying the genetic causes may also help with developing treatments in the future.

Why have I been invited to join the study?

Overgrowth disorders are rare, so we need to ask as many people as possible with these conditions to take part. We will also ask your parents to participate as this will help in interpreting the study data.

Do I have to take part in the study?

It is completely up to you to decide whether you wish to join. You may decide to take part in only some parts of the study. If you decide not to join, your decision will not affect the health care you receive in any way. You will be free to withdraw at any time without having to give a reason.

What will happen if I decide to participate in the study?

If you agree to join, we will ask you to sign a consent form.

Your doctor will ask you about your medical history in detail and perform a routine clinical examination. We will ask to take a set of clinical photographs of you. You may choose to opt out of clinical photography.

You will be asked to provide a small blood sample (up to 20ml or about three teaspoons) which will be taken at your outpatient clinic appointment. If this is not possible, a saliva sample may be given instead. The saliva sample kit, instructions for how to give the sample and an addressed prepaid envelope will be provided for you to post the saliva sample back to us.

We will ask individuals with overgrowth affecting only part of the body for a skin sample taken by skin biopsy.

Individuals who have previously had an operation or biopsy of an area of overgrowth or a tumour will be asked to give consent for us to access any stored tissue samples.

You may choose to decline to provide any or all of these samples.

Your parents will be asked to sign a consent form and provide a blood sample or saliva sample. If it is not possible for one or both of your parents to participate you will still be eligible to join the study.

What will happen next?

The clinical information, test results, and information from your medical notes and other health records will be entered into a secure research database. Your medical data will be identified by a unique study number, so researchers looking at the data will not be able to see your personal identifiable information. Your personal identifiable information will be stored on a separate database with the unique study identifier to enable us to contact you in the future. Anonymised medical data and samples may be shared with commercial companies in the future for ethically approved research but commercial companies will not have access to your personal identifiable data.

What will happen to the samples I give?

We will extract and store the genetic material from your sample. We may determine the DNA sequence of your genetic code. Part of the sample will be stored for future ethically approved studies.

How often will you contact me?

We will contact you a maximum of four times during the period of the study. After the study has ended we may contact you again up to a maximum of four times per year to give you information about future studies of rare conditions, potentially including trials of treatment. You are under no obligation to participate in any of these future studies.. You are welcome to contact us to ask any questions you may have about these studies.

Who will know about me taking part in the study?

The information collected about you during the course of the research project will be kept strictly confidential. Your clinical details, blood samples and information from genetic tests will be given a unique sample study number. The database linking unique sample study numbers to personal details will only be accessed by authorised members of the database team. You will not be identified personally in any report or publication. We will ask you to give consent for us to inform your GP that you are taking part in the study.

What are the benefits of joining the study?

It is likely that there will be no direct benefit to you or your family by joining but you will make a contribution to science and future improvements in the care of individuals with overgrowth conditions.

It is possible that the research could identify a cause of overgrowth in you or your family. This could enable your health care team to provide you with the best possible care. With your prior consent we would let your doctor and your clinical care team know about this result. All

research results would need to be confirmed in an accredited diagnostic laboratory before being used in the clinical management of you and your family members.

There is a voluntary agreement between the government and insurance companies that individuals are not required to disclose any genetic test results acquired as part of clinical research. This agreement has been extended until 2019.

If this research leads to the development of a new treatment or medical test, you will not benefit financially from this.

What other information may be produced by the research?

You will not be told about any genetic results that may be identified as a by-product of this research ('incidental findings') and are not relevant to an overgrowth condition.

What are the risks and disadvantages of participating in the research?

Donation of a small blood sample may cause brief discomfort and occasionally a small bruise. Infection at the site is very rare.

Individuals who have overgrowth affecting only one part of the body will be asked for a skin sample obtained via a skin biopsy. Skin biopsy may cause a small scar and there is a small risk of bleeding or infection. This procedure is performed under local anaesthetic. If an individual happens to be having a procedure or operation under general anaesthetic for clinical reasons, they may opt to have the skin biopsy taken at the same time.

What will happen if I don't want to carry on with the study?

You are free to with withdraw from the study at any time without giving a reason. Any stored samples will be destroyed and we will not contact you again. It will not be possible to destroy samples that have already been prepared for testing, the results of any information obtained from your samples or the information held on the research database.

Who is funding this study?

This study is funded by the NHS National Institute for Health Research (NIHR) Rare Disease Translational Research Collaboration (RD-TRC) and is sponsored by the University of Birmingham.

Who do I contact with any concerns?

If you have any questions or concerns about this project, please contact Dr Alison Foster via [redacted] or [redacted], or the Patient Advise & Liaison Service (PALS) on 0121 627 2747. Alternatively you can write to the following contact address:

Dr Alison Foster

Clinical Genetics Unit

Birmingham Women's Hospital

Edgbaston

Birmingham

B15 2TG

Thank you for reading this information and considering participation in the study.

B.2.2 Child age 6-10

INFORMATION LEAFLET
FOR CHILDREN AGE 6-10

POD STUDY

Version 1.1 16/6/15

Please read this carefully. If there is anything you do not understand please ask your doctor, nurse or parents. Thank you ☺

What is this study about?

Sometimes children grow more than other children of the same age. These children might need to see doctors and nurses to help keep them healthy, or need extra help at school.

We want to find out more about the genes that cause this. Our genes are the building blocks of our bodies. The more we know about our genes and how they can make us unwell the better we can help these children.

Why me?

We are asking you to take part because you are tall or have a large head size for your age. You might also see doctors and nurses or have extra help at school.

What will I have to do?

Your doctor will talk to you and your parents. They will examine you, for example measure how tall you are. We will also ask if we can take some photos of you

We will ask you to give a small amount of blood or saliva. Your doctor or nurse

Do I have to join?

No, it is up to you. It's ok if you don't want to take part. If you want to join but don't want your photo taken or to give a blood sample that's ok too.

Where will the study take place?

Your doctor will see you at your next hospital appointment. The blood sample will be taken then too.

Will anyone know I'm doing this?

No-one apart from your family and our doctors and nurses will know you are taking part. The scientists who study your blood sample will not know who it came from.

What if I change my mind?

You can change your mind at any time and don't have to tell us why. You will still see your doctor for hospital appointments if you need to.

What happens now?

If you decide to take part we will ask you to sign a form. Your parents will also sign a form to say you are allowed to take part.

Thank you!

Thank you for reading. If you have any questions please ask your doctor, nurse or parents.

B.2.3 Young people age 11-15

INFORMATION LEAFLET
FOR YOUNG PEOPLE AGE 11-15

POD STUDY

Version 1.1 16/06/15

Sometimes children and young people grow more than others of the same age. This can be due to an overgrowth condition. These are life-long conditions that in some people cause health problems. We want to learn more about the problems associated with overgrowth conditions and the genes that cause them. Our genes are the building blocks of our bodies. The more we understand about overgrowth the more chance we have of giving people with these conditions the best possible health care.

We would like to invite you to take part in our research. Before you decide if you would like to join, please read this information leaflet which explains why we are asking for your help and what this will involve. We ask that you discuss this with your parent or guardian and if you have any questions please ask your doctor or nurse. Thank you.

Why are we doing this research?

Overgrowth conditions are rare. Often we are not able to identify the underlying genetic cause in an individual with overgrowth, or know which health problems might occur. This makes it difficult to give information about the condition, what it means for that person and what medical management is best. We would like to study a large number of people with overgrowth to improve our understanding of these conditions.

Why me?

We are asking you to take part because you may have an overgrowth condition. We need as many people as possible to take part in the study.

What will I have to do?

Your doctor will ask you and your parents about your medical and family history. They will perform a routine examination, for example measuring how tall you are. We will also ask if we can take some photographs of you.

We will ask you to give a small blood sample (about three teaspoons) or saliva sample. Your doctor or nurse will explain how this is done. We can put cream or spray on your arm to so it doesn't hurt as much.

We will ask some individuals for a tiny piece of skin (skin biopsy). Your doctor or nurse will tell you if this is needed from you and will explain how this is done. We numb the skin with an injection so it does not hurt.

Do I have to join?

No, it is up to you. It's ok if you don't want to take part and won't affect the medical care you receive. If you want to join but don't want your photograph taken or to give a blood sample that's ok too.

What happens next?

Your information will be kept on a research database. This will only be accessed by people approved by the study. Your information will be stored under a special number rather than your name and personal details.

In the future we may contact you to invite you to take part in other studies that might be of interest to you. We will explain the study and what is involved to you and your family. You will be able to decide if you want to take part or not.

How often will I be contacted?

The maximum number of invitations would be four times each year.

Where will the study take place?

Your doctor will see you at your next hospital appointment. The blood sample will be taken at the hospital. Saliva samples can be taken at home and posted to us.

Will anyone know I'm doing this?

No-one apart from your family, your own doctors including your GP, and our doctors and nurses will know you are taking part. Your sample will be given a special number and the scientists who study your blood sample will not know who it came from.

What will happen to any samples I give?

We will store your samples in the laboratory and may run a number of tests, including looking at your genetic material (DNA). Samples will be stored for future research.

What if I change my mind?

You can change your mind at any time and don't have to tell us why. You can tell your doctor or nurse or ask your parents or guardian to let us know. You will still see your doctor for hospital appointments if you need to.

What are the possible benefits of taking part?

The study may not help you, but the information we obtain will help improve management of other children and adults with overgrowth conditions in the future.

It is possible that we might identify a cause for your overgrowth condition. This could help your doctors and nurses give you the best possible care. It might also be helpful for other members of your family.

What happens now?

If you decide to take part we will ask you to sign a form giving your assent to joining the study. Your parents will also sign a form to given their consent. You will be given copy of this leaflet and your signed form to keep.

Thank you for reading this. If you have any questions please ask your doctor or nurse.

B.2.4 Consultee

INFORMATION LEAFLET FOR CONSULTEES

PHENOTYPING OF OVERGROWTH DISORDERS (POD)

Version 1.1 16/6/15

We feel your relative/friend is unable to decide for himself/herself whether to participate in this research.

To help decide if he/she should join the study, we'd like to ask your opinion whether or not they would want to be involved. We'd ask you to consider what you know of their wishes and feelings, and to consider their interests. Please let us know of any advance decisions they may have made about participating in research. These should take precedence.

If you decide your relative/friend would have no objection to taking part we will ask you to read and sign a consultee declarationform. We'll then give you a copy to keep. We will keep you fully informed during the study so you can let us know if you have any concerns or you think your relative/friend should be withdrawn.

If you decide that your friend/relative would not wish to take part it will not affect the standard of care they receive in any way.

If you are unsure about taking the role of consultee you may seek independent advice.

We will understand if you do not want to take on this responsibility.

Your friend/relative is invited to take part in a research project. Before you decide whether or not he/she would wish to take part it is important for you to understand why the project is being done and what it will involve. Please take the time to read the following information carefully. Feel free to discuss it with your friend's/relative's family or close friends and ask us if there is anything that is not clear, or if you would like more information. You will be given as much time as you like to make a decision.

What is the purpose of this study?

We would like to increase our understanding of the clinical and genetic features of rare genetic overgrowth disorders. At present we do not fully understand the genetic causes of these conditions and the medical problems that are associated with each condition. Looking at the clinical features (the 'phenotype') of individuals with overgrowth will increase our knowledge of these conditions.

Improving our understanding of these disorders will enable health care professionals to provide more accurate information and the best possible care to individuals with overgrowth conditions. Identifying the genetic causes may also help with developing treatments in the future.

Why has my friend/relative been invited to join the study?

Overgrowth disorders are rare, so we need to ask as many people as possible with these conditions to take part.

Does my friend/relative have to take part in the study?

It is up to you to decide whether your friend/relative would wish to join. You may decide he/she would want to take part in only some parts of the study. If you decide he/she would not wish to join, this decision will not affect the health care your friend/relative receives in any way. If your friend/relative indicates or shows any sign they do not want to participate in the study they will not be included in the study. If you feel that your friend/relative does not want to continue in the study, or he/she shows any signs that they do not wish to continue, they will be withdrawn from the study. He/she will be free to withdraw at any time without you or him/her having to give a reason.

What will happen if I decide my friend/relative would want to participate in the study?

If you decide your friend/relative would agree to join, we will ask you to sign a declaration form.

Your friend's/relative's doctor will ask about his/her medical history in detail and perform a routine clinical examination. We will ask to take a set of clinical photographs of your friend/relative. You may choose to opt out of clinical photography if you feel this is what he/she would want

Your friend/relative will be asked to provide a small blood sample (up to 20ml or about three teaspoons) which will be taken at his/her outpatient clinic appointment. If this is not possible, a saliva sample may be given instead. The saliva sample kit, instructions for how to give the

sample and an addressed prepaid envelope will be provided for you to post the saliva sample back to us.

We will ask individuals with overgrowth affecting only part of the body for a skin sample taken by skin biopsy.

If your friend/relative has previously had an operation or biopsy of an area of overgrowth or a tumour, you will be asked to sign a declaration to allow us access to any stored tissue samples.

You may choose to decline any or all of these samples if you feel your friend/relative would not want to provide these.

What will happen next?

The clinical information, test results, and information from your friend's/relative's medical notes and other health records will be entered into a secure research database. Your friend's/relative's medical data will be identified by a unique study number, so researchers looking at the data will not be able to see his/her personal identifiable information. Your friend's/relative's personal identifiable information will be stored on a separate database with the unique study identifier to enable us to contact you and your friend/relative in the future. Anonymised medical data and samples may be shared with commercial companies in the future for ethically approved research but commercial companies will not have access to your friend's/relative's personal identifiable data.

What will happen to the samples my friend/relative gives?

We will extract and store the genetic material from your friend's/relative's sample. We may determine the DNA sequence of his/her genetic code. Part of the sample will be stored for future ethically approved studies.

How often will you contact me and my friend/relative?

We will contact you a maximum of four times during the period of the study. After the study has ended, we may contact you again up to a maximum of four times per year to give you information about future studies of overgrowth disorders, potentially including trials of treatment. . Your friend/relative is under no obligation to participate in any of these future studies. You are welcome to contact us to ask any questions you may have about these studies.

Who will know about my friend/relative taking part in the study?

The information collected about your friend/relative during the course of the research project will be kept strictly confidential. His/her clinical details, blood samples and information from genetic tests will be given a unique sample study number. The database linking unique sample study numbers to personal details will only be accessed by authorised members of the database team. Your friend/relative will not be identified personally in any report or publication. We will ask you to declare that your friend/relative would want us to inform his/her GP that he/she is taking part in the study.

What are the benefits of joining the study?

It is likely that there will be no direct benefit to your friend/relative or his/her family by joining but he/she will make a contribution to science and future improvements in the care of individuals with overgrowth conditions.

It is possible that the research could identify a cause of overgrowth in your friend/relative or his/her family. This could enable your health care team to provide your friend/relative with the best possible care. With your prior declaration we would let his/her doctor and clinical care team know about this result. All research results would need to be confirmed in an accredited diagnostic laboratory before being used in the clinical management of your friend/relative and his/her family members.

There is a voluntary agreement between the government and insurance companies that individuals are not required to disclose any genetic test results acquired as part of clinical research. This agreement has been extended until 2019.

If this research leads to the development of a new treatment or medical test, you and your friend/relative will not benefit financially from this.

What other information may be produced by the research?

You and your friend/relative will not be told about any genetic results that may be identified as a by-product of this research (‘incidental findings’) and are not relevant to an overgrowth condition.

What are the risks and disadvantages of participating in the research?

Donation of a small blood sample may cause brief discomfort and occasionally a small bruise. Infection at the site is very rare.

Individuals who have overgrowth affecting only one part of the body will be asked for a skin sample obtained via a skin biopsy. Skin biopsy may cause a small scar and there is a small risk of bleeding or infection. This procedure is performed under local anaesthetic. If an individual happens to be having a procedure or operation under general anaesthetic for clinical reasons, they may opt to have the skin biopsy taken at the same time.

What will happen if I feel my friend/relative does not wish to carry on with the study or he/she shows signs they do not wish to continue?

Your friend/relative is free to withdraw from the study at any time without giving a reason. Any stored samples will be destroyed and we will not contact you or your friend/relative again. It will not be possible to destroy samples that have already been prepared for testing, the results of any information obtained from your friend's/relative's samples or the information held on the research database.

Who is funding this study?

This study is funded by the NHS National Institute for Health Research (NIHR) Rare Disease Translational Research Collaboration (RD-TRC) and is sponsored by the University of Birmingham.

Who do I contact with any concerns?

If you have any questions or concerns about this project, please contact Dr Alison Foster via [redacted] or [redacted] or the Patient Advise & Liaison Service (PALS) on 0121 627 2747. Alternatively you can write to the following contact addresses:

Dr Alison Foster

Clinical Genetics Unit

Birmingham Women's Hospital

Edgbaston

Birmingham

B15 2TG

Thank you for reading this information and considering whether your friend/relative would wish to participate in the study.

B.2.5 Parent of child/young person

INFORMATION LEAFLET FOR PARENTS OF CHILDREN AND YOUNG PEOPLE

PHENOTYPING OF OVERGROWTH DISORDERS (POD)

Version 1.1 16/6/15

Your family is invited to take part in a research project. Before you decide whether or not you wish to take part it is important for you to understand why the project is being done and what it will involve. Please take the time to read the following information carefully. Feel free to discuss it with your family or close friends and ask us if there is anything that is not clear, or if you would like more information. You will be given as much time as you like to make a decision.

What is the purpose of this study?

We would like to increase our understanding of the clinical and genetic features of rare genetic overgrowth disorders. At present we do not fully understand the genetic causes of these conditions and the medical problems that are associated with each condition. Studying the clinical features (the 'phenotype') of individuals with overgrowth disorders will increase our knowledge of these conditions.

Improving our understanding of these disorders will enable health care professionals to provide more accurate information and the best possible care to individuals with overgrowth conditions. Identifying the genetic causes may also help with developing treatments in the future.

Why has our family been invited to join the study?

Overgrowth disorders are rare, so we need to ask as many families as possible with these conditions to take part.

Do we have to take part in the study?

It is completely up to you and your child to decide whether either or both of you wish to join. You may decide to take part in only some parts of the study. If you, your child or both of you decide not to join, your decision will not affect the health care you receive in any way. You, your child or both of you will be free to withdraw at any time without having to give a reason.

What will happen if my child and I decide to participate in the study?

If you and your child agree to join, we will ask you to sign two consent forms, one for yourself and another for your child. Children age between 6 and 15 years of age will be asked to sign an assent form if appropriate.

Your doctor will ask about you and your child's medical histories and perform a routine clinical examination of your child. If you have features suggestive of an overgrowth disorder the doctor will ask to perform a routine clinical examination of you. If you do not have any features of an overgrowth disorder we will ask to measure your head circumference and height only. We will ask to take a set of clinical photographs of your child, and yourself if you have features of an overgrowth condition. You and/or your child may choose to opt out of clinical photography.

You and your child will be asked to provide a small blood sample (up to 20ml or about three teaspoons) which will be taken at your outpatient clinic appointment. If this is not possible, a saliva sample may be given instead. The saliva sample kit, instructions for how to give the sample and an addressed prepaid envelope will be provided for you to post the saliva sample back to us.

We will ask individuals with overgrowth affecting only part of the body for a skin sample taken by skin biopsy.

If your child has previously had an operation or biopsy of an area of overgrowth or a tumour, we will ask you for permission to access any stored tissue samples.

You and/or your child may choose to decline to provide any or all of these samples.

What will happen next?

The clinical information, test results, and information from the medical notes and other health records from you and your child will be entered into a secure research database. You and your child's medical data will be identifiable by a unique study number, so researchers looking at

the data will not be able to see your personal identifiable information You and your child's personal identifiable information will be stored on a separate research database with the unique study identifier to enable us to contact you and your child in the future. Anonymised medical data and samples may be shared with commercial companies in the future for ethically approved research but commercial companies will not have access to you or your child's personal identifiable data.

What will happen to the samples I give?

We will extract and store the genetic material from you and your child's samples. We may determine the DNA sequence of you and your child's genetic code. Part of the sample will be stored for future ethically approved studies.

How often will my child and I be contacted?

We will contact you a maximum of four times during the period of the study. After the study has ended we may contact you again up to a maximum of four times per year to give you information about future studies of rare conditions, potentially including trials of treatment. You and your child are under no obligation to participate in any of these future studies. You are welcome to contact us to ask any questions you may have about these studies.

Who will know about me taking part in the study?

The information collected about you and your child during the course of the research project will be kept strictly confidential. The clinical details, blood samples and information from genetic tests will be given a unique sample study number. The database linking unique sample study numbers to personal details will only be accessed by authorised members of the database team. You and your child will not be identified personally in any report or publication. We will ask you to give consent for us to inform your GP that you are taking part in the study.

What are the benefits of joining the study?

It is likely that there will be no direct benefit to you or your family by joining but you will make a contribution to science and future improvements in the care of individuals with overgrowth conditions.

It is possible that the research could identify a cause of overgrowth in you or your family. This could enable your health care team to provide you with the best possible care. With your prior consent we would let your doctor and your clinical care team know about this result. All

research results would need to be confirmed in an accredited diagnostic laboratory before being used in the clinical management of you and your family members.

There is a voluntary agreement between the government and insurance companies that individuals are not required to disclose any genetic test results acquired as part of clinical research. This agreement has been extended until 2019.

If this research leads to the development of a new treatment or medical test, you will not benefit financially from this.

What other information may be produced by the research?

You will not be told about any genetic results that may identified as a by-product of this research ('incidental findings') and are not relevant to an overgrowth condition.

What are the risks and disadvantages of participating in the research?

Donation of a small blood sample may cause brief discomfort and occasionally a small bruise. Infection at the site is very rare.

Individuals who have overgrowth affecting only one part of the body will be asked for a skin sample obtained via a skin biopsy. Skin biopsy may cause a small scar and there is a small risk of bleeding or infection. This procedure is performed under local anaesthetic. If an individual happens to be having a procedure or operation under general anaesthetic for clinical reasons, they may opt to have the skin biopsy taken at the same time.

What will happen if my child or I don't want to carry on with the study?

You and/or your child are free to withdraw from the study at any time without giving a reason. Any stored samples will be destroyed and we will not contact you again. It will not be possible to destroy samples that have already been prepared for testing, the results of any information obtained from the samples or the information held on the research database.

Who is funding this study?

This study is funded by the NHS National Institute for Health Research (NIHR) Rare Disease Translational Research Collaboration (RD-TRC) and is sponsored by the University of Birmingham.

Who do I contact with any concerns?

If you have any questions or concerns about this project, please contact Dr Alison Foster via [redacted] or [redacted] or the Patient Advise & Liaison Service (PALS) on 0121 627 2747. Alternatively you can write to the following contact addresses:

Dr Alison Foster

Clinical Genetics Unit

Birmingham Women's Hospital

Edgbaston

Birmingham

B15 2TG

Thank you for reading this information and considering participation in the study.

B.2.6 Parent of adult unable to give consent

INFORMATION LEAFLET FOR PARENTS OF ADULT PARTICIPANTS UNABLE TO GIVE CONSENT

PHENOTYPING OF OVERGROWTH DISORDERS (POD)

Version 1.1 16/6/15

Your family is invited to take part in a research project. Before you decide whether or not you wish to take part it is important for you to understand why the project is being done and what it will involve. Please take the time to read the following information carefully. Feel free to discuss it with your family or close friends and ask us if there is anything that is not clear, or if you would like more information. You will be given as much time as you like to make a decision.

What is the purpose of this study?

We would like to increase our understanding of the clinical and genetic features of rare genetic overgrowth disorders. At present we do not fully understand the genetic causes of these conditions and the medical problems that are associated with each condition. Studying at the clinical features (the 'phenotype') of individuals with overgrowth will increase our knowledge of these conditions.

Improving our understanding of these disorders will enable health care professionals to provide more accurate information and the best possible care to individuals with overgrowth conditions. Identifying the genetic causes may also help with developing treatments in the future.

Why has our family been invited to join the study?

Overgrowth disorders are rare, so we need to ask as many families as possible with these conditions to take part.

Do we have to take part in the study?

It is completely up to you whether you choose to join. You may decide to take part in only some parts of the study. If you decide not to join, your decision will not affect the health care you receive in any way. You will be free to withdraw at any time without having to give a reason.

What will happen if I decide to participate in the study?

If you agree to join, we will ask you to sign a consent form

Your doctor will ask about your medical history. If you have features suggestive of an overgrowth disorder the doctor will ask to perform a routine clinical examination of you. If you do not have any features of an overgrowth disorder we will ask to measure your head circumference and height only. We will ask to take a set of clinical photographs if you have features of an overgrowth condition. You may choose to opt out of clinical photography.

You will be asked to provide a small blood sample (up to 20ml or about three teaspoons). If this is not possible, a saliva sample may be given instead. The saliva sample kit, instructions for how to give the sample and an addressed prepaid envelope will be provided for you to post the saliva sample back to us. You may choose to decline to provide these samples.

What will happen next?

The clinical information, test results, and information from the medical notes and other health records from you and your son/daughter will be entered into a secure research database. You and your son's/daughter's medical data will be identified by a unique study number, so researchers looking at the data will not be able to see your personal identifiable information. You and your son's/daughter's personal identifiable information will be stored on a separate database with the unique study identifier to enable us to contact you and your son/daughter in the future. Anonymised medical data and samples may be shared with commercial companies in the future for ethically approved research but commercial companies will not have access to you or your son's/daughter's personal identifiable data.

What will happen to the samples I give?

We will extract and store the genetic material from your sample. We may determine the DNA sequence of your genetic code. Part of the sample will be stored for future ethically approved studies.

How often will my son/daughter and I be contacted?

We will contact you a maximum of four times during the period of the study. After the study has ended, we may contact you again up to a maximum of four times per year to give you information about future studies of rare conditions. You are under no obligation to participate in any of these future studies. You are welcome to contact us to ask any questions you may have about these studies.

Who will know about me taking part in the study?

The information collected about you and your son/daughter during the course of the research project will be kept strictly confidential. The clinical details, blood samples and information from genetic tests will be given a unique sample study number. The database linking unique sample study numbers to personal details will only be accessed by authorised members of the database team. You and your son/daughter will not be identified personally in any report or publication. We will ask you to give consent for us to inform your GP that you are taking part in the study.

What are the benefits of joining the study?

It is likely that there will be no direct benefit to you or your family by joining but you will make a contribution to science and future improvements in the care of individuals with overgrowth conditions.

It is possible that the research could identify a cause of overgrowth in you or your family. This could enable your health care team to provide you with the best possible care. With your prior consent we would let your doctor and your clinical care team know about this result. All research results would need to be confirmed in an accredited diagnostic laboratory before being used in the clinical management of you and your family members.

There is a voluntary agreement between the government and insurance companies that individuals are not required to disclose any genetic test results acquired as part of clinical research. This agreement has been extended until 2019.

If this research leads to the development of a new treatment or medical test, you will not benefit financially from this.

What other information may be produced by the research?

You will not be told about any genetic results that may identified as a by-product of this research (‘incidental findings’) and are not relevant to an overgrowth condition.

What are the risks and disadvantages of participating in the research?

Donation of a small blood sample may cause brief discomfort and occasionally a small bruise. Infection at the site is very rare.

What will happen if I don't want to carry on with the study?

You are free to withdraw from the study at any time without giving a reason. Any stored samples will be destroyed and we will not contact you again. It will not be possible to destroy samples that have already been prepared for testing, the results of any information obtained from the samples or the information held on the research database.

Who is funding this study?

This study is funded by the NHS National Institute for Health Research (NIHR) Rare Disease Translational Research Collaboration (RD-TRC) and is sponsored by the University of Birmingham.

Who do I contact with any concerns?

If you have any questions or concerns about this project, please contact Dr Alison Foster via [redacted] or [redacted], or the Patient Advise & Liaison Service (PALS) on 0121 627 2747. Alternatively you can write to the following contact addresses:

Dr Alison Foster

Clinical Genetics Unit

Birmingham Women's Hospital

Edgbaston

Birmingham

B15 2TG

Thank you for reading this information and considering participation in the study.

B.3 Consent forms

B.3.1 Competent adult

ADULT PARTICIPANT CONSENT FORM

NIHR Rare Diseases Translational Research Collaboration (RD-TRC): Phenotyping of Overgrowth Disorders (POD)

Version 1.1 16/6/2015

Please initial boxes

1. I confirm that I have read and understood the information leaflet dated __/__/__ (version 1) for the POD Study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my participation in this study is voluntary and that I am free to withdraw at any time, without giving any reason and without my medical care or legal rights being affected. ☐

3. I agree for details about me and any samples I provide to be kept on a secure database; I understand that my medical data and personal identifiable data will be kept on separate databases and linked by a unique study identifier ☐

4. I understand that my medical notes and health records may be looked at by individuals from the University of Birmingham, the Birmingham Women's Hospital, the NIHR RD-TRC, or from the NHS Trust responsible for my care; I give permission for these individuals to have access to my records for the purposes of this study. ☐

5. I agree that any anonymised samples I provide may be moved between partners involved in the study, including the Hospital Trust, Higher Education Institute and commercial partners based in the UK. ☐

6. I agree that the information gathered about me and any samples I donate can be stored for use in future ethically approved research studies. ☐

7. I understand that this research may include the participation of commercial companies and that I will not benefit financially if this research leads to new treatments or medical tests. ☐

8. I agree that my anonymised data can be shared with other disease registries and research projects relevant to my condition. ☐

9. I agree that I can be contacted and invited to participate in medical research studies based on the results obtained from any samples I have provided and information about me which has been retrieved from databases. I will be provided with full information about these studies, when and if I am contacted. I understand that I am free to decide whether or not to take part in these studies. ☐

10. I give consent for my GP to be informed that I am taking part in this study. ☐

Optional – please circle yes if you agree, no if you do not agree

11. I give consent for clinical photographs of me to be kept on a secure database.

Y N

12. I agree to give a sample of blood or saliva for medical research including genetic analysis.

Y N

13. I agree to give a skin sample for medical research including genetic analysis (participants with regional overgrowth only).

Y N

14. I agree that stored samples of my tissue taken at the time of previous surgery or biopsy may be used for medical research including genetic analysis.

Y N

15. I give consent for the research team to feedback the results of genetic tests relating to the cause of my disease to my extended clinical care team. I agree that the clinical care team can feedback this information to me.

Y N

Name of Participant (BLOCK CAPITALS) Date of Birth Date Signature

---

Name of Researcher (BLOCK CAPITALS) Date Signature

B.3.2 Parent on behalf of child/young person

PARENT/GUARDIAN CONSENT FOR CHILD

NIHR Rare Diseases Translational Research Collaboration (RD-TRC): Phenotyping of Overgrowth Disorders (POD)

Version 1.1 16/06/2015

Please initial boxes

1. I, the undersigned, am the parent or legal guardian of the child named below, and I have the authority to execute this Consent Form on behalf of the child. ☐

2. I confirm that I have read and understood the information leaflet dated __/__/__ (version 1) for the POD study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

3. I understand and agree that my child's participation in this study is voluntary and that he/she is free to withdraw at any time, without giving any reason and without his/her medical care or legal rights being affected. ☐

4. I agree for details about my child and any samples he/she provides to be kept on a secure database; I understand that his/her medical data and personal identifiable data will be kept on separate databases and linked by a unique study identifier. ☐

5. I agree that any anonymised samples my child provides may be moved between partners involved in the study, including the Hospital Trust, Higher Education Institute and commercial partners based in the UK. ☐

6. I understand that my child's medical notes and health records may be looked at by individuals from University of Birmingham, the Birmingham Women's Hospital, the NIHR RD-TRC, or from the NHS Trust responsible for my child's care. I give permission for these individuals to have access to my child's records for the purposes of this study ☐

7. I agree that any samples donated by my child and the information provided can be stored for use in future ethically approved research studies. ☐

8. I understand that this research may include the participation of commercial companies and I understand that neither I nor my child will benefit financially if this research leads to new treatments or medical tests. ☐

9. I agree to be contacted and my child invited to participate in medical research studies based on the results obtained from any samples he/she has provided and the informationretrieved from the databases. Both my child and I will be provided with age appropriate information about these studies, when and if we are contacted. I understand that my child and I are free to decide whether or not my child will take part in these studies. ☐

10. I agree that my child's anonymised data can be shared with other disease registries and research projects relevant to their condition. ☐

11. I give consent for my child's GP to be informed that my child is taking part in this study. ☐

Optional – please circle yes if you agree, no if you do not agree

12. I give consent for clinical photographs of my child to be kept on a secure database.

    Y N

13. I agree to my child providing a blood or saliva for medical research including genetic analysis.

    Y N

14. I agree to my child providing a skin sample for medical research including genetic analysis (participants with regional overgrowth only)

    Y N

15. I agree that samples of my child's tissue taken at the time of surgery or biopsy may be used for medical research including genetic analysis.

    Y N

16. I give consent for the research team to feedback the results of genetic tests that may be linked to the cause of my child's disease to his or her extended clinical care team. I agree that the clinical care team can feedback this information to my child and myself .

    Y N

---------- ---------- ----------

First name and surname of parent (PRINT) Date Signature

---------- ---------- ---------- ----------

Name of child (BLOCK CAPITALS) Date of Birth Date Signature

---------- ---------- ----------

Researcher Date Signature

B.3.3 Child/young person assent

ASSENT FORM FOR CHILDREN AND YOUNG PEOPLE AGED 6-15 YEARS

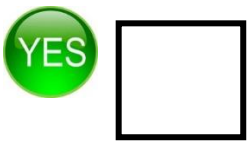
POD Study

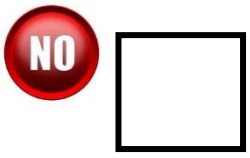
Version 1.1 16/06/2015

Young person to circle all they agree with please:

Have you read (or had read to you) about this project?	Yes	No
Has somebody else explained this project to you?	Yes	No
Do you understand what this project is about?	Yes	No
Have you asked all the questions you want?	Yes	No
Have you had your questions answered in a way you understand?	Yes	No
Do you understand it's OK to stop taking part at any time?	Yes	No

Are you happy to take part?





If any answers are ‘no’ or you **don’t** want to take part, **don’t** sign your name.

If you do want to take part, please write your name and today’s date.

Name of child or young person (PRINT)……

Date of Birth.....

Signature…….. Date……...

Name of mother* (PRINT)……..

Signature…….. Date……...

Name of father* (PRINT)……..

Signature…….. Date……...

Name of Guardian(s)
(PRINT)……..

Signature…….. Date……...

Name of person obtaining consent

(PRINT)……..

Signature…….. Date……...

*Only one of the parents has to sign the form to validate it, but if parents wish to they can both sign.

B.3.4 Parent of child/young person participant

PARENT PARTICIPANT CONSENT FORM

NIHR Rare Diseases Translational Research Collaboration (RD-TRC): Phenotyping of Overgrowth Disorders (POD)

Version 1.1 16/06/2015

Please initial boxes

1. I confirm that I have read and understood the information leaflet dated __/__/__ (version_) for the POD study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2 .I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason and without my medical care or legal rights being affected. ☐

3. I agree for details about me and any samples I provide to be kept on a secure database; I understand that my medical data and personal identifiable data will be kept on separate databases and linked by a unique study identifier. ☐

4. I understand that my medical notes and health records may be looked at by individuals from the University of Birmingham, the Birmingham Women's Hospital, the NIHR RD-TRC, or from the NHS Trust responsible for my care; I give permission for these individuals to have access to my records now and in the future for the purposes of this and other ethically approved research. ☐

5. I agree that any anonymised samples I provide may be moved between partners involved in the study including the Hospital Trust, Higher Education Institutes and commercial partners based in the UK. ☐

6. I agree that the information gathered about me and any samples I donate can be stored for use in future ethically approved research studies. ☐

7. I understand that this research may include the participation of industry and commercial companies and that I will not benefit financially if this research leads to new treatments or medical tests. ☐

8. I agree that I can be contacted and invited to participate in medical research studies based on the results obtained with the samples and information from me which has been retrieved from databases. I will be provided with full information about these studies, when and if I am contacted. I understand that I am free to decide whether or not to take part in these studies. ☐

9. I give consent for my GP to be informed that I am taking part in this study. ☐

Optional Please circle yes if you agree and no if you do not agree

10. I agree to give a sample of blood or saliva for medical research including genetic analysis.

Y N

11. I give consent for the research team to feedback the results of genetic tests relating to the cause of disease in my child to his or her extended clinical care team. I agree that the clinical care team can feedback this information to me.

Y N

---------- ---------- ---------- ----------

Name of Participant (BLOCK CAPITALS) Date of Birth Date Signature

---------- ---------- ----------

Name of Researcher (BLOCK CAPITALS) Date Signature

B.3.5 Parent of adult participant

PARENT OF ADULT PARTICIPANT CONSENT FORM

NIHR Rare Diseases Translational Research Collaboration (RD-TRC): Phenotyping of Overgrowth Disorders (POD)

Version 1.1 16/6/2015

Please initial boxes

1. I confirm that I have read and understood the information leaflet dated __/__/__ (version_) for the POD study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason and without my medical care or legal rights being affected. ☐

3. I agree for details about me and any samples I provide to be kept on a secure database; I understand that my medical data and personal identifiable data will be kept on separate databases and linked by a unique study identifier. ☐

4. I understand that my medical notes and health records may be looked at by individuals from the University of Birmingham, the Birmingham Women's Hospital, the NIHR RD-TRC, or from the NHS Trust responsible for my care; I give permission for these individuals to have access to my records for the purposes of this study. ☐

5. I agree that any anonymised samples I provide may be moved between partners involved in the study, including the Hospital Trust, Higher Education Institute and ☐

commercial partners based in the UK.

6. I agree that the information gathered about me and any samples I donate can be stored for use in future ethically approved research studies. ☐

7. I understand that this research may include the participation of industry and commercial companies and that I will not benefit financially if this research leads to new treatments or medical tests. ☐

8. I agree that I can be contacted and invited to participate in medical research studies based on the results obtained from any samples and information about me which has been retrieved from databases. I will be provided with full information about these studies, when and if I am contacted. I understand that I am free to decide whether or not to take part in these studies. ☐

9. I give consent for my GP to be informed that I am taking part in this study. ☐

Optional – please circle yes if you agree, no if you do not agree

10. I agree to give a sample of blood or saliva for medical research including genetic analysis.

Y N

11. I give consent for the research team to feedback the results of genetic tests that are causative of the overgrowth in the family if they are relevant to my health to my

extended clinical care team. I agree that the clinical care team can feedback this information to me.

Y N

Name of Participant (BLOCK CAPITALS) Date of Birth Date Signature

---

Name of Researcher (BLOCK CAPITALS) Date Signature

B.3.6 Consultee declaration

CONSULTEE DECLARATION FORM

NIHR Rare Diseases Translational Research Collaboration (RD-TRC): Phenotyping of Overgrowth Disorders (POD)

Version 1.0 16/06/2015

Please initial boxes

1. I [ ] have been consulted about [ ]'s participation in this research project. In my opinion he/she would have no objection to taking part in this study. I have answered the questions below accordingly. ☐

2. I confirm that I have read and understood the information leaflet dated __/__/__ (version 1) for the POD study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily. ☐

3. I understand that my friend's/relative's participation in this study is voluntary and that he/she is free to withdraw at any time, without giving any reason and without my medical care or legal rights being affected. ☐

4. I agree for details about my friend/relative and any samples he/she provides to be kept on a secure database; I understand that his/her medical data and personal identifiable data will be kept on separate databases and linked by a unique study identifier. ☐

5. I understand that my friend's/relative's medical notes and health records may be looked at by individuals from the University of Birmingham, the Birmingham ☐

Women's Hospital, the NIHR RD-TRC, or from the NHS Trust responsible for my care; I agree these individuals may have access to these records for the purposes of this study.

6. I agree that anyanonymised samples my friend/relative provides may be moved between partners involved in the study, including the Hospital Trust, Higher Education Institute and commercial partners based in the UK ☐

7. I agree that the information gathered about my friend/relative and any samples he/she donates can be stored for use in future ethically approved research studies. ☐

8. I understand that this research may include the participation of commercial companies and that my friend/relative will not benefit financially if this research leads to new treatments or medical tests. ☐

9. I agree that my friend/relative can be contacted and invited to participate in medical research studies based on the results obtained from any samples and information about him/her which has been retrieved from databases. He/she, and an appropriate consultee, will be provided with full information about these studies, when and if he/she is contacted. I understand that he/she is free to decide whether or not to take part in these studies. ☐

10. I agree that my friend's/relative's anonymised data can be shared with other disease registries and research projects relevant to my condition. ☐

11. I agree for my friend's/relative's GP to be informed that he/she is taking part in this study. ☐

Optional – please circle yes if you agree, no if you do not agree

12. I agree for clinical photographs of my friend/relative to be kept on a secure database.

Y N

13. I agree for my friend/relative to give a sample of blood or saliva for medical research including genetic analysis.

Y N

14. I agree for my friend/relative to give a skin sample for medical research including genetic analysis (participants with regional overgrowth only)

Y N

15. I agree that samples of my friend's/relative's tissue taken at the time of surgery or biopsy may be used for medical research including genetic analysis.

Y N

16. I agree for the research team to feedback the results of genetic tests relating to the cause of my friend's/relative's disease to his/her extended clinical care team. I agree that the clinical care team can feedback this information.

Y N

---------- ---------- ----------

Name of Participant (BLOCK CAPITALS) Date of Birth Date

---------- ---------- ---------- ----------

Name of Consultee (BLOCK CAPITALS) Date of Birth Date Signature

---------- ---------- ----------

Name of Researcher (BLOCK CAPITALS) Date Signatur

B.4 Clinical Record Form (CRF)

POD Study: Case Report Form V1.4

Visit 1

Centre Name:……Name of reporting health professional:……

Date of visit: D D M M Y Y Y Y Date of completing CRF: D D M M Y Y Y Y

Patient Identifier: ☐ ☐ ☐ ☐

Name:……NHS Number: ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

Date of Birth: D D M M Y Y Y Y Hospital Number: ……

Age at recruitment……years

Sex: M ☐ F ☐

Known genetic overgrowth disorder: Yes / No

If yes please give condition name…… gene name……

Eligibility criteria met:

OFC, height, and/or weight >2 SD (>98th centile) plus additional feature of overgrowth syndrome (intellectual disability, congenital anomaly or childhood tumour) ☐

OFC or height > 3 SD (>99.9th centile) ☐

Lateralised overgrowth (segmental/regional/hemihyperplasia/hemihypertrophy) ☐

Genetic or genomic variant known to be associated with overgrowth ☐

Genetic Investigations:

Any genetic investigations performed: Yes ☐ No ☐ If yes, please circle from the following list:

Karyotype CGH microarray FRAX NSD1 EZH2 PTEN PIK3CA

Overgrowth gene panel test Clinical exome Other (please list) .....

Results.....

.....

.....

.....

Enrolled in other study: please circle: COG/ DDD/GOOS/100K Genomes/ other

Results.....

.....

.....

.....

Pregnancy, Birth, Neonatal History

Conception: Natural ☐ Assisted ☐

Singleton: Yes ☐ No ☐ Details……

Medical problems in pregnancy? Yes ☐ No ☐ Details……

……..

Medications/drugs in pregnancy? Yes ☐ No ☐ Details ……

……

Abnormality on antenatal scans? Yes ☐ No ☐ Details: ……

……

Antenatal testing? Yes ☐ No ☐ Details……

Gestation……

Breech presentation? Yes ☐ No ☐

Delivery (circle as appropriate) Spontaneous/ NVD / Instrumental / Elective LSCS / Emergency LSCS

Resuscitation required? Yes ☐ No ☐ Details:……

Admission to NNU required? Yes ☐ No ☐ Length of stay ……. Days

Features at birth (circle as appropriate)

Large for gestational age (macrosomia): Yes / No / Don't know
Macroglossia: Yes / No / Don't know
Omphalocoele: Yes / No / Don't know
Umbilical hernia: Yes / No/ Don't know
Birthmarks: Yes / No/ Don't know
Hypotonia: Yes / No/ Don't know
Hypoglycaemia: Yes/ No/Don't know
Feeding Difficulty: Yes / No / Don't know
Jaundice: Yes / No/ Don't know

Details and any other features

……

……..

……..

……..

Growth

Birth length ☐☐.☐ cm

Birth weight ☐.☐☐☐ kg

Head circumference at birth ☐☐.☐ cm

Current measurements (date): D D M M Y Y

Height/ Length ☐☐☐.☐ cm

Weight ☐☐☐.☐ kg

Head circumference ☐☐☐.☐ cm

Shoe size: . . . .

Other measurements

Date: D D M M Y Y

Height/ Length ☐☐☐.☐ cm

Weight ☐☐☐.☐ kg

Head circumference ☐☐☐.☐ cm

Date: D D M M Y Y

Height/ Length ☐☐☐.☐ cm

Weight ☐☐☐.☐ kg

Head circumference ☐☐☐.☐ cm

Stature

Proportionate/ Disproportionate (circle as appropriate)

If disproportionate, please complete further measurements:

Arm span ☐☐☐.☐ cm

Upper segment ☐☐☐.☐ cm

Lower segment ☐☐☐.☐ cm

Sitting height ☐☐☐.☐ cm

Bone age done Yes ☐ No ☐ Actual age years....Months..... Bone age years.....months.....

Asymmetry/ hemihypertrophy? Yes ☐ No ☐

Circle affected: Upper limb / Lower limb / Face

Undergoing regular tumour screening? Yes ☐ No ☐

Screening site.....

Age commenced (years).....

Modality.....

Frequency: Circle applicable: One-off / regularly

Age of puberty (menarche/ voice breaking): .....Years .....Months

MEDICAL HISTORY

History of any medical problems in the following areas:

Endocrine: Yes ☐ No ☐ If yes, please complete the following;

Diabetes mellitus: Yes ☐ No ☐

Thyroid problems: Yes ☐ No ☐

Hypothalamus-pituitary axis: Yes ☐ No ☐

Adrenal: Yes ☐ No ☐

Reproductive: Yes ☐ No ☐

Details.....

Cardiac: Yes ☐ No ☐ If yes, please complete the following:

Congenital heart disease: Yes ☐ No ☐

Echocardiogram performed: Yes ☐ No ☐ If Yes: normal/abnormal

ECG performed: Yes ☐ No ☐ If Yes: normal/abnormal

Details.....

Respiratory: Yes ☐ No ☐

Details.....

Neurological: Yes ☐ No ☐ If yes, please complete the following:

Seizures: Yes ☐ No ☐

Hydrocephalus: Yes ☐ No ☐

Hypotonia Yes ☐ No ☐

Other neurological features: Yes ☐ No ☐

EEG performed: Yes ☐ No ☐ If Yes: normal/abnormal

Cranial imaging performed: Yes ☐ No ☐ If Yes: normal/abnormal

Details.....

Patient ID ☐☐☐ 6

Renal: Yes ☐ No ☐ If yes, please complete the following:

Structural abnormalities: Yes ☐ No ☐

Vesico-ureteric reflux: Yes ☐ No ☐

Renal imaging: Yes ☐ No ☐

Details……

Gastrointestinal: Yes ☐ No ☐

Details……

Musculoskeletal: Yes ☐ No ☐ If yes, please complete the following:

Contractures: Yes ☐ No ☐

Hypermobility: Yes ☐ No ☐ If yes, Beighton score……

Scoliosis: Yes ☐ No ☐ Please give details……

Details ……

Dermatological: Yes ☐ No ☐

Abnormalities of dentition: Yes ☐ No ☐

Details……

Ophthalmic: Yes ☐ No ☐

Details……

Audiological: Yes ☐ No ☐

Details……

Immunology: Yes ☐ No ☐

Details……

Malignancy: Yes ☐ No ☐ If yes, please complete the following:

Age of diagnosis: ……..….years ……months Type: ……..

Psychiatric: Yes ☐ No ☐

Details……

Any other information……

……

……

……

……

Patient ID ☐☐☐ 8

DEVELOPMENT AND BEHAVIOUR

Developmental delay: Y / N

Global: Yes ☐ No ☐ Mild / Moderate / Severe

Gross motor: Yes ☐ No ☐ Mild /Moderate /Severe

Age sat independently: ……..months or N/A

Age walked independently: ….months or N/A

Fine motor: Yes ☐ No ☐ Mild/ Moderate /Severe

Speech and language: Yes ☐ No ☐ Mild / Moderate / Severe

Age first word: ….. months or N/A

Age two word sentences: ……. months or N/A

Social & behavioural: Yes ☐ No ☐ Mild /Moderate / Severe

Formal developmental assessment performed: Yes ☐ No ☐

Details……

Primary Schooling: Mainstream/ Mainstream with assistance/ Mainstream with statement/ Special school/ NA

Secondary schooling: Mainstream/ Mainstream with assistance/ Mainstream with statement/ Special school/ NA

Behavioural difficulties: Yes ☐ No ☐ (Please circle)

Aggression / Emotional Lability / Temper Tantrums / Anxiety / Features of autism spectrum disorder

Hyperactivity / Short attention span / Obsessive-compulsive behaviour / Self-injury / Hypersensitivity to pain /

Pain insensitivity/ Abnormality of temperature regulation / Increased sweating/ / Light hypersensitivity/

Hyperacusis/ Oral aversion / Feeding problems / Polyphagia / Sleep problems

Diagnosis of autism spectrum disorder: Yes ☐ No ☐

Diagnosis of ADHD: Yes ☐ No ☐

Details:……

……

Patient ID ☐☐☐ 9

ADULTHOOD (if over 16 years)

Living: Independently/ with family member/ in sheltered accommodation/ in residential care

Activities of daily living: Independent / requires assistance / dependent

Education: GCSEs/ A-levels/ Degree (or equivalents)

Employment: Yes ☐ No ☐ Details……

Fertility problems Yes ☐ No ☐ Details……

Children? Yes ☐ No ☐ If yes, does the child have an overgrowth disorder? Yes ☐ No ☐

Family History

Mothers' birthweight……kg

Current height……cm OFC……cm

Father's birthweight……kg

Current height……cm OFC……cm

Siblings: Yes ☐ No ☐ If yes:

Sibling 1: Sex: M / F Age (years): . . . . . birth length……cm weight……kg OFC……cm

Current height……cm weight……kg OFC……cm

Sibling 2: Sex: M / F Age (years): . . . . . birth length……cm weight……kg OFC……cm

Current height……cm weight……kg OFC……cm

Sibling 3: Sex: M / F Age (years): . . . . . birth length……cm weight……kg OFC……cm

Current height……cm weight……kg OFC……cm

Any family member with overgrowth disorder: Yes ☐ No ☐ Details……

……

Consanguinity: Yes ☐ No ☐

Pedigree:

Examination findings/ Dysmorphology:

Clinical photography taken Y / N

Consent for proband Y / N
Samples from proband:
Blood Y / N
Saliva Y / N
Buccal Y / N
Skin biopsy Y / N
Other Y / N

Consent for mother Y / N
Samples from mother:
Blood Y / N
Other Y / N

Consent for father Y / N
Samples from father:
Blood Y / N
Other Y / N

Form Completed by:

.....Signature: .....

Date: D D M M Y Y Y Y

Patient ID ☐☐☐ 12

C. Human Phenotype Ontology terms in the OpenClinica database

Table: HPO terms selected for coding dysmorphology

Abnormality of the cranium	**Brachycephaly**	HP:0000248
	Dolichocephaly	HP:0000268
	Macrocephaly	HP:0000256
	Microcephaly	HP:0000252
	Flat occiput	HP:0005469
	Prominent occiput	HP:0000269
	Plagiocephaly	HP:0001357
	Cloverleaf skull	HP:0002676
	Trigonocephaly	HP:0000243
	Turricephaly	HP:0000262
Abnormality of the hair	Frontal balding	HP:0002292
	Frontal upsweep	HP:0002236
	Abnormal number hair whorls	HP:0010813
	Abnormal position hair whorl	HP:0010814
	High anterior hairline	HP:0009890
	Low anterior hairline	HP:0000294
	Low posterior hairline	HP:0002162
	Sparse scalp hair	HP:0002209
	Widow's peak	HP:0000349
Abnormality of the face	Broad face	HP:0000283
	Coarse face	HP:0000280
	Expressionless face	HP:0000298

	Flat face	HP:0012368
	Long face	HP:0000276
	Narrow face	HP:0000275
	Round face	HP:0000311
	Short face	HP:0011219
	Square face	HP:0000321
	Triangular face	HP:0000325
Abnormality of the forehead	Broad forehead	HP:0000337
	Narrow forehead	HP:0000341
	Prominent forehead	HP:0011220
	Sloping forehead	HP:0000340
	Vertical forehead creases	HP:0011221
	Frontal bossing	HP:0002007
	Depressed glabella	HP:0011222
	Prominent glabella	HP:0002057
	Metopic depression	HP:0011223
	Prominent metopic ridge	HP:0005487
	Prominent supraorbital ridges	HP:0000336
	Underdeveloped supraorbital ridges	HP:0009891
Abnormality of the maxilla and midface	Cheekbone prominence	HP:0010620
	Cheekbone underdevelopment	HP:0010669
	Full cheeks	HP:0000293
	Sunken cheeks	HP:0009938
	Malar flattening	HP:0000272
	Malar prominence	HP:0012370
	Midfacial prominence	HP:0012371
	Midface retrusion	HP:0011800
	Prominent nasolabial fold	HP:0005272
	Underdeveloped nasolabial fold	HP:0010801

	Premaxillary prominence	HP:0430029
	Premaxillary underdevelopment	HP:0010650
Abnormality of the mandible	Broad jaw	HP:0012802
	Narrow jaw	HP:0012801
	Cleft mandible	HP:0010752
	Micrognathia	HP:0000308
	Prognathia	HP:0000303
	Retrognathia	HP:0000278
Abnormality of the chin	Broad chin	HP:0011822
	Chin dimple	HP:0010751
	Horizontal crease in chin	HP:0011823
	H-shaped crease in chin	HP:0011824
	Pointed chin	HP:0000307
	Short chin	HP:0000331
	Tall chin	HP:0400000
	Vertical crease in chin	HP:0400001
Abnormality of the neck	Broad neck	HP:0000475
	Long neck	HP:0000472
	Short neck	HP:0000470
	Neck webbing	HP:0000465
	Redundant nuchal skin	HP:0005989
Abnormality of the periorbital region	Ablepharon	HP:0011224
	Ankyloblepharon	HP:0009755
	Blepharochalasis	HP:0010749
	Blepharophimosis	HP:0000581
	Cryptophthalmos	HP:0001126
	Ectropion	HP:0000656

Entropion	HP:0000621
Epiblepharon	HP:0011225
Epicanthus	HP:0000286
Epicanthus inversus	HP:0000537
Closely spaced eyes	HP:0000601
Deeply set eyes	HP:0000490
Widely spaced eyes	HP:0000316
Broad eyebrows	HP:0011229
Highly arched eyebrows	HP:0002553
Horizontal eyebrows	HP:0011228
Laterally extended eyebrows	HP:0011230
Sparse eyebrows	HP:0100840
Thick eyebrows	HP:0000574
Absent eyelashes	HP:0000561
Long eyelashes	HP:0000527
Prominent eyelashes	HP:0011231
Sparse eyelashes	HP:0000653
Cleft eyelid	HP:0000625
Infra-orbital crease	HP:0100876
Infra-orbital fold	HP:0011232
Absent lacrimal punctum	HP:0001092
Ectopic lacrimal punctum	HP:0010748
Lagophthalmos	HP:0030001
Almond-shaped palpebral fissure	HP:0007874
Downslanted palpebral fissure	HP:0000494
Long palpebral fissure	HP:0000637
Short palpebral fissure	HP:0012745
Upslanted palpebral fissure	HP:0000582
Proptosis	HP:0000520

	Ptosis	HP:0000508
	Synophrys	HP:0000664
	Telecanthus	HP:0000506
	Fullness of upper eyelid	HP:0012724
Abnormality of the Nose	Cleft ala nasi	HP:0003191
	Thick ala nasi	HP:0009928
	Underdeveloped ala nasi	HP:0000430
	Broad columella	HP:0010761
	High insertion columella	HP:0012807
	Low hanging columella	HP:0009765
	Low insertion columella	HP:0010763
	Short columella	HP:0002000
	Anteverted nares	HP:0000463
	Enlarged naris	HP:0009931
	Narrow nares	HP:0009933
	Single naris	HP:0009932
	Supernumery naris	HP:0009934
	Narrow nasal base	HP:0012809
	Wide nasal base	HP:0012810
	Depressed nasal bridge	HP:0005280
	Narrow nasal bridge	HP:0000446
	Prominent nasal bridge	HP:0000426
	Wide nasal bridge	HP:0000431
	Absent nasal cartilage	HP:0030028
	Concave nasal ridge	HP:0011120
	Convex nasal ridge	HP:0000444
	Depressed nasal ridge	HP:0000457
	Narrow nasal ridge	HP:0000418

	Wide nasal ridge	HP:0012811
	Bifid nasal tip	HP:0000456
	Broad nasal tip	HP:0000455
	Depressed nasal tip	HP:0000437
	Deviated nasal tip	HP:0011831
	Narrow nasal tip	HP:0011832
	Overhanging nasal tip	HP:0011833
	Absent nose	HP:0009927
	Bifid nose	HP:0011803
	Bulbous nose	HP:0000414
	Long nose	HP:0003189
	Narrow nose	HP:0000460
	Prominent nose	HP:0000448
	Short nose	HP:0003196
	Wide nose	HP:0000445
	Fullness paranasal tissue	HP:0012812
	Proboscis	HP:0012806
Abnormality of the philtrum	Maligned philtral ridges	HP:0011827
	Broad philtrum	HP:0000289
	Deep philtrum	HP:0002002
	Long philtrum	HP:0000343
	Midline raphe of philtrum	HP:0011826
	Midline sinus of philtrum	HP:0011828
	Narrow philtrum	HP:0011829
	Short philtrum	HP:0000322
	Smooth philtrum	HP:0000319
	Tented philtrum	HP:0011825
Abnormality of the lips	Commissural pit	HP:0002710
	Absent Cupid’s bow	HP:0010800
	Exaggerated Cupid’s bow	HP:0002263

	Lip freckle	HP:0010798
	Lip pit	HP:0100267
	Prominent nasolabial fold	HP:0005272
	Underdeveloped nasolabial fold	HP:0010801
	Perioral hyperpigmentation	HP:0010802
	Everted lower lip vermilion	HP:0000232
	Thick lower lip vermilion	HP:0000179
	Thin lower lip vermilion	HP:0010282
	Everted upper lip vermilion	HP:0010803
	Tented upper lip vermilion	HP:0010804
	Thick upper lip vermilion	HP:0000215
	Thin upper lip vermilion	HP:0000219
Abnormality of the mouth	Fibrous syngnathia	HP:0009754
	Intra-oral hyperpigmentation	HP:0010284
	Downturned corners of mouth	HP:0002714
	Narrow mouth	HP:0000160
	Upturned corners of mouth	HP:0010805
	Wide mouth	HP:0000154
	Accessory oral frenulum	HP:0000191
	Oral synechia	HP:0010285
	U-shaped upper lip vermilion	HP:0010806
Abnormality of the oral cavity	Alveolar ridge overgrowth	HP:0009085
	Ankyloglossia	HP:0010296
	Single maxillary central incisor	HP:0006315

Dental crowding	HP:0000678
Diastema	HP:0000699
Advanced eruption	HP:0006337
Delayed eruption	HP:0000684
Gingival overgrowth	HP:0000212
Glossoptosis	HP:0000162
Macrodontia	HP:0001572
Microdontia	HP:0000691
Oligodontia	HP:0000677
Open bite	HP:0010807
Short hard palate	HP:0010290
High palate	HP:0000218
Narrow palate	HP:0000189
Submucous cleft palate	HP:0000176
Prominent palatine ridges	HP:0010291
Fused teeth	HP:0011090
Widely spaced teeth	HP:0000687
Bifid tongue	HP:0010297
Furrowed tongue	HP:0000221
Large tongue	HP:0000158
Lobulated tongue	HP:0000180
Protruding tongue	HP:0010808
Small tongue	HP:0000171
Smooth tongue	HP:0010298
Natal tooth	HP:0000695
Premature tooth loss	HP:0006323
Supernumerary tooth	HP:0011069
Absent uvula	HP:0010292
Broad uvula	HP:0010809
Cleft uvula	HP:0000193
Long uvula	HP:0010810
Narrow uvula	HP:0010811
Short uvula	HP:0010812

Abnormality of the ears	Anotia	HP:0009892
	Antihelical shelf	HP:0011233
	Absent antihelix	HP:0011234
	Additional crus antihelix	HP:0011235
	Angulated antihelix	HP:0011236
	Broad inferior crus antihelix	HP:0011237
	Prominent inferior crus antihelix	HP:0011238
	Underdeveloped crus antihelix	HP:0011239
	Prominent stem antihelix	HP:0011240
	Serpiginous stem antihelix	HP:0011241
	Underdeveloped stem antihelix	HP:0011242
	Prominent superior crus antihelix	HP:0011247
	Underdeveloped superior crus antihelix	HP:0011246
	Absent antitragus	HP:0011249
	Bifid antitragus	HP:0011250
	Everted antitragus	HP:0011248
	Prominent antitragus	HP:0008593
	Underdeveloped antitragus	HP:0011272
	Extra fold concha	HP:0400002
	Cryptotia	HP:0011252
	Crumpled ear	HP:0009901
	Cupped ear	HP:0000378
	Focal absence ear	HP:0400003
	Long ear	HP:0400004
	Low-set ear	HP:0000369
	Increased posterior angulation ear	HP:0000358
	Protruding ear	HP:0000411

Short ear	HP:0400005
Cleft helix	HP:0009902
Crimped helix	HP:0011262
Absent crus helix	HP:0011255
Crus helix connected to antihelix	HP:0011256
Expanded terminal portion crus of helix	HP:0011259
Horizontal crus of helix	HP:0009897
Prominent crus of helix	HP:0009899
Serpiginous crus of helix	HP:0011257
Tragal bridge of crus of helix	HP:0011258
Underdeveloped crus of helix	HP:0009898
Darwin notch of helix	HP:0011260
Darwin tubercle of helix	HP:0011261
Discontinuous ascending root of helix	HP:0011264
Overfolded helix	HP:0000396
Posterior helix pit	HP:0008523
Squared superior portion of helix	HP:0030026
Underfolded helix	HP:0008577
Absent lobe	HP:0000387
Anterior creases of earlobe	HP:0009908
Attached lobe	HP:0009907
Cleft earlobe	HP:0011265
Forward facing earlobe	HP:0011263
Large earlobe	HP:0009748
Small earlobe	HP:0000385
Uplifted earlobe	HP:0009909

	Lop ear	HP:0000394
	Macrotia	HP:0000400
	Microtia, first degree	HP:0011266
	Microtia, second degree	HP:0008569
	Microtia, third degree	HP:0011267
	Auricular pit	HP:0030025
	Preauricular pit	HP:0100277
	Pretragal ectopia	HP:0030024
	Quelprud Nodule	HP:0030023
	Question mark ear	HP:0030022
	Satyr ear	HP:0100015
	Shell ear	HP:0008569
	Stahl ear	HP:0100015
	Auricular tag	HP:0030021
	Preauricular tag	HP:0000384
	Absent tragus	HP:0011268
	Bifid tragus	HP:0011269
	Duplicated tragus	HP:0011270
	Prominent tragus	HP:0011271
	Underdeveloped tragus	HP:0011251
Abnormality of the hands or feet	Adactyly	HP:0009776
	Camptodactyly	HP:0012385
	Clinodactyly	HP:0030084
	Clubbing	HP:0001217
	Prominent digit pad	HP:0011298
	Digital constriction ring	HP:0010491
	Absent finger	HP:0009380
	Broad finger	HP:0001500
	Cutaneous syndactyly of fingers	HP:0010554
	Long fingers	HP:0100807
	Overlapping fingers	HP:0010557

Partial absence of finger	HP:0011299
Radial deviation of finger	HP:0009466
Short finger	HP:0009381
Short distal phalanx of finger	HP:0009882
Slender finger	HP:0001238
Small finger	HP:0030033
Splayed finger	HP:0030029
Tapered finger	HP:0001182
Ulnar deviation of finger	HP:0009465
Broad fingertip	HP:0011300
Absent foot	HP:0011301
Broad foot	HP:0001769
Long foot	HP:0001833
Narrow foot	HP:0001786
Osseous syndactyly of toes	HP:0010717
Partial absence of foot	HP:0030032
Postaxial polydactyly of foot	HP:0001830
Preaxial polydactyly of foot	HP:0001841
Rocker bottom foot	HP:0001838
Short foot	HP:0001773
Split foot	HP:0001839
Absent hallux	HP:0012386
Broad hallux	HP:0010055
Hammertoe	HP:0001765
Absent hand	HP:0004050
Clenched hand	HP:0001188
Osseous syndactly of the fingers	HP:0010492
Postaxial polydactyly of hand	HP:0001162
Preaxial polydactyly	HP:0001177

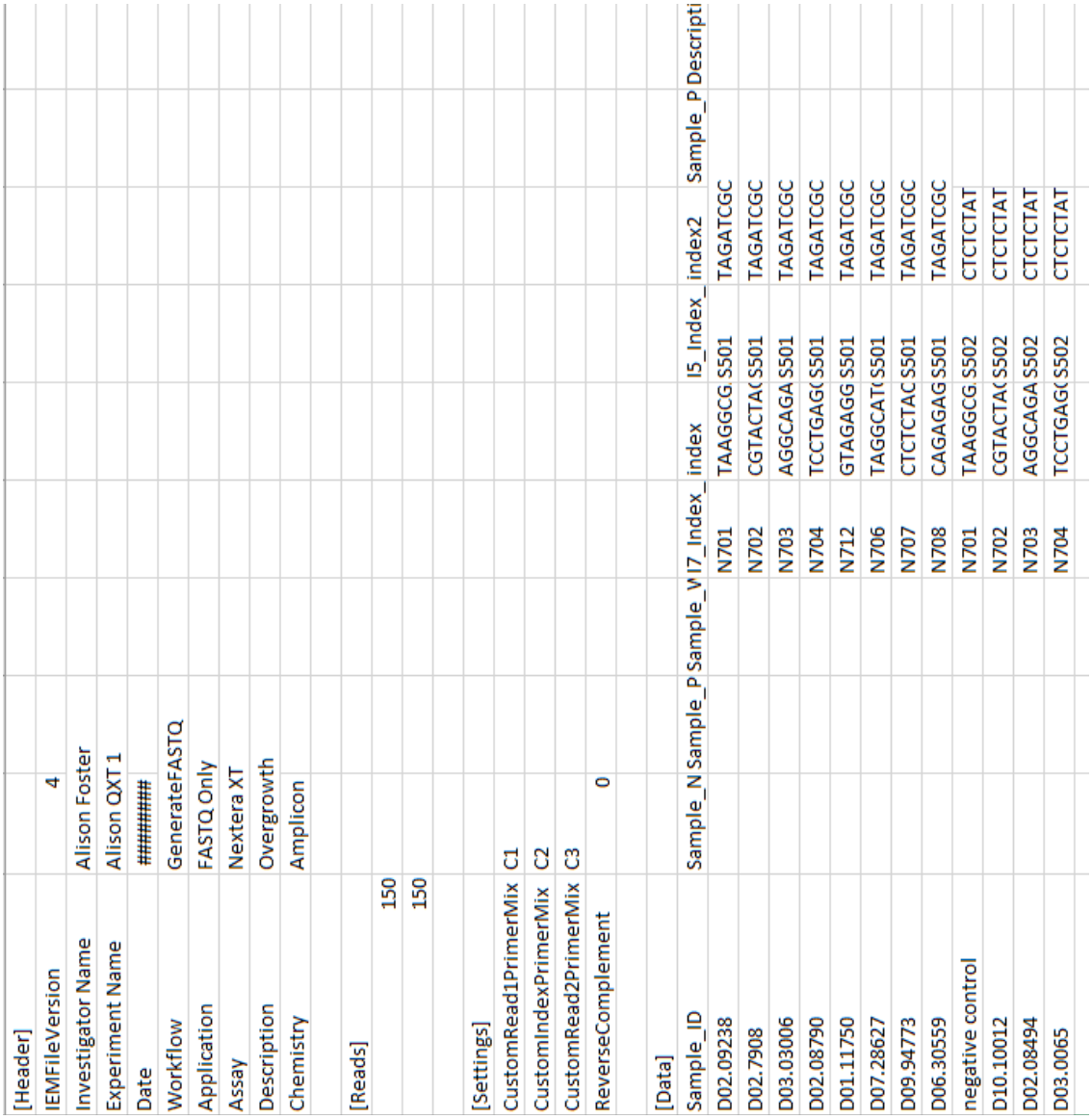
of hand	
Radial deviation of hand	HP:0009486
Small hand	HP:0200055
Split hand	HP:0001171
Trident hand	HP:0004060
Ulnar deviation of hand	HP:0009487
Prominent heel	HP:0012428
Small hypothenar eminence	HP:0010487
Macrodactyly	HP:0004099
Short metacarpal	HP:0010049
Short metatarsal	HP:0010743
Metatarsus adductus	HP:0001840
Broad palm	HP:0001169
Long palm	HP:0011302
Narrow palm	HP:0004283
Short palm	HP:0004279
Pes cavus	HP:0001761
Pes planus	HP:0001763
Mesoaxial polydactyly	HP:0100260
Mirror image polydactyly	HP:0010689
Absent ray	HP:0030030
Sandal gap	HP:0001852
Convex contour of sole	HP:0011303
Small thenar eminence	HP:0001245
Absent thumb	HP:0009777
Adducted thumb	HP:0001181
Broad thumb	HP:0011304
Hitchhiker thumb	HP:0001234
Partial absence of thumb	HP:0009659
Proximal placement	HP:0009623

	of thumb	
	Triphalangeal thumb	HP:0001199
	Absent toe	HP:0010760
	Broad toe	HP:0001837
	Cutaneous syndactyly of toes	HP:0010621
	Long toe	HP:0010511
	Overlapping toes	HP:0001845
	Partial absence of toe	HP:0011305
	Short toe	HP:0001831
	Short distal phalanx of toe	HP:0001857
	Slender toe	HP:0011308
	Small toe	HP:0030031
	Splayed toes	HP:0011307
	Tapered toe	HP:0011309
	Widely spaced toes	HP:0008094
Abnormality of the creases	Absent palmar crease	HP:0010489
	Bridged palmar crease	HP:0011310
	Decreased palmar creases	HP:0006184
	Deep palmar crease	HP:0006191
	Single transverse palmar crease	HP:0000954
	Deep longitudinal palmer crease	HP:0004681
	Sydney crease	HP:0011311
Abnormality of the nails	Bifid nail	HP:0010793
	Concave nail	HP:0001598
	Fused nails	HP:0011312
	Hyperconvex nail	HP:0001795
	Narrow nail	HP:0011313
	Nail pits	HP:0001803
	Ridged nail	HP:0001807
	Short nail	HP:0001799

Small nail	HP:0001792
Split nail	HP:0001809
Thick nail	HP:0001805
Thin nail	HP:0001816

D. Sample sheets

D.1 Example of sample sheet for upload to MiSeq (QXT)



[Header]										
IEMFileVersion	4									
Investigator Name	Alison Foster									
Experiment Name	Alison QXT 1									
Date	########									
Workflow	GenerateFASTQ									
Application	FASTQ Only									
Assay	Nextera XT									
Description	Overgrowth									
Chemistry	Amplicon									
[Reads]										
	150									
	150									
[Settings]										
CustomRead1PrimerMix	C1									
CustomIndexPrimerMix	C2									
CustomRead2PrimerMix	C3									
ReverseComplement	0									
[Data]										
Sample_ID	Sample_N	Sample_P	Sample_W	I7_Index_	index	I5_Index_	index2	Sample_P	Descripti	
D02.09238				N701	TAAGGCG	S501	TAGATCGC			
D02.7908				N702	CGTACTA	S501	TAGATCGC			
D03.03006				N703	AGGCAGA	S501	TAGATCGC			
D02.08790				N704	TCCTGAG	S501	TAGATCGC			
D01.11750				N712	GTAGAGG	S501	TAGATCGC			
D07.28627				N706	TAGGCAT	S501	TAGATCGC			
D09.94773				N707	CTCTCTAC	S501	TAGATCGC			
D06.30559				N708	CAGAGAG	S501	TAGATCGC			
negative control				N701	TAAGGCG	S502	CTCTCTAT			
D10.10012				N702	CGTACTA	S502	CTCTCTAT			
D02.08494				N703	AGGCAGA	S502	CTCTCTAT			
D03.0065				N704	TCCTGAG	S502	CTCTCTAT			

D.2 Example of sample sheet for upload to MiSeq (TSCA)

ɔn	4										
Jame	AF										
lame	Run1										
	#########										
	Custom Amplicon										
	TruSeq Amplicon										
	TruSeq Amplicon										
	Overgrowth										
	Amplicon										
	TruSeq CAT Manifest TC0069601-CAT.txt										
150											
150											
Quality	30										
ıevcf	FALSE										
	Sample_N	Sample_P	Sample_W	I7_Index_	index	I5_Index_	index2	Manifest	GenomeF	Sample_P	Description
238		Run1	A01	A702	ACAGTGG	A503	TGTTCTCT	A	Homo_sapiens\UCSC\hg19\Sequence\Whol		
908		Run1	A02	A710	TGTGACC/	A503	TGTTCTCT	A	Homo_sapiens\UCSC\hg19\Sequence\Whol		
906		Run1	A03	A711	AGGGTCA	A503	TGTTCTCT	A	Homo_sapiens\UCSC\hg19\Sequence\Whol		
790		Run1	A04	A712	AGGAGTG	A503	TGTTCTCT	A	Homo_sapiens\UCSC\hg19\Sequence\Whol		
750		Run1	B01	A702	ACAGTGG	A505	CTAATCG/	A	Homo_sapiens\UCSC\hg19\Sequence\Whol		
527		Run1	B02	A710	TGTGACC/	A505	CTAATCG/	A	Homo_sapiens\UCSC\hg19\Sequence\Whol		
248		Run1	B03	A711	AGGGTCA	A505	CTAATCG/	A	Homo_sapiens\UCSC\hg19\Sequence\Whol		
773		Run1	B04	A712	AGGAGTG	A505	CTAATCG/	A	Homo_sapiens\UCSC\hg19\Sequence\Whol		
389		Run1	C01	A702	ACAGTGG	A506	CTAGAAC	A	Homo_sapiens\UCSC\hg19\Sequence\Whol		
559		Run1	C02	A710	TGTGACC/	A506	CTAGAAC	A	Homo_sapiens\UCSC\hg19\Sequence\Whol		
244		Run1	C03	A711	AGGGTCA	A506	CTAGAAC	A	Homo_sapiens\UCSC\hg19\Sequence\Whol		
012		Run1	C04	A712	AGGAGTG	A506	CTAGAAC	A	Homo_sapiens\UCSC\hg19\Sequence\Whol		
547		Run1	D02	A710	TGTGACC/	A508	TAGACCT/	A	Homo_sapiens\UCSC\hg19\Sequence\Whol		
494		Run1	D03	A711	AGGGTCA	A508	TAGACCT/	A	Homo_sapiens\UCSC\hg19\Sequence\Whol		
re__		Run1	D08	A706	AACCCCT(	A508	TAGACCT/	A	Homo_sapiens\UCSC\hg19\Sequence\Whol		

E. Lists of variants

E.1 Example of Agilent SureCall test sample 14

117	0.998	565	39	INTRON(M	INTRON|NM_001195573				CODING	NM_001195	
189	0.106	122	39	INTRON(M	INTRON|NM_001195573				CODING	NM_001195	
117	0.995	214	39	INTRON(M	INTRON|NM_001195573				CODING	NM_001195	
255	0.504	482	39	INTRON(M	INTRON|NM_001014431				CODING	NM_001014	
255	0.465	554	39	INTRON(M	INTRON|NM_001014431				CODING	NM_001014	
255	0.5	975	38	INTRON(M	INTRON|NM_001014431				CODING	NM_001014	
255	0.515	997	39	INTRON(M	INTRON|NM_001014431				CODING	NM_001014	
255	0.481	681	39	SYNONYM	SYNONYM	SILENT	gaG/gaA	E242	CODING	NM_0010	NI
255	0.514	515	39	INTRON(M	INTRON|NM_001014431				CODING	NM_001014	
255	0.451	223	39	INTRON(M	INTRON|NM_001014431				CODING	NM_001014	
255	0.503	746	39	INTRON(M	INTRON|NM_001014431				CODING	NM_001014	
255	0.994	678	39	INTRON(M	INTRON|NM_001014431				CODING	NM_001014	
117	0.996	691	39	INTRON(M	INTRON|NM_001014431				CODING	NM_001014	
255	0.994	349	39	INTRON(M	INTRON|NM_017831				CODING	NM_017831	
255	0.491	698	39	NON_SYN(	NON_SYN(	MISSENSE	cGg/cAg	R165C	CODING	NM_0178:	NI
114	0.98	48	38	NON_SYN(	NON_SYN(	MISSENSE	Agc/Cgc	S234R	CODING	NM_0050:	NI
255	0.996	838	39	NON_SYN(	NON_SYN(	MISSENSE	Tcc/Ccc	S313P	CODING	NM_0050:	NI
117	0.998	444	39	INTRON(M	INTRON|NM_005027				CODING	NM_005027	
117	0.999	862	39	INTRON(M	INTRON|NM_005027				CODING	NM_005027	
255	0.519	927	39	SYNONYM	SYNONYM	SILENT	atC/atA	I556	CODING	NM_0050:	NI
255	0.503	947	39	SYNONYM	SYNONYM	SILENT	agT/agC	S637	CODING	NM_0050:	NI
255	0.476	642	39	SYNONYM	SYNONYM	SILENT	acC/acT	T709	CODING	NM_0050:	NI
255	0.521	817	39	UTR_3_PR	UTR_3_PRIME|NM_001243027				CODING	NM_001243	
255	0.469	786	39	UTR_3_PR	UTR_3_PRIME|NM_001243027				CODING	NM_001243	
255	0.485	883	39	UTR_3_PR	UTR_3_PRIME|NM_001243027				CODING	NM_001243	
255	0.522	988	39	INTRON(M	INTRON|NM_001243027				CODING	NM_001243	
255	0.996	1118	39	INTRON(M	INTRON|NM_001243027				CODING	NM_001243	
255	0.47	304	39	INTRON(M	INTRON|NM_001243027				CODING	NM_001243	
114	0.997	299	38	INTRON(M	INTRON|NM_001243027				CODING	NM_001243	
255	0.461	450	39	INTRON(M	INTRON|NM_001164617				CODING	NM_001164	

I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y
Pos	Ref /	Alt A	Quality	Allele Frec	Read Dept	Ma	Effect	Primary El	Function C	Codon	AA	Coding	Transcript	Exon ID	SIFT	PolyF
1.1E+07	TTA	T	255	0.464	407	40	INTRON(M	INTRON|NM_004958				CODING	NM_004958			
1.1E+07	C	T	255	0.494	1046	39	SYNONYM	SYNONYM	SILENT	ctG/ctA	L2303	CODING	NM_0049!	NM_0049!		
1.1E+07	T	C	117	0.999	1044	39	INTRON(M	INTRON|NM_004958				CODING	NM_004958			
1.1E+07	G	A	255	0.48	841	39	SYNONYM	SYNONYM	SILENT	agC/agT	S1851	CODING	NM_0049!	NM_0049!		
1.1E+07	C	T	255	0.493	944	39	SYNONYM	SYNONYM	SILENT	gcG/gcA	A1577	CODING	NM_0049!	NM_0049!		
1.1E+07	G	A	255	0.525	652	39	INTRON(M	INTRON|NM_004958				CODING	NM_004958			
1.1E+07	G	C	255	0.527	760	39	INTRON(M	INTRON|NM_004958				CODING	NM_004958			
1.1E+07	G	A	255	0.506	1193	39	SYNONYM	SYNONYM	SILENT	aaC/aaT	N999	CODING	NM_0049!	NM_0049!		
1.1E+07	T	A	255	0.498	433	39	INTRON(M	INTRON|NM_004958				CODING	NM_004958			
1.1E+07	A	G	255	0.467	166	39	INTRON(M	INTRON|NM_004958				CODING	NM_004958			
1.1E+07	T	A	255	0.507	407	39	INTRON(M	INTRON|NM_004958				CODING	NM_004958			
1.1E+07	T	C	255	0.447	509	39	INTRON(M	INTRON|NM_004958				CODING	NM_004958			
1.1E+07	T	C	255	0.505	197	39	INTRON(M	INTRON|NM_004958				CODING	NM_004958			
1.1E+07	A	G	255	0.504	951	39	SYNONYM	SYNONYM	SILENT	gaT/gaC	D479	CODING	NM_0049!	NM_0049!		
1.1E+07	C	T	255	0.541	402	39	INTRON(M	INTRON|NM_004958				CODING	NM_004958			
1.1E+07	C	G	255	0.51	680	40	INTRON(M	INTRON|NM_004958				CODING	NM_004958			
1.1E+07	A	G	255	0.487	303	39	INTRON(M	INTRON|NM_004958				CODING	NM_004958			
2.4E+08	TACT	T	255	0.568	512	40	INTRON(M	INTRON|NM_001206729				CODING	NM_001206729			
2.4E+08	G	A	255	0.463	228	40	INTRON(M	INTRON|NM_001206729				CODING	NM_001206729			
2.5E+07	G	T	255	0.488	815	39	INTRON(M	INTRON|NM_022552				CODING	NM_022552			
2.5E+07	C	T	255	0.477	853	39	SYNONYM	SYNONYM	SILENT	ctG/ctA	L233	CODING	NM_1537!	NM_1537!		
2.5E+07	C	T	255	0.496	933	39	INTRON(M	INTRON|NM_022552				CODING	NM_022552			
2.5E+07	C	T	255	0.481	269	39	INTRON(M	INTRON|NM_022552				CODING	NM_022552			
2.3E+08	T	G	120	0.987	76	40	INTRON(M	INTRON|NM_001257281				CODING	NM_001257281			
2.3E+08	CTT	C	255	0.521	624	40	INTRON(M	INTRON|NM_001257281				CODING	NM_001257281			
2.3E+08	A	G	255	0.521	1002	40	SYNONYM	SYNONYM	SILENT	caA/caG	Q152	CODING	NM_0012!	NM_0012!		
2.3E+08	A	G	255	0.49	415	39	INTRON(M	INTRON|NM_001257281				CODING	NM_001257281			
2.3E+08	C	T	255	0.48	559	39	INTRON(M	INTRON|NM_001257281				CODING	NM_001257281			
2.3E+08	G	A	255	0.525	891	39	INTRON(M	INTRON|NM_001257281				CODING	NM_001257281			
2.3E+08	A	G	255	0.495	571	39	INTRON(M	SYNONYM	SILENT	ctA/ctG	L981	CODING	NR_04647	NR_04647		
2.3E+08	A	G	255	0.469	1089	39	DOWNSTR	INTRON|NM_001257281				CODING	NM_001257281			
2.3E+08	C	T	255	0.534	312	39	DOWNSTR	INTRON|NM_001257281				CODING	NM_001257281			
2.3E+08	A	G	255	0.451	799	39	DOWNSTR	INTRON|NM_001257281				CODING	NM_001257281			
2.3E+08	C	G	255	0.523	834	39	INTRON(M	INTRON|NM_001257281				CODING	NM_001257281			
4.7E+07	C	T	255	0.5	445	39	DOWNSTR	INTRON|NM_014159				CODING	NM_014159			
4.7E+07	G	A	255	0.516	544	39	INTRON(M	INTRON|NM_014159				CODING	NM_014159			
4.7E+07	T	A	255	0.333	134	39	INTRON(M	INTRON|NM_014159				CODING	NM_014159			
4.7E+07	TTGT	T	255	0.528	519	39	INTRON(M	INTRON|NM_014159				CODING	NM_014159			
4.7E+07	G	A	255	0.513	983	40	NON_SYNC	NON_SYN(	MISSENSE	cCc/cTc	P1962	CODING	NM_0141!	NM_0141!	0.6	B
4.7E+07	ACTT	A	255	0.529	461	39	INTRON(M	INTRON|NM_014159				CODING	NM_014159			
4.7E+07	G	A	255	0.513	983	40	NON_SYNC	NON_SYN(	MISSENSE	cCc/cTc	P1962	CODING	NM_0141!	NM_0141!	0.6	B
4.7E+07	ACTT	A	255	0.529	461	39	INTRON(M	INTRON|NM_014159				CODING	NM_014159			
4.7E+07	C	T	120	0.993	140	40	INTRON(M	INTRON|NM_014159				CODING	NM_014159			
4.7E+07	TCC	T	255	0.451	264	39	INTRON(M	INTRON|NM_014159				CODING	NM_014159			
4.7E+07	AGT(	A	255	0.379	359	39	INTRON(M	INTRON|NM_014159				CODING	NM_014159			
4.7E+07	CCA/	C	255	0.731	525	39	INTRON(M	INTRON|NM_014159				CODING	NM_014159			
4.7E+07	T	C	255	0.512	325	39	INTRON(M	INTRON|NM_014159				CODING	NM_014159			
4.7E+07	A	G	255	0.466	928	39	SYNONYM	SYNONYM	SILENT	aaT/aaC	N1155	CODING	NM_0141!	NM_0141!		
4.7E+07	G	C	255	0.438	136	39	INTRON(M	INTRON|NM_014159				CODING	NM_014159			
4.7E+07	C	G	255	0.49	595	39	INTRON(M	INTRON|NM_014159				CODING	NM_014159			
4.7E+07	G	A	255	0.459	171	39	INTRON(M	INTRON|NM_014159				CODING	NM_014159			
4.7E+07	CGC(	C	255	0.683	249	38	UPSTREAM	UPSTREAM|NR_033373				NON_COD	NR_033373			
1.8E+08	C	A	255	0.51	611	40	INTRON(M	INTRON|NM_006218				CODING	NM_006218			
1.8E+08	C	G	255	0.377	158	40	INTRON(M	INTRON|NM_006218				CODING	NM_006218			
1.8E+08	C	G	255	0.514	349	39	INTRON(M	INTRON|NM_022455				CODING	NM_022455			
1.8E+08	TTCT	T	255	0.47	366	39	INTRON(M	INTRON|NM_022455				CODING	NM_022455			
1.8E+08	C	T	187	0.101	226	39	INTRON(M	INTRON|NM_022455				CODING	NM_022455			
1.8E+08	ACTT	A	255	0.39	123	39	INTRON(M	INTRON|NM_022455				CODING	NM_022455			
1.8E+08	G	A	117	0.998	429	39	INTRON(M	INTRON|NM_022455				CODING	NM_022455			
1.8E+08	T	C	255	0.995	1045	39	SYNONYM	SYNONYM	SILENT	Ttg/Ctg	L2008	CODING	NM_1723	NM_1723		
1.8E+08	G	C	255	0.995	1094	39	SYNONYM	SYNONYM	SILENT	ggG/ggC	G2032	CODING	NM_1723	NM_1723		
1.5E+08	AG	A	255	0.478	563	38	UTR_3_PR	UTR_3_PRIME|NM_001203247				CODING	NM_001203247			
1.5E+08	T	G	255	0.482	959	39	INTRON(M	INTRON|NM_001203247				CODING	NM_001203247			
1.5E+08	C	T	255	0.484	549	39	INTRON(M	INTRON|NM_001203247				CODING	NM_001203247			
1.5E+08	A	G	255	0.544	460	39	INTRON(M	INTRON|NM_001203247				CODING	NM_001203247			
1.5E+08	TAA	T	255	0.501	732	39	INTRON(M	INTRON|NM_001203247				CODING	NM_001203247			
3.6E+07	T	C	255	0.496	733	39	SYNONYM	SYNONYM	SILENT	ctT/ctC	L6	CODING	NM_0039	NM_0039		
3.6E+07	T	C	255	0.501	798	39	INTRON(M	INTRON|NM_003995				CODING	NM_003995			
3.6E+07	C	A	255	0.433	977	40	INTRON(M	INTRON|NM_003995				CODING	NM_003995			
3.6E+07	G	C	255	0.529	700	39	INTRON(M	INTRON|NM_003995				CODING	NM_003995			
3.6E+07	C	T	255	0.474	876	39	SYNONYM	SYNONYM	SILENT	taC/taT	Y598	CODING	NM_0039	NM_0039		
3.6E+07	C	A	255	0.417	241	39	DOWNSTR	INTRON|NM_003995				CODING	NM_003995			
3.6E+07	AGC'	A	255	0.428	638	38	DOWNSTR	INTRON|NM_003995				CODING	NM_003995			
3.6E+07	C	T	255	0.396	239	39	DOWNSTR	INTRON|NM_003995				CODING	NM_003995			
3.6E+07	C	G	255	0.473	200	40	DOWNSTR	INTRON|NM_003995				CODING	NM_003995			
3.6E+07	A	G	255	0.477	1016	39	DOWNSTR	INTRON|NM_003995				CODING	NM_003995			
9E+07	T	G	255	0.411	274	40	NON_SYNC	NON_SYN(	MISSENSE	Tac/Gac	Y68D	CODING	NM_0003	NM_0003		D
906196	ACC(	A	255	0.88	75	38	CODON_C	CODON_CHANGE_PLU		gctccggtcg	APVAV	CODING	NM_0011	NM_0011		
9.6E+07	A	T	255	0.38	423	40	UTR_3_PR	UTR_3_PRIME|NM_001195573				CODING	NM_001195573			
9.6E+07	TAC/	T	255	0.503	147	39	INTRON(M	INTRON|NM_001195573				CODING	NM_001195573			
9.6E+07	C	A	255	0.517	285	39	INTRON(M	INTRON|NM_001195573				CODING	NM_001195573			

E.2 Example of Illumina VariantStudio test sample 14

	32406.01	99.6	812	809	NM_0224!	synonymc	6967	2277	L	Ttg/Ctg		NP_0
	22212.01	52.7	1227	641	NM_0044!	intron_variant		0				NP_0
	2646.97	33.3	703	97	NM_0044!	intron_variant		0				NP_0
	38132.01	52	2125	1103	NM_0044!	intron_variant		0				NP_0
	2447.01	48	150	72	NM_0039!	synonymc	18	6	L	ctT/ctC		NP_0
	28172.01	50.4	1600	806	NM_0039!	intron_variant		0				NP_0
	55970.01	49.5	3262	1613	NM_0039!	intron_variant		0				NP_0
	13806.01	45.5	899	407	NM_0039!	intron_variant		0				NP_0
	45982.01	54.6	2688	1463	NM_0039!	synonymc	1794	598	Y	taC/taT		NP_0
	139.01	42.8	744	318	NM_1723!	downstream_gene_		0				NP_7
	24.02	28.3	778	220	NM_1723!	downstream_gene_		0				NP_7
	54.01	26.5	778	206	NM_1723!	downstream_gene_		0				NP_7
'ariar	129.01	9.2	769	70	NM_1723!	downstream_gene_		0				NP_7
	16.11	21.5	778	167	NM_1723!	downstream_gene_		0				NP_7
	896.01	69.4	36	25	NM_0039!	intron_variant		0				NP_0
	23340.01	48.7	1385	673	NM_0003!	missense_	1234	68	Y/D	Tac/Gac	delete proba	NP_0
	5163	31.3	1457	303	NM_0003!	intron_variant, featu		0				NP_0
	10032	22.6	3127	597	NM_0003!	3_prime_l	2309	0				NP_0
	2776.96	95	20	19	NM_0000!	inframe_c	788-799		APV	gCTCCGGTCGCGGtc/gtc		NP_0
'ariar	476.97	0.5	251	1	NM_0306!	intron_variant, featu		0				NP_0
'ariar	7579	0	251	0	NM_0306!	intron_variant, featu		0				NP_0
	3425	53	251	133	NM_0306!	intron_variant, featu		0				NP_0
	11798	31.4	2683	647	NM_0306!	intron_variant, featu		0				NP_0
'ariar	1253.97	16.2	698	103	NM_0306!	splice_region_variar		0				NP_0
	2490.97	32.6	431	100	NM_0306!	intron_variant, featu		0				NP_0
	2647.97	43.3	431	122	NM_0306!	intron_variant, featu		0				NP_0
	436.01	39.1	1187	456	NM_0010!	intron_variant		0				NP_0
	3451.01	74.9	1192	881	NM_0010!	intron_variant		0				NP_0
	3731.01	72.7	1192	853	NM_0010!	intron_variant		0				NP_0
	16700.01	43.3	1194	515	NM_0010!	intron_variant		0				NP_0
	10679.01	45.5	696	316	NM_0010!	intron_variant		0				NP_0
	9263.01	44.4	658	291	NM_0010!	intron_variant		0				NP_0
	7796.01	99	198	194	NM_0010!	intron_variant		0				NP_0
	7918.01	99.5	198	196	NM_0010!	intron_variant		0				NP_0
iQ;Lo	11.34	100	1	1	NM_0323!	intron_variant		0				NP_1
	4706	22.8	1300	258	NM_0323!	intron_variant, featu		0				NP_1

1113.97	19.8	182	35	NM_0178!	intron_variant, featu		0			
14.18	12.8	180	23	NM_0178!	intron_variant		0			
1203.94	20.2	182	4	NM_0178!	intron_variant		0			
1203.94	20.2	182	5	NM_0178!	intron_variant		0			
53645.01	53.4	2845	1517	NM_0178!	missense_	876	165	R/Q	cGg/cAg	tolerat be
206.97	76.9	21	10	NM_0178!	intron_variant, featu		0			
1076.96	100	25	25	NM_0050!	splice_acc	?-1111				
118.96	100	25	25	NM_0050!	frameshif	1116	202			
769	100	25	25	NM_0050!	synonymc	1118	202	A	gcG/gcT	
95.96	100	25	25	NM_0050!	frameshif	1122-1123	204			
6698.01	100	169	168	NM_0050!	missense_	1449	313	S/P	Tcc/Ccc	tolerat be
74726.01	99.8	1830	1824	NM_0050!	intron_variant		0			
31.01	19.6	97	19	NM_0050!	synonymc	2099	529	K	aaG/aaA	
68.01	8.4	97	8	NM_0050!	missense_	2112	534	E/Q	Gag/Cag	delete pc
14971.01	49.8	867	430	NM_0050!	synonymc	2423	637	S	agT/agC	
9507.01	51.6	538	277	NM_0063!	upstream_gene_vari		0			
13471.01	52.6	740	389	NM_0016!	3_prime_l	1754	0			
19505.01	47.1	1228	578	NM_0016!	intron_variant		0			
18928	47.7	1148	547	NM_0016!	intron_variant, featu		0			
72.01	56.7	93	51	NM_0016!	missense_	960	221	T/P	Acc/Ccc	delete pr
8291	13.2	646	48	NM_0016!	intron_variant, featu		0			
30201.01	47.8	1817	867							
75048.01	99.9	1832	1830							
24778.01	52.4	1340	702							
17417.01	45	1144	514	NM_0011!	intron_variant		0			
1013.01	17.1	1528	241	NM_0011!	missense_	1034	279	N/K	aaT/aaA	tolerat be
3675.01	24.2	1639	366	NM_0011!	missense_	1032	279	N/H	Aat/Cat	tolerat be
1153.01	15.9	1662	253	NM_0011!	missense_	1030	278	C/F	tGc/tTc	delete pr
1851.01	37.2	1678	605	NM_0011!	missense_	1027	277	Y/F	tAc/tTc	tolerat be
4763	30.6	1170	247	NM_0011!	splice_region_variar		0			

G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
nic	Filters	Quality	Alt Varian	Read Dep	Alt Read D	Transcript	Conseque	cDNA Posi	Protein Pc	Ami	Codons	Sift	PolyF	ENSP
	PASS	59485	51.3	2468	1267	NM_0049	intron_variant, featu		0					NP_0
	PASS	24124.01	56.1	1225	686	NM_0049	synonymc	5674	1851	S	agC/agT			NP_0
	PASS	35717.01	50.3	2041	1027	NM_0049	synonymc	4852	1577	A	gcG/gcA			NP_0
	PASS	4504.01	45	291	131	NM_0049	synonymc	3118	999	N	aaC/aaT			NP_0
	PASS	13666.01	46.5	877	406	NM_0049	intron_variant		0					NP_0
	PASS	20889.01	50.3	1205	600	NM_0049	intron_variant		0					NP_0
	PASS	30845.01	52.9	1674	885	NM_0049	intron_variant		0					NP_0
	PASS	19422.01	44.4	1282	568	NM_0049	intron_variant		0					NP_0
	PASS	46647.01	45.8	2969	1356	NM_0049	intron_variant		0					NP_0
	PASS	14965.01	53.2	817	434	NM_0049	intron_variant		0					NP_0
	PASS	107399	99.9	2627	2619	NM_0054	intron_variant		0					NP_0
	R8	2476.97	50.2	338	108	NM_0054	intron_variant, featu		0					NP_0
	PASS	19571.01	47.9	1193	572	NM_0054	intron_variant		0					NP_0
	PASS	19629.01	99.8	507	505	NM_0054	intron_variant		0					NP_0
	PASS	12848.01	50.8	733	372	NM_0225	synonymc	1533	422	L	ctG/ctA			NP_0
	PASS	25182.01	52.9	1390	735	NM_0225	splice_region_variar		0					NP_0
	PASS	24140.01	49.1	1421	697	NM_0225	intron_variant		0					NP_0
	LowVariar	363.01	7.9	747	59	NM_0225	intron_variant		0					NP_0
	PASS	53.01	60.3	747	450	NM_0225	intron_variant		0					NP_0
	R8	11240	24.1	3495	687	NM_1523	intron_variant, featu		0					NP_6
	PASS	7218.01	56.5	375	212	NM_1523	intron_variant		0					NP_6
	PASS	26896.01	48.6	1587	771	NM_1523	intron_variant		0					NP_6
	LowGQ	126.99	47.2	53	25	NM_1523	3_prime_l	2958-2960	0					NP_6
	LowGQ;Lo	50.76	100	2	2									
	PASS	559.01	69.6	23	16	NM_0141	intron_variant		0					NP_0
	R8	236.97	48.1	44	13	NM_0141	intron_variant, featu		0					NP_0
	PASS	4479.01	100	110	110	NM_0141	intron_variant		0					NP_0
	PASS	34159	53.6	1408	752	NM_0141	intron_variant, featu		0					NP_0
	PASS	11.34	38.2	220	84	NM_0141	synonymc	620	189	S	tcT/tcC			NP_0
	PASS	205.01	52.1	220	114	NM_0141	synonymc	608	185	S	tcA/tcC			NP_0
	R8	7911	30.8	1706	409	NM_0062	intron_variant, featu		0					NP_0
	LowVariar	238.01	14.4	90	13	NM_0062	splice_reg	2068	637	Q	caG/caA			NP_0
	LowVariar	6775	18.2	3139	488	NM_0062	intron_variant, featu		0					NP_0
	PASS	7133.01	49.7	433	214	NM_0062	intron_variant		0					NP_0
	PASS	278471	97.6	3617	3401	NM_0062	intron_variant, featu		0					NP_0
	PASS	30694.01	46.1	1967	906	NM_0224	intron_variant		0					NP_0
	PASS	5841.01	99.3	145	144	NM_0224	intron_variant		0					NP_0

F. Variant confirmation form

<table>
<tr><th colspan="2">Variant Confirmation Form</th></tr>
<tr><td>Lab Number</td><td></td></tr>
<tr><td>Name/ DOB</td><td></td></tr>
<tr><td>Referring Clinician</td><td></td></tr>
<tr><td>Chromosome/genomic location</td><td></td></tr>
<tr><td>HGVS: c.DNA and protein</td><td></td></tr>
<tr><td>NM number/transcript</td><td></td></tr>
<tr><td>Gene/exon</td><td></td></tr>
<tr><td>Inheritance pattern</td><td></td></tr>
<tr><td>Any other key variants present</td><td></td></tr>
<tr><td>Key referral information</td><td></td></tr>
<tr><td>Summary of findings and Variant Classification</td><td>
<table>
<tr><th>Class</th><th>Description</th><th>Class</th></tr>
<tr><td>1</td><td>Clearly not pathogenic</td><td></td></tr>
<tr><td>2</td><td>Unlikely to be pathogenic</td><td></td></tr>
<tr><td>3</td><td>Unknown significance</td><td></td></tr>
<tr><td>4</td><td>Likely to be pathogenic</td><td></td></tr>
<tr><td>5</td><td>Clearly pathogenic</td><td></td></tr>
</table>
</td></tr>
<tr><td>Variant form created by/date</td><td></td></tr>
<tr><td>Variant form checked by/date</td><td></td></tr>
</table>

G. Summary of participants

n	POD	Sex	Age	Dev delay/ID	Height SD	Weight SD	OFC SD
1	1.0	F	13	no	3.6	3.2	0.5
2	2.0	F	10	yes	1.3	0.8	2
3	3.0	M	11	no	3.5	3.7	0.6
4	4.0	F	4	yes	3.7	4.8	1.7
5	5.0	F	10	no	5.5	3.2	
6	6.0	M	13	yes	3	3.8	1.1
7	8.0	M	15	no	3.9	1.6	-0.4
8	9.0	M	4	yes	3	0.7	-0.2
9	10.0	M	14	yes	1.7	3.4	3.9
10	11.0	F	18	yes	2.8	2	2
11	12.0	M	11	yes	0.7	1	1.4
12	13.0	F	7	no	-0.6	-1.4	-0.6
13	14.0	M	6	yes	3.1	4	0.9
14	15.0	F	2	yes	-0.7	0.1	3.2
15	16.0	M	6	yes	1.3	2.4	2.5
16	17.0	F	5	yes	1.9	1.7	1.2
17	18.0	M	9	no	4.5	3.1	1.4
18	19.0	F	3	yes	2.2	4.1	3.7
19	20.0	M	4	no	5.4	5	3.6
20	21.0	M	13	no	3.9	2.6	2.3
21	22.0	F	10	no	5	4	
22	28.0	M	4	yes	4.5	3.5	2.2
23	29.0	M	6	no	4.5	2.5	2.2
24	30.0	M	7	yes	0.8	1	3.4
25	31.0	F	47	yes	0.3	4	2.5
26	32.0	F	4	yes	4.3	0.8	-0.2
27	33.0	M	3	yes	3.3	4	5.3
28	34.0	M	4	yes	-0.4	4.8	1.6
29	35.0	M	10	yes	2.7	3.3	3.2

30	36.0	M	8	yes	2.8	3	2.1
31	37.0	F	8	yes	3.4	3.1	0.7
32	38.0	M	5	yes	1.3	2.2	1.2
33	39.0	M	6	yes	2.9	2.6	0.5
34	40.0	M	5	yes	1.9	2.3	1.3
35	41.0	M	9	no	4.6	4.3	3.2
36	42.0	F	2	no	0	-0.7	-1.6
37	43.0	M	3	yes	3	3.2	0.6
38	45.0	M	2	no	2.5	1.6	0.4
39	46.0	M	8	yes	7	4.2	
40	47.0	F	15	yes	1.9	3.7	6.2
41	48.0	M	12	yes	-0.9	0	2.5
42	49.0	M	1	yes	1.8	0.9	0.2
43	50.0	M	13	yes	1.9	2.7	2.9
44	51.0	M	10	yes	2.4	2.4	2.7
45	52.0	F	5	yes	0.4	3.7	2.6
46	53.0	M	3	yes	-0.4	0.4	1.9
47	54.0	F	2	yes	3.5	2.7	1.2
48	55.0	M	18	yes	1.8	4.5	3.6
49	56.0	F	2	yes	2.9	2	1.5
50	57.0	M	4	yes	2.1	0.2	1.3
51	60.0	M	11	yes	1.6	3	2.6
52	61.0	M	1	no	1.4	0.7	0.4
53	62.0	M	5	yes	3	2.4	1
54	63.0	M	8	yes	0.2	0.3	4.1
55	64.0	M	6	yes	2.8	1.7	-1.5
56	65.0	F	15	yes	3.1	1.4	5
57	66.0	F	25	yes	0.5	3.2	3.2
58	67.0	M	4	yes	-3.4	-0.7	1.2
59	68.0	F	11	yes	5	3.7	0.5
60	69.0	F	12	yes	0.9	3.3	1.5
61	70.0	F	10	yes	-0.3	-0.8	4.4

62	71.0	M	12	yes	2.8	3.5	1.3
63	72.0	M	4	yes	1.7	2.4	0.5
64	73.0	M	11	yes	3.3	2.9	1.6
65	74.0	F	4	yes	1.3	-0.7	2
66	75.0	M	14	yes	2.9	3.8	2.5
67	76.0	M	5	yes	2.3	2.8	0
68	77.0	M	11	yes	2.6	2.3	1.6
69	78.0	M	78	no	3.4	2.1	0.5
70	79.0	M	2	yes	2.8	5	0.3
71	80.0	M	12	yes	3.9	2.4	2
72	81.0	M	81	yes	1.1	1.5	4
73	82.0	F	9	yes	3.9	3.4	0.8
74	83.0	M	83	yes	2.4	2.4	0.6
75	84.0	M	6	yes	2.7	5.4	1.8
76	85.0	M	3	yes	2.5	1.6	0.8
77	87.0	F	55	no	-1.1	3.2	2.5
78	88.0	M	13	yes	1.2	2.2	
79	88.2	M	88		0.4		1.6
80	89.0	M	13	yes	-0.3	1.3	2.6
81	89.1	F	36	no	1	6.4	6.5
82	90.0	M	12	yes	-0.5	2.5	2.1
83	92.0	M	5	yes	2.3	3.6	2.1
84	93.0	F	22	yes	1.3	3.1	0.4
85	94.0	M	3	no	-0.5	0.7	0.1
86	95.0	F	1	yes	3.4	0.8	2.6
87	96.0	M	9	yes	1.7	3	-1.2
88	97.0	M	1	yes	2	4.9	0.7
89	98.0	M	4	yes	2.2	1.9	0.5
90	99.0	F	17	no	2.1	1.3	-0.3
91	99.3	M	14	yes	3.7	4	3.6
92	100.0	M	17	yes	2.7	2.3	0.1
93	101.0	F	5	yes	2.1	1.1	1

94	103.0	M	11	yes	2.1	1.7	0.2
95	103.1	F	39	no	1.4	4.7	0
96	103.3	M	6	yes	2.4	1.5	-1.8
97	104.0	M	25	yes	2.4	3.6	0.4
98	104.1	F	45	yes	1.7	4.5	-0.3
99	104.3	F	12	yes	2.8	2.4	1.5
100	104.4	F	14	yes	3.3	3	1.3

H. Output arising from the study

H.1 Publications

Book chapter:

Cole, T.R.P. and **Foster, A.C.** (2021). 'Sotos Syndrome' in Cassidy, S.B. and Allanson, J.E. (eds.) *Management of Genetic Syndromes*, 4th Edition. Wiley-Blackwell. 2020.

Journal articles:

Mulder PA, van Balkom IDC, Landlust AM, Priolo M, Menke LA, Acero IH, Alkuraya FS, Arias P, Bernardini L, Bijlsma EK, Cole T, Coubes C, Dapia I, Davies S, Di Donato N, Elcioglu NH, Fahrner JA, **Foster A**, González NG, Huber I, Iascone M, Kaiser AS, Kamath A, Kooblall K, Lapunzina P, Liebelt J, Lynch SA, Maas SM, Mammì C, Mathijssen IB, McKee S, Mirzaa GM, Montgomery T, Neubauer D, Neumann TE, Pintomalli L, Pisanti MA, Plomp AS, Price S, Salter C, Santos-Simarro F, Sarda P, Schanze D, Segovia M, Shaw-Smith C, Smithson S, Suri M, Tatton-Brown K, Tenorio J, Thakker RV, Valdez RM, Van Haeringen A, Van Hagen JM, Zenker M, Zollino M, Dunn WW, Piening S, Hennekam RC. Development, behaviour and sensory processing in Marshall-Smith syndrome and Malan syndrome: phenotype comparison in two related syndromes. *J Intellect Disabil Res*. 2020

Walker H, **Foster A**, Cole T, Jester, A. Carpal tunnel syndrome in paediatric patients: A novel association with Kosaki overgrowth syndrome. *JPRAS Open*. 2020

Foster A, Chalot B, Antoniadi T, Schaefer E, Keelagher R, Ryan G, Thomas Q, Philippe C, Bruel AL, Sorlin A, Thauvin-Robinet C, Bardou M, Luu M, Quenardelle V, Wolff V, Woodley J, Vabres P, Lim D, Igbokwe R, Joseph A, Walker H, Jester A, Ellenbogen J, Johnson D, Rooke B, Moss C, Cole T, Faivre L. Kosaki overgrowth syndrome: a novel pathogenic variant in *PDGFRB* and expansion of the phenotype including cerebrovascular complications. *Clinical Genetics*. 2020.

Ostrowski PJ, Zachariou A, Loveday C, Beleza-Meireles A, Bertoli M, Dean J, Douglas AGL, Ellis I, **Foster A**, Graham JM, Hague J, Hilhorst-Hofstee Y, Hoffer M, Johnson D, Josifova D, Kant SG, Kini U, Lachlan K, Lam W, Lees M, Lynch S, Maitz S, McKee S, Metcalfe K, Nathanson K, Ockeleon CW, Parker MJ, Pierson TM, Rahikkala E, Sanchez-Lara PA, Spano A, Van Maldergem L, Cole T, Douzgou S, Tatton-Brown K. The *CHD8* overgrowth syndrome: A detailed evaluation of an emerging overgrowth phenotype in 27 patients. *American Journal of Medical Genetics Part C*. 2019;1-8.

Foster A, Zachariou A, Loveday C, Ashraf T, Blair E, Clayton-Smith J, Dorkins H, Fryer A, Gener B, Goudie D, Henderson A, Irving M, Joss S, Keeley V, Lahiri N, Lynch SA, Mansour S, McCann E, Morton J, Motton N, Murray A, Riches K, Shears D, Stark Z, Thompson E, Vogt J, Wright M, Cole T, Tatton-Brown K. The phenotype of Sotos syndrome in adulthood: A review of 44 individuals. *American Journal of Medical Genetics Part C*. 2019; 1-7.

Griffiths S, Loveday C, Zachariou A, Behan L-A, Chandler K, Cole T, D'Arrigo S, Dieckmann A, **Foster A**, Gibney J, Hunter M, Milani D, Pantaleoni C, Roche E, Sherlock M, Springer A, White SM, Childhood Overgrowth Collaboration, Tatton-Brown K. EED and EZH2 constitutive variants: A study to expand the Cohen-Gibson phenotype and contrast it with Weaver syndrome. *American Journal of Medical Genetics Part A*. 2019;179A;588-594.

Priolo M, Schanze D, Tatton-Brown K, Mulder PA, Tenorio J, Kooblall K, Hernandez Acero I, Alkuraya, FS, Arias P, Bernardini L, Bijlsma EK, Cole T, Coubes C, Dapia I, Davies S, Di Donato N, Elcioglu NH, Fahrner JA, **Foster A**, Garcia Gonzalez NG, Huber I, Iascone M, Kaiser A-S, Kamath A, Liebelt J, Lynch SA, Maas SM, Mammi C, Mathijssen IB, McKee S, Menke LA, Mirzaa GM, Montgomery T, Neubauer D, Neumann TE, Pintomalli L, Pisanti MA, Plomp AS, Price S, Slater C, Santos-Simarro, Sarda P, Segovia M, Shaw-Smith A, Smithson S, Suri M, Valdez RM, Van Haeringen A, Van Hagen JM, Zollino M, Lapunzina P, Thakker R, Zenker M, Hennekam R. Further delineation of Malan syndrome. *Human Mutation*. 2018. June;39:1226-1237.

H.2 Poster presentations

B Rooke,S Taibjee, **A Foster**, D Lim, C Moss, M-L Lovgren. Myofibroma-like skin nodules are part of the cutaneous phenotype in Kosaki overgrowth syndrome. World Congress of Paediatric Dermatology, Edinburgh, 2021.

B Chalot, **A Foster**, E Schaeffer, CF Rustad, K Tveten, T Cole, C Thauvin-Robinet, J Woodley, A-L Bruel, R Keelagher, C Philippe, T Antoniadi, P Vabres, D Lim, L Faivre. Expansion of the phenotype of Kosaki overgrowth syndrome, and description of the long-term outcome in the oldest case. European Human Genetics Conference (ESGH), Gothenburg, Sweden 2019.

A Foster, T Cole, J Woodley, R Keelagher, T Antoniadi, D Lim. Expanding the phenotype of Kosaki overgrowth syndrome: the first UK patient. Manchester Dysmorphology Conference, Manchester, 2018.

H.3 Presentations

Sotos syndrome. Sotos Syndrome Support Association (SSSA) virtual conference, 2021

An ultra-rare overgrowth disorder. Virtual Midlands Dysmorphology Meeting, 2021. Awarded the Louise Brueton Memorial prize.

Sotos syndrome Child Growth Foundation (CGF) virtual convention. Spoken presentations to Sotos group, 2020

Sotos syndrome in adulthood: a review of 44 individuals. Sotos Syndrome Support Association (SSSA) annual conference, Boston USA, 2019

The Phenotyping of Overgrowth Disorders (POD) study. NIHR Clinical Research Network East Midlands and West Midlands Inaugural Genetics Specialty Event, Lichfield, 2019

Overgrowth Disorders. Rare Diseases Symposium, Medical School, University of Birmingham, 2018

The Phenotyping of Overgrowth Disorders (POD) Study. British Society for Paediatric Endocrinology and Diabetes (BSPED) conference, Newcastle, 2017

The Phenotyping of Overgrowth Disorders (POD) Study. Birmingham Children's Hospital R&D Showcase, Birmingham, 2017

Overgrowth Disorders. Sheffield Rare Disease Study Day, Sheffield, 2017

The Phenotyping of Overgrowth Disorders (POD) Study. Midlands Paediatric Endocrine Group (MPEG) meeting, Birmingham 2017

The Phenotyping of Overgrowth (POD) study. NIHR CRN West Midlands Children's and Genetics Joint Specialty meeting, Birmingham, 2016