SCIENTIFIC PROJECT

STUDY TITLE:

Late-Onset and Paediatric Presentations of CHKB-Related Megaconial Myopathy: A Multicentre Retrospective Survey

RUNNING TITLE: CHKB-Myo survey

STUDY SCOPE AND DOMAIN

Multicentre, non-interventional, retrospective survey collecting clinical and paraclinical data issued from the medical records of patients with CHKB-related megaconial myopathy, with a particular focus on late-onset phenotypes.

SCIENTIFIC CONTEXT

CHKB encodes choline kinase beta, a 395-amino acid enzyme that catalyses the first step of the CDP-choline (Kennedy) pathway — the phosphorylation of choline to phosphocholine — required for de novo synthesis of phosphatidylcholine (PC), the most abundant phospholipid in eukaryotic cell membranes. Two choline kinase genes exist in humans: *CHKA* (choline kinase alpha) and *CHKB*. The tissue expression of the two isoforms is differential: CHKβ is the predominant isoform in the proximal skeletal muscle of the lower limbs, while CHKα predominates in the brain, liver, heart, and upper limb muscles. (Sher, Roger B et al. The Journal of biological chemistry 2006; Tavasoli, Mahtab et al. The Journal of biological chemistry, 2020)

Biallelic loss-of-function variants in CHKB cause megaconial congenital muscular dystrophy (Online Mendelian Inheritance in Man, OMIM 602541), a rare autosomal recessive disease characterized on muscle biopsy by the presence of giant peripheral mitochondria (megacones) with central mitochondrial depletion. Since the initial description by Nishino et al. in 1998, and the genetic identification by Mitsuhashi et al. in 2011, fewer than 50 patients have been reported worldwide.

The predominant phenotype is a severe congenital muscular dystrophy (CMD) with neonatal hypotonia, early progressive motor impairment and intellectual disability in most cases (Chan SHS, Nishino I., 2023). Congenital structural heart defects and ichthyosiform cutaneous lesions have also been associated. However, the phenotypic spectrum is broader than initially recognized, namely with the description of three late onset-cases that presented with a limb-girdle muscular dystrophy (LGMD) phenotype, normal or mild cognitive involvement, no cardiomyopathy, and rhabdomyolysis as the presenting feature (Brady et al., Muscle Nerve, 2016; De Fuenmayor-Fernández De La Hoz, Carlos Pablo et al. Muscle & nerve 2016).

The phenotypic variability of CHKB-related disease raises several unsolved questions. No formal genotype-phenotype correlation has been established: patients carrying identical biallelic variants have been reported with both CMD and LGMD phenotype (Magri, Francesca et al. Skeletal muscle, 2022). In addition it is known that in CHKB patients and Chkb−/− knockout mice, not all tissues are affected, suggesting that residual choline kinase activity encoded by CHKA may rescue these unaffected cell types (Sayed-Zahid, Ambreen A et al. Human molecular genetics, 2019). Nowadays the mechanisms underlying phenotypic heterogeneity — including the role of variant type, residual enzymatic activity, and modifier genes — remain largely unknown in human patients.

Furthermore, several disease features of potential clinical relevance have not been systematically evaluated such as the degree of respiratory chain dysfunction, histopathological features, brain magnetic resonance imaging (MRI) findings, muscle MRI pattern.

This study will systematically characterize both the late-onset and paediatric phenotypes in a large cohort of CHKB patients, to better understand genotype-phenotype associations and provide a comprehensive description of the clinical phenotype to reduce diagnostic delay and optimise patient management. A better understanding of the natural history of this disease, across its full phenotypic spectrum, is a prerequisite for the development of therapeutic trials and the identification of therapeutic targets — including peroxisome proliferator-activated receptor (PPAR) agonists, which have recently demonstrated a proof-of-concept effect in the Chkb−/− mouse model (Tavasoli et al., Nature Communications 2022).

PRIMARY AND SECONDARY RESEARCH OBJECTIVES

Primary objective: To describe the clinical and paraclinical phenotype of patients with CHKB-related megaconial myopathy, including available muscle biopsy findings, genotype, muscle MRI, brain MRI, electromyography (EMG), creatine kinase (CK) levels, cardiac assessment and pulmonary function tests.

Secondary objectives:

- Evaluation of disease course: loss of motor milestones (e.g. loss of independent ambulation), occurrence of respiratory or cardiac events, tracheostomy, pacemaker implantation and death, where applicable.
- Correlation of clinical phenotype and disease course with genotype.
- Systematic comparison of the late-onset (LGMD) phenotype with the paediatric-onset (CMD) phenotype, with a focus on cognitive involvement, cardiac involvement, rhabdomyolysis, and respiratory function.

Given the ultra-rare nature of the condition, this study is expected to be primarily descriptive and exploratory. The main objective is providing a comprehensive description of the clinical phenotype — both paediatric and late-onset — to reduce diagnostic delay, improve patient management, and increase survival by proposing standardized cardiac, respiratory, and orthopaedic follow-up recommendations. If recurrent variants are identified across the international cohort, exploratory correlations between specific genotypes and clinical outcomes will be attempted, subject to sample size and data completeness

PROTOCOL AND MATERIALS

Regulatory framework

This is a retrospective, observational case-series study based exclusively on the reuse of data already collected during routine clinical care. No study-specific intervention, additional procedure, visit, or investigation will be carried out, and patient management will not be modified in any way.

As the study relies solely on the retrospective reuse of existing health data, it does not constitute research involving the human person within the meaning of the French Jardé law. For the French coordinating site and any French participating centres, data processing falls under the CNIL MR-004 reference methodology, validated by the institutional Data Protection Officer.

For non-French participating centres, the local investigator is responsible for ensuring compliance with applicable local ethical, institutional, and data-protection requirements before transmitting any data to the coordinating centre.

No directly identifying patient data will be transmitted to the coordinating centre.

Population concerned

This study focuses on an ultra-rare disease with an estimated prevalence of approximately 1 case per 1,000,000 individuals for the congenital *CHKB* phenotype (Orphanet). Given the exceptional rarity of the condition, a conventional sample size calculation based on statistical power is neither feasible nor appropriate.

The objective is not to obtain a statistically representative sample of a broader population, but rather to capture as comprehensively as possible the existing clinical experience and available cases across expert centres.

To maximise case ascertainment and reduce selection bias, the study invitation and CRF will be distributed to identified European Reference Centres, centres of expertise and specialised clinical units involved in the diagnosis and management of neuromuscular disorders. Potential centres will be identified through European Reference Network for Rare Neuromuscular Diseases (ERN EURO-NMD) channels and Filnemus (the French national health network for rare neuromuscular diseases). The target population therefore corresponds to the accessible population of expert centres rather than to a predefined sample.

Data sources

Data will be extracted from the medical records of patients with CHKB-related megaconial myopathy followed in participating neuromuscular reference centres or expert centres in France and Europe. Data extraction will be performed locally by the treating physician or an authorised member of the local clinical team.

Data collected — Case Report Form (CRF)

- Study ID composed of: [ERN reference centre number, patient number] The correspondence table between the study ID and the patient's identity will remain at the participating centre and will not be transmitted to the coordinating centre.
- Age at relevant time points (e.g. age at first symptoms, age at muscle biopsy, age at last follow-up), sex, mode of inheritance, affected family members and consanguinity.
- Molecular genetic diagnosis and annotation of identified variants (Human Genome Variation Society – HGVS- nomenclature, American College of Medical Genetics and Genomics -ACMG- classification), type of genetic analysis performed (neuromuscular disease -NMD- panel, Whole Exome Sequencing -WES-, Whole Genome Sequencing -WGS-, Sanger)
- Clinical data: first symptom, age at symptom onset, distribution of weakness, associated symptoms, orthopaedic deformities, cardiac and/or respiratory symptoms, functional assessment
- Cognitive assessment: intellectual disability (presence and severity), developmental delay, speech/language delay
- Epilepsy: presence, seizure type, electroencephalogram findings
- Skin abnormalities (ichthyosis, hyperkeratosis) and hearing loss
- CK levels
- Results of EMG and nerve conduction studies
- Results of electrocardiogram and cardiac echocardiography
- Results of pulmonary functional tests (FVC)
- Muscle MRI: affected muscles, rostro-caudal gradient, pattern description (if available)

- Brain MRI: structural findings, MR spectroscopy results if available (choline/N-acetylaspartate [NAA] ratio)
- Results of muscular biopsy (if performed during the diagnostic process): histological features, megaconial pattern, electron microscopy); Western blot of CHKβ from muscle biopsy (if available and optional)

Planned statistical analysis

All analyses will be primarily descriptive and exploratory, given the expected small sample size and the rarity of the disease. Sex, molecular confirmation of CHKB-related disease and type of clinical presentation (CMD versus LGMD / paediatric-onset versus late-onset) are considered key variables for patient inclusion and analysis.

Categorical variables will be presented as counts and percentages. Continuous variables will be presented as medians with interquartile ranges or as means with standard deviations, as appropriate. Normality may be assessed using the Shapiro–Wilk test when the sample size is sufficient.

Comparisons between groups may be performed using non-parametric or parametric tests as appropriate, including the Mann–Whitney U test for two independent groups. Correlations may be explored using Spearman or Pearson correlation coefficients, as appropriate. Multivariable analyses will only be performed if sample size, data completeness and the number of events is sufficient; otherwise, analyses will remain descriptive and exploratory.

Where relevant and feasible, Kaplan–Meier survival curves with log-rank tests may be used to explore survival or time-to-event outcomes in paediatric-onset versus late-onset groups. Statistical significance will be set at $p < 0.05$, recognising the exploratory nature of the analyses

PARTICIPANT RECRUITMENT MODALITIES

Eligible patients will be identified retrospectively from the medical records of participating neuromuscular reference centres and expert centres in France and Europe.

Inclusion criteria:

- Patients of any age, alive or deceased, followed at a neuromuscular reference centre or expert centre up to the date of study closure
- Clinical evidence of neuromuscular involvement, including muscle weakness, myopathy, congenital muscular dystrophy or limb-girdle muscular dystrophy phenotype.
- Molecular confirmation of CHKB-related disease, defined as biallelic pathogenic or likely pathogenic variants in CHKB, or biallelic CHKB variants considered compatible with the phenotype after expert review, particularly when supported by characteristic muscle biopsy findings.

Exclusion criteria:

- Patients, or where applicable the holder(s) of parental authority, for whom an objection to the reuse of the patient's health data for research purposes has been recorded.
- Patients with megaconial myopathy attributed to a non-CHKB aetiology, including immune-mediated megaconial myopathy or another alternative diagnosis.
- Patients for whom the available medical records are insufficient to confirm eligibility or extract the minimum required dataset.

DATA PROCESSING CONDITIONS

Patients will be excluded if an objection to the reuse of their health data for research purposes has been recorded, either by the patient where applicable or by the holder(s) of parental authority.

Minors may be included in the study if they meet the eligibility criteria. As this is a retrospective study based exclusively on the reuse of data collected during routine clinical care, no study-specific intervention or additional procedure will be performed.

For patients who are minors at the time of inclusion, information regarding the study and the right to object to the reuse of health data for research purposes will be provided in accordance with applicable regulations, including to the holder(s) of parental authority where required. Where appropriate, information will also be provided to the minor in a form adapted to their age and level of understanding.

For patients who are adults at the time of data reuse but whose relevant clinical data were collected during childhood, the patient's current legal status will be considered for information and exercise of rights, while allowing retrospective review of paediatric medical records where permitted.

Deceased patients may be included if they meet the eligibility criteria, unless an objection to the reuse of their health data for research purposes was recorded during their lifetime, subject to applicable legal and institutional requirements.

Pseudonymisation and data minimisation

The treating physician or an authorised member of the local clinical team will complete the CRF using a study-specific pseudonymised code. The code will consist of a centre code followed by a sequential patient number (for example: FR01-001).

The correspondence table linking the study code to the patient's identity will be kept locally at each participating centre, under the responsibility of the local investigator, and will not be transmitted to the coordinating centre.

The CRF will not include directly identifying data. Particular attention will be paid to the risk of indirect re-identification in an ultra-rare disease context; therefore, exact dates, excessive geographical detail and unnecessary free-text information will be avoided wherever possible.

Data transfer, storage and access

The final data-collection and transfer procedure will be validated by the coordinating institution and/or data protection officer before data collection begins. Preferably, data will be collected through a secure institutional electronic CRF (eCRF) or another approved secure research data-collection tool.

If password-protected CRF files are used, transmission will occur only through institutional secure messaging or another approved secure transfer method, and passwords will be transmitted separately from the files. The use of password-protected Excel files will be subject to prior validation by the data protection officer or institutional data governance body.

The study dataset will be stored on secure institutional storage with restricted access. Access to the dataset for statistical analyses will be limited to the principal investigator and authorised co-investigators of the coordinating centre. Access rights will be documented and reviewed as necessary.

Data retention, archiving and deletion procedures will follow the requirements validated by the coordinating institution, data protection officer and applicable regulations. The planned retention period is 10 years.

STUDY CALENDAR AND DISSEMINATION PLAN

- Clinical data collection: [Month 1-2]
- Statistical analysis: [Month 3]
- Manuscript preparation: [Month 4-6]
- Presentation of results at an international neuromuscular meeting

Results will be disseminated in aggregate form only. No directly identifying patient data will be included in presentations or publications. Particular care will be taken to avoid indirect identification of patients in an ultra-rare disease context, especially when reporting combinations of age, sex, genotype, centre/country and highly specific clinical features.