Journal Pre-proof

GenIDA, a participatory patient registry for genetic forms of intellectual disability provides detailed caregiver reported information on 237 individuals with Koolen-de Vries syndrome

Florent Colin, Pauline Burger, Timothée Mazzucotelli, Axelle Strehle, Joost Kummeling, Nicole Collot, Elyette Broly, Angela T. Morgan, Kenneth A. Myers, Agnès Bloch-Zupan, Charlotte W. Ockeloen, Bert B.A. de Vries, Tjitske Kleefstra, Pierre Parrend, David A. Koolen, Jean-Louis Mandel

PII: S2949-7744(23)00826-9

DOI: https://doi.org/10.1016/j.gimo.2023.100817

Reference: GIMO 100817

To appear in: *Genetics in Medicine Open*

Received Date: 4 April 2023

Revised Date: 3 May 2023

Accepted Date: 4 May 2023

Please cite this article as: Colin F, Burger P, Mazzucotelli T, Strehle A, Kummeling J, Collot N, Broly E, Morgan AT, Myers KA, Bloch-Zupan A, Ockeloen CW, de Vries BBA, Kleefstra T, Parrend P, Koolen DA, Mandel JL, GenIDA, a participatory patient registry for genetic forms of intellectual disability provides detailed caregiver reported information on 237 individuals with Koolen-de Vries syndrome, *Genetics in Medicine Open* (2023), doi: https://doi.org/10.1016/j.gimo.2023.100817.

Author Details

GenIDA, a participatory patient registry for genetic forms of intellectual disability provides detailed caregiver reported information on 237 individuals with Koolen-de Vries syndrome

Florent Colin[1,2#], Pauline Burger[1#], Timothée Mazzucotelli[1], Axelle Strehle[1], Joost Kummeling[3], Nicole Collot[1], Elyette Broly[1,4,5], Angela T. Morgan[6,7], Kenneth A. Myers[8,9,10], Agnès Bloch-Zupan[1,4,5], Charlotte W. Ockeloen[3], Bert B. A. de Vries[3], Tjitske Kleefstra[3,11,12], Pierre Parrend[13], David A. Koolen[3,*] & Jean-Louis Mandel[1,14,*]

[1]Institute of Genetics and Molecular and Cellular Biology (IGBMC), Université de Strasbourg, INSERM U1258, CNRS UMR7104, Illkirch, France.

[2]Current address: INSERM UMR_S1109, Tumor Biomechanics Lab, University of Strasbourg, Fédération de Médecine Translationnelle de Strasbourg (FMTS), Strasbourg, France.

[3]Department of Human Genetics, Donders Institute for Brain, Cognition and Behavior, Radboud university medical center, Nijmegen, The Netherlands.

[4]Faculté de Chirurgie Dentaire, Université de Strasbourg, Strasbourg, France.

[5]Centre de référence des maladies rares orales et dentaires, O-Rares, Filière Santé Maladies rares Tête Cou, European Reference Network ERN CRANIO, Hôpitaux Universitaires de Strasbourg (HUS), Strasbourg, France.

[6]Murdoch Children's Research Institute, Parkville, VIC 3052, Australia.

[7]Faculty of Medicine, Dentistry and Health Sciences, University of Melbourne, Parkville, 3052, Melbourne, Australia.

[8]Department of Pediatrics, Montreal Children's Hospital, McGill University, Montreal, Quebec, Canada.

[9]Department of Neurology & Neurosurgery, Montreal Children's Hospital, McGill University, Montreal, Quebec, Canada.

[10]Research Institute of the McGill University Health Centre, Montreal, Quebec, Canada.

[11]Centre of Excellence for Neuropsychiatry, Vincent van Gogh Institute for Psychiatry, Venray, The Netherlands.

[12]Department Clinical Genetics, ErasmusMC Rotterdam, The Netherlands.

[13]ICube laboratory (Laboratoire des sciences de l'ingénieur, de l'informatique et de l'imagerie), UMR 7357, Université de Strasbourg, CNRS, Strasbourg, France ; EPITA, 5, Rue Gustave Adolphe Hirn, Strasbourg, France.

[14]University of Strasbourg Institute for Advanced Studies (USIAS), Strasbourg, France.

[#]These authors contributed equally to this work.

*These authors contributed equally to this work.

*Corresponding author: Pauline BURGER (burgerp@igbmc.fr; 0033 3 88 65 57 46)

ORCID numbers

Florent Colin 0000-0002-0527-2879

Pauline Burger 0000-0003-2109-5451

Timothée Mazzucotelli 0000-0003-1282-5088

Axelle Strehle – no ORCID n°

Joost Kummeling 0000 0003 4870 0666

Nicole Collot – no ORCID n°

Elyette Broly 0000-0001-9065-4697

Angela T. Morgan 0000-0002-1117-8782

Kenneth A. Myers 0000-0001-7831-4593

Agnès Bloch-Zupan 0000-0002-6511-2615

Charlotte W. Ockeloen 0000-0003-0329-1520

Bert B. A. de Vries 0000-0002-0000-2917

Tjitske Kleefstra 0000-0002-0956-0237

Pierre Parrend 0000-0002-1680-1182

David A. Koolen 0000-0002-2152-4626

Jean-Louis Mandel 0000-0002-0535-6589

Data availability

Data specific to a cohort of individuals diagnosed with a NDD in GenIDA are accessible in processed form via the visualization tab (https://genida.unistra.fr/visualizations/), item by item, on the website to any caregiver registered in GenIDA in the said cohort (https://genida.unistra.fr/register/) as soon as the threshold of 10 completed files for the said condition is reached (defined for statistical analysis). They are also available to any professional user upon simple registration in GenIDA.

Each participating caregiver can download its relevant PDF file, sort of anonymous clinical records containing all questionnaire responses for a single individual diagnosed with a NDD.

Registered investigators/professionals wishing to have access to these PDFs must justify their request (short statement explaining their project) to the GenIDA administration by email (genida@igbmc.fr) and commit to respecting the confidentiality of the study participants and therefore not to disseminate the PDFs.

Acknowledgements

The authors wish to thank the Institute for Advanced Studies of the University of Strasbourg (USIAS), the University of Strasbourg Foundation (Roche Fund for Personalized Medicine), the Fondation Jérôme Lejeune and the National Research Network “Groupement d’Intérêt Scientifique Autisme et Troubles du Neuro- Développement” (GIS Autisme et TND) for their financial support.

The GenIDA project is supported by the French RaDiCo (Rare Disease Cohorts) research program for national and European cohorts of rare disease patients. This work is generated within the European Reference Network for Intellectual disability, TeleHealth, Autism and Congenital Anomalies (ERN ITHACA). The authors wish to thank Doulaye Dembele, Genomeast, IGBMC, for his help in the statistical analysis of the GenIDA data. Finally, the authors wish to thank all the GenIDA participants and the patients’ associations (and notably https://supportingkdvs.org/; http://www.koolkidalliance.com/; https://www.koolendevriesfrance.org/).

Funding Statement

This research was funded by the Institute for Advanced Studies of the University of Strasbourg (USIAS), by the University of Strasbourg Foundation (Roche Fund for Personalized Medicine), by the Fondation Jérôme Lejeune and by the GIS “Autisme & TND”.

Author Contributions

Conceptualization: FC; TM; TK; PP; DAK; JLM; Data curation: FC; PB; JLM; Formal analysis: FC; PB; JK; NC; EB; JLM; Funding acquisition: FC; PB; JLM; Investigation: FC; PB; JK; JLM; Project administration: FC; PB; AS; JLM; Software (Design and development of the registry and confidentiality model): TM; PP; Supervision: JLM; Visualization: FC; TM; JLM; Writing—original draft: FC; PB; JK; NC; TK; DAK; JLM

Writing—review & editing: FC; PB; JK; NC; ATM; KAM; ABZ; CWO; BBAdV; TK; PP; DAK; JLM

Ethics Declaration

GenIDA satisfies all ethical requirements: the research is carried out in accordance with the provisions of the French Data Protection Act (French law of 6 January 1978, amended by the law of 6 August 2004 on the protection of individuals with regard to the processing of personal data), and complies with the General Data Protection Regulation 2016/679 (GDPR). The study has been declared to the French “Commission Nationale Informatique et Libertés” (CNIL) on 27/11/2015, number 1907912v0. It has been reviewed and approved by the Ethics Evaluation Committee (IORG0003254, FWA00005831), the Institutional Review Board (IRB00003888) of the French Institute of Medical Research and Health on 15/11/2016, number 16-338, and on 04/09/2019, number 16-338 bis. It was developed in collaboration with the French Rare Disease Cohort program (RaDiCo: https://www.radico.fr/en/accueil) and information security and data management specialist from ICube laboratory in Strasbourg (PP). Potential risks concerning data storage, security, privacy, type of data asked, or structure of the website have been considered during the development of this project and are detailed in the validated research protocol of the project (available on request).

Individuals registered to the study were asked to read the information note and give their informed consent during the registration, before accessing the questionnaire. Whether because of their age or their ability to participate, only a minority of people with ID participate in GenIDA on their own behalf and therefore give their consent to participate in our study. In the vast majority of cases, it is the caregivers (usually the parents) who register and therefore give their consent to participate in the study. They are sole actors in the data submission process and have the right to edit and delete their answers. At any time and without any justification, they can withdraw his/her participation agreement. Consent is not sought from the person for whom a record is completed, even in cases where that person is classified as "without ID” or is diagnosed as having a mild ID or as not (yet) needing special

education. Indeed, an individual listed as "without ID" may either not have an ID, may be borderline, or may not have been tested for IQ or other tests and thus not have a formal diagnosis at the time of participation in the study (especially children and youth who may be developmentally delayed but have not yet been formally tested, or adults who have not been formally tested or who were tested long ago but have forgotten the test results). In addition, most of the people with ID reported in GenIDA are minors (and may not yet need formal special education) and are still in their parents' care. However, if a person for whom a file has been completed by a parent caregiver wishes to withdraw consent, he or she is fully entitled to ask us at any time to delete his or her record and answers that concern him or her.

Informed consent was obtained from all participants (primarily caregivers, as well as 9/237 individuals with KdVS participating on their own behalf) enrolled to this specific Koolen-de Vries study as required. All the data collected were de-identified. This research was conducted in accordance with the principles set out in the Declaration of Helsinki.

Conflict of interest

The authors declare that they have no competing interests.

GenIDA, a participatory patient registry for genetic forms of intellectual disability provides detailed caregiver reported information on 237 individuals with Koolen-de Vries syndrome

Abstract

Purpose: GenIDA is an international patient registry for individuals diagnosed with intellectual disability, autism spectrum disorder and/or epilepsy based on an online questionnaire completed by parent caregivers. In this study the GenIDA data on Koolen-de Vries syndrome (KdVS) was analyzed illustrating the value of GenIDA and patient/caregiver participation in rare genetic neurodevelopmental disorders (NDDs).

Methods: Recruitment was done on the GenIDA website from November 2016 to February 2022. Clinical information on individuals with KdVS was extracted for in-depth analysis and for comparison with the GenIDA data of individuals diagnosed with other NDDs.

Results: 1,417 patients/caregivers across 35 genetic conditions answered to the GenIDA questionnaire, including caregivers of 237 individuals with KdVS. GenIDA findings on KdVS were consistent with the existing literature and there were no significant differences between individuals with a 17q21.31 microdeletion and those with a pathogenic variant in the *KANSL1* gene. GenIDA provided detailed clinical information including features that are overrepresented in KdVS compared to other NDDs (e.g., laryngomalacia). Modeling of the natural history showed a positive development of speech and language over time and relatively good reading ability in KdVS. Valproate and oxcarbazepine were reported as effective antiepileptic drugs and responses to open-ended questions indicated that childhood recurrent pneumonia and asthma are clinically relevant comorbidities that were not described in KdVS before.

Conclusion: GenIDA is a powerful registry to collect and harness valuable data on rare NDDs. The study shows that caregiver-driven data collection is effective in terms of global recruitment and centralization of clinical data.

Keywords

Autism spectrum disorder; GenIDA; Intellectual disability; *KANSL1*; Koolen-de Vries syndrome; Neurodevelopmental disorders; Patient registry

Introduction

GenIDA is a global patient registry for individuals diagnosed with genetic forms of intellectual disability (ID), with or without autism spectrum disorder (ASD) and/or epilepsy (https://genida.unistra.fr/)[1]. Data on clinical features, comorbidities, progression over time, and responses to treatments are entered and updated by parents and other caregivers *via* a structured online questionnaire maximizing the enormous potential of their experience with these conditions.

Neurodevelopmental disorders (NDDs) including ID and ASD, affect approximately 2% of the population[2] and are characterized by a wide clinical and genetic variability. Advances in sequencing techniques have improved the diagnostic yield in clinical practice for individuals with NDDs in the last decade[3] showing that *de novo* variants are key contributors to such disorders[4–8]. There is an ever-growing list of up to 1,500 genes or recurrent copy number variants involved in NDDs[9–11], corresponding to an equivalent number of rare and distinct conditions. For many of these conditions the knowledge of the clinical spectrum and natural history is limited, although this information is essential for the medical management of individuals diagnosed with a NDD and for addressing the legitimate concerns of parents and families[12]. In order to address these knowledge gaps GenIDA was launched in 2016. Caregivers can fill out a structured questionnaire that is available in 8 languages (French, English, Dutch, German, Spanish, Italian, Portuguese and Greek) and contains 5 numerical questions, 36 multiple choice questions (MCQs) and their 92 sub-questions, and 5 additional open-ended questions[1]. The data are visualized online by numerical graphs and natural history for specific features is modelled by cross-sectional data from individuals at different ages. By February 2022, the GenIDA registry contained over 71,200 answers to questions in the survey providing clinical information on an international cohort of 1,417 individuals diagnosed with a NDD associated with more than 35 different genes or CNVs (**Supplementary Figure 1**). In this study, the data on 237 individuals with Koolen-de Vries syndrome (KdVS, OMIM#610443), the largest cohort in the registry is analyzed. KdVS is a rare multi-system NDD including hypotonia, developmental delay, mild to moderate ID, speech and language disorder, epilepsy, characteristic facial features, musculoskeletal anomalies, congenital heart anomalies and urogenital malformations[13–17]. The syndrome can be caused by a deletion of 17q21.31 that includes the *KANSL1* gene or by a heterozygous intragenic pathogenic variant in *KANSL1*[18–21]. This study is (by far) the largest clinical study to research

KdVS and the first to use a dataset solely built by caregivers. The study shows that GenIDA data are reliable and valuable and can provide useful knowledge about KdVS and rare genetic NDDs in general.

Methods

GenIDA. The setup and design of the GenIDA registry including ethical approval, confidentiality, and the registration / recruitment process are described in Burger *et al.* 2023[1]. Mandatory information is collected during the online registration procedure on the individual with the NDD and on the participating parent/caregiver. The mandatory information includes the name of the gene or CNV causing the NDD without specification of the exact pathogenic variant or duplication/deletion borders as this is potentially identifying information. No copy of the genetic report is required, as this would raise issues of potential patient identification (GenIDA does not have the authorization to collect such data). The full online questionnaire can be found in **Supplementary Material 1** and daily operational aspects are described in **Supplementary Material 2**. The recruitment of individuals with KdVS in the current study was international (**Supplementary Figure 2**). Recruitment was launched in November 2015 for a year-long beta-testing period during which voluntary families were included to test the website features. Open recruitment was done exclusively on the GenIDA website after November 2016. In February 2022, KdVS data from the questionnaire, along with demographic information collected at registration were locked and extracted for in-depth analysis.

Curation process and data analysis. To maximize data quality, six steps of curation were performed: (i) check for double accounts, (ii) check for fake and empty or very low participating (<10 answers) accounts, (iii) curation for wrong “year of birth” (since some caregivers were confused during registration and specified their own year of birth). In addition, caregivers were (iv) re-contacted when the genetic cause was unclear (e.g., both 17q21.31 deletion and a *KANSL1* pathogenic variant reported), and (v) data re-filling for lower participation (>8 to 30 responses) was done by using the answers to the 5 open-ended questions (which some families filled in exhaustively) and moving them to the appropriate MCQ. Finally, (vi) broad curation of numerical data was performed, to remove mistakes (e.g., wrong units for weight / height).

Caregivers are encouraged to update data on a regular basis (preferably annually). Approximately 10% of the KdVS records are updated per year, generally describing the occurrence of a novel clinical feature or adverse drug reaction (**Supplementary Figure 3**).

Statistical analyses of MCQs' and numerical data included means, standard deviations, percentages, and absolute values/sizes for quantifiable items (i.e., comorbidities reported through MCQs). The analysis and construction of the graphs were done according to the nature of the data collected[22]:

- In case of cross-sectional data (collected at a single point in time), only the most recent data concerning a given individual was used for data analysis as parent caregivers can actualize their answers over time. It should however be stressed that all serial responses for a given individual were conserved in the database and can be found in the corresponding individual PDF file.
- In case of data collected for different age groups (0-2; >2-4; >6-13; >13-19; >19y), each group was treated individually, meaning that if a caregiver responded multiple times to the survey (i.e., when the child was 1, 4 and 8), the answers were used in the relevant age group of the graph. If the caregiver responded multiple times for a single age group (i.e., when the child was 6, 7 and 8), only the most recent data was used in the analysis for the given age group.

Comparison between cohorts included frequencies, calculation of ratios, and nominal chi-square tests-based *p*-values calculation, corrected for multiple-testing. Textual data mining was done manually. For cohorts' comparisons, the answers to MCQs ('Yes' / 'No') were extracted excluding the 'I don't know' ('IDK') answers (**Supplementary Figure 4** and **Supplementary Table 1** and **2**). Unless specified in captions, frequencies for sub-questions were based on the total number of individuals for whom answers ('Yes' / 'No') were available at the corresponding main question.

In the KdVS cohort, the majority of participants (>75%) was French, English or Dutch speaking (**Supplementary Figure 2**), and their textual responses were analysed and interpreted by one of the co-authors of this article in whose native language the answers were written. The other subset of answers was analysed using the free version of the DeepL software (https://www.deepl.com/translator/). To remove inconsistencies in some translations, clinicians and researchers who participated in the specific analysis of the data, were asked to provide input.

Results

Overview and validation of GenIDA clinical data on KdVS. In total information on 237 KdVS individuals was documented in GenIDA by February 2022 (**Supplementary Figure 2**). The most represented countries are the USA (42.1%), followed by France (13.2%), Australia (7.9), the Netherlands (7.9%), the UK (7.0%) and Germany (5.4%) (**Supplementary Figure 2A**). The cohort includes 116 males and 121 females (mean age 14.0y; 1.8y to 47.6y). In total, 197 persons had KdVS due to the 17q21.31 deletion and 40 persons had a *KANSL1* pathogenic variant (**Supplementary Figure 2B**). The survey was fully completed in 54.0%, and at least half of the questions were answered for 90.7% of the KdVS individuals (**Supplementary Figure 2C**).

The most commonly reported clinical features in the KdVS cohort are summarized in **Table 1** showing 36 different clinical features divided into 10 categories or organ systems. Comparison of the data with data from two previous KdVS cohort studies[20,21] for major 'Yes/No' MCQs yielded no significant differences (χ^2 tests with 2 degrees of freedom; **Supplementary Figure 4**).

Genotype phenotype correlation. Comparison of the full dataset containing 196 clinical items between the individuals with a 17q21.31 microdeletion and those with a pathogenic variant in *KANSL1* did not reveal significant clinical differences according to the χ^2 test in agreement with previous observations[20] (*p*-value<0.00025; **Supplementary Table 1**).

Clinical features overrepresented in the KdVS cohort. The reported frequencies of the clinical features in the KdVS cohort were compared with data from the other NDDs included in the GenIDA registry (**Supplementary Table 2**). Clinical features that were statistically more frequent (*p*-value<*0.00025*) in KdVS compared to other NDDs in the registry included low amniotic fluid (3.8-fold), lack of fetal movement (3.3-fold), hypotonia at birth (2.9-fold), jaundice (2.5-fold), neonatal feeding difficulties (1.5-fold), atrial septal defect (ASD) (2.5-fold), laryngomalacia (4.4-fold), tracheomalacia (5.2-fold), epilepsy (1.5-fold), joint laxity (2.1-fold), hip dysplasia (4.1-fold), high number of moles (5.6-fold), hypodontia (4.8-fold), and good reading ability (1.7-fold). Interestingly, behavioral problems were significantly less common in KdVS compared to the GenIDA data in general, according to the χ^2 test (0.8-fold).

A summary of the main KdVS features and systems involved is given below. The reported clinical features in the KdVS cohort are summarized in **Table 1** and details and answers to sub-questions are shown in **Supplementary Table 2**.

Prenatal and neonatal history.

Problems during pregnancy, labor and/or delivery were common in KdVS (77.9%) and mainly included low amniotic fluid (17.2%), reduced fetal movement (19.9%) or lack of fetal movement (14.5%). Abnormal ultrasound results were reported in 22.0% of the cases. The proportion of Caesarean-section deliveries (53.2%) was twice as high as in the general population[23] (**Figure 1A**). At birth, on average, the APGAR scores at 1, 5 and 10 minutes were 7, 9 and 9 respectively. At 1 minute the Apgar score was below 6 in 25% of the individuals with KdVS (1st quartile) (**Figure 1B**). The neonatal period was complicated in 86.4%, including neonatal hypotonia (62.9%), feeding difficulties (61.5%) and jaundice (49.3%) (**Figure 1C**).

Psychomotor development.

The developmental motor milestones in KdVS were delayed when compared to the median of the World Health Organization – WHO[24]: ability to sit without support at 10.7 months (versus 6m), to stand without support at 17.5 months (vs 11m) and to walk without support at 23.4 months (versus 12m) (**Figure 1D**). Fine motor problems or clumsiness was reported for 42.0% of the individuals with KdVS (**Supplementary Table 2**). The onset of first words was delayed by more than 12 months (average onset of words at 2.2y) compared to typical development (**Figure 2A**). Interestingly, speech ability increased over time, as over 43.8% (14/32) of the individuals with KdVS over 13 years of age were able to use full and correct sentences (**Figure 2B**). 95.4% of the individuals with KdVS over the age of 6 years could be understood by the family, whereas 48.1% could be understood by people outside the direct family circle, confirming previous studies on speech and language impairment in KdVS[16,25]. To evaluate the specificity of these findings, a similar analysis was performed on two other cohorts in GenIDA, i.e., the Kleefstra syndrome (KS) (M/F 47%/53%; mean age 13.9y) and KBG syndrome (KBGS) (M/F 61%/39%; mean age 14.4y).

The age of first words and understandability were similar for KdVS and KS (**Figures 2A** and **2C**), whereas compared to KBGS the "age of first words" is higher in KdVS (avg. 2.2y, n=139 versus avg. 1.6y, n=35), and “understandability by strangers” is reduced (48.1%, n=108 versus 61.5%, n=26) (**Figure 2C**).
A favorable evolution of reading ability was observed over time, as 20 of 31 individuals with KdVS over 13 years of age had good reading skills. Writing skills were limited even at a later age as only 6 of 30 persons over 13 years of age had good writing ability according to their caregivers. In 89.2% of the individuals with KdVS over 6 years of age, a formal diagnosis of ID was made (mild 16.9%; moderate 60.2%; severe 2.4%; profound 20.5%). Many individuals with KdVS over 6 years of age were able to dress themselves (89.4%), to brush their teeth (82.1%), to actively assist with personal care (90.2%), but were considerably less capable to tie their shoelaces (21.2%) (**Supplementary Table 2**). Children over 5 years of age had special educational needs in 85.1% of the cases.

Neurological and psychiatric features.

***Epilepsy*:** Epilepsy was present in 47.3% of individuals with KdVS (**Figure 3A**) which is in line with data from the medical literature[20,21]. Seizure types reported the most were tonic-clonic seizures (37/97, 38.1%), absence seizures (28/97, 28.9%), and complex partial seizures (e.g., focal impaired awareness seizures, 22/97, 22.7%). The average age at the first seizure was 3.4 years, with a median of 2.0 years (**Figure 3B**), supporting the findings in a previous study on 31 KdVS individuals with epilepsy[26]. Multiple reports of adverse effect of anti-epileptic drug treatment in the open answers’ fields for the individuals with KdVS prompted us to perform an additional analysis on the perceived efficacy of the treatments by drugs, and the frequency of adverse events for each (data freeze: Nov. 2016 – Aug. 2019; **Figure 3C**). Of the 81 KdVS individuals with epilepsy, 14 reported no medication as the epilepsy spontaneously resolved or seizures were rare enough that treatment was not initiated. There was no information regarding the specific medications used for an additional 13 individuals with KdVS. The drugs used most frequently were levetiracetam (30/54), valproate (24/54) and oxcarbazepine (13/54) with perceived efficacy ranging from 67-85%. Reports of significant adverse events were relatively rare for oxcarbazepine (1/13, 8.0%), but higher for levetiracetam (14/30, 47.0%) and valproate (7/24, 29.0%; 3 additional individuals reported mild secondary events). Four major adverse events were reported: severe behavioral changes with

levetiracetam, Stevens-Johnson syndrome with carbamazepine, Cushing syndrome and mood change with adrenocorticotropic hormone, and respiratory distress requiring hospitalization with combined diazepam and clonazepam treatment.

***Behavioral problems*:** The frequency of behavioral problems was reported across 13 sub-questions. The most frequent behavioral problems reported for KdVS were repetitive behavior/stereotypes (35.2%), attention deficit (32.7%), anxiety (31.2%), obsessive behavior (29.6%), and hyperactivity (27.6%). Attention deficit was rated most frequently as a major problem (**Figure 4A**).
Overall, individuals with KdVS (over 6 years of age) were described as average to highly sociable with familiar children in 88.6% and in 98.1% with familiar adults. The individuals in the KdVS cohort were less likely to experience behavioral problems compared to other individuals with NDDs (odds ratio 0.8, **Supplementary Table 1**) and behavioral problems were also reported less frequently in KdVS (54.8%) compared to KS and KBGS, two other specific NDDs for which sufficient data was available (respectively 71.4% and 83.7%) (**Figure 4B**).

Vision and hearing: Hypermetropia is present in 38.8% of individuals in the GenIDA KdVS cohort, and strabismus was reported in 34.7% (**Supplementary Table 2**). Other clinical features involving the eyes were nystagmus (4.6%) and 4 cases of infantile cataract (2.0%). Hearing problems (40.8%) mainly included deafness (12.2%) and recurrent ear infections (17.3%).

Cardiovascular: Congenital heart anomalies in the KdVS cohort mainly included ASD (18.9%) or ventricular septal defect (10.9%), but other cardiac anomalies were also reported, including bicuspid aortic valve, cardiomyopathy, pulmonary stenosis, broadening of the aorta/aorta root, patent ductus arteriosus, tetralogy of Fallot and cardiac rhythm problems (**Supplementary Table 2**).

Respiratory and pulmonary system: In the open questions of the GenIDA questionnaire, respiratory problems were reported as one of the major medical problems that affect the persons' health and quality of life. Overall, respiratory problems were reported in 39.8% (80/201) of the KdVS individuals (**Figure 5A**).

Tracheo- (17/201, 8.5%) and laryngomalacia (31/201, 15.4%) were reported previously[20,21], but the relatively high frequency of respiratory infections was unknown. There was no MCQ sub-question on pulmonary respiratory infections and these problems were self-reported by the caregivers in the open-ended questions (**Figure 5B**). They often reported 'croup' or 'chest infection' without further specification, and notably cited pneumonia, a condition described as recurrent in 13 persons and often led to hospitalization. Pneumonia was generally of bacterial or viral origin with two reported cases of "aspiration" pneumonia. In the majority, the first pneumonia occurred within the first 18 months (average 15.5m), and most did not have pneumonia after age 10 although some still had pneumonia in adolescence or adulthood (latest event on average 8.4y) (**Figure 5C**). Other reported respiratory problems included asthma (33/201, 16.4%). In 12 cases, caregivers also reported immune system problems including lymphopenia, neutropenia, common variable immune deficiencies, and IgG deficiencies.

Gastrointestinal tract and urogenital system. The most frequently reported clinical feature of the gastrointestinal tract in the KdVS cohort was constipation (31.1%) followed by hypersalivation (27.3%) and swallowing difficulties (16.1%; **Supplementary Table 2**). Regurgitation (7.8%), repeated vomiting episodes (8.3%) and diarrhea (8.8%) were also reported. Renal and urogenital anomalies were present in 38.3%. Cryptorchidism was present in 22.6 % of the male individuals with KdVS. Other important anomalies of the urogenital tract included vesicoureteral reflux (6.2%), and recurrent urinary infections in 20.0% of the female individuals with KdVS and 3.2% of the male individuals with KdVS.

Musculoskeletal system. Musculoskeletal anomalies were present in 75.5% including hypermobility (joint laxity, 50.0%), hip dislocation/dysplasia (18.0%), pes planus (22.0%) and scoliosis (25.5%; **Supplementary Table 2**). Other musculoskeletal problems were reported in less than 10.0%, e.g., club feet (3.0%) and contractures (2.5%).

Miscellaneous. Skin, nail and hair abnormalities were present in 51.7% of the cases, for example multiple nevi (21.9%). Dental problems were reported in 65.1%. A detailed analysis of the orodental features is available in **Supplementary Material 3.** It is of note that 'tongue tie' (= ankyloglossia) was

mentioned by 7 families, which is relevant to address in the overall management of individuals with KdVS. In the responses to the open-ended questions, 4 cases of double cataract at a young age were reported, and 4 cases of Hashimoto disease. Sleep problems occurred in approximately 42.6% of KdVS cases. There is no significant difference according to the χ^2 test (**Supplementary Table 2**) in the occurrence of sleep problems in KdVS compared to all other cohorts in GenIDA. Still, the high number of individuals with KdVS experiencing a sleep problem warrants follow-up analysis and research considering how interventions that improve sleep can be incorporated into KdVS health care management.

Discussion

Detailed and extensive clinical data. The GenIDA KdVS data analysis in this study shows that caregivers provide valuable information about KdVS and rare NDDs in general. They are motivated to complete a relatively long structured questionnaire and thereby contribute to research on rare NDDs. Parents are experts on their child and know him/her better than any health care professional. For rare diseases this is a huge source of information and GenIDA can play an important role in collecting and summarizing these data. The involvement of clinicians in such registries is essential to promote recruitment and conduct data analysis and interpretation. In this study. caregivers provided extensive clinical information by answering an extensive online questionnaire allowing assessments on 196 items, with additional numeric and textual elements. The data are reliable as they give frequencies of phenotypes that are comparable to previously published observations from clinicians-driven studies[20,21] (**Supplementary Figure 4**), although, it might be possible that data on some individuals are included in both GenIDA and one of the studies in the comparison.

Importantly, the GenIDA dataset adds to the body of knowledge about KdVS and its management. The data is in many aspects much more extensive than previously published and the cohort is considerably larger than cohorts described in the medical literature[20,21]. This allows new findings to be made, but it also confirms previous findings in the smaller cohorts. Most answers to the questions and sub-questions of the questionnaire concerning the KdVS cohort (n=237) are presented in **Figures 1-5** and in **Supplementary Tables 1** and **2**, with the exception of textual answers to open-ended questions. Only a sampling of the textual answers is presented for respiratory problems (**Figure 5**) and responses to anti-epileptic

treatments (**Supplementary Material 4**). This data provides an overview of the clinical spectrum of the syndrome which is valuable in daily clinical practice. Many doctors caring for a child with KdVS have only little experience with this rare condition and this information can help them to counsel families and provide appropriate care for the patient.

No significant genotype phenotype relation. The current dataset contains 197 persons diagnosed with KdVS due to a 17q21.31 deletion and 40 persons with a pathogenic *KANSL1* variant. Comparison of the full dataset between these two groups did not show significant clinical differences according to the χ^2 test (**Supplementary Table 1**) which is in agreement with previous observations[20] stressing the previous conclusion that haploinsufficiency of *KANSL1* is sufficient to cause the core phenotype of KdVS[18,19].

Clinical features that are overrepresented compared to other NDDs. Comparison of the KdVS data to the overall GenIDA data shows clinical features that are significantly more frequent in KdVS according to the χ^2 test (**Supplementary Table 2**). This demonstrates an interesting potential for GenIDA, because many clinical features in individuals with NDD are relatively unspecific, such as hypotonia, developmental delay or constipation. This type of comparison may reveal clinical features that require extra attention in a specific condition. Clinical features that were statistically more frequent (*p*-value<*0.00025*) in KdVS compared to other NDDs mainly include clinical signs in the perinatal period such as low amniotic fluid, lack of fetal movement, hypotonia at birth and neonatal feeding difficulties. ASD and epilepsy are also reported more often. There is a >4-fold greater odds of tracheo-/laryngomalacia for KdVS than for other NDDs stressing the importance of upper airway evaluation in infants & children with KdVS and signs or symptoms suspicious of tracheo-/laryngomalacia.

A high number of moles is also reported more often (5.6-fold). This is something specific to look for at physical examination and individuals with KdVS who have multiple naevi or a skin type that is at greater risk for developing melanoma should be evaluated periodically to assess ectodermal findings and cutaneous changes.

Modelled natural history. A great potential value of GenIDA is the collection of longitudinal data on KdVS and other rare NDDs. This is important for counseling but also for determining outcome measures for possible treatment and interventions. Clinical information on individual patients at different time points is still limited, but cross-sectional data from individuals diagnosed with KdVS of different ages allowed the modeling of the natural history of 4 important aspects of NDDs: speech, language, reading- and writing ability. The results regarding speech and language in KdVS were encouraging: children continued to develop verbal speech and language after the delayed onset of first words, with the majority of respondents noting that children could speak using well-formed sentences as teenagers. Yet the clarity of speech or "understandability" of individuals with KdVS (**Figure 2**) was significantly reduced with only 48.1% of the persons over 6 years of age could be understood by strangers. These findings are in line with detailed speech and language phenotyping studies of children with KdVS that used standardized assessments which also show a high percentage of speech production disorders associated with poor speech intelligibility[16,25]. Writing skills appear limited even at a later age and could be impacted by the fine motor skills problems reported in the KdVS (42.0%), a finding which is again in line with the previous phenotyped cohorts[16,25].

***Information on adverse effects of drugs*.** For epilepsy, GenIDA confirms earlier data based on smaller cohorts with respect to the proportion of KdVS individuals with epilepsy[26]. GenIDA also shows data on the perceived efficacy of drugs and information on their adverse effects stressing that the open answers from caregivers are a rich source of information (**Figure 3** and **Supplementary Material 4**). Based on the current data, it might be cautiously suggested that valproate and oxcarbazepine (the latter used less often) exhibit better efficacy and lower adverse effects than levetiracetam, a trend already noted by Myers *et al.* in a smaller cohort[26].

Clinical features that were previously not been reported. Importantly, clinical problems were identified that were previously not been reported in KdVS studies such as childhood respiratory disorders, including recurrent pneumonia and childhood asthma, affecting about 40% of individuals with KdVS (**Figure 5**). These features were often perceived as a major health concern by caregivers, based on the open text

answers (**Figure 5**). Children with neurocognitive impairment often present with chronic or recurrent respiratory problems which can have an important impact on the quality of life. Possible underlying causes are risk of aspiration, insufficient cough, upper airway obstruction and progressive kyphoscoliosis [27,28]. Other contributing factors might be an impaired immune response (which is reported by some caregivers). It is of interest that GWAS studies have identified an association of the *KANSL1* locus with parameters of lung functions[29–31].

Limitations of GenIDA. Some caregivers indicated that the online questionnaire took too long to complete which in turn might have resulted in higher drop off rates. In addition, some GenIDA users encountered problems resulting from poor website ergonomics, mainly in the process of registration and in answering the survey questions. The GenIDA study also faces the limitations inherent in questionnaire studies, including recall bias, missing responses, and variable personal interpretation of questions by respondents.

Conclusion. The analysis of GenIDA information on 237 individuals with KdVS shows that caregiver-driven data collection is an approach that is effective in terms of both global recruitment and centralization of clinical data in rare NDDs. The caregiver derived data is in line with previous KdVS cohort studies and there was no significant genotype-phenotype relation (CNV deletion vs. *KANSL1* pathogenic variant). Comparison of the KdVS cohort with other NDD cohorts in the registry revealed clinical features that are overrepresented in the KdVS cohort including tracheo/laryngomalacia. Modeling of natural history showed a positive development of speech and language over time and relatively good reading ability. In case of epilepsy valproate and oxcarbazepine were reported as effective antiepileptic drugs, whereas levetiracetam may have a relative lack of efficacy. Finally, childhood respiratory disorders, including recurrent pneumonia, is a group of novel clinical conditions that was not reported in KdVS before.
In conclusion, GenIDA is a powerful registry to collect and harness valuable data on rare NDDs, which is essential for genetic counseling, clinical decision making and determining clinical outcome measures. The registry may serve as a starting point for follow-up research into the clinically relevant observations which might result in better management and personalized treatment strategies in NDDs.

Supplementary Material

Supplementary Figure 1: Participation to the GenIDA project.

Supplementary Figure 2: Characteristics of the KdVS cohort in GenIDA.

Supplementary Figure 3: Details on high level of participation to the GenIDA 46-questions questionnaire.

Supplementary Figure 4: Comparison of the GenIDA KdVS data and published clinical data[20,21].

Supplementary Table 1: Reported frequencies of clinical items compared between individuals with the 17q21.31 microdeletion and those with a pathogenic variant in *KANSL1*.

Supplementary Table 2: Reported frequencies of clinical items compared between individuals with KdVS and other individuals diagnosed with a NDD in the GenIDA registry.

Supplementary Material 1: GenIDA questionnaire (version: October 2016).

Supplementary Material 2: Daily operational aspects of the GenIDA registry.

Supplementary Material 3: Orodental issues reported for individuals with KdVS in GenIDA.

Supplementary Material 4: Examples of open-ended responses reporting problems following the use of antiepileptic drugs.

References

1. Burger P, Colin F, Strehle A, et al. GenIDA: an international participatory database to gain knowledge on health issues related to genetic forms of neurodevelopmental disorders. *J Neural Transm*. 2023;130(3):459-471. doi:10.1007/s00702-022-02569-3

2. Maulik PK, Mascarenhas MN, Mathers CD, Dua T, Saxena S. Prevalence of intellectual disability: A meta-analysis of population-based studies. *Res Dev Disabil*. 2011;32(2):419-436. doi:10.1016/j.ridd.2010.12.018

3. Srivastava S, Love-Nichols JA, Dies KA, et al. Meta-analysis and multidisciplinary consensus statement: exome sequencing is a first-tier clinical diagnostic test for individuals with neurodevelopmental disorders. *Genet Med*. 2019;21(11):2413-2421. doi:10.1038/s41436-019-0554-6

4. Iossifov I, O'Roak BJ, Sanders SJ, et al. The contribution of de novo coding mutations to autism spectrum disorder. *Nature*. 2014;515(7526):216-221. doi:10.1038/nature13908

5. Krumm N, Turner TN, Baker C, et al. Excess of rare, inherited truncating mutations in autism. *Nat Genet*. 2015;47(6):582-588. doi:10.1038/ng.3303

6. Coe BP, Stessman HAF, Sulovari A, et al. Neurodevelopmental disease genes implicated by de novo mutation and copy number variation morbidity. *Nat Genet*. 2019;51(1):106-116. doi:10.1038/s41588-018-0288-4

7. Kaplanis J, Samocha KE, Wiel L, et al. Evidence for 28 genetic disorders discovered by combining healthcare and research data. *Nature*. 2020;586(7831):757-762. doi:10.1038/s41586-020-2832-5

8. Satterstrom FK, Kosmicki JA, Wang J, et al. Large-Scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell*. 2020;180(3):568-584.e23. doi:10.1016/j.cell.2019.12.036

9. Popp B, Ekici AB, Thiel CT, et al. Exome Pool-Seq in neurodevelopmental disorders. *Eur J Hum Genet*. 2017;25(12):1364-1376. doi:10.1038/s41431-017-0022-1

10. Kochinke K, Zweier C, Nijhof B, et al. Systematic Phenomics Analysis Deconvolutes Genes Mutated in Intellectual Disability into Biologically Coherent Modules. *Am J Hum Genet*. 2016;98(1):149-164. doi:10.1016/j.ajhg.2015.11.024

11. SysNDD database. SysNDD database. Accessed September 7, 2022. https://sysndd.dbmr.unibe.ch/

12. Garbade SF, Zielonka M, Komatsuzaki S, et al. Quantitative retrospective natural history modeling for orphan drug development. *J Inherit Metab Dis*. 2021;44(1):99-109. doi:10.1002/jimd.12304

13. Koolen DA, Vissers LELM, Pfundt R, et al. A new chromosome 17q21.31 microdeletion syndrome associated with a common inversion polymorphism. *Nat Genet*. 2006;38(9):999-1001. doi:10.1038/ng1853

14. Sharp AJ, Hansen S, Selzer RR, et al. Discovery of previously unidentified genomic disorders from the duplication architecture of the human genome. *Nat Genet*. 2006;38(9):1038-1042. doi:10.1038/ng1862

15. Shaw-Smith C, Pittman AM, Willatt L, et al. Microdeletion encompassing MAPT at chromosome 17q21.3 is associated with developmental delay and learning disability. *Nat Genet*. 2006;38(9):1032-1037. doi:10.1038/ng1858

16. Morgan AT, Haaften L van, Hulst K van, et al. Early speech development in Koolen de Vries syndrome limited by oral praxis and hypotonia. *Eur J Hum Genet*. 2018;26(1):75-84. doi:10.1038/s41431-017-0035-9

17. Koolen DA, Morgan A, de Vries BB. Koolen-de Vries Syndrome. In: Adam MP, Ardinger HH, Pagon RA, et al., eds. *GeneReviews*®. University of Washington, Seattle; 2010. Accessed March 23, 2022. http://www.ncbi.nlm.nih.gov/books/NBK24676/

18. Koolen DA, Kramer JM, Neveling K, et al. Mutations in the chromatin modifier gene KANSL1 cause the 17q21.31 microdeletion syndrome. *Nat Genet*. 2012;44:639-641.

19. Zollino M, Orteschi D, Murdolo M, et al. Mutations in KANSL1 cause the 17q21.31 microdeletion syndrome phenotype. *Nat Genet*. 2012;44(6):636-638. doi:10.1038/ng.2257

20. Zollino M, Marangi G, Ponzi E, et al. Intragenic KANSL1 mutations and chromosome 17q21.31 deletions: broadening the clinical spectrum and genotype–phenotype correlations in a large cohort of patients. *J Med Genet*. 2015;52(12):804-814. doi:10.1136/jmedgenet-2015-103184

21. Koolen DA, Pfundt R, Linda K, et al. The Koolen-de Vries syndrome: a phenotypic comparison of patients with a 17q21. 31 microdeletion versus a KANSL1 sequence variant. *Eur J Hum Genet*. 2016;24(5):652-659.

22. Cross-sectional vs. longitudinal studies. Accessed January 11, 2023. https://www.iwh.on.ca/what-researchers-mean-by/cross-sectional-vs-longitudinal-studies

23. Boerma T, Ronsmans C, Melesse DY, et al. Global epidemiology of use of and disparities in caesarean sections. *The Lancet*. 2018;392(10155):1341-1348. doi:10.1016/S0140-6736(18)31928-7

24. WHO Multicentre Growth Reference Study Group. WHO Motor Development Study: windows of achievement for six gross motor development milestones. *Acta Paediatr Oslo Nor 1992 Suppl*. 2006;450:86-95. doi:10.1111/j.1651-2227.2006.tb02379.x

25. St John M, van Reyk O, Koolen DA, de Vries BBA, Amor DJ, Morgan AT. Expanding the speech and language phenotype in Koolen-de Vries syndrome: late onset and periodic stuttering a novel feature. *Eur J Hum Genet EJHG*. Published online December 19, 2022. doi:10.1038/s41431-022-01230-7

26. Myers KA, Mandelstam SA, Ramantani G, et al. The epileptology of Koolen-de Vries syndrome: Electro-clinico-radiologic findings in 31 patients. *Epilepsia*. 2017;58(6):1085-1094. doi:10.1111/epi.13746

27. Morgan AT, Dodrill P, Ward EC. Interventions for oropharyngeal dysphagia in children with neurological impairment. *Cochrane Database Syst Rev*. 2012;10:CD009456. doi:10.1002/14651858.CD009456.pub2

28. Proesmans M. Respiratory illness in children with disability: a serious problem? *Breathe Sheff Engl*. 2016;12(4):e97-e103. doi:10.1183/20734735.017416

29. Wain LV, Shrine N, Artigas MS, et al. Genome-wide association analyses for lung function and chronic obstructive pulmonary disease identify new loci and potential druggable targets. *Nat Genet*. 2017;49(3):416-425. doi:10.1038/ng.3787

30. Wain LV, Shrine N, Miller S, et al. Novel insights into the genetics of smoking behaviour, lung function, and chronic obstructive pulmonary disease (UK BiLEVE): a genetic association study in UK Biobank. *Lancet Respir Med*. 2015;3(10):769-781. doi:10.1016/S2213-2600(15)00283-0

31. Degenhardt F, Ellinghaus D, Juzenas S, et al. New susceptibility loci for severe COVID-19 by detailed GWAS analysis in European populations. Published online July 23, 2021:2021.07.21.21260624. doi:10.1101/2021.07.21.21260624

32. Phenodent.org - The project. Accessed September 23, 2021. http://www.phenodent.org/indexgb.php

Captions for Figures and Tables

Figures

Figure 1: Pregnancy, newborn and early life problems (data collection period: Nov. 2016 – Feb. 2022). A. Frequencies of issues / complications during pregnancy, labor and delivery for individuals with KdVS and their mother [percentages were calculated for those with a reported problem, i.e., response = 'Yes]. **B.** APGAR scores at 1, 5 and 10min. **C.** Frequencies of newborn issues. **D.** Developmental milestones in months (sitting, standing, walking alone). For comparison, medians from the WHO (World Health Organization) are indicated in the table in brackets[24].

Figure 2: Speech ability (data collection period: Nov. 2016 – Feb. 2022). A. Reported age at first word in years for individuals with KdVS (M/F 49%/51%; mean age 14.0y), compared to individuals with Kleefstra (M/F 47%/53%; mean age 13.9y) and KBG syndrome (M/F 61%/39%; mean age 14.4y) respectively. **B.** Speech ability for individuals with KdVS (age ≥2y) as reported by families. The 25% and 50% thresholds for 'full correct sentences' are indicated by arrows in the longitudinal representation (below): 10% of the data points derived from answers concerning a given individual diagnosed with a NDD at 2 age ranges. **C.** Reported understandability (age ≥6y) of individuals diagnosed with a NDD by family members and by strangers.

Figure 3: Epilepsy (seizures) in KdVS (data collection period: Nov. 2016 – Feb. 2022). A. Frequency of epilepsy in individuals with KdVS (n=204) and reported types of seizures and their perceived severity [percentages were calculated as follows: number of events reported (major + moderate + minor) / number of individuals who answered 'yes' to the question 'Epilepsy', i.e., n=97]. **B.** Age at first seizure (age of onset) compared to current age of 80 well described epileptic individuals with KdVS. **C.** Summary of drugs' use, perceived efficacy, and corresponding frequencies of adverse events [percentages were calculated only when drug use was reported for at least 5 persons]. [Anterior data freeze (Nov. 2016 – Aug. 2019): 'no medication' was reported for 14 persons (as epilepsy resolved on its own) or mentioned only 'sporadic events'; treatments were not specified for 13 persons who were thus excluded from the table] (see also examples of open text answers in **Supplementary Material 4**).

Figure 4: Behavioral problems and their perceived severity in KdVS, and comparison with KS and KBGS (data collection period: Nov. 2016 – Feb. 2022). A. Frequencies of various behavioral problems

in individual with KdVS and their perceived severity [percentages were calculated as follows: number of events reported (major + moderate + minor event) / number of individuals who answered to the question 'behavior problem', i.e., n=199]. **B.** Overall frequencies of behavioral problems reported for individuals with KdVS, KS and KBGS.

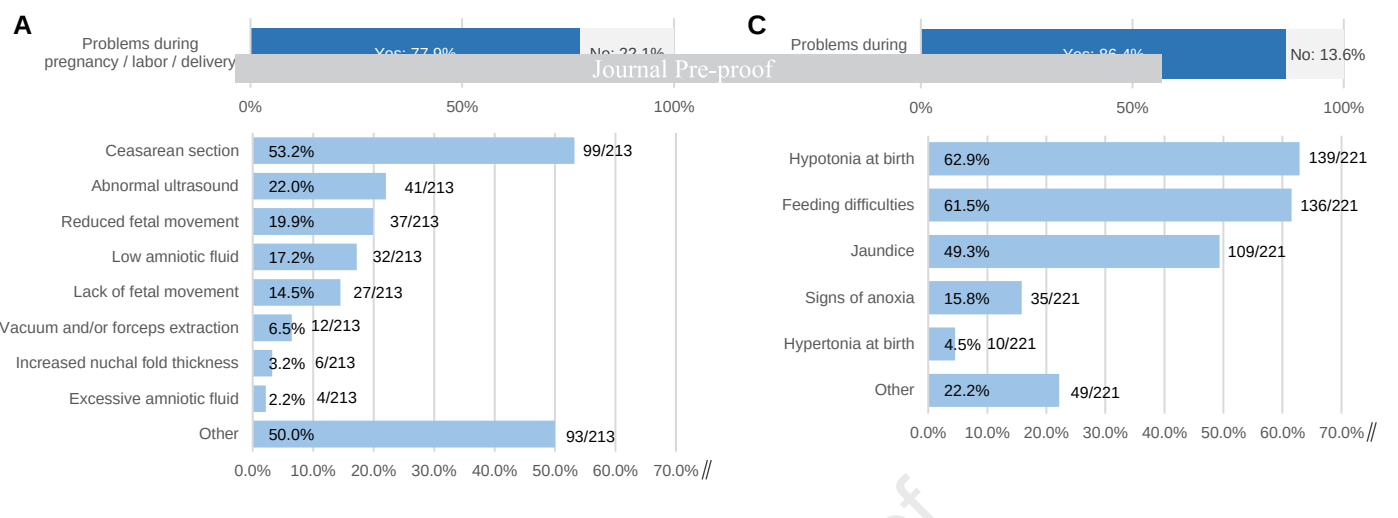
Figure 5: Respiratory problems reported by individuals with KdVS (data collection period: Nov. 2016 – Feb. 2022). A. Types and frequencies of respiratory problems reported (laryngo- and tracheomalacia were the only manifestations previously reported in literature). **B.** Examples of open-ended answers reporting respiratory problems. **C.** Age at earliest and latest pneumonia events from well documented individuals diagnosed with KdVS (data collected after re-contacting families).

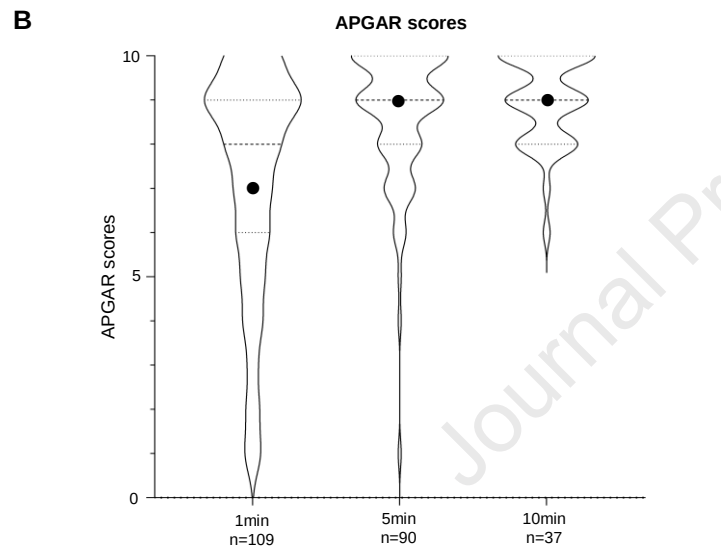
Tables

Table 1: Most reported clinical features for KdVS in GenIDA (all "I don't know" answers were excluded; data collection period: Nov. 2016 – Feb. 2022).

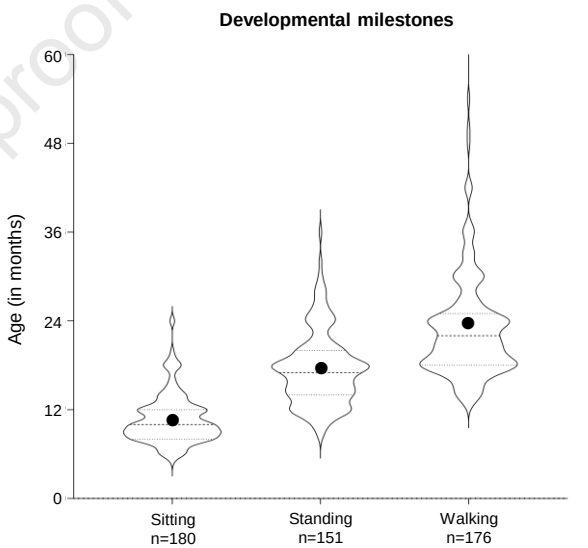
Table 1: Most reported clinical features for KdVS in GenIDA (all "I don't know" answers were excluded; data collection period: Nov. 2016 – Feb. 2022).

	Yes	Total	%
Prenatal and neonatal history			
Low amniotic fluid	32	213	15.0
Caesarean section	99	213	46.5
Abnormal ultrasound results	41	213	19.2
Jaundice	109	226	48.2
Signs of anoxia	35	226	15.5
Hypotonia	139	226	61.5
Feeding difficulties	136	226	60.2
Psychomotor development			
Fine motor problems, clumsiness	68	162	42.0
Walking problems (children over 2y)	60	153	39.2
Speech and language delay (children over 2y)	128	174	73.6
Diagnosis of intellectual disability (children over 6y)	165	185	89.2
Neurologic and psychiatric features			
Epilepsy	97	204	47.5
Behavioral problems	109	199	54.8
Vision & hearing			
Hypermetropia	76	196	38.8
Strabismus	68	196	34.7
Hearing impairment	24	196	12.2
Recurrent ear infections	34	196	17.3
Cardiovascular system	81	201	40.3
Atrial septal defect (ASD)	38	201	18.9
Ventricular septal defect (VSD)	22	201	10.9
Respiratory and pulmonary system	80	201	39.8
Asthma	33	201	16.4
Laryngomalacia	31	201	15.4
Tracheomalacia	17	201	8.5
Digestive / excretory system			
Swallowing difficulties	23	143	16.1
Hypersalivation	39	143	27.3
Constipation	60	193	31.1
Urogenital system			
Renal / kidney, bladder and urogenital system problems	74	193	38.3
Musculoskeletal system	151	200	75.5
Joint laxity	100	200	50.0
Hip dysplasia	36	200	18.0
Pes planus	44	200	22.0
Scoliosis	51	200	25.5
Miscellaneous			
Sleeping disorders	75	176	42.6
Skin, nails and hair problems	104	201	51.7
High number of moles	44	201	21.9
Dental problems	112	172	65.1
Endocrine and metabolic systems' problems	31	177	17.5
Blood and immune system problems	88	206	42.7





	APGAR1	APGAR5	APGAR10
Average (●)	**7**	**9**	**9**
Median	**8**	**9**	**9**
Upper quartile	9	10	10
Lower quartile	6	8	8
Group size	**109**	**90**	**37**



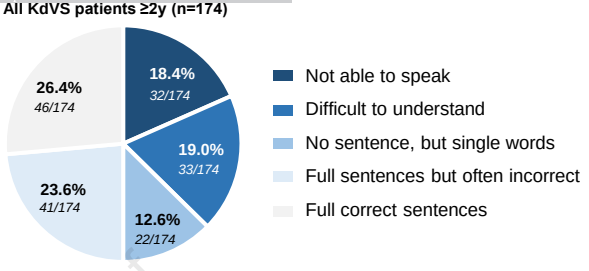
	Sitting	Standing	Walking
Average (●)	**10.7** (6.0)	**17.5** (11.0)	**23.4** (12.0)
Median	**10.0**	**17.0**	**22.0**
Upper quartile	12.0	20.0	25.0
Lower quartile	8.0	14.0	18.0
Group size	**180**	**151**	**176**

Figure 1: Pregnancy, newborn and early life problems (data collection period: Nov. 2016 – Feb. 2022). **A.** Frequencies of issues / complications during pregnancy, labor and delivery for individuals with KdVS and their mother [percentages were calculated for those with a reported problem, i.e., response = 'Yes]. **B.** APGAR scores at 1, 5 and 10min. **C.** Frequencies of newborn issues. **D.** Developmental milestones in months (sitting, standing, walking alone). For comparison, medians from the WHO (World Health Organization) are indicated in the table in brackets[24].

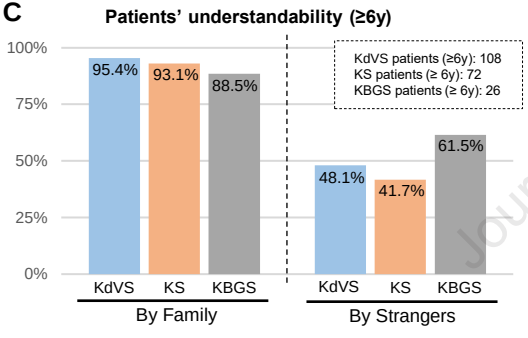
A

Age at first word (y)	KdVS	KS	KBGS
Average	**2.2**	**2.2**	**1.6**
Median	2.0	2.0	1.5
Upper quartile	3.0	3.0	2.0
Lower quartile	1.2	1.2	1.0
Number of answers	**n=139**	**n=95**	**n=35**

B



C



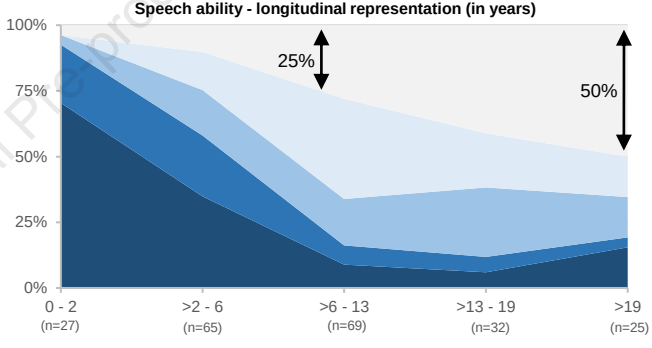
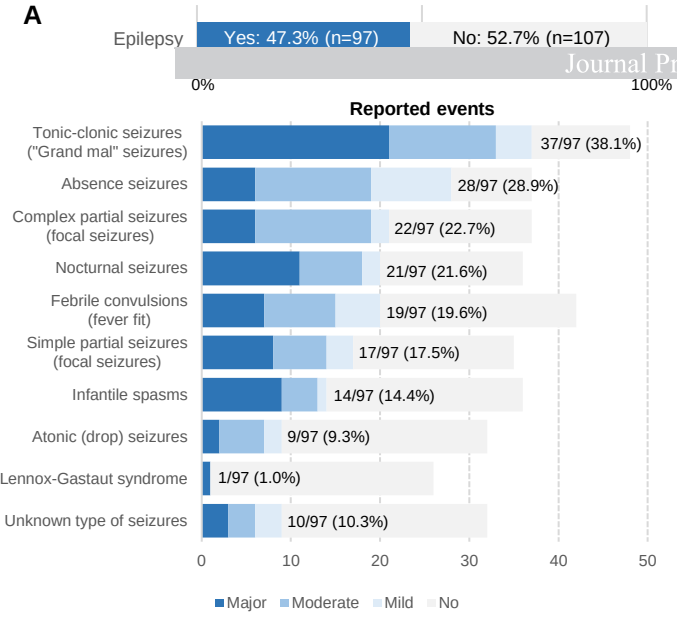
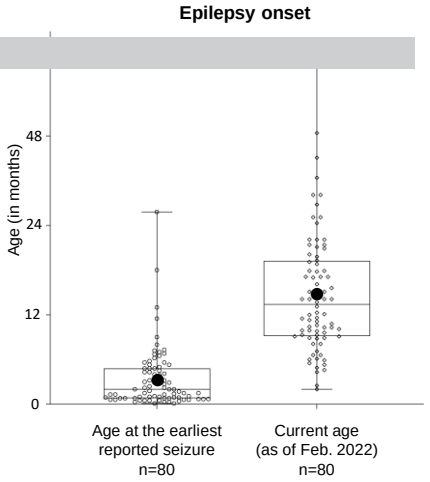


Figure 2: Speech ability (data collection period: Nov. 2016 – Feb. 2022). A. Reported age at first word in years for individuals with KdVS (M/F 49%/51%; mean age 14.0y), compared to individuals with Kleefstra (M/F 47%/53%; mean age 13.9y) and KBG syndrome (M/F 61%/39%; mean age 14.4y) respectively. **B.** Speech ability for individuals with KdVS (age ≥2y) as reported by families. The 25% and 50% thresholds for 'full correct sentences' are indicated by arrows in the longitudinal representation (below): 10% of the data points derived from answers concerning a given individual diagnosed with a NDD at 2 age ranges. **C.** Reported understandability (age ≥6y) of individuals diagnosed with a NDD by family members and by strangers.





	First event	Current age
Average (●)	**3.4y**	**14.8y**
Median	2.0y	13.4y
Maximum	25.8y	47.6y
Minimum	0.1y	2.0y

C

	Drug use reports	Good efficiency		Adverse events	
	n	n/tot	%	n/tot	%
Anti-epilectic drug					
Levetiracetam	**30**	20/30	67%	14/30*	**47%**
Valproate	**24**	20/24	**83%**	10/24	**42%**
Oxcarbazepine	**13**	11/13	**85%**	1/13	8%
Topiramate	6	3/6	50%	2/6	27%
Carbamazepine	5	2/5	40%	3/5*	**60%**
Lamotrigine	5	4/5	**80%**	0/5	0%
Zonisamide	4	4/4		0/4	
Phenobarbital	3	2/3		0/3	
Lacosamide	2	2/2		0/2	
Phenytoin	2	2/2		0/2	
ACTH	1	0/1		1/1	
Vigabatrin	1	1/1		0/1	
Benzodiazepine					
Clobazam	7	6/7	**86%**	4/7	**57%**
Diazepam	4	4/4		1/4*	
Midazolam	3	3/3		0/3	
Clonazepam	2	1/2		1/2*	
Lorazepam	2	2/2		1/2	
Others / Add-on					
Alimemazine	1	0/1		0/1	
Hydrocortisone	1	1/1		0/1	
Prednisone	1	0/1		0/1	
Ketogenic diet	3	1/3		0/3	

** Rare major adverse events reported*

Figure 3: Epilepsy (seizures) in KdVS (data collection period: Nov. 2016 – Feb. 2022). A. Frequency of epilepsy in individuals with KdVS (n=204) and reported types of seizures and their perceived severity [percentages were calculated as follows: number of events reported (major + moderate + minor) / number of individuals who answered 'yes' to the question 'Epilepsy', i.e., n=97]. **B.** Age at first seizure (age of onset) compared to current age of 80 well described epileptic individuals with KdVS. **C.** Summary of drugs' use, perceived efficacy, and corresponding frequencies of adverse events [percentages were calculated only when drug use was reported for at least 5 persons]. [Anterior data freeze (Nov. 2016 – Aug. 2019): 'no medication' was reported for 14 persons (as epilepsy resolved on its own) or mentioned only 'sporadic events'; treatments were not specified for 13 persons who were thus excluded from the table] (see also examples of open text answers in **Supplementary Material 4**).

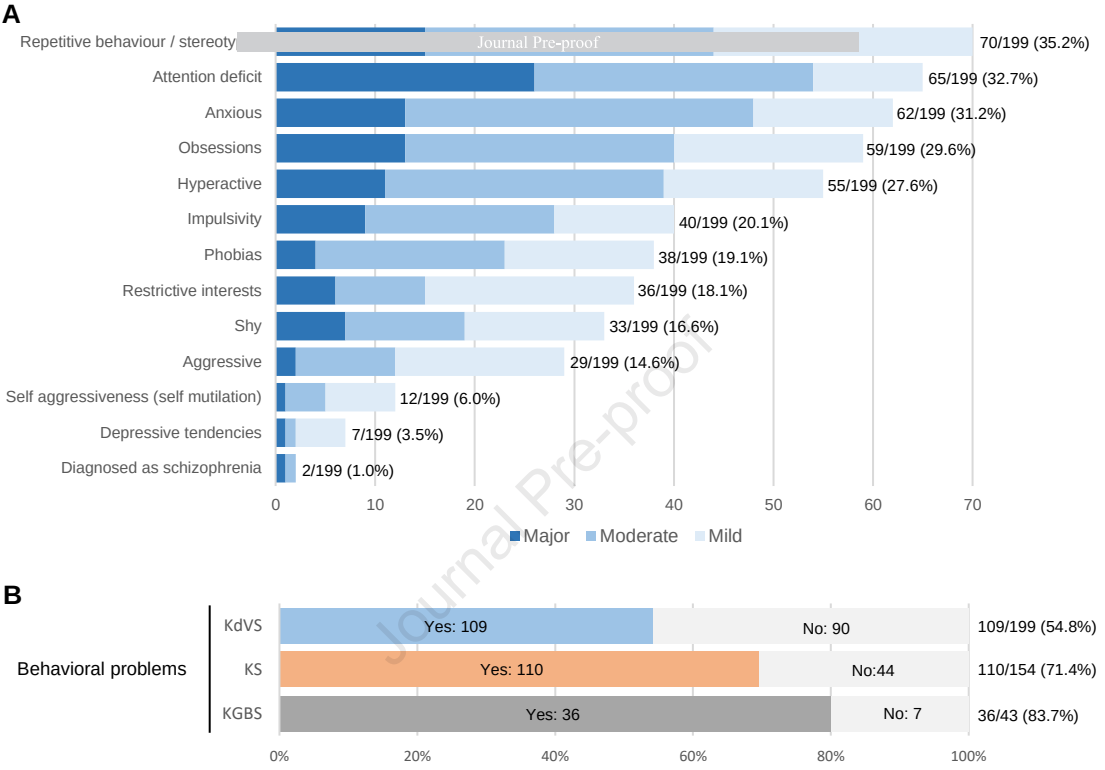
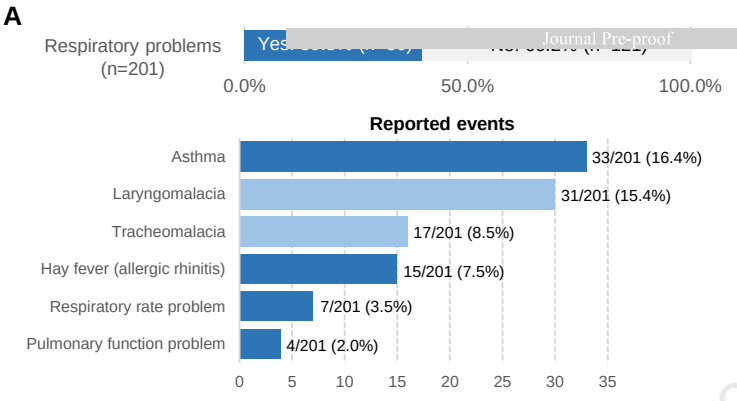
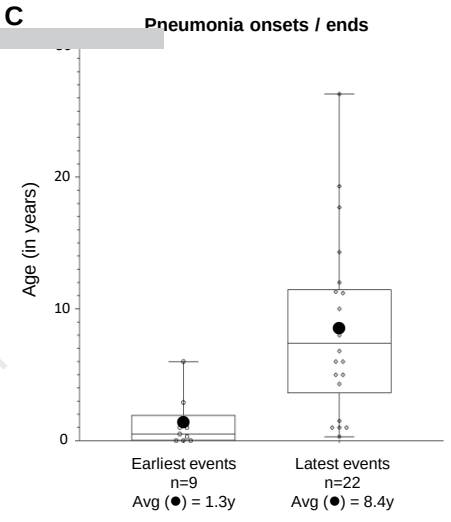


Figure 4: Behavioral problems and their perceived severity in KdVS, and comparison with KS and KBGS (data collection period: Nov. 2016 – Feb. 2022). A. Frequencies of various behavioral problems in individual with KdVS and their perceived severity [percentages were calculated as follows: number of events reported (major + moderate + minor event) / number of individuals who answered to the question 'behavior problem', i.e., n=199]. **B.** Overall frequencies of behavioral problems reported for individuals with KdVS, KS and KBGS.





B

Male – 2.0yo	RSV virus bronchiolitis/**pneumonia** and hospitalized for a week at 3 months old. Dx **Restrictive Airway Disease (asthma for babies)**
Female – 6.9yo	**Laryngomalacia (0-2 years old) Croup/bronchitis/ pneumonia (0-ongoing in 2018)**
Female – 7.7yo	1-5 years old: Immune problems, **recurring bouts of pneumonia**, a few which ended up with her in the hospital
Male – 8.0yo	**Many bouts of aspiration pneumonia**. Always hospitalised for **severe croup requiring nebuliser adrenaline and other medications**
Female – 8.3yo	Last year a **pneumonia with hospitalization at 6**
Female – 10.0yo	Until the age of 6 she had **several pneumonias and bronchial problems**. She has recovered now.
Male – 10.4yo	[…] **asthma (infant), repeated pneumonias**, Otitis media (birth), […]
Male – 12.9yo	**Pulmonology Asthma w/acute exacerbation, Obstructive Sleep Apnea and re-occuring bacterial and viral Pneumonia**
Female – 18.2yo	[…] a de l'épilepsie, des **problèmes respiratoires (bronchites, laryngites, pneumonie, asthme)**

Figure 5: Respiratory problems reported by individuals with KdVS (data collection period: Nov. 2016 – Feb. 2022). A. Types and frequencies of respiratory problems reported (laryngo- and tracheomalacia were the only manifestations previously reported in literature). **B.** Examples of open-ended answers reporting respiratory problems. **C.** Age at earliest and latest pneumonia events from well documented individuals diagnosed with KdVS (data collected after re-contacting families).