

Original Article

Reproducible Coding of Admission Complaints in Pediatric Rehabilitation: A Dictionary-Based Study

Syed Gufran Sadiq Zaidi¹, Saule Karibzhanova¹¹Nazarbayev University, School of Medicine, Astana, Kazakhstan

Received: Oct 13, 2026

Accepted: Oct 18, 2026

Corresponding author's email:

sgs.zaidi@nu.edu.kz

Citation: Zaidi SGS, Karibzhanova S. Reproducible Coding of Admission Complaints in Pediatric Rehabilitation: A Dictionary-Based Study. *Epidemiology and Health Data Insights*. 2026;2(5):ehdi059. <https://doi.org/10.63946/ehdi/19672>

Copyright: By the author(s).

License: A non-exclusive license by the publisher. Published by Australasia Publishing Group LLP.

Open Access: This article is an open access article distributed under the terms and conditions of the CC-BY Creative Commons Attribution license <https://creativecommons.org/licenses/by/4.0>



ABSTRACT

Background: Family-reported complaints at admission shape rehabilitation planning, but they are usually stored as free text and are rarely analyzed.

Methods: We analyzed the admission complaint text of 318 children admitted to a national pediatric rehabilitation center in one admission course (22 August–8 September 2025), written in Russian in discharge summaries. We coded the text into ten complaint categories with a rule-based dictionary, compared the coding with hand coding of 209 children by a single clinician (including a test set of 40 children not used to build the dictionary), compared the categories by diagnosis group with permutation tests, and explored complaint profiles with non-negative matrix factorization of lemmatized text.

Results: On the test set the dictionary agreed with the clinician's hand coding on 98.3% of category decisions (95% CI 97.0 to 99.5%). Speech and language complaints were most frequent (73%), followed by motor (59%) and social and behavioral (36%) complaints. Motor complaints were recorded in 98% of children with cerebral palsy, and 61% of these children also had speech or language complaints. Categories differed by diagnosis (all $p < 0.05$ except learning and memory), but motor complaints were also frequent in Down syndrome and other diagnoses. Three complaint profiles were found.

Conclusions: Family-reported complaints in routine records can be coded reproducibly, and speech and language concerns were frequent in children admitted for motor disorders. Agreement was measured against a single clinician and may overestimate agreement with other clinicians.

Keywords: Parents; Child; Neurodevelopmental Disorders; Rehabilitation; Natural Language Processing; Data Accuracy

Introduction

When a child is admitted to a rehabilitation center, the first thing the physician records is why the family has come. These complaints about admission constitute the family's description of their primary concerns and form the initial treatment plan and the objectives to be developed. Modern pediatric rehabilitation is based on the premise of family-centered service, which is founded on three principles: first, the family knows the child best and secondly, services should reflect the families' priorities [1]. However, these complaints are generally recorded as an unstructured paragraph in the medical record. They are read by treating team and rarely counted, compared or analyzed so a center might know what diagnosis it sees, but might not have a clear picture of what families ask for.

This matters because the diagnosis and the complaint are not the same thing. Cerebral palsy is primarily characterized by the disorder of movement and posture but the consensus-based definition emphasizes that the motor disorder also often manifests as disorders of sensation, perception, cognition, communication and behavior and that epilepsy is also a feature [2]. Register data show that communication impairments and oromotor dysfunction are common in children with cerebral palsy [3]. Systematic reviews of interventions for cerebral palsy likewise cover many goals beyond motor function, including communication and behavior [4,5]. Other diagnoses show the same overlap. Speech impairment is a prominent feature of Down syndrome [6], and children with autism spectrum disorder have more feeding problems than their peers [7]. A family's account of its main concerns may therefore tell a rehabilitation service something that the coded diagnosis does not, for example how much speech and language therapy will be needed in a ward where cerebral palsy is the largest diagnosis group. The same information could also help centers decide how many speech, behavioral and feeding specialists a rehabilitation team needs, and could help families and referring clinicians understand what the center usually treats beyond the headline diagnosis. For these uses the complaints must first be turned into counts, and the way this is done must be clear enough for others to judge.

Analyzing complaints at scale is limited by the language of the records. In Russian and other highly inflected languages a single concept appears in many word forms, because endings change with grammatical case, number and gender, so counting words splits one complaint into many. Extracting information from clinical free text is a well-developed field, and reviews describe both rule-based and machine-learning

systems, with most work done in English [8,9]. Clinical language processing in other languages has fewer shared resources, and the authors of a review on this topic point to the need for openly available methods that other groups can reuse and check [10]. Rule-based classifiers have a long record in coding free-text chief complaints for surveillance [11]. Dictionary-based coding with lemmatization is a simple option for this situation. It is transparent, every decision can be traced to a word, it needs no training data, and it can be shared and rerun. Its weakness is that it depends on the dictionary, so its agreement with a reference must be measured and reported. A further weakness is that physicians write complaints in their own style, with abbreviations, informal wording and, in this setting, a mixture of the parents' words and the doctor's, so any dictionary will miss some phrasings and must be checked on text it has not seen.

Establishing that agreement is complicated. The conventional approach is to choose standard values of chance-corrected agreement typically Cohen's kappa [12] and reference them to the benchmarks of Landis and Koch [13]. The level of kappa used depends on the proportion of the category [14] and may be low even if almost all children were answered by the same two coders if one answer predominated, thus it should always be accompanied by the raw proportions in each direction. Both solution providers and language interpreters must understand the language to provide human double coding, which is optimal.

Finally, a description of complaint categories can be complemented by another unsupervised look at the text itself. Topic models such as non-negative matrix factorization find groups of words that occur together, and they allow the question of whether the complaints naturally form a small number of profiles to be asked without imposing categories in advance. The number of topics is a choice that is often made informally, but it can be guided by how stable the solutions are across random starts.

We therefore studied the admission complaints of all children admitted to a national pediatric rehabilitation center in one admission course in August–September 2025. We asked four questions: which complaints do families report; how do they differ by diagnosis; how closely the complaints follow the coded diagnosis; and what complaint profiles appear in the text without supervision. We also aimed to provide a fully reproducible, openly shared coding method for Russian-language complaint text, with its agreement against a hand-coded reference set reported in full.

Materials and Methods

Setting and data: We used the discharge summaries of 318 children admitted to a national pediatric rehabilitation center in one admission course between 22 August and 8 September 2025 (282 of the 318 children, 89%, were admitted on 22 August, and most were discharged on 15 September). Each summary contains a section on complaints at admission, written by the physician in Russian, often prefaced by who gave the history (for example, the mother), so the text is the physician's record of what the family reported and not a verbatim transcript. Diagnosis group (cerebral palsy, speech/language delay, autism spectrum disorder, Down syndrome, other), age and sex were taken from a cleaned analysis dataset built from the same summaries; sex was blank in the source summaries for 66 children and was left as missing. The study used anonymized records and was approved by an institutional ethics committee (see Ethics).

Coding of complaints: We defined ten categories including speech and language; motor; attention and activity; social and behavior; seizures; feeding and growth; vision and hearing; continence; fatigue and learning and memory. Each category is a list of Russian word stems and phrases (the English meanings are in Supplementary Table S1). A child was coded 1 for a category if any term occurred in the complaint text. Before matching, the phrase "according to" (со слов), which contains a word stem used for speech, and phrases describing stereotyped or repetitive movements were removed so that they would not be counted as speech or motor complaints.

Reference coding and validation: One Russian-speaking clinician did the reference coding by reading the Russian complaint text of 209 children and assigned categories by hand, each child being coded before the dictionary output for that child was shown. This was done in stages: 169 children were used to build and refine the dictionary (the development set), and a final set of 40 children, drawn at random from children not yet read, was coded before the dictionary was used on it and was not used to change the dictionary (the test set). We report precision, recall and Cohen kappa [12] for each category, and overall agreement across all child-category decisions (Supplementary Tables S2 and S3). Because most child-category decisions are negative, which inflates raw agreement, we also report kappa, macro-averaged F1 (the mean of the ten per-category F1 scores) and balanced accuracy (the mean of sensitivity and specificity), with 95% percentile bootstrap confidence intervals over children (2,000 resamples; Supplementary Table S6). No second human reader

checked the reference coding; this is a limitation (see Limitations).

Statistical analysis: We described category frequency overall and by diagnosis group. Differences between diagnosis groups were tested with a permutation chi-square test (10,000 permutations, seed 2025; $p = (b + 1)/(10,000 + 1)$, so no p value is zero [15]), and p values were adjusted for ten categories with the Benjamini–Hochberg procedure [16]. To see how well complaints agree with the coded diagnosis, we calculated for three diagnosis–complaint pairs (cerebral palsy and motor, speech/language delay and speech and language, autism spectrum disorder and social and behavior) the proportion of children with the diagnosis who had the complaint, the proportion of children without it who did not have the complaint, the proportion of children with the complaint who had the diagnosis, and kappa with a 95% percentile bootstrap confidence interval (2,000 resamples) [17,18]. This kappa measures overlap between diagnosis and complaint, not coding quality. Informants were identified from phrases in the text (mother, father, grandmother, caregiver; when more than one was named, the first in this order was used), and the number of categories per child was compared across stated informants with the Kruskal–Wallis test [19]. Children with and without a recorded sex were compared for age (Mann–Whitney test), diagnosis group (permutation chi-square test) and complaint categories (Fisher exact test with Benjamini–Hochberg adjustment) (Supplementary Table S7). All tests are two-sided and exploratory.

Complaint profiles: The text was lemmatized with the pymorphy3 Russian morphological analyzer [20], converted to term frequency–inverse document frequency features [21] (single words and word pairs, minimum document frequency 6), and decomposed with non-negative matrix factorization [22]. Word forms that the analyzer lemmatized wrongly because they are ambiguous (for example, "стоит", 'stands', read as 'costs') were corrected with a short list of overrides, and physicians' surnames were removed from the vocabulary. The number of topics (2 to 7) was chosen using a stability criterion adapted from Greene et al. [23]. Stability was defined as the mean best-match Jaccard overlap between the top ten terms of a reference solution and those of ten solutions started from random initializations (stability for each solution is in Supplementary Table S4). Only solutions with at least three topics were eligible, so that the profiles could be more than a single speech-versus-motor contrast; the two-topic solution was therefore excluded whatever its stability. Among the eligible solutions the one with the

highest stability was chosen, and the smallest number of topics if solutions were tied. As a sensitivity analysis the two-, three- and four-topic solutions are compared side by side (Supplementary Table S8). Each child was assigned to the topic with the highest weight.

Reproducibility: All steps run from two input files (the extracted summaries and the cleaned dataset) with one command; random seeds are fixed, and the package versions are listed. All text analysis was done directly on the Russian text, without machine translation of the complaints.

Results

The 318 children had a median age of 5.8 (3.7 to 10.0) years; the diagnosis groups were cerebral palsy (121), speech/language delay (79), autism spectrum disorder (36), Down syndrome (18) and other (64) (Table 1). Sex was not recorded for 66 children (21%);

Data and code availability: The child-level clinical dataset and the complaint texts concern minors and are confidential, so they are not publicly available. They can be requested from the corresponding author, subject to the approval of the center and the ethics committee. The analysis code, the term dictionary with its English glossary and the hand-coded reference labels can also be requested from the corresponding author.

these children did not differ from the others in age (median 5.4 vs 5.9 years), diagnosis group ($p=0.99$) or any complaint category (all adjusted $q>0.9$; Supplementary Table S7).

Table 1. Children included.

Item	Value
Children	318
Age, years, median (IQR)	5.8 (3.7 to 10.0)
Male	214
Female	38
Sex not recorded	66
Complaint categories per child, median (IQR)	2 (2 to 3)
Diagnosis: Cerebral palsy	121
Diagnosis: Speech/language delay	79
Diagnosis: Autism spectrum disorder	36
Diagnosis: Down syndrome	18
Diagnosis: Other	64

Validation: On the development set the dictionary agreed with hand coding on 99.6% of category decisions; this is optimistic, because the dictionary was built on these children; on the test set it agreed on 393 of 400 decisions (98.3%, 95% CI 97.0 to 99.5%), and for 34 of 40 children (85%) all ten categories matched (Table 2). Because 300 of the 400 test-set decisions were negative, we also calculated kappa (0.95, 95% CI 0.91 to 0.99), macro-averaged F1 (0.97, 0.91 to

0.99) and balanced accuracy (0.97, 0.93 to 0.99) (Supplementary Table S6); per-category counts are in Supplementary Table S3. Precision was 1.00 or close to it in every category. Recall was lowest for feeding and growth (0.70) and vision and hearing (0.83), where the test set contained wording not in the dictionary (for example, a child who did not chew, and a child who did not react to loud sounds). In the full cohort, four children had no category matched.

Table 2. Agreement of the dictionary with hand coding in the test set (40 children not used to build the dictionary).

Category	Hand-coded (n)	Dictionary (n)	Precision	Recall	Kappa
Speech and language	27	27	1.00	1.00	1.00
Motor	27	27	0.96	0.96	0.89
Attention and activity	4	4	1.00	1.00	1.00

Category	Hand-coded (n)	Dictionary (n)	Precision	Recall	Kappa
Social and behavior	14	13	1.00	0.93	0.94
Seizures	1	1	1.00	1.00	1.00
Feeding and growth	10	7	1.00	0.70	0.78
Vision and hearing	6	5	1.00	0.83	0.89
Continence	6	6	1.00	1.00	1.00
Fatigue	2	2	1.00	1.00	1.00
Learning and memory	3	3	1.00	1.00	1.00

Development set (169 children): 99.6% of category decisions agree; test set: 98.3% (393 of 400 decisions; 95% CI 97.0 to 99.5%). Hand coding was done by a Russian-speaking clinician. True- and false-positive counts, the development-set results and cell-level totals are in Supplementary Tables S2 and S3.

Frequency: Speech and language complaints were recorded for 233 children (73%), motor complaints for 188 (59%), social and behavior complaints for 115 (36%), attention for 77 (24%), feeding and growth for 60 (19%), vision and hearing for 34 (11%), continence for 29 (9%), fatigue for 27 (8%), learning and memory for 19 (6%) and seizures for 9 (3%). The most common combinations were motor alone (39 children), speech and language with motor (37), and speech and language with attention and social and behavior (31).

By diagnosis (Table 3, Figure 1), categories differed between groups after correction for multiple

testing, except learning and memory. Motor complaints were recorded for 98% of children with cerebral palsy, but also for 67% of children with Down syndrome and 73% of the other group. Speech and language complaints were recorded for 61% of children with cerebral palsy (74 of 121). Attention complaints were frequent in speech/language delay (58%) and autism spectrum disorder (50%) and rare in cerebral palsy (3%). Feeding and growth complaints were most frequent in autism spectrum disorder (44%).

Table 3. Complaint categories by diagnosis group, n (%).

Category	All (n=318)	Cerebral palsy (n=121)	Speech/language delay (n=79)	Autism spectrum disorder (n=36)	Down syndrome (n=18)	Other (n=64)	p	q_BH
Speech and language	233 (73%)	74 (61%)	77 (97%)	34 (94%)	15 (83%)	33 (52%)	<0.001	<0.001
Motor	188 (59%)	119 (98%)	9 (11%)	1 (3%)	12 (67%)	47 (73%)	<0.001	<0.001
Attention and activity	77 (24%)	4 (3%)	46 (58%)	18 (50%)	0 (0%)	9 (14%)	<0.001	<0.001
Social and behavior	115 (36%)	16 (13%)	51 (65%)	30 (83%)	7 (39%)	11 (17%)	<0.001	<0.001
Seizures	9 (3%)	3 (2%)	0 (0%)	0 (0%)	0 (0%)	6 (9%)	0.015	0.018
Feeding and growth	60 (19%)	22 (18%)	12 (15%)	16 (44%)	2 (11%)	8 (12%)	<0.001	0.001
Vision and hearing	34 (11%)	19 (16%)	1 (1%)	1 (3%)	4 (22%)	9 (14%)	0.003	0.005
Continence	29 (9%)	11 (9%)	2 (3%)	2 (6%)	2 (11%)	12 (19%)	0.017	0.019
Fatigue	27 (8%)	13 (11%)	0 (0%)	0 (0%)	3 (17%)	11 (17%)	0.001	0.002

Category	All (n=318)	Cerebral palsy (n=121)	Speech/language delay (n=79)	Autism spectrum disorder (n=36)	Down syndrome (n=18)	Other (n=64)	p	q_BH
Learning and memory	19 (6%)	10 (8%)	2 (3%)	0 (0%)	1 (6%)	6 (9%)	0.164	0.164

p, permutation chi-square test (10,000 permutations); *q*, Benjamini–Hochberg adjusted. Zero counts are shown as 0 (0%); no cell is missing.

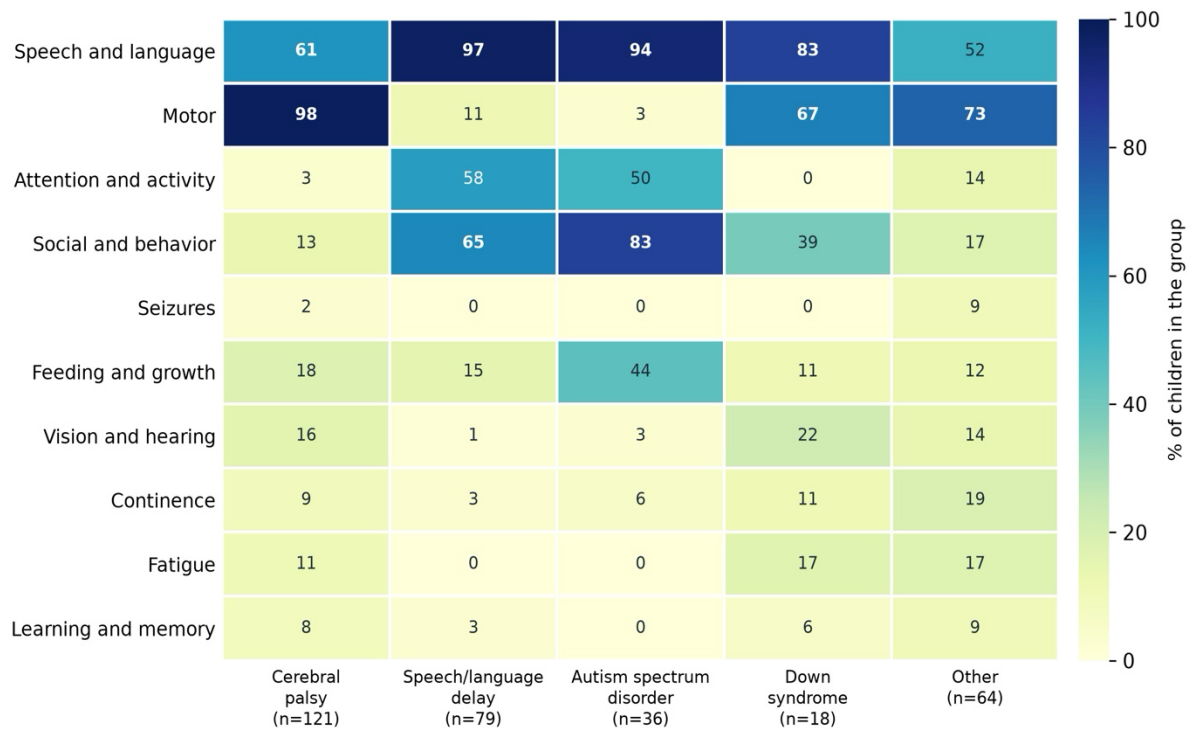


Figure 1. Percentage of children with each complaint category, by diagnosis group.

Agreement with diagnosis (Table 4): The expected complaint was recorded for 98% of children with cerebral palsy, 97% with speech/language delay and 83% with autism spectrum disorder, but the complaints were not specific: only 63%, 33% and 26% of

children with the complaint had the corresponding diagnosis. Kappa was 0.57 for cerebral palsy and motor complaints, 0.19 for speech/language delay and speech and language complaints, and 0.27 for autism spectrum disorder and social and behavior complaints.

Table 4. Agreement between expected complaint and coded diagnosis.

Diagnosis	Complaint	n with diagnosis	Complaint among children with the diagnosis (%)	No complaint among children without the diagnosis (%)	Diagnosis among children with the complaint (%)	kappa (95% CI)
Cerebral palsy	Motor	121	98	65	63	0.57 (0.49 to 0.65)
Speech/language delay	Speech and language	79	97	35	33	0.19 (0.14 to 0.25)
Autism spectrum disorder	Social and behavior	36	83	70	26	0.27 (0.17 to 0.37)

Informants: The informant was stated for 133 children (42%): mother 84, caregiver 37, father 8 and grandmother 4. The number of categories per child did not differ by stated informant ($p=0.39$).

Complaint profiles (Table 5, Figure 2; the Russian top terms of each topic are in Supplementary Table S5): Stability was 1.00 for the solutions with 2 and 3 topics, 0.91 with 4 topics and 0.81 with 5 topics (Supplementary Table S4). The two-topic solution was not eligible, because the rule requires at least three topics, and the three-topic solution had the highest stability among the eligible solutions, so 3 topics were kept. In the four-topic solution the developmental topic

split into a speech-and-behavior topic (90 children) and a limited-vocabulary topic (54), and the two motor topics were essentially unchanged (Supplementary Table S8). Topic 1 (143 children) was a developmental profile of speech delay with hyperactivity and social and behavior complaints. Topic 2 (103) was a profile of limb weakness, limited movement and gait disturbance. Topic 3 (72) was a profile of severe motor delay, in which children did not sit, stand, walk or crawl; feeding and growth complaints (36%) and seizures (11%) were most frequent in this topic. Topics were strongly related to diagnosis group (chi-square $p<0.001$).

Table 5. Complaint topics from non-negative matrix factorization: size, main terms and % of children with each complaint category.

Topic	n children	Main terms (lemmas)	Speech and language	Motor	Attention and activity	Social and behavior	Seizures	Feeding and growth	Vision and hearing	Continence	Fatigue	Learning and memory
1	143	speech and development delay, hyperactivity (речевой развитие, задержка, гиперактивность)	97	13	50	64	1	21	6	6	4	5
2	103	limb weakness, limited movement, gait (конечность, слабость, ограничение, походка)	47	98	5	13	0	4	16	11	18	12
3	72	motor development; walking, standing, sitting, crawling (моторный развитие, ходить, стоять, сидеть, ползать)	65	96	1	14	11	36	12	14	3	0

Russian terms: see Supplementary Table S5.

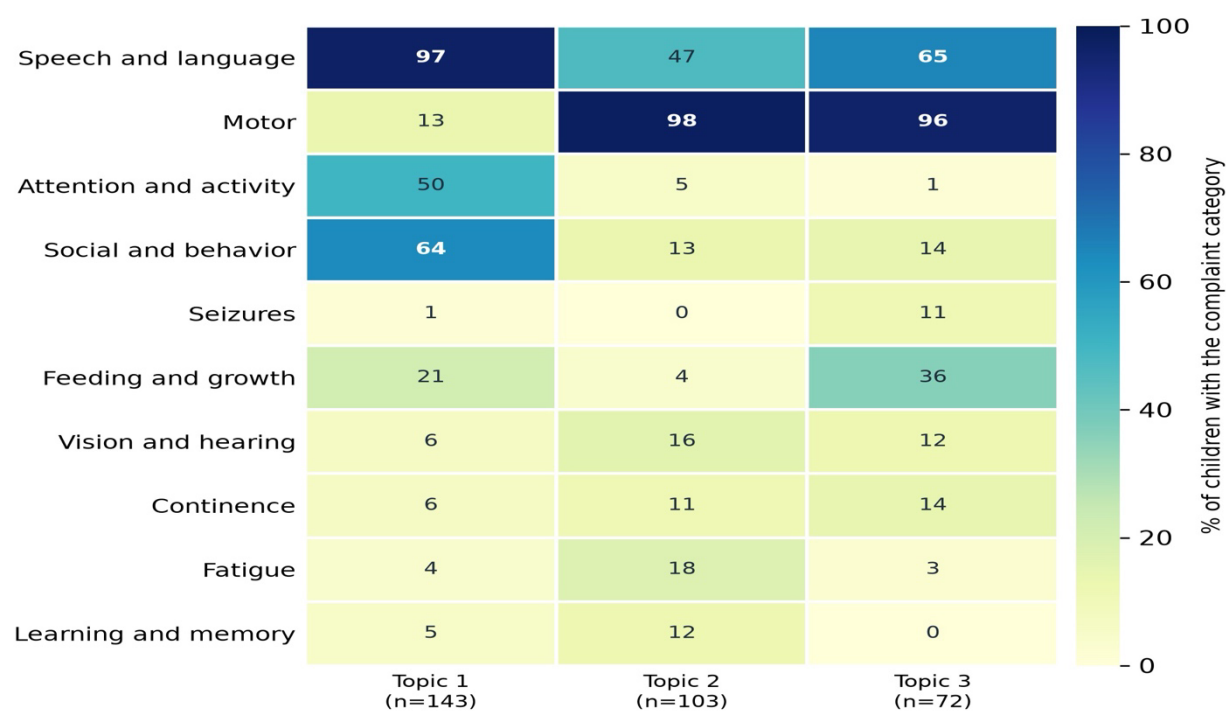


Figure 2. Percentage of children with each complaint category within each complaint topic.

Discussion

In this cohort, the complaints families reported fell into a small number of categories that a transparent dictionary could code with high agreement with hand coding. Speech and language was the most common concern and appeared in children whose main diagnosis was motor. In particular, 61% of children with cerebral palsy had a speech or language complaint, which suggests that speech therapy need in these children is large and that a motor diagnosis alone is a poor guide to the therapy mix a family expects.

The complaints followed the diagnosis in the expected way but were not specific. Motor complaints were almost universal in cerebral palsy yet also common in children with Down syndrome and in the other group; speech complaints were close to universal in speech/language delay and autism spectrum disorder. The low kappa between diagnosis and complaint is not a measure of coding quality. It reflects the very high frequency of speech and language complaints across diagnoses, which limits the value of kappa, so we also report the proportions in each direction. The unsupervised profiles reproduced the main distinction between developmental, limb weakness and gait, and severe motor presentations, and the severe motor profile carried most of the feeding, growth and seizure complaints, which is clinically plausible.

These results agree with, and add to, what is known for single diagnoses. Communication impairment is common in cerebral palsy [3], and in our

cohort speech and language complaints were recorded for 61% of children admitted with this diagnosis, a figure that comes from the family’s own account and not from a speech assessment. Speech complaints were also recorded for 83% of children with Down syndrome, in line with the prominence of speech impairment in this condition [6], although the group is small. Feeding and growth complaints were most frequent in autism spectrum disorder (44%), which fits the higher rate of feeding problems reported for these children [7]. That a transparent dictionary reproduces such expected patterns is a useful sign that the coding is capturing real content, but it is not a substitute for validation against independent clinicians. Because the dictionary was refined against one clinician’s coding, the agreement reported here is probably an upper bound for the agreement that would be found against a different clinician or a consensus of several. The disagreements in feeding and growth and in vision and hearing could reflect gaps in the dictionary or one reader’s judgements, and a second reader would be needed to tell the two apart.

On the strength of agreement, the kappa for cerebral palsy and motor complaints (0.57) is moderate on the Landis and Koch benchmarks [13], and the kappas for speech/language delay (0.19) and autism spectrum disorder (0.27) are slight to fair. This is not a coding failure: kappa depends on how common a category is, and it can be low when one answer dominates even if agreement on the individual

decisions is high [14]. These complaints both occur across diagnosis, a large number of children have these complaints without diagnoses, and only a small number of children with the complaint are diagnosed. Reporting in the two directions as we do, provides added clarity compared to reporting just the kappa. The practical implication for service planning is that the diagnosis indicates which service is likely to be the primary complaint, rather than all complaints, and therefore the therapy mix cannot be planned from the diagnosis list. The hand codes are shared, so any Russian-speaking clinician can recode a sample and replace the reference without changing any other part of the pipeline.

These findings have practical uses. A structured complaint checklist at admission, matched to the ten categories, would let services record the same information in a form that can be counted. Structured interviews such as the Canadian Occupational Performance Measure already elicit family-identified problems and priorities in this way [24]. The dictionary and the code can be reused by other centers working with Russian-language records after checking the wording used locally.

Limitations: The data come from one national center and one admission course (all children were

admitted between 22 August and 8 September 2025, 282 of 318 on 22 August), and the numbers in some diagnosis groups (Down syndrome $n=18$, autism spectrum disorder $n=36$) are small. The complaint text is the physician's record of what the family said, not a verbatim report, so the wording reflects both, and "family-reported" here means reported by the family as recorded by the physician. The dictionary was developed on the same cohort; the test set of 40 children guards against over-fitting but is small, and the lowest recall was for feeding and growth and for vision and hearing. Sex was not recorded for 21% of children; these children did not differ from the others in the variables we measured, but sex differences could not be examined. The topic model is exploratory, depends on the choices made for preprocessing, and the stability criterion is a heuristic; the two- to four-topic solutions are compared in Supplementary Table S8. Finally, the reference coding was done by a single clinician and no second reader checked it, so the agreement figures reflect one reader's judgement, probably overstate agreement with other clinicians, and are not an estimate of inter-rater reliability.

Conclusion

Family-reported admission complaints in routine Russian-language records can be coded reproducibly with a transparent dictionary, which agreed with clinician hand coding on 98.3% of category decisions in an unseen test set. Speech and language complaints were the most common (73%) and were recorded for 61% of children with cerebral palsy, so the diagnosis alone did not predict the full set of family concerns. Complaints followed the coded diagnosis but were not specific to it, and three complaint profiles (developmental, limb weakness and gait, and severe motor delay) matched clinical presentation more

closely than diagnosis group alone. For services, a structured admission checklist built on these ten categories would make family priorities countable and comparable between children and over time. These findings come from one center and one admission course, and the reference coding was done by a single clinician, so the agreement estimates and the profiles should be confirmed in other centers and with independent coders. The dictionary, code and hand codes are available from the corresponding author on request, so that other groups can recode a sample and test the method on their own records.

Acknowledgements

Author Contributions (CRediT): Syed Gufran Sadiq Zaidi: conceptualization; methodology; validation; formal analysis; investigation; resources; data curation; writing – original draft; visualization; supervision. Saule Karibzhanova: validation; investigation; resources; data curation; writing-review and editing. All authors have read and agreed to the published version of the manuscript.

Funding and Acknowledgements: The research work presented in this paper was supported by Grant BR27199433 from the Science Committee of the Ministry of Science and Higher Education of the

Republic of Kazakhstan. We are sincerely thankful to the National Center for Children's Rehabilitation, Kazakhstan for their cooperation in the provision of data.

Ethics: The study was approved by the Nazarbayev University Institutional Review Ethics Committee Board (IRB No:1081/21.05. 2025)

Disclosure of AI Use: Generative AI is used only for language improvement purposes using the prompt "Proofread:". We confirm that the conceptualization, data curation, analyses, and manuscript writing are original.

References

- Rosenbaum P, King S, Law M, King G, Evans J. Family-centred service: a conceptual framework and research review. *Phys Occup Ther Pediatr.* 1998;18(1):1-20. doi:[10.1080/J006v18n01_01](https://doi.org/10.1080/J006v18n01_01)
- Rosenbaum P, Paneth N, Leviton A, Goldstein M, Bax M, Damiano D, et al. A report: the definition and classification of cerebral palsy April 2006. *Dev Med Child Neurol Suppl.* 2007;109:8-14. doi:[10.1111/j.1469-8749.2007.tb12610.x](https://doi.org/10.1111/j.1469-8749.2007.tb12610.x)
- Parkes J, Hill N, Platt MJ, Donnelly C. Oromotor dysfunction and communication impairments in children with cerebral palsy: a register study. *Dev Med Child Neurol.* 2010;52(12):1113-9. doi:[10.1111/j.1469-8749.2010.03765.x](https://doi.org/10.1111/j.1469-8749.2010.03765.x)
- Novak I, McIntyre S, Morgan C, Campbell L, Dark L, Morton N, et al. A systematic review of interventions for children with cerebral palsy: state of the evidence. *Dev Med Child Neurol.* 2013;55(10):885-910. doi:[10.1111/dmcn.12246](https://doi.org/10.1111/dmcn.12246)
- Novak I, Morgan C, Fahey M, Finch-Edmondson M, Galea C, Hines A, et al. State of the evidence traffic lights 2019: systematic review of interventions for preventing and treating children with cerebral palsy. *Curr Neurol Neurosci Rep.* 2020;20(2):3. doi:[10.1007/s11910-020-1022-z](https://doi.org/10.1007/s11910-020-1022-z)
- Kent RD, Vorperian HK. Speech impairment in Down syndrome: a review. *J Speech Lang Hear Res.* 2013;56(1):178-210. doi:[10.1044/1092-4388\(2012/12-0148\)](https://doi.org/10.1044/1092-4388(2012/12-0148))
- Sharp WG, Berry RC, McCracken C, Nuhu NN, Marvel E, Saulnier CA, et al. Feeding problems and nutrient intake in children with autism spectrum disorders: a meta-analysis and comprehensive review of the literature. *J Autism Dev Disord.* 2013;43(9):2159-73. doi:[10.1007/s10803-013-1771-5](https://doi.org/10.1007/s10803-013-1771-5)
- Wang Y, Wang L, Rastegar-Mojarad M, Moon S, Shen F, Afzal N, et al. Clinical information extraction applications: a literature review. *J Biomed Inform.* 2018;77:34-49. doi:[10.1016/j.jbi.2017.11.011](https://doi.org/10.1016/j.jbi.2017.11.011)
- Kreimeyer K, Foster M, Pandey A, Arya N, Halford G, Jones SF, et al. Natural language processing systems for capturing and standardizing unstructured clinical information: a systematic review. *J Biomed Inform.* 2017;73:14-29. doi:[10.1016/j.jbi.2017.07.012](https://doi.org/10.1016/j.jbi.2017.07.012)
- Névél A, Dalianis H, Velupillai S, Savova G, Zweigenbaum P. Clinical natural language processing in languages other than English: opportunities and challenges. *J Biomed Semantics.* 2018;9(1):12. doi:[10.1186/s13326-018-0179-8](https://doi.org/10.1186/s13326-018-0179-8)
- Conway M, Dowling JN, Chapman WW. Using chief complaints for syndromic surveillance: a review of chief complaint based classifiers in North America. *J Biomed Inform.* 2013;46(4):734-43. doi:[10.1016/j.jbi.2013.04.003](https://doi.org/10.1016/j.jbi.2013.04.003)
- Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas.* 1960;20(1):37-46. doi:[10.1177/001316446002000104](https://doi.org/10.1177/001316446002000104)
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics.* 1977;33(1):159-74. doi:[10.2307/2529310](https://doi.org/10.2307/2529310)
- Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol.* 1993;46(5):423-9. doi:[10.1016/0895-4356\(93\)90018-V](https://doi.org/10.1016/0895-4356(93)90018-V)
- Phipson B, Smyth GK. Permutation P-values should never be zero: calculating exact P-values when permutations are randomly drawn. *Stat Appl Genet Mol Biol.* 2010;9(1):Article39. doi:[10.2202/1544-6115.1585](https://doi.org/10.2202/1544-6115.1585)
- Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Series B Stat Methodol.* 1995;57(1):289-300. doi:[10.1111/j.2517-6161.1995.tb02031.x](https://doi.org/10.1111/j.2517-6161.1995.tb02031.x)
- Efron B. Bootstrap methods: another look at the jackknife. *Ann Stat.* 1979;7(1):1-26. doi:[10.1214/aos/1176344552](https://doi.org/10.1214/aos/1176344552)
- DiCiccio TJ, Efron B. Bootstrap confidence intervals. *Stat Sci.* 1996;11(3):189-228. doi:[10.1214/ss/1032280214](https://doi.org/10.1214/ss/1032280214)
- Kruskal WH, Wallis WA. Use of ranks in one-criterion variance analysis. *J Am Stat Assoc.*

- 1952;47(260):583-621.
doi:[10.1080/01621459.1952.10483441](https://doi.org/10.1080/01621459.1952.10483441)
20. Korobov M. Morphological analyzer and generator for Russian and Ukrainian languages. In: Khachay MY, Konstantinova N, Panchenko A, Ignatov DI, Labunets VG, editors. Analysis of Images, Social Networks and Texts. AIST 2015. Communications in Computer and Information Science, vol 542. Cham: Springer; 2015. p. 320-32. doi:[10.1007/978-3-319-26123-2_31](https://doi.org/10.1007/978-3-319-26123-2_31)
21. Spärck Jones K. A statistical interpretation of term specificity and its application in retrieval. J Doc. 1972;28(1):11-21. doi:[10.1108/eb026526](https://doi.org/10.1108/eb026526)
22. Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. Nature. 1999;401(6755):788-91. doi:[10.1038/44565](https://doi.org/10.1038/44565)
23. Greene D, O'Callaghan D, Cunningham P. How many topics? Stability analysis for topic models. In: Calders T, Esposito F, Hüllermeier E, Meo R, editors. Machine Learning and Knowledge Discovery in Databases. ECML PKDD 2014. Lecture Notes in Computer Science, vol 8724. Berlin: Springer; 2014. p. 498-513. doi:[10.1007/978-3-662-44848-9_32](https://doi.org/10.1007/978-3-662-44848-9_32)
24. Law M, Baptiste S, McColl M, Opzoomer A, Polatajko H, Pollock N. The Canadian Occupational Performance Measure: an outcome measure for occupational therapy. Can J Occup Ther. 1990;57(2):82-7. doi:[10.1177/000841749005700207](https://doi.org/10.1177/000841749005700207)