## Methods

We performed multiple layers of QC before working on this dataset. First, we removed all cases with high missingness (either null/0 values or with absolutely no symptoms present). We also excluded RareX-named disease entities if there were fewer than five cases with non-missing data as there would not be enough cases from which to draw conclusions. We dropped Symptoms not found in at least one case in the RareX data (most of these were cancer-related Symptoms). We normalized the CSHQ subscales although did not end up analyzing this data numerically. Instead we transformed the CSHQ features into a binary phenotype by treating all positive values as the symptom being present and all 0 values as the symptom being absent.

We semi-automatically mapped each Symptom to a corresponding term in the Human Phenotype Ontology so we could take advantage of that reference data. We extracted all unique Symptoms from the survey dataset. Every Symptom had “\_Symptom\_Present” removed and underscores were replaced with spaces, saved as “Symptom Name” in our final datasheet. Any Symptoms that were marked as not present were removed as handling/inclusion of absence of symptoms in HPO is not yet common available in most entries. Symptom names were then run through the biolark HPO annotate package (performs concept recognition and matches symptoms to HPO terms) and the matching HPO label/codes were added to the (“HPO\_Code” and “HPO\_label”) in the final dataset. Any Symptom names that did not return a result were added to an “Unmatched terms” column for visibility. We then manually reviewed these automated matches, picking a more appropriate HPO term for the Symptom if necessary. For unmatched terms, we manually curated a suitable HPO term. In our final dataset, we mapped the original symptom to its curated HPO term and added these as the “HPO\_Code” and “HPO\_Label” columns.

We mapped diseases to reference databases that provide summary frequency data for clinical findings (OrphaNet [ON] and HPO, where HPO draws data from either OMIM or OrphaNet or both) using the gene name if only a single disease entity has been reported in either HPO or ON. If multiple disease entities were reported in HPO or ON, we compared the RareX Symptom frequencies to all disease entities as it was not clear which HPO/ON disease entities most closely matched each RareX case. HPO and ON both report the frequency of clinical findings in a hypothetical “reference” set of affected individuals. ON classifies frequencies into broad categories (e.g. Very Frequent, Frequent, Rare, etc.) while HPO either uses ON’s categories or may report exact percentages. We converted all HPO/ON frequencies to the ON categories to unify on a single set of expected clinical finding frequencies. We had to exclude the FAM177A1 disease entity from analysis as the gene discovery is unpublished, so there are no reference frequencies to which that data can be compared. Additionally, the symptom frequency data currently discoverable by google searching derives from a largely overlapping set of families in the RareX dataset, so there would be no point in trying to identify differences in frequencies of clinical findings.

We generated one final result Excel workbook per gene and one sheet if applicable for each claimed/named unique disease entity in HPO/ON. Each row contains a symptom/HPO term associated with the disease entity in either RareX, HPO, or ON.

We calculated a number of quantities, including the proportion of cases with a disease with a given symptom/clinical finding, IDF (inverse document frequency), and TF-IDF (term frequency/inverse document frequency). IDF measures the relative rarity of a term in the set of diseases while TF-IDF measures importance of a term which is inversely related to its frequency across diseases within a given dataset (e.g. across all of the RareX survey data or across all of ON or HPO).

RareX has many terms that were not included in the ON/HPO lists of expected clinical findings for each disease and vice versa. To improve matching between clinical finding lists, we attempted to use both a ontology-based similarity metric, reported in best\_match\_terms\_pyhpo[\_gt\_06], to identify terms that might be referring to the same clinical finding between RareX and HPO/ON, and a text/meaning-based similarity algorithm, reported in best\_match\_meaning\_terms. In both cases, we found that the scoring method was not very reliable, with some poor term matches receiving high scores (e.g. 1.0) while other good term matches could be assigned middling scores (e.g. 0.6). We chose a minimum cutoff of 0.6 (max 1.0) to more efficiently winnow through poor scoring matches, but the results are still noisy.

## Final dataset column names and definitions

|  |  |
| --- | --- |
| Column Name | Explanation |
| Disease\_Name | Disease name in RareX survey data |
| Symptom | Symptom as labeled by RareX for this disease |
| Count | Number of the times a symptom is mentioned among RareX cases of a disease |
| term\_proportion | count of cases with symptom/total number of RareX cases with a disease |
| survey\_frequency | ‘Symptom Count’/total number of the terms in a disease, d |
| survey\_idf | IC/IDF of a term in the survey data  survey\_idf: (h|d), the survey IDF of a term for a disease (d)  survey\_idf (h|d) = (np.log (total diseases in survey / (number of the diseases, the term (h) is mentioned + 1)) + 1) |
| survey\_tf\_idf | tf \* idf of a term in the survey data for the disease  ‘survey\_frequency’ \* ‘survey\_idf’ |
| HPO\_Code | mapped HPO code for the symptom |
| HPO\_label | phenotype name/HPO label of a term |
| hpo\_dataset\_idf | IDF information of a term in the HPO data  hpo\_dataset\_idf(h|d), the IDF of a term for a disease (d) in the HPO dataset  hpo\_dataset\_idf (h|d) = (np.log(total diseases in HPO dataset / (number of the diseases, the term (h) is mentioned in HPO dataset + 1)) + 1) |
| survey\_hpo\_idf | (survey\_idf + hpo\_dataset\_idf): combination of the survey and HPO idf |
| survey\_hpo\_tf\_idf: | (tf\* survey\_hpo\_idf): a combined tf-idf of a term from survey and HPO data  term frequency (tf) = Survey\_frequency |
| Survey\_OrphaClass | survey term proportion formatted to Orphanet style |
| HPO\_Frequency: | If the term frequency is sourced from the Orphanet dataset it is labeled as “sourced from OrphaNet, otherwise the frequency is sourced from the HPO dataset and this is represented as a blank entry in this column. |
| HPO\_Frequency\_OrphaClass | HPO\_Count formatted to Orphanet style |
| HPO\_Count | Conversion of the mean proportion of reference cases with a symptom per the HPO dataset to a count (e.g. out of 100 cases, the proportion HPO reports would have the symptom). |
| hpo\_dataset\_tf\_idf | tf \* idf of a term in the HPO dataset for a disease, d  hpo\_tf = HPO\_Count/Total HPO\_Count in a disease, d  hpo\_dataset\_tf\_idf = hpo\_tf \* hpo\_dataset\_idf |
| orpha\_dataset\_idf | IDF information of a term in the HPO data  orpha\_dataset\_idf(h|d), the IDF of a term for a disease (d) in the Orphanet dataset  orpha\_dataset\_idf (h|d) = (np.log(total diseases in Orphanet dataset / (number of the diseases, the term (h) is mentioned in Orphanet dataset + 1)) + 1) |
| orpha\_hpo\_dataset\_idf | (orpha\_dataset\_idf + hpo\_dataset\_idf): combination of the Orphanet and HPO idf |
| Orpha\_Frequency | Frequency of a term for a disease in the Orphanet |
| Orpha\_Count | Mean of Orpha\_Frequency transformed into a raw count (assuming 100 individuals with disease) |
| orpha\_dataset\_tf\_idf | tf \* idf of a term in the Orphanet dataset for a disease, d  orpha\_tf = Orpha\_Count/Total Orpha\_Count in a disease, d  orpha\_dataset\_tf\_idf = orpha\_tf \* orpha\_dataset\_idf |
| hpo\_orpha\_dataset\_tf\_idf | (tf\* orpha\_hpo\_dataset\_idf): a combined tf-idf of a term from Orphanet and HPO data  term frequency (tf) = HPO\_Count/Total HPO\_Count in a disease HPO dataset if available  or term frequency = Orpha\_Count/Total Orpha\_Count in a disease Orpha dataset |
| Mapped\_Disease\_name | Disease name is mapped to Orphanet or HPO disease name |
| Matched | Look up the mapped HPO term from the survey data in the orphanet/hpo dataset. This column has the following classifications:  'New in RareX' represents when a term of the survey disease does not exist in the orphanet/HPO  'Matched in Orpha/HPO' represents when a term already exists in the orphanet/HPO for a disease  'Not in survey, Found in Orpha/HPO' represents when a term was in the Orphanet/HPO dataset but was not in the survey data. |
| Matched\_no\_HPO\_frequency | Look up the mapped HPO term from the survey data in the Orphanet/HPO dataset that have zero or null frequency. Classifications are the same as Matched. |
| best\_match\_terms\_pyhpo | This column defines the best matched/similar HPO labels from Orphanet/HPO for a term of the survey data. The data has the following classifications  'not match': no related HPO term exists in the Orphanet/HPO for a term of a disease in survey data  Phenotype with similarity score and Orphanet/HPO frequency: best HPO terms have highest similarity score to the important term of a disease  Found only in HPO/Orphanet: represents when the term exists in the Orphanet dataset for a disease but not in the survey data  We applied the default similarity method from the ‘pyhpo’ python package, available at ttps://pypi.org/project/pyhpo/ |
| best\_match\_terms\_pyhpo\_gt\_06 | This column defines the best matched terms which have a pyhpo similarity score of more than 0.6 Classifications are same as ‘best\_match\_terms\_pyhpo’ |
| best\_match\_meaning\_terms | This column defines the best matched HPO labels based on HPO label meaning from Orphanet/HPO for a term of the survey data using bert similarity method. Classifications are same as ‘best\_match\_terms\_pyhpo’ |
| abs\_hpo\_term\_prop\_difference | This column defines the absolute difference between survey and HPO term proportion. If the term is new in rareX, we provide the difference of each term in ‘best\_terms\_match’ column and the selected term.  abs\_hpo\_term\_prop\_difference = ‘term\_proportion’- HPO\_Count/100 |
| abs\_orphanet\_term\_prop\_difference | This column defines the absolute difference between survey and orphanet term proportion. If the term is new in rareX, we provide difference of each term in ‘best\_terms\_match’ column and the selected term  abs\_orphanet\_term\_prop\_difference = ‘term\_proportion’- Orpha\_Count/100 |
| Fisher\_pvalue | This column defines the pvalue of the Fisher exact test of difference in proportions of individuals with/without a clinical finding between the RareX survey data and HPO. We calculate the pvalue for the count of a term in the survey dataset and HPO dataset. |
| top terms | This column defines if a term is top of the list of the survey\_hpo\_tf\_idf. We take the mean survey\_hpo\_tf\_idf for each disease, and find the top terms if it is above the mean value of the survey\_hpo\_tf\_idf. The data has the following classifications  1: a term is in the top terms list for a disease  0: not in the top term list |
| NewRareX | This column defines if a term is new and is in the top terms list in survey for a disease. The data has the following classifications  1: a term is a new for a disease  0: already exists in the Orpha/HPO dataset for a disease. |
| top\_terms\_not\_match\_but\_frequent | This column defines if a term is a new finding for a disease. We say a term is a new finding for a disease if a term is in a top term list of the survey\_hpo\_tf\_idf and not matched or similar to any term in Orphanet/HPO dataset, and frequent in survey data. The data represent has the following classifications.  1: a term is a new finding for a disease  0: a term is not new |
| matched\_terms\_frequent\_both | This column defines if a term is common/frequent to both the survey and HPO datasets. The data has the following classifications.  1: a term is common for both dataset  0: a term is not common for both dataset |
| top\_matched\_terms\_frequent\_survey\_only: | This column defines if an existing term which is not frequent in Orphanet/HPO dataset but is frequent in survey, and in the top term list of survey\_hpo\_tf\_idf. The data has the following classifications.  1: a top existing term is common/frequent to survey not frequent/common to HPO dataset  0: a term does not fall into the condition. |
| matched\_terms\_count: | This column defines the number of the times a matched term from the best\_match\_terms\_pyhpo\_gt\_06 has been mentioned within the disease. |
| Found\_only\_in\_HPO | This column defines if a term found only in hpo/orphanet but not in the survey dataset. |

**Results are available in Excel spreadsheets.**

* Filename: disease name\_orpha means the survey disease mapped to one OrphaNet disease
* Filename: disease name\_HPO\_ONLY means the survey disease mapped to one HPO disease
* Filename: disease name\_RX\_MULT\_DIS means the survey disease mapped to multiple HPO diseases, each disease data results are in the different sheet/tab in same Excel workbook

## Results

For all diseases considered, there were far more clinical findings that were unique to either RareX or to HPO/ON than there were clinical findings overlapping. This is not just a matter of, e.g. HPO/ON reporting on frequencies of differences in facial features while RareX does not. The discrepancies were often about findings that are potentially clinically-meaningful. For example, for 4H Leukodystrophy, HPO/ON report delayed puberty in ~half of affected individuals while RareX does not include this clinical finding or any findings that are seemingly related.

As a result of the sheer number differences in clinical findings included in each dataset, we did not have time to review every difference but are choosing to highlight some interesting and potentially clinically-meaningful differences. We particularly focused our discussion on clinical findings that were identified as relatively frequent in RareX but are rare in the HPO/ON list of common clinical findings (and there were also no good matches with an HPO/ON finding by ontology or meaning-based similarity) for the disease, the clinical finding is “important” (as defined by TF-IDF; top\_terms\_not\_match\_but\_frequent == 1), and the difference in frequencies between RareX and HPO/ON is statistically significant (by Fisher p-value). It is possible to do the converse analysis and identify terms that are common in HPO/ON but not in RareX, which would be interesting as the frequencies of features in HPO/ON might not be representative of the typical range of features in a larger cohort. However we chose to focus on the terms that are missing/rare in HPO/ON but common in RareX because it is important to document that there is a high likelihood of observing such findings versus what is recorded in reference databases and we also opted to focus on clinical findings that are more likely to impact a family’s or patient’s quality of life / daily living than, e.g. differences in facial appearance, even though the latter are useful for clinicians when assessing a potential diagnosis. We manually reviewed the filtered list of clinical findings to confirm that there were no relevant matching terms in HPO/ON (e.g., intellectual disability may have been labeled with a number of alternative/related terms such as developmental delay or cognitive impairment, and these alternate terms may not have been considered a match by our similarity metrics).

One notable pattern across all diseases analyzed is the detailed list of sleep-related symptoms frequently reported by RareX families that are largely not represented in the data summarized by HPO/ON. This likely reflects a set of features that might not be routinely recorded by clinicians but are notable for families perhaps because the symptoms are particularly challenging on a day-to-day basis. Additionally, we were unable to map a number of the specific sleep symptoms in the RareX dataset to HPO terms (these had to be record as generic sleep abnormalities), so there are additional specific symptoms that are likely to be frequently-reported by RareX families that are not yet recognized as common features of the disorders by HPO/ON. Finally, a future improvement to our analysis approach would be to consider any feature reported by HPO/ON but without a frequency category as Very Frequent, assuming that otherwise the features would not be included in those databases. This would reduce the false positive rate of findings with frequency differences between RareX and HPO/ON.

* 4H Leukodystrophy
  + A number of clinical findings related to sleep disturbances are common in the RareX cohort, but sleep disturbances are not reported in the list of typical findings in HPO/ON. These findings include somnambulism, insomnia, drowsiness, sleep apnea, and parasomnia.
  + Delayed puberty was reported by approximately half of all RareX families but it is not included as a common finding in HPO/ON.
* 8p-related disorders
  + A number of clinical findings related to sleep disturbances are common in the RareX cohort, but sleep disturbances are not reported in the list of typical findings in HPO/ON. These findings include somnambulism, insomnia, drowsiness, and sleep apnea.
  + RareX families also report challenges with feeding (non-specific “feeding difficulties” but also constipation, dysphagia, and non-specific esophagus issues) which are not mentioned in HPO/ON.
  + Interestingly, ~40% of families report some sort of maternal health problems (Maternal\_Health\_Problem\_Symptom\_Present) and it would be informative to investigate whether this reflects a history of health issues while pregnant with an affected child.
* AHC (Alternating Hemiplegia of Childhood)
  + A number of clinical findings related to sleep disturbances are common in the RareX cohort, but sleep disturbances are not reported in the list of typical findings in HPO/ON. These findings include somnambulism, insomnia, drowsiness (however sleep apnea was reported in HPO/ON).
  + Hyperactivity, abnormal repetitive movements, short attention span, and anxiety were also common in RareX and not in HPO/ON, however it is hard to assess if these differences are meaningful because HPO/ON does report “atypical behavior” as a finding and all of these might potentially fall under that umbrella.
* *CACNA1A*-related disorders
  + Multiple disease entities have been associated with mutations in *CACNA1A*. After reviewing the overlap in present clinical findings and considering the likely ascertainment bias in the RareX dataset, we suspect most individuals in the RareX *CACNA1A* cohort probably have a disease most similar to epileptic encephalopathy, so we focused on comparison of clinical finding frequencies to that disease entity. We did not identify any notable differences aside from sleep disturbances.
* *CASK*-related disorders
  + The RareX cohort overlapped with both *CASK*-related disease entities in HPO/ON so the following summary applies to both.
  + A number of clinical findings related to sleep disturbances are common in the RareX cohort, but sleep disturbances are not reported in the list of typical findings in HPO/ON. These findings include somnambulism, insomnia, drowsiness, sleep apnea, and parasomnia.
  + Hypohidrosis is a notable clinical finding that was reported in ~1/3 of the RareX cohort and was noted in HPO/ON but no reference frequency was provided.
  + Coordination Issues and “Issues with Muscles of the Hand and Feet” were common in RareX, but it is unclear if these features were simply a result of, e.g. hypotonia, which is already a documented common clinical finding in HPO/ON.
* *CHAMP1*-related disorders
  + A number of clinical findings related to sleep disturbances are common in the RareX cohort while only obstructive sleep apnea is reported in the list of typical findings in HPO/ON. Clinical findings present in the RareX cohort but absent from HPO/ON include somnambulism, insomnia, and drowsiness.
  + Behavioral features including abnormal aggressive behavior, tantrums, and autistic behavior were reported by RareX families. HPO/ON does include abnormal repetitive behavior, which has some overlap with autistic behavior, but the other behavior features may be challenging for families to handle and are not mentioned in HPO/ON. That said, an association between *CHAMP1* mutations and aggressive behavior has been noted in recent literature (see <https://www.mdpi.com/2073-4425/14/8/1546> for references), so this simply might reflect a need to update the common clinical findings listed in HPO/ON.
  + Additional notable features unique to the data from the RareX cohort include headaches/migraines, visual impairment, and anxiety.
* *CHD2*-related disorders
  + A number of clinical findings related to sleep disturbances are common in the RareX cohort, but sleep disturbances are not reported in the list of typical findings in HPO/ON. These findings include somnambulism, insomnia, drowsiness, and sleep apnea.
  + Behavioral features including abnormal aggressive behavior, emotional lability, impulsivity, hyperactivity, anxiety, and a short attention span were also notable.
  + Constipation was reported by ~31% of RareX families and is not part of the HPO/ON records.
* Classic homocystinuria
  + A number of clinical findings related to sleep disturbances are common in the RareX cohort, but sleep disturbances are not reported in the list of typical findings in HPO/ON. These findings include somnambulism, insomnia, drowsiness, and sleep apnea.
  + This cohort is notable because there were relatively few symptoms noted by families in the RareX survey while the HPO/ON list of clinical findings was far mor extensive. This may be because the RareX survey covered a limited range of findings and/or the categories/labels used in the survey were too broad for families to recognize that their affected members may have a symptom in that category (e.g. osteoporosis is found in >80% of cases according to HPO/ON but RareX only asks about Skeletal dysplasia)
* *DYRK1A* Syndrome
  + A number of clinical findings related to sleep disturbances are common in the RareX cohort, but sleep disturbances are not reported in the list of typical findings in HPO/ON. These findings include somnambulism, insomnia, drowsiness, and sleep apnea.
  + Hypertonia (~44%) and issues with coordination (37.5%) were moderately common symptoms in the RareX cohort and neither have a matching term in HPO/ON.
  + Impulsivity and short attention span were reported in RareX but not in HPO/ON, however HPO/ON do document hyperactivity and atypical behavior as part of this syndrome, and these terms may all be referring to the same behaviors in affected individuals.
* *FOXP1* Syndrome
  + A number of clinical findings related to sleep disturbances are common in the RareX cohort, but sleep disturbances are not reported in the list of typical findings in HPO/ON. These findings include somnambulism, insomnia, drowsiness, and sleep apnea.
  + Behavioral features including abnormal aggressive behavior and emotional lability were also notable although HPO/ON does include “Atypical behavior” which is sufficiently broad that it might encompass both.
  + 31% of RareX families reported bladder symptoms (Abnormality of the bladder) but HPO/ON reported abnormality of the kidney as only an occasional finding (5-29% frequency). We would need to do a literature review to determine whether this is a real difference in frequency as the upper range of the HPO/ON “Occasional” category is close to the 31% frequency in RareX.
* *HUWE1*-related disorders
  + A number of clinical findings related to sleep disturbances are common in the RareX cohort while HPO/ON does report sleep disturbances (no specifics) as an occasional feature.
  + Autism and self-injurious behavior were common in RareX families, but these are not specifically included as part of the HPO/ON list, although HPO/ON does include “abnormal repetitive mannerisms” which might have some overlap in practice.
  + ~50% of RareX families report failure to thrive, short stature, feeding difficulties, and/or constipation. Of the symptoms, only short stature is included in the HPO/ON list but without a frequency annotation.
* *KDM5C*-related disorders
  + A number of clinical findings related to sleep disturbances are common in the RareX cohort. These findings include somnambulism, insomnia, drowsiness, and sleep apnea. HPO/ON do report sleep disturbances (no specifics) as an occasional feature.
  + Interestingly, ~30-40% of families report some sort of maternal health problems (Maternal\_Health\_Problem\_Symptom\_Present or Delivery Complication) and it would be informative to investigate whether these reflects a history of health issues while pregnant with an affected child or just complications during delivery.
  + 30% of families also report Bladder symptoms, which are not reflected in HPO/ON.
* Kleefstra syndrome
  + 42% of RareX families report hearing impairment issues, while only 5-29% of cases referenced by HPO/ON report hearing impairment. This suggests hearing impairment is a more common feature in Kleefstra syndrome than expected based on HPO/ON data.
* Koolen-de Vries Syndrome
  + A number of clinical findings related to sleep disturbances are common in the RareX cohort, but sleep disturbances are not reported in the list of typical findings in HPO/ON. These findings include somnambulism, insomnia, drowsiness, and sleep apnea.
  + 57% of RareX families report hearing impairment issues, which is noted in HPO/ON. This suggests hearing impairment is a far more common feature than expected based on HPO/ON data.
  + Abnormal delivery and decrease fetal movements were reported in almost half (43% of RareX cases and this is not included in that HPO/ON entry, although HPO/ON does report that individuals were often small for gestational age, which might prompt an early delivery (i.e. abnormal).
* Malan Syndrome
  + A number of clinical findings related to sleep disturbances are common in the RareX cohort, but sleep disturbances are not reported in the list of typical findings in HPO/ON. These findings include somnambulism, insomnia, drowsiness, and sleep apnea.
  + 31% of RareX families reported myopia which is not in HPO/ON although other clinical findings likely to be associated with visual impairment are included, such as optic disc abnormalities.
  + 44% RareX families also report temper tantrums as a challenge and no similar terms were noted in HPO/ON.
* *MSL3*-related disorders
  + A number of clinical findings related to sleep disturbances are common in the RareX cohort, but sleep disturbances are not reported in the list of typical findings in HPO/ON. These findings include somnambulism, insomnia, drowsiness, and sleep apnea.
  + Abnormal delivery and abnormality of amniotic fluids were reported in almost half (43% of RareX cases and this is not included in that HPO/ON entry.
* Ogden syndrome (*NAA10*)
  + Absent speech was noted by many RareX families and not in HPO/ON, but in the context of severe global developmental delay as recorded in HPO/ON, we are not convinced this is a meaningful difference.
  + 44% of RareX families report premature birth, however the HPO/ON entries report a number of clinical findings that might have led to premature birth (e.g. intrauterine growth retardation, oligohydramnios, and decreased fetal movements), so it is not clear that this is a meaningful difference.
* Pallister-Killian mosaic syndrome
  + A number of clinical findings related to sleep disturbances are common in the RareX cohort, but sleep disturbances are not reported in the list of typical findings in HPO/ON. These findings include somnambulism, insomnia, and drowsiness. Apneic episodes are included in HPO/ON.
  + Most other terms had a high degree of overlap between the RareX cohort and HPO/ON.
* Ring14 and related disorders
  + Sleep disturbances are common in the RareX cohort, but sleep disturbances are not reported in the list of typical findings in HPO/ON.
  + 50% of RareX families report gingivitis
  + 2/3 of RareX families report constipation and gastrointestinal issues such as gastrointestinal dysmotility
* *SETPBP1*-related disorders
  + 40-60% of RareX families reported an abnormal delivery and abnormality of the amniotic fluid during pregnancy, which are not documented in HPO/ON
* *STXBP1*-related disorders
  + A number of clinical findings related to sleep disturbances are common in the RareX cohort, but sleep disturbances are not reported in the list of typical findings in HPO/ON. These findings include somnambulism, insomnia, drowsiness, and sleep apnea.
  + Feeding difficulties was also a notable absence from the HPO/ON data.
* *SYNGAP1*-related disorders
  + A number of clinical findings related to sleep disturbances are common in the RareX cohort, but sleep disturbances are not reported in the list of typical findings in HPO/ON. These findings include somnambulism, insomnia, and drowsiness. Apneic episodes are included in HPO/ON.
  + Aggressive behavior, self-injurious behavior, and hyperactivity are also commonly reported by RareX families (50-66%) but not included in HPO/ON although it is possible these are included in the more broad category of Atypical behavior that is listed in HPO/ON.
* Wiedemann-Steiner Syndrome
  + A number of clinical findings related to sleep disturbances are common in the RareX cohort, but sleep disturbances are not reported in the list of typical findings in HPO/ON. These findings include somnambulism, insomnia, drowsiness, and sleep apnea.