

# Phenotype Analysis

03/03/2021

## PentaTech

### Group Members

1. Manreet Kaur
2. Chirag Arora
3. Birti Kaur
4. Nafis Ahmed
5. Yajing Zhou

```
library(ggplot2)
library(tibble)
library(tidyr)
library(tidyverse)
library(utils)
```

```
Phenotypes_subset <- read.csv("Phenotypes_subset.csv")
head(Phenotypes_subset)
```

```
##   OrphaCode      DisorderName DisorderType      HPOTerm0
## 1      58      Alexander disease      Disease      Macrocephaly
## 2      61      Alpha-mannosidosis      Disease      Macroglossia
## 3      93      Aspartylglucosaminuria      Disease      Inguinal hernia
## 4     585      Multiple sulfatase deficiency      Disease      Hydrocephalus
## 5     141      Canavan disease      Disease      Macrocephaly
## 6     213      Cystinosis      Disease      Proteinuria
##           HPOFrequency0      HPOTerm1      HPOFrequency1
## 1 Very frequent (99-80%) Intellectual disability Very frequent (99-80%)
## 2 Very frequent (99-80%) Coarse facial features Very frequent (99-80%)
## 3 Occasional (29-5%)      Macroorchidism      Frequent (79-30%)
## 4 Frequent (79-30%)      Microcephaly      Occasional (29-5%)
## 5 Frequent (79-30%)      Hearing impairment      Frequent (79-30%)
## 6 Very frequent (99-80%) Nephropathy Very frequent (99-80%)
##           HPOTerm2      HPOFrequency2      HPOTerm3
## 1 Seizures Very frequent (99-80%)      Spasticity
## 2 Hearing impairment Very frequent (99-80%)      Cataract
## 3 Macroglossia      Frequent (79-30%) Abnormality of the dentition
## 4 Macrocephaly      Frequent (79-30%)      Coarse facial features
## 5 Visual impairment      Frequent (79-30%)      Blindness
## 6 Renal tubular dysfunction Very frequent (99-80%)      Photophobia
##           HPOFrequency3      HPOTerm4      HPOFrequency4
## 1 Very frequent (99-80%) Agenesis of corpus callosum Very frequent (99-80%)
## 2 Very frequent (99-80%) Intellectual disability Very frequent (99-80%)
## 3 Frequent (79-30%)      Gingival overgrowth Very frequent (99-80%)
## 4 Frequent (79-30%)      Smooth philtrum      Frequent (79-30%)
```

```

## 5      Frequent (79-30%)      Optic atrophy Very frequent (99-80%)
## 6 Very frequent (99-80%)      Stereotypy Very frequent (99-80%)
##
##      HPOTerm5      HPOFrequency5
## 1      Hyperreflexia Very frequent (99-80%)
## 2      Global developmental delay Very frequent (99-80%)
## 3      Coarse facial features      Frequent (79-30%)
## 4      Sensorineural hearing impairment      Frequent (79-30%)
## 5 Abnormality of visual evoked potentials      Frequent (79-30%)
## 6      Hypothyroidism Very frequent (99-80%)
##      HPOTerm6      HPOFrequency6      HPOTerm7
## 1      Megalencephaly Very frequent (99-80%)      Failure to thrive
## 2      Splenomegaly Very frequent (99-80%)      Hepatomegaly
## 3 Mandibular prognathia Very frequent (99-80%)      Hypertelorism
## 4      Anteverted nares      Frequent (79-30%)      Visual impairment
## 5      Seizures      Occasional (29-5%)      Muscular hypotonia
## 6      Delayed puberty Very frequent (99-80%)      Muscle weakness
##      HPOFrequency7      HPOTerm8      HPOFrequency8
## 1 Very frequent (99-80%)      Frontal bossing Very frequent (99-80%)
## 2 Very frequent (99-80%)      Skeletal dysplasia Very frequent (99-80%)
## 3 Very frequent (99-80%)      Chronic otitis media      Occasional (29-5%)
## 4 Very frequent (99-80%)      Cataract      Frequent (79-30%)
## 5      Frequent (79-30%)      Global developmental delay Very frequent (99-80%)
## 6 Very frequent (99-80%)      Failure to thrive Very frequent (99-80%)
##      HPOTerm9      HPOFrequency9
## 1      Nausea and vomiting Very frequent (99-80%)
## 2 Delayed skeletal maturation Very frequent (99-80%)
## 3      Wide nasal bridge Very frequent (99-80%)
## 4      Thick eyebrow      Frequent (79-30%)
## 5      Hypertonia      Frequent (79-30%)
## 6      Dehydration Very frequent (99-80%)
##      HPOTerm10      HPOFrequency10      HPOTerm11
## 1 Neurological speech impairment Very frequent (99-80%)      Clonus
## 2      Craniofacial hyperostosis Very frequent (99-80%)      Depressed nasal bridge
## 3      Carious teeth      Frequent (79-30%)      Behavioral abnormality
## 4      Optic atrophy      Frequent (79-30%)      Intellectual disability
## 5      Flexion contracture      Occasional (29-5%)      Gastroesophageal reflux
## 6      Polydipsia Very frequent (99-80%)      Vomiting
##      HPOFrequency11
## 1 Very frequent (99-80%)
## 2 Very frequent (99-80%)
## 3      Occasional (29-5%)
## 4 Very frequent (99-80%)
## 5      Frequent (79-30%)
## 6 Very frequent (99-80%)

```

```

# Frequencies phenotypic abnormalities in human diseases

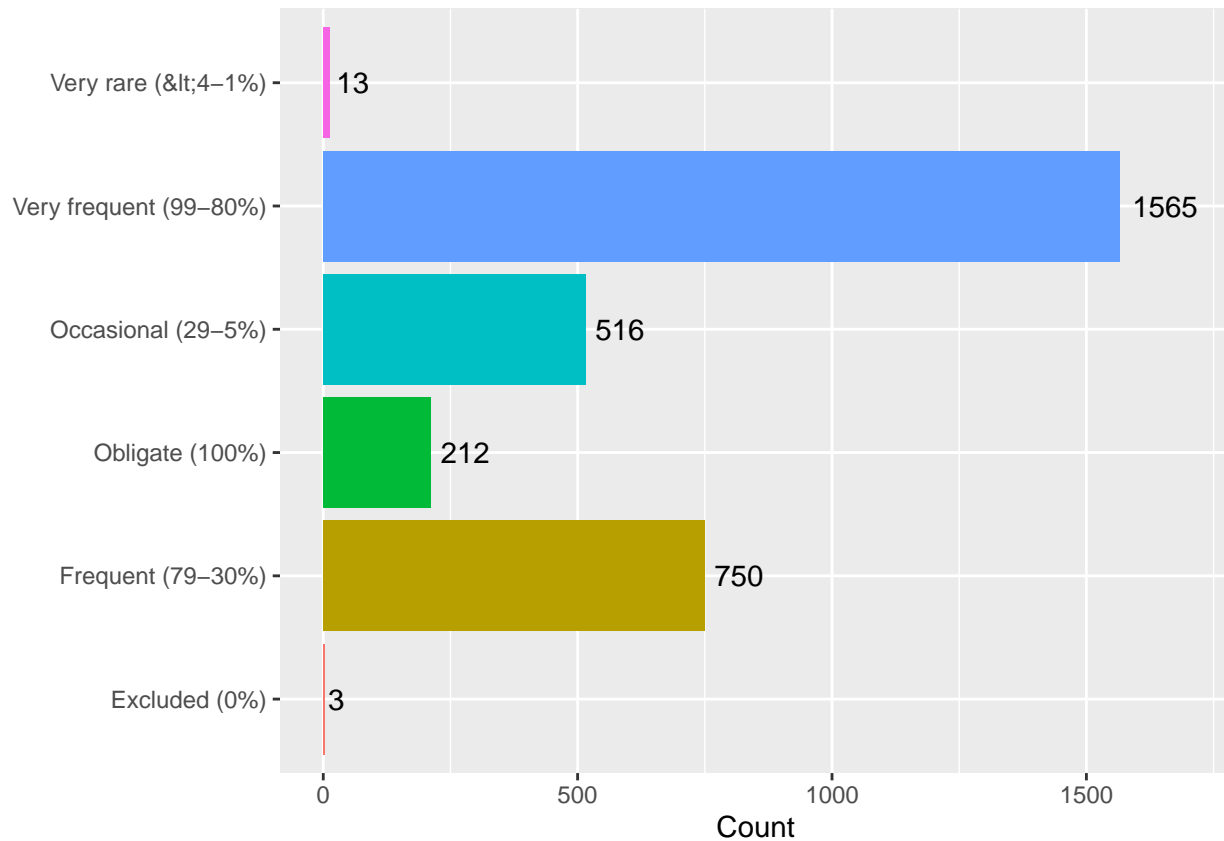
```

```

Frequency_of_abnormalities <- group_by(Phenotypes_subset, HPOFrequency0) %>%
  summarize(N = n(), .groups = "drop_last")
Plot <- ggplot(data = Frequency_of_abnormalities,
  mapping = aes(x = HPOFrequency0,
    y = N,
    fill = HPOFrequency0))
Plot + geom_col(position = "dodge") +
  labs(x = NULL, y = "Count") +

```

```
guides(fill = FALSE) +
  coord_flip() +
  geom_text(aes(label = N, hjust = -0.2)) +
  ylim(NA, 1700)
```



```
library(dplyr)
AlexanderDisease <- read.csv("Associated_Phenotypes_With_AlexanderDisease.csv")
AlexanderDisease <- AlexanderDisease %>% slice(1:61)
head(AlexanderDisease)
```

```
##           HPOTerms      HPOTerms      HPOTerms
## 1      Macrocephaly Very frequent (99-80%)
## 2      Intellectual disability Very frequent (99-80%)
## 3      Seizures Very frequent (99-80%)
## 4      Spasticity Very frequent (99-80%)
## 5      Agnesis of corpus callosum Very frequent (99-80%)
## 6      Hyperreflexia Very frequent (99-80%)
```

```
# Frequencies phenotypic abnormalities in human diseases
```

```
Frequency_of_abnormalities <- group_by(AlexanderDisease, HPOTerms) %>%
  summarize(N = n(), .groups = "drop_last")
```

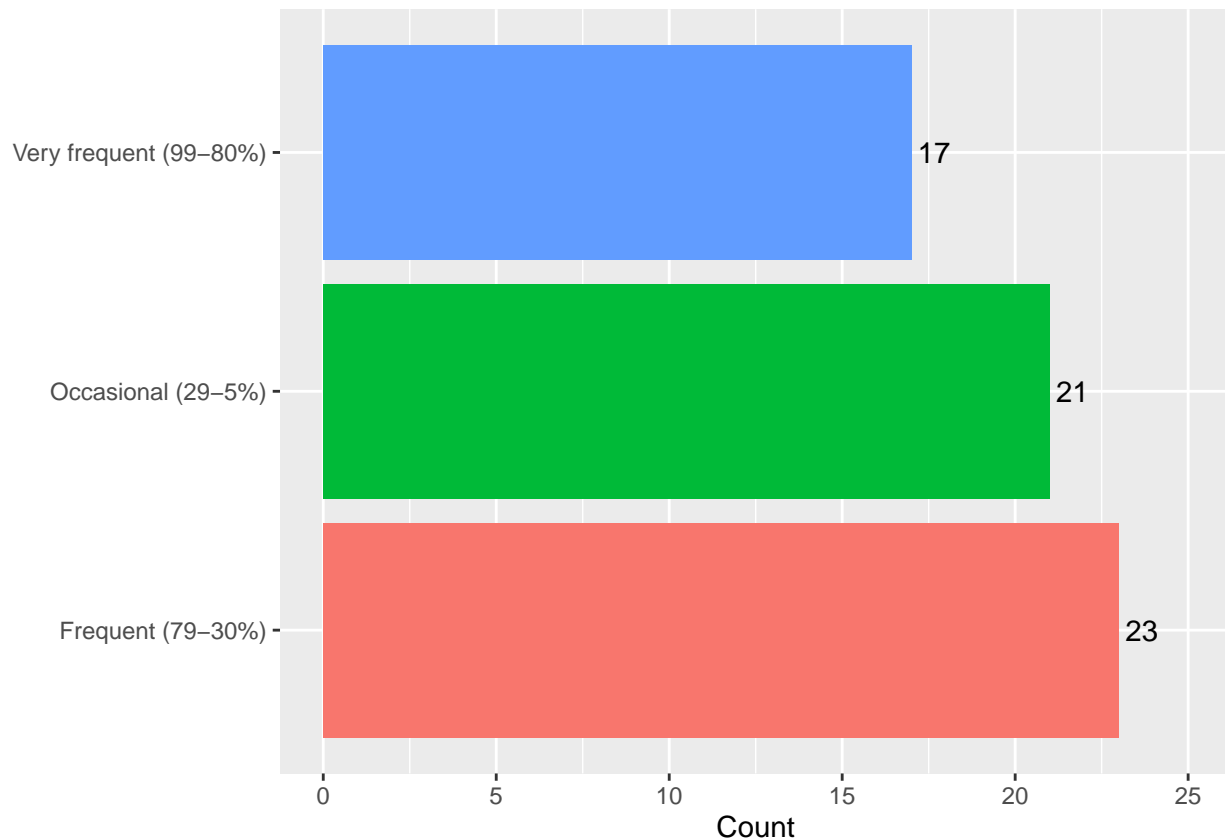
```
Frequency_of_abnormalities
```

```
## # A tibble: 3 x 2
##   HPOTerms      N
##   <chr>      <int>
## 1 Frequent (79-30%)      23
```

```
## 2 Occasional (29-5%)      21
## 3 Very frequent (99-80%)  17

# Frequencies phenotypic abnormalities in human diseases
Plot <- ggplot(data = Frequency_of_abnormalities,
               mapping = aes(x = HPOFrequencies,
                             y = N,
                             fill = HPOFrequencies))

Plot + geom_col() +
  labs(x = NULL, y = "Count") +
  guides(fill = FALSE) +
  coord_flip() +
  geom_text(aes(label = N, hjust = -0.2)) +
  ylim(NA, 25)
```



```
# Frequencies phenotypic abnormalities in human diseases
Plot <- ggplot(data = AlexanderDisease,
               mapping = aes(x = HPOFrequencies, y = HPOTerms, color = HPOFrequencies))

Plot + geom_point() + geom_text(mapping = aes(label = HPOTerms)) +
  theme(axis.text.y = element_blank(), axis.ticks.y = element_blank())
```

