

Pilot study

Preliminary investigation into using HPO for integrating Ayurvedic and Modern clinical terminologies

The ancient texts of Ayurveda provide descriptions of diseases in Sanskrit and classify the clinical conditions based on perturbations of the Vata (V), Pitta (P), and Kapha (K) doshas (1, 2). The diseases that have predominance of a single dosha are called Nanatmaj Vikaras and 80,40 and 20 conditions have been ascribed to V, P and K respectively (See the boxes below) . However, in contemporary times, Ayurveda clinicians examine patients who present with a gamut of symptoms described by modern medical terminologies. The clinician has to diagnose these diseases based on the closest matching descriptions from their texts for evolving treatment strategies.

Some of the important considerations during mapping are highlighted below

1.1 Mapping of HPO-ID to dosha attributes : Clinical aspects

Each dosha bears specific attributes that influence the phenotypic presentation during states of imbalance. For instance, attributes of dryness are a hallmark of Vata imbalance, manifesting across multiple organ systems as symptoms like dry skin (HP:0000958), dry mouth (HP:0000217), vaginal dryness (HP:0031088), dry cough (HP:0031246), and dry eyes (HP:0001097). These and other related signs to dryness are thus categorized under Vata-related phenotypes. Dosha classifications were further refined by the descriptions that reflect deviation of phenotypic expressions from a person's baseline health. Elevated Vata, for example, is characterized by phenotypes associated with pain (HP:0012514, HP:0003418) and muscle cramps (HP:0003394, HP:0032155). On the other hand, an excess of Pitta may present with bleeding (HP:0000421, HP:0040242) and burning sensations (HP:0032143, HP:6000420), while Kapha imbalance might be reflected in symptoms of anorexia (HP:0002039), excessive salivation (HP:0003781), and lethargy (HP:0001254, HP:0011973). Also phenotypic attributes related to sweat, stool, urine etc and dosha involvement in symptoms from different tissues like lymph, blood, muscle, adipose, bones, marrow, reproductive tissue, skin, nervous were considered.

1.2 Mapping of Nanatmaj Vikara to HPO IDs : clinical aspects

There are key differences between Ayurveda and modern medicine for defining diseases. This was taken into consideration during annotation of all Nanatmaj Vikara. For example

- Brittle nails (Nakhabheda): In Ayurveda, this is considered a disorder associated with the V dosha (nV), while HPO classifies it as a phenotypic abnormality of nails (HPO ID: HP:0001808) related to 28 different diseases.
- Excessive thirst (Polydipsia, trishna-adhikya): Ayurveda describes this as a P-related disorder (nP), whereas HPO lists it as a symptom (HPO ID: HP:0001959) found in over 62 diseases.
- Excessive sleepiness (Nidra adhikya): This is categorized as a K-related disorder (nK) in Ayurveda. In contrast, HPO associates it with hypersomnia (HPO ID: HP:0100786) and links it to more than 28 diseases.

To create a more objective and interoperable framework, NLP can be employed to bridge the medical terminologies of both the streams. Given that the Human Phenotype Ontology (HPO) has been integrated

into various cohorts and languages (3). Utilizing HPO could connect these Ayurvedic and modern clinical descriptions, thereby expanding access to numerous existing cohort data for Ayurveda based stratification.

Vattaj Nanatamja Vikara- The 80 conditions of Vata Perturbations

तत्रादौवातविकाराननुव्याख्यास्यामः; तद्यथा -
नखभेदश्च, विपादिका च, पादशूलं च, पादभंशश्च, पादसुप्तता च, वातखुड़ता च, गुल्फग्रहश्च, पिण्डिकोद्देष्टनं च, गृध्रसी च, जानुभेदश्च, जानुविश्लेषश्च, ऊरुस्तम्भश्च, ऊरुसादश्च, पाहङ्गुल्यं च, गुदभंशश्च, गुदर्तिश्च, वृषणा क्षेपश्च, शोफस्तम्भश्च, वङ्घ्नाणानाहश्च, श्रोणिभेदश्च, विङ्गभेदश्च, उदावर्तश्च, खञ्जत्वं च, कुञ्जत्वं च, वामनत्वं च, त्रिकग्रहश्च, पृष्ठग्रहश्च, पाश्वाविमर्दश्च, उदरावेष्टश्च, हन्मोहश्च, हृदद्रवश्च, वक्षौदघर्षश्च, वक्षौपरोधश्च, वक्षस्तोदश्च, बाहुशोषश्च, ग्रीवास्तम्भश्च, मन्यास्तम्भश्च, कण्ठोदध्वंसश्च, हनुभेदश्च, ओषुभेदश्च, अक्षिभेदश्च, दन्तभेदश्च, दन्तशैथिल्यं च, मूकत्वं च, वाक्सङ्घ-श्च, कण्ठायास्यता च, मुखशोषश्च, अरसज्जता च, ग्राणनाशश्च, कर्णशूलं च, अशब्दश्वरणं च, उच्चैःश्रुतिश्च, बाधिर्यं च, वर्त्मस्तम्भश्च, वर्त्मसङ्कोचश्च, तिमिरं च, अक्षिशूलं च, अक्षिव्युदासश्च, भृत्युदासश्च, शङ्खभेदश्च, ललाटभेदश्च, शिरोरुकं च, केशभूमिस्फुटनं च, अर्दितं च, एकाङ्गरोगश्च, सर्वाङ्गरोगश्च, पक्षवधश्च, आक्षेपकश्च, दण्डकश्च, तमश्च, भ्रमश्च, वेपथुश्च, जृम्भा च, हिकका च, विषादश्च, अतिप्रलापश्च, रौक्षयं च, पारुष्यं च, श्यावारुणावभासता च, अस्वप्रश्च, अनवस्थितचित्तत्वं च; इत्यशीतिर्वार्तविकारा वात-विकाराणामपरिसङ्ख्येयानामाविष्कृततमा व्याख्याताः।

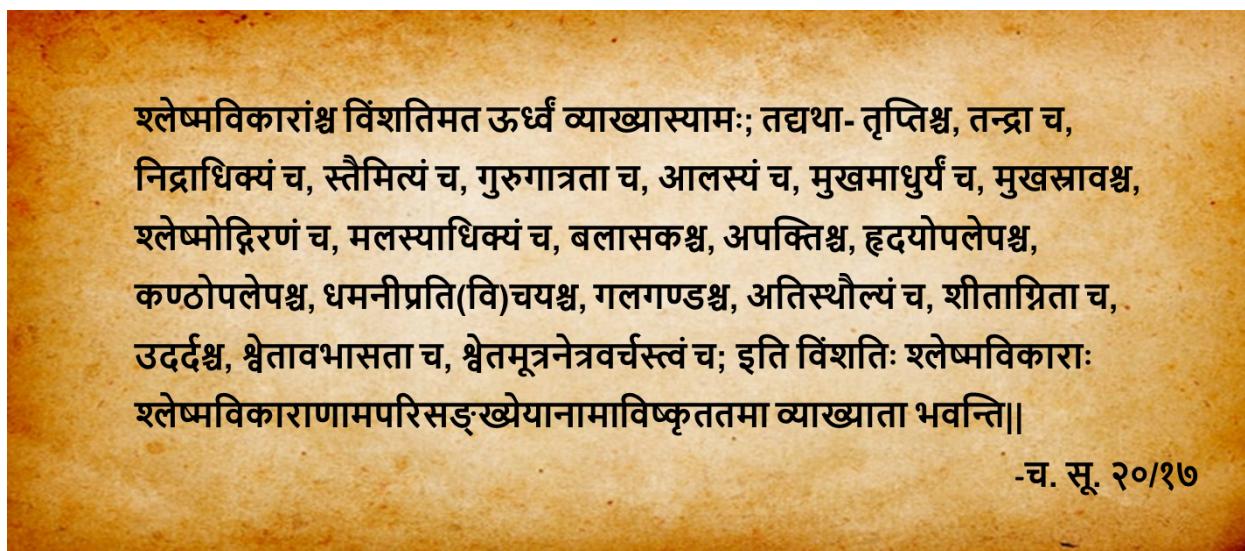
-च. सू. २०/११

Pittaj Nanatamja Vikara- The 40 conditions of Pitta Perturbations

पित्तविकारांश्वत्वारिंशतमत ऊर्ध्वमनुव्याख्यास्यामः-
ओषश्च, प्लोषश्च, दाहश्च, दवथुश्च, धूमकश्च, अम्लकश्च, विदाहश्च, अन्तर्दाहश्च, अंसदाहश्च, ऊष्माधिक्यं च, अतिस्वेदश्च(अङ्गस्वेदश्च), अङ्गन्धश्च, अङ्गावदरणं च, शोणितक्लेदश्च, मांसक्लेदश्च, त्वग्दाहश्च, (मांसदाहश्च), त्वगवदरणं च, चर्मदलनं च, रक्तकोठश्च, रक्त विस्फोटश्च, रक्तपित्तं च, रक्तमण्डलानि च, हरितत्वं च, हारिद्रत्वं च, नीलिका च, कक्षा (क्ष्या)च, कामला च, तिक्तास्यता च, लोहितगन्धास्यता च, पूतिमुखता च, तृष्णाधिक्यं च, अतृप्तिश्च, आस्यविपाकश्च, गलपाकश्च, अक्षिपाकश्च, गुदपाकश्च, मेढपाकश्च, जीवादानं च, तमःप्रवेशश्च, हरितहारिद्रनेत्रमूत्रवर्चस्त्वं च; इति चत्वारिंशतिपित्तविकाराः पित्तविकाराणामपरिसङ्ख्येयानामाविष्कृततमा व्याख्याताः।

-च. सू. २०/१४

Kaphaj Nanatamja Vikara- The 20 conditions of Kapha Perturbations



In order to explore the ontological links between Ayurveda and modern medical terminologies we conducted a pilot study where we explored three possibilities of linkages

1. Nanatmaj Vikara in HPO diseases
2. Associations of clinical features of dosha with modern diseases
3. Links to OMIM diseases from molecular links of doshas

Case Study 1- Nanatmaj Vikara in HPO diseases

A condition of bleeding risk (Von Willebrand disease) that is also referred to amongst one of the Pittaj Nanatmaj Vikara (Raktapitta) was first considered. In an earlier study variation in a gene VWF linked to this disease was also associated with Pitta phenotypes (4). A search term for “Abnormality of von Willebrand factor” in HPO yielded 19 diseases listed below (Table 1).(version: May version, dated 2022-10-05).

Table 1: List of syndromes with a search for “Abnormality of von Willebrand factor” term in the HPO database

HPO Disease Id	HPO Disease Name
ORPHA:251061	7q31 microdeletion syndrome
ORPHA:99147	Acquired von Willebrand syndrome
ORPHA:274	Bernard-Soulier syndrome
OMIM:231200	Bernard-Soulier syndrome

OMIM:153670	Bernard-Soulier syndrome, type A2, autosomal dominant
OMIM:614201	Bleeding disorder, platelet-type, 11
OMIM:618462	Bleeding disorder, platelet-type, 22
OMIM:619271	Bleeding disorder, platelet-type, 24, autosomal dominant
ORPHA:70591	Chronic thromboembolic pulmonary hypertension
OMIM:273800	Glanzmann thrombasthenia
ORPHA:849	Glanzmann thrombasthenia
OMIM:619267	Glanzmann thrombasthenia 2
ORPHA:79259	Glycogen storage disease due to glucose-6-phosphatase deficiency type Ib
OMIM:139090	Gray platelet syndrome
ORPHA:169802	Severe hemophilia A
OMIM:619130	Thrombocytopenia, autosomal dominant, 7
ORPHA:903	Von Willebrand disease
OMIM:193400	von Willebrand disease, type 1
OMIM:277480	von Willebrand disease, type 3

Each of the above diseases are associated with many clinical features each of which are assigned an HP ID. For example one of the diseases listed above Glanzmann thrombasthenia ([OMIM ID 273800](#)) in HPO lists features with HP IDs that are given in Table 2.

Table 2: List of phenotypes associated with Glanzmann thrombasthenia ([OMIM ID 273800](#))

	HPO term id	HPO term name
1	HP:0011873	Abnormal platelet count
2	HP:0000007	Autosomal recessive inheritance
3	HP:0000978	Bruising susceptibility
4	HP:0001975	Decreased platelet glycoprotein IIb-IIIa
5	HP:0031364	Ecchymosis
6	HP:0000421	Epistaxis
7	HP:0030138	Excessive bleeding from superficial cuts
8	HP:0002239	Gastrointestinal hemorrhage
9	HP:0000225	Gingival bleeding
10	HP:0004866	Impaired ADP-induced platelet aggregation
11	HP:0031126	Impaired clot retraction
12	HP:0008320	Impaired collagen-induced platelet aggregation
13	HP:0008148	Impaired epinephrine-induced platelet aggregation
14	HP:0003540	Impaired platelet aggregation

HPO Disease Id	HPO Disease Name	HPO Term Id	HPO Term Name	VPK Mapping
		HP:0000132	Menorrhagia	V/P
		HP:0030138	Excessive bleeding from superficial cuts	P
OMIM:273800	Glanzmann thrombasthenia	HP:0003540	Impaired platelet aggregation	V/K
		HP:0003010	Prolonged bleeding time	V
		HP:0031364	Ecchymosis	P
		HP:0002170	Intracranial hemorrhage	P
		HP:0031126	Impaired clot retraction	P/K
		HP:0011873	Abnormal platelet count	V
		HP:0011871	Impaired ristocetin-induced platelet aggregation	K
		HP:0000979	Purpura	K/nP
		HP:0008320	Impaired collagen-induced platelet aggregation	V
		HP:0000007	Autosomal recessive inheritance	V
		HP:0100309	Subdural hemorrhage	P
		HP:0000978	Bruising susceptibility	P
		HP:0008148	Impaired epinephrine-induced platelet aggregation	V
		HP:0002239	Gastrointestinal hemorrhage	P
		HP:0000421	Epistaxis	P
		HP:0004866	Impaired ADP-induced platelet aggregation	K
		HP:0001975	Decreased platelet glycoprotein IIb-IIIa	V
		HP:0003623	Neonatal onset	V
		HP:0000225	Gingival bleeding	P
15	HP:0011871	Impaired ristocetin-induced platelet aggregation		
16	HP:0002170	Intracranial hemorrhage		
17	HP:0000132	Menorrhagia		
18	HP:0003623	Neonatal onset		
19	HP:0003010	Prolonged bleeding time		
20	HP:0000979	Purpura		
21	HP:0100309	Subdural hemorrhage		

All the HP IDs (149 features) related to the above syndromes were provided to domain experts/Ayurveda clinicians for assignment into V, P and K based on their clinical description and matches to Ayurveda texts. These features were then mapped back to each of the OMIM diseases. Table 3 illustrates an example of such a labeling with their corresponding mapping to Glanzmann thrombasthenia.

Table 3: Labeling of the V/P/K phenotypes in one of the disease

The cumulative count of V/P/K features for each disease was aggregated in a matrix as shown below (Table 4) and plotted to observe the trend of labeled phenotypes (Figure 1). The results reveal significant differences in the number of V/P/K features. V features are highest (as described in SI Methods) followed by P and K. Bleeding is associated with imbalance of P and the proportion of P seems to be reflected in all the diseases associated with bleeding. This would be more evident when we compare it with the example below. These diseases to an Ayurveda doctor in a clinical setting would be considered as an imbalance of Pitta and termed as Raktapitta .

Table 4: Cumulative count of phenotypes labeled as V, P and K for the diseases

HPO Disease Id	HPO Disease Name	Total Count	V Count	P Count	K Count
OMIM:273800	Glanzmann thrombasthenia	9	12	5	
OMIM:619267	Glanzmann thrombasthenia 2	10	6	3	
ORPHA:79259	Glycogen storage disease due to glucose-6-phosphatase deficiency type Ib	40	25	11	
OMIM:139090	Gray platelet syndrome	13	6	2	
ORPHA:169802	Severe hemophilia A	14	16	4	
OMIM:619130	Thrombocytopenia, autosomal dominant, 7	6	3	2	
OMIM:193400	Von willebrand disease, type 1	13	11	2	
OMIM:277480	Von willebrand disease, type 3	10	10	1	
ORPHA:99147	Acquired von Willebrand syndrome	14	18	2	
OMIM:231200	Bernard-Soulier syndrome	11	9	2	
OMIM:619271	Bleeding disorder, platelet-type, 24, autosomal dominant	6	6	3	
ORPHA:70591	Chronic thromboembolic pulmonary hypertension	28	9	16	

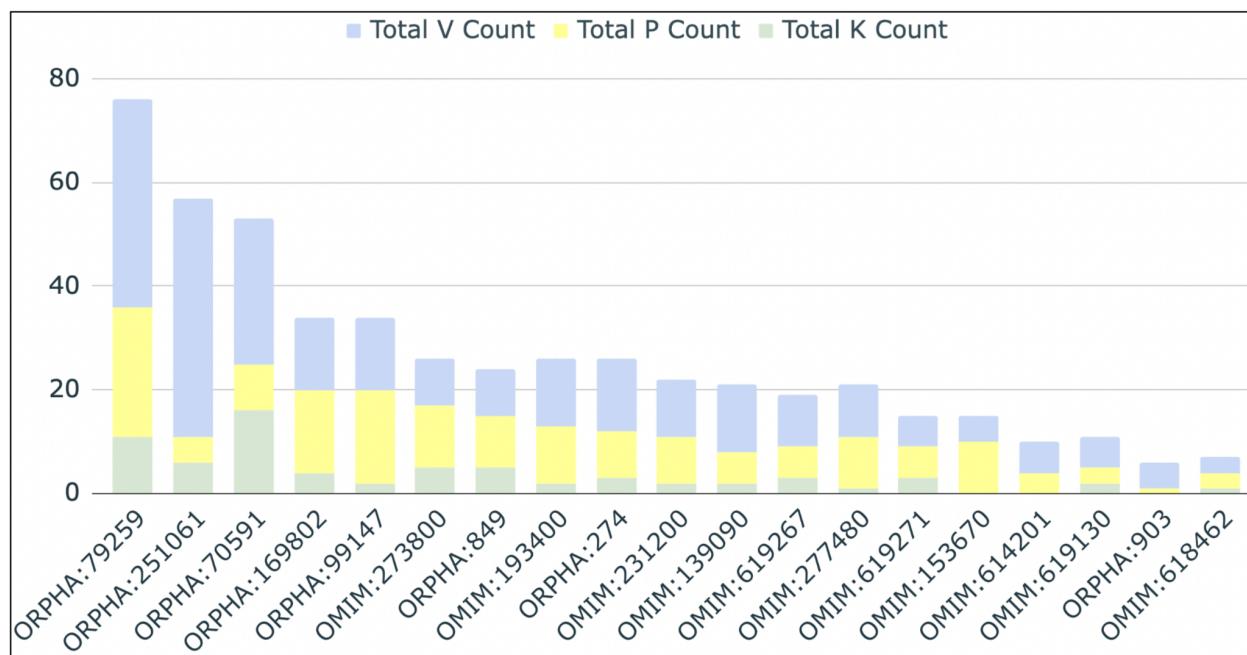
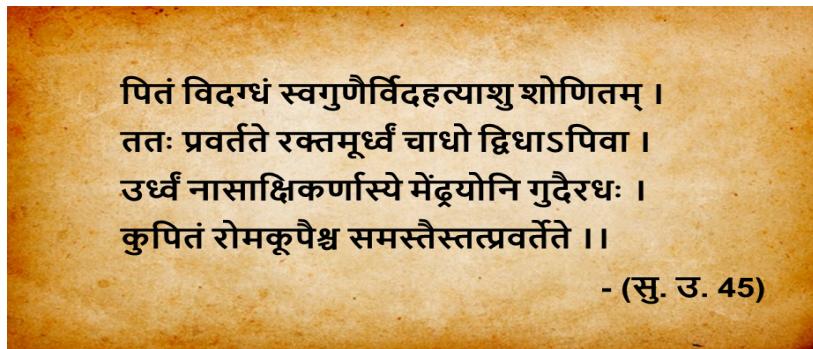


Figure 1: The barplot depicts the total count of V, P and K phenotypes across all the 19 diseases

The clinical features of bleeding disorder described as RaktaPitta from ancient text is given below



Case Study 2- Associations of clinical features of dosha with modern diseases

We carried out a similar exercise starting from a clinical condition of Ataxia (HP:0001251). Ataxia HPO yields 1255 diseases along with a set of 4029 non-redundant phenotypes (version: May version, dated 2022-10-05). Despite its heterogeneity, an Ayurveda clinician in practice would ascribe ataxia as a disease of Vata manifestation (1). A similar exercise as above was conducted on a set of 25 diseases associated with ataxia. As is evident, in contrast to the bleeding associated diseases, this set had a more frequent and significant presence of V features.

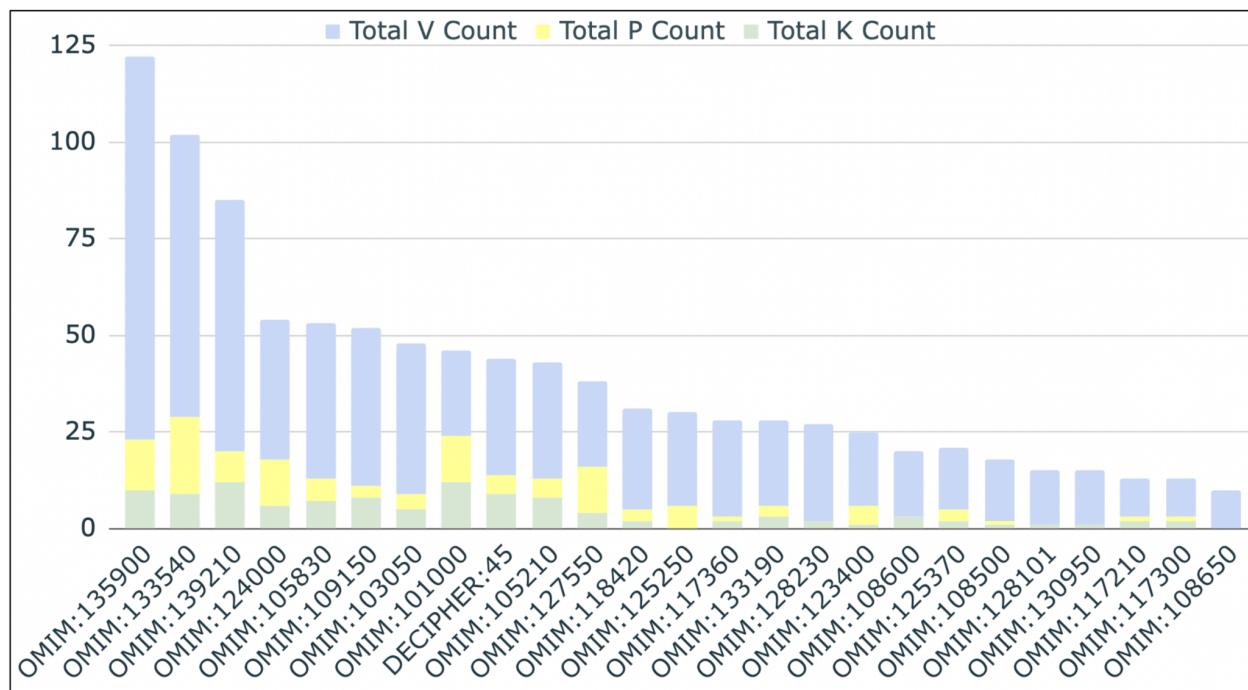


Figure 2: The barplot depicts the total count of V, P and K phenotypes across 25 diseases (subset). The cumulative count of V is observed high for this subset.

Case Study 3 - Associations of clinical features of dosha with modern diseases

Ayurveda also has descriptions of dosha specific cellular functions that have been correlated in our earlier studies (5). For example Vata is described in Ayurveda to be responsible for cell division and morphogenesis. Enhanced rates of cell proliferation rates have been reported in the transcriptome as well as cell lines derived from Vata individuals. (Cell Cycle). A query for “abnormality of cell cycle” in HPO provides four diseases associated with Fanconi anemia (Table 5). Annotation of Fanconi anemia with Ayurveda doshas reveal enrichment of Vata (Table 6, Figure 3)

Table 5: List of syndromes with a search for “abnormality of cell cycle” term in the HPO database

HPO Disease Id	HPO Disease Name
OMIM:227650	Fanconi anemia
OMIM:227645	Fanconi anemia, complementation group C
OMIM:227646	Fanconi anemia, complementation group D2
OMIM:600901	Fanconi anemia, complementation group E

Table 6: Labeling of the phenotypes in one of the disease

HPO Disease Id	HPO Disease Name	HPO Term Id	HPO Term Name	VPK Mapping
OMIM:227650	Fanconi Anemia	HP:0030680	Abnormal cardiovascular system morphology	V
		HP:0001627	Abnormal heart morphology	V
		HP:0012210	Abnormal renal morphology	V
		HP:0001000	Abnormality of skin pigmentation	P
		HP:0003974	Absent radius	V
		HP:0009777	Absent thumb	V
		HP:0001903	Anemia	P
		HP:0001017	Anemic pallor	P
		HP:0000007	Autosomal recessive inheritance	V
		HP:0000978	Bruising susceptibility	P
		HP:0000957	Cafe-au-lait spot	V/P
		HP:0003221	Chromosomal breakage induced by crosslinking agents	V
		HP:0009943	Complete duplication of thumb phalanx	V
		HP:0000028	Cryptorchidism	V
		HP:0003213	Deficient excision of UV-induced pyrimidine dimers in DNA	V
		HP:0000081	Duplicated collecting system	K
		HP:0000086	Ectopic kidney	V

		HP:0000365	Hearing impairment	V/K
		HP:0000085	Horseshoe kidney	V
		HP:0000815	Hypergonadotropic hypogonadism	V
		HP:0001249	Intellectual disability	P
		HP:0001909	Leukemia	V/P/K
		HP:0003251	Male infertility	V
		HP:0000252	Microcephaly	V
		HP:0000568	Microphthalmia	V
		HP:0001875	Neutropenia	V
		HP:0001876	Pancytopenia	V
		HP:0003214	Prolonged G2 phase of cell cycle	V
		HP:0000104	Renal agenesis	V
		HP:0001896	Reticulocytopenia	V
		HP:0004322	Short stature	V/nV
		HP:0009778	Short thumb	V
		HP:0001518	Small for gestational age	V
		HP:0000486	Strabismus	V
		HP:0001873	Thrombocytopenia	V

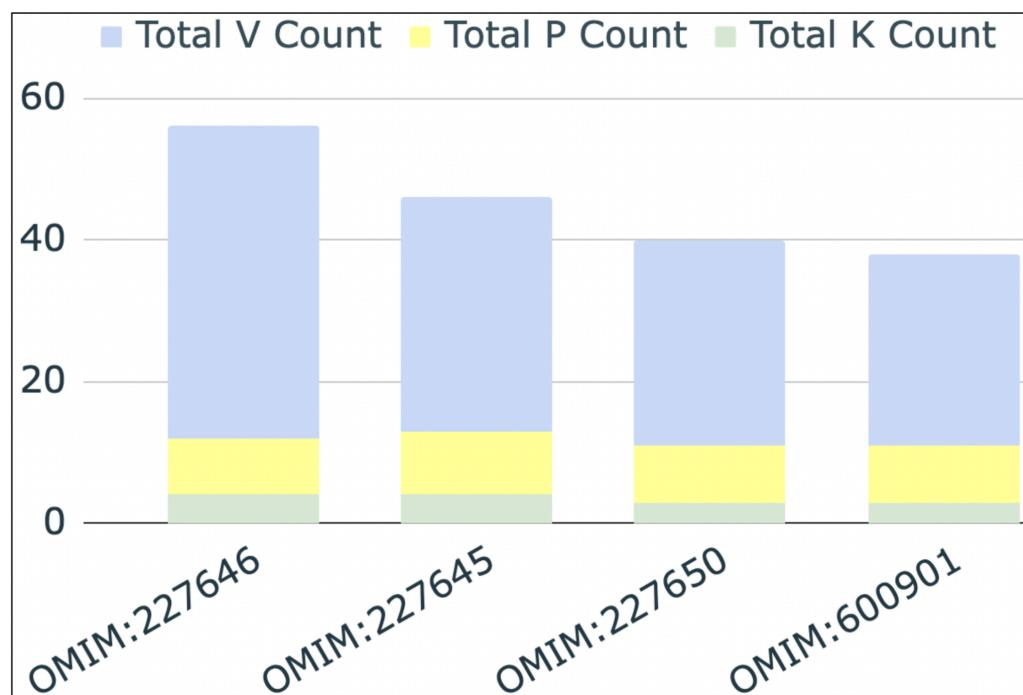


Figure 3: The barplot depicts the total count of V, P and K phenotypes across all the 4 diseases. The cumulative count for V as a major phenotype is evident from the figure. Since cell division and morphogenesis is associated with imbalance of V

Based on the above exercise we surmised that there is a possibility for integrating Ayurveda and modern medicine clinical descriptions using the interface of HPO to probe the ontological links. This could be either through clinical features from ayurveda or modern medicine as well as functional attributes described for dosha. An extensive effort was undertaken to assign 12000 diseases associated with a non redundant set of over 10000 phenotypes to understand the rare diseases from Ayurveda perspective.

References

1. P. V. Sharma, Charaka Samhita Text With English Translation P. V. Sharma (Chaukambha Orientalia Publisher, 1981).
2. B. Prasher, et al., Whole genome expression and biochemical correlates of extreme constitutional types defined in Ayurveda. *J Transl Med* 6, 48 (2008).
3. M. A. Gargano, et al., The Human Phenotype Ontology in 2024: phenotypes around the world. *Nucleic Acids Res* 52, D1333–D1346 (2024).
4. S. Aggarwal, et al., Combined genetic effects of EGLN1 and VWF modulate thrombotic outcome in hypoxia revealed by Ayurgenomics approach. *J Transl Med* 13, 1–11 (2015).
5. S. Chakraborty, et al., Baseline cell proliferation rates and response to UV differ in lymphoblastoid cell lines derived from healthy individuals of extreme constitution types. *Cell Cycle* 20, 903–913 (2021).