We considered multiple numerical attributes that could descried how different input data could influence strength of PathFX predictions. We were particularly interested in how side-effect-associated genes in drug networks ("neighborhood gene lists") and how side-effect pathway genes ("interactome gene lists") changed. We pursued multiple metrics that quantified changes in neighborhood and interactome gene lists (**Table X**).

Table X. Numerical attribute quantified changes in underlying PathFX data – specifically changes in neighborhood and interactome gene lists and relationships between these data.

Single Version – Numerical Attributes

What's the neighborhood size relative to the pathway size?

Difference between Interactome − Neighborhood Gene List Size for PathFX Version (X) = Interactome Gene List Size Version (X) − Neighborhood Gene List Size Version (X)
interactome_gene_lst_neighborhood_gene_lst_difference_v1
interactome_gene_lst_neighborhood_gene_lst_difference_v2

What's the ratio of the neighborhood size relative to the pathway size?

$$\text{Ratio of Neighborhood/Interactome Gene List Size (Neighborhood Gene List Percentage) for PathFX Version (X)} = \frac{\text{Neighborhood Gene List Size Version (X)}}{\text{Interactome Gene List Size Version (X)}}$$

neighborhood_gene_lst_percentage_v1
neighborhood_gene_lst_percentage_v2

$$\text{Ratio of Unique Neighborhood/Interactome Gene List Size (Unique Neighborhood Gene List Percentage) for PathFX Version (X)} = \frac{\text{Unique Neighborhood Gene List Size Version (X)}}{\text{Interactome Gene List Size Version (X)}}$$

unique_neighborhood_gene_lst_percentage_v1
unique_neighborhood_gene_lst_percentage_v2

$$\text{Ratio of Intersecting Neighborhood/Interactome Gene List Size for PathFX Version (X)} = \frac{\text{Intersecting Neighborhood Gene List Size Version}}{\text{Interactome Gene List Size Version (X)}}$$

intersecting_neighborhood_gene_lst_interactome_gene_lst_percentage_v1
intersecting_neighborhood_gene_lst_interactome_gene_lst_percentage_v2

$$\text{Unique Interactome Gene List Percentage for PathFX Version (X)} = \frac{\text{Unique Interactome Gene List Size Version (X)}}{\text{Interactome Gene List Size Version (X)}}$$

unique_interactome_gene_lst_percentage_v1
unique_interactome_gene_lst_percentage_v2

$$\text{Intersecting Interactome Gene List Percentage for PathFX Version (X)} = \frac{\text{Intersecting Interactome Gene List Size}}{\text{Interactome Gene List Size Version (X)}}$$

intersecting_interactome_gene_lst_percentage_v1
intersecting_interactome_gene_lst_percentage_v2

Drug Target Gene List Size for PathFX Version (X) (for each PathFX Version) = Unique Drug Targets Gene List Size Version (X) + Intersecting Drug Targets Gene List Size
drug_targets_v1_size
drug_targets_v2_size

$$\text{Ratio of Drug Targets to Neighborhood Genes for PathFX Version (X) (for each PathFX Version)} = \frac{\text{Drug Targets Gene List Size Version (X)}}{\text{Neighborhood Gene List Size Version (X)}}$$

ratio_drug_target_gene_over_neigh_gene_lst_v1_size
ratio_drug_target_gene_over_neigh_gene_lst_v2_size

Single Version – Numerical Attributes Across all 355 Records Statistics Table:

	count	mean	std	min	max	sum
Difference between Interactome-Neighborhood Gene List Size for PathFX Version 1	355.0	447.611268	324.592549	29.000000	1015.000000	158902.000000
Difference between Interactome-Neighborhood Gene List Size for PathFX Version 2	355.0	609.095775	630.910616	31.000000	1771.000000	216229.000000
Ratio of Neighborhood/Interactome Gene List Size (Neighborhood Gene List Percentage) for PathFX Version 1	355.0	0.071785	0.053877	0.003610	0.296875	25.483677
Ratio of Neighborhood/Interactome Gene List Size (Neighborhood Gene List Percentage) for PathFX Version 2	355.0	0.056050	0.042022	0.002008	0.220000	19.897870
Ratio of Unique Neighborhood/Interactome Gene List Size (Unique Neighborhood Gene List Percentage) for PathFX Version 1	355.0	0.029199	0.028248	0.000000	0.187500	10.365789
Ratio of Unique Neighborhood/Interactome Gene List Size (Unique Neighborhood Gene List Percentage) for PathFX Version 2	355.0	0.010320	0.014166	0.000000	0.138298	3.663653
Ratio of Intersecting Neighborhood/Interactome Gene List Size for PathFX Version 1	355.0	0.042586	0.031047	0.000000	0.156250	15.117888
Ratio of Intersecting Neighborhood/Interactome Gene List Size for PathFX Version 2	355.0	0.045730	0.038052	0.000000	0.200000	16.234217
Unique Interactome Gene List Percentage for PathFX Version 1	355.0	0.287294	0.138312	0.142857	0.758209	101.989193
Unique Interactome Gene List Percentage for PathFX Version 2	355.0	0.301335	0.155418	0.027778	0.535996	106.973798
Intersecting Interactome Gene List Percentage for PathFX Version 1	355.0	0.712706	0.138312	0.241791	0.857143	253.010807
Intersecting Interactome Gene List Percentage for PathFX Version 2	355.0	0.698665	0.155418	0.464004	0.972222	248.026202
Drug Target Gene List Size for PathFX Version 1	355.0	12.760563	10.081399	1.000000	44.000000	4530.000000
Drug Target Gene List Size for PathFX Version 2	355.0	13.177465	10.095283	1.000000	42.000000	4678.000000
Ratio of Drug Targets to Neighborhood Genes for PathFX Version 1	355.0	0.747682	0.784428	0.032258	5.666667	265.427276
Ratio of Drug Targets to Neighborhood Genes for PathFX Version 1	355.0	1.134804	1.220876	0.052632	10.000000	402.855320

Single Version – Numerical Attributes Across Grouped by Adverse Events Statistics Table:

	Difference between Interactome-Neighborhood Gene List Size for PathFX Version 1						Difference between Interactome-Neighborhood Gene List Size for PathFX Version 2				...	Ratio of Drug Targets to Neighborhood Genes for PathFX Version 1									
	count	mean	std	min	max	sum	count	mean	std	min	...	std	min	max	sum	count	mean	std	min	max	sum
Phenotype																					
Cerebral Infarction	5	284.800000	7.661593	273	290	1424	5	274.600000	4.979960	267	...	0.522716	0.312500	1.571429	4.235119	5	1.430000	0.857030	0.600000	2.750000	7.150000
Deep Vein Thrombosis	2	107.000000	4.242641	104	110	214	2	105.000000	2.828427	103	...	0.707107	0.500000	1.500000	2.000000	2	1.333333	0.942809	0.666667	2.000000	2.666667
Delirium	11	29.727273	1.009050	29	32	327	11	38.545455	1.293340	37	...	1.480076	1.200000	5.666667	38.200000	11	3.833117	1.547186	1.571429	6.166667	42.164286
Edema	18	315.277778	7.135706	300	329	5675	18	84.777778	1.664705	83	...	0.417205	0.181818	1.571429	14.705533	18	2.183752	1.146850	0.333333	4.333333	39.307540
Gastric ulcer	26	103.192308	4.578377	93	111	2683	26	84.038462	3.736102	77	...	0.587928	0.125000	2.555556	25.098734	26	1.238274	0.795099	0.333333	3.857143	32.195120
Hemorrhage	12	69.833333	1.642245	68	72	838	12	33.500000	1.243163	31	...	0.616774	0.250000	2.000000	12.350000	12	2.011111	1.160140	0.500000	4.000000	24.133333
Hyperlipidemia	19	272.894737	10.413105	251	291	5185	19	218.263158	5.762350	205	...	0.491208	0.222222	2.333333	17.511839	19	1.494304	0.754164	0.437500	3.875000	28.391781
Myocardial Infarction	80	976.600000	20.483151	923	1015	78128	80	1737.525000	18.607182	1699	...	0.182010	0.047619	1.250000	19.862715	80	0.270876	0.180799	0.058824	1.222222	21.670068
Pancreatitis	24	262.291667	8.201162	249	276	6295	24	275.416667	5.792924	266	...	0.581504	0.352941	2.857143	24.365227	24	1.525234	0.879540	0.600000	3.750000	36.605624
Peripheral Neuropathy	1	739.000000	NaN	739	739	739	1	301.000000	NaN	301	...	NaN	0.294118	0.294118	0.294118	1	0.857143	NaN	0.857143	0.857143	0.857143
Pneumonia	70	437.957143	10.711548	422	469	30657	70	471.471429	7.777190	460	...	0.263758	0.032258	1.083333	28.517235	70	0.529216	0.392362	0.052632	2.000000	37.045109
Proteinuria	4	198.750000	5.377422	192	205	795	4	194.250000	3.862210	190	...	0.217858	0.250000	0.727273	2.290998	4	1.066288	0.583788	0.333333	1.750000	4.265152
Pulmonary Edema	3	69.666667	2.309401	67	71	209	3	70.333333	0.577350	70	...	0.608276	1.900000	3.000000	6.900000	3	2.992063	1.319729	1.833333	4.428571	8.976190
Sepsis	33	574.454545	11.803312	550	599	18957	33	528.969697	9.970221	508	...	0.262272	0.043478	1.037037	12.877782	33	0.565489	0.458520	0.083333	2.000000	18.661145
Tardive Dyskinesia	24	50.416667	4.221237	45	59	1210	24	42.833333	2.200132	39	...	0.527849	0.538462	2.315789	34.409235	24	2.804012	0.935086	1.400000	5.571429	67.296284
Thrombocytopenia	23	242.000000	6.640099	231	252	5566	23	329.304348	7.594464	316	...	0.708297	0.142857	3.333333	21.808741	23	1.368256	1.950232	0.166667	10.000000	31.469879

16 rows × 96 columns

Two Versions– Numerical Attributes

Set Symmetric Difference in Neighborhood Gene List Size = Neighborhood Gene List Size Version 1 + Unique Neighborhood Gene List Size Version 2
neigh_genes_v1_and_neigh_genes_v2_set_symmetric_difference_size

Set Symmetric Difference in Interactome Gene List Size = Unqiue Interactome Gene List Size Version 1 + Unique Interactome Gene List Size Version 2
intom_genes_v1_and_intom_genes_v2_set_symmetric_difference_size

Set Symmetric Difference in Drug Target Gene List Size = Unqiue Drug Target Gene List Size Version 1 + Unqiue Drug Target Gene List Size Version 2
drug_targets_v1_and_drug_targets_v2_set_symmetric_difference_size

Are there differences in the number of side effect neighborhood genes?
Difference in Neighborhood Gene List Size = Neighborhood Gene List Size Version 2 − Neighborhood Gene List Size Version 1
version_difference_gene_neighborhood_lst_size

How many unique neighborhood genes are associated in each version?
Difference in Unique Neighborhood Gene List Size = Unique Neighborhood Gene List Size Version 2 − Unique Neighborhood Gene List Size Version 1
version_difference_unique_gene_neighborhood_lst_size

Are the sizes of the pathway different between versions?
Difference in Interactome Gene List Size = Interactome Gene List Size Version 2 − Interactome Gene List Size Version 1
version_difference_gene_interactome_lst_size

How many unique side-effect genes are in each version?
Difference in Unique Interactome Gene List = Unique Interactome Gene List Size Version 2 − Unique Interactome Gene List Size Version 1
version_difference_unique_gene_interactome_lst_size

How much of the pathway is disengaged?
Difference in NonNeighborhood Gene List Size = (Interactome Gene List Size Version 2 − Neighborhood Gene List Size Version 2) − (Interactome Gene List Size Version 1 − Neighborhood Gene List Size Version 1)
version_difference_interactome_gene_lst_neighborhood_gene_lst_difference

Are there differences in the number of side effect neighborhood genes? (Normalized to interactome list size)
Difference in Neighbrohood Gene List Percentage = $\frac{\text{Neighborhood Gene List Size Version 2}}{\text{Interactome Gene List Size Version 2}} - \frac{\text{Neighborhood Gene List Size Version 1}}{\text{Interactome Gene List Size Version 1}}$
version_difference_neighborhood_gene_lst_percentage

How many unique neighborhood genes are associated in each version? (Normalized to interactome list size)
Difference in Unique Neighborhood Gene List Percentage = $\frac{\text{Unique Neighborhood Gene List Size Version 2}}{\text{Interactome Gene List Size Version 2}} - \frac{\text{Unique Neighborhood Gene List Size Version 1}}{\text{Interactome Gene List Size Version 1}}$
version_difference_unique_neighborhood_gene_lst_percentage

Difference in Neighborhood Gene List Percentage with Intersecting Interactome Gene List Size = $\frac{\text{Neighborhood Gene List Size Version 2} - \text{Neighborhood Gene List Size Version 1}}{\text{Interactome Gene Interaction List Size}}$
version_difference_neighborhood_gene_lst_percentage_intersecting_interactome_gene_lst

How many unique neighborhood genes are associated in each version? (normalized to intersection of interactome lists between versions)
Difference in Unique Neighborhood Gene List Percentage with Intersecting Interactome Gene List Size = $\frac{\text{Unique Neighborhood Gene List Size Version 2} - \text{Unique Neighborhood Gene List Size Version 1}}{\text{Interactome Gene Interaction List Size}}$
version_difference_unique_neighborhood_gene_lst_percentage_intersecting_interactome_gene_lst

Of the common genes, how much do these overlap with the interactome gene lists?
Difference in Intersecting Neighborhood Gene List Percentage = $\frac{\text{Intersecting Neighborhood Gene List Size}}{\text{Interactome Gene List Size Version 2}} - \frac{\text{Intersecting Neighborhood Gene List Size}}{\text{Interactome Gene List Size Version 1}}$

version_difference_intersecting_neighborhood_gene_lst_interactome_gene_lst_percentage

$$\text{Difference in Unique Interactome Gene List Percentage} = \frac{\text{Unique Interactome Gene List Size Version 2}}{\text{Interactome Gene List Size Version 2}} - \frac{\text{Unique Interactome Gene List Size Version 1}}{\text{Interactome Gene List Size Version 1}}$$

version_difference_unique_interactome_gene_lst_percentage

$$\text{Difference in intersecting Interactome Gene List Percentage} = \frac{\text{Intersecting Interactome Gene List Size}}{\text{Interactome Gene List Size Version 2}} - \frac{\text{Intersecting Interactome Gene List Size}}{\text{Interactome Gene List Size Version 1}}$$

version_difference_intersecting_interactome_gene_lst_percentage

Are the number of targets per drug changing?

$$\text{Difference in Drug Target Gene List Size} = \text{Drug Targets Gene List Size Version 2} - \text{Drug Targets Gene List Size Version 1}$$

version_difference_drug_target_gene_lst

$$\text{Difference in Unique Drug Target Gene List Size} = \text{Unique Drug Targets Gene List Size Version 2} - \text{Unique Drug Targets Gene List Size Version 1}$$

version_difference_unique_drug_target_gene_lst

Do the number of drug targets overlap with the pathway more in one version?

$$\text{Difference in Drug Target Gene List Percentage} = \frac{\text{Drug Targets Gene List Size Version 2}}{\text{Interactome Gene List Size Version 2}} - \frac{\text{Drug Targets Gene List Size Version 1}}{\text{Interactome Gene List Size Version 1}}$$

version_difference_drug_target_gene_lst_percentage

$$\text{Difference in Unqiue Drug Target Gene List Percentage} = \frac{\text{Unique Drug Targets Gene List Size Version 2}}{\text{Interactome Gene List Size Version 2}} - \frac{\text{Unique Drug Targets Gene List Size Version 1}}{\text{Interactome Gene List Size Version 1}}$$

version_difference_unique_drug_target_gene_lst_percentage

$$\text{Difference in intersecting Drug Target Gene List Percentage} = \frac{\text{intersecting Drug Targets Gene List Size}}{\text{Interactome Gene List Size Version 2}} - \frac{\text{intersecting Drug Targets Gene List Size}}{\text{Interactome Gene List Size Version 1}}$$

version_difference_intersecting_drug_target_gene_lst_percentage

$$\text{Difference in Ratio of Drug Targets to Neighborhood Genes} = \frac{\text{Drug Targets Gene List Size Version 2}}{\text{Neighborhood Gene List Size Version 2}} - \frac{\text{Drug Targets Gene List Size Version 1}}{\text{Neighborhood Gene List Size Version 1}}$$

version_difference_ratio_drug_target_gene_over_neigh_gene_lst_size

Two Versions – Numerical Attributes Across all 355 Records Statistics Table:

	count	mean	std	min	max	sum
Set Symmetric Difference in Neighborhood Gene List Size	355.0	14.478873	12.057371	0.000000	69.000000	5140.000000
Set Symmetric Difference in Interactome Gene List Size	355.0	393.867606	417.333839	19.000000	1147.000000	139823.000000
Set Symmetric Difference in Drug Target Gene List Size	355.0	3.295775	3.948090	0.000000	14.000000	1170.000000
Difference in Neighborhood Gene List Size	355.0	-4.490141	7.159348	-27.000000	47.000000	-1594.000000
Difference in Unique Neighborhood Gene List Size	355.0	-4.490141	7.159348	-27.000000	47.000000	-1594.000000
Difference in Interactome Gene List Size	355.0	156.994366	332.030791	-448.000000	759.000000	55733.000000
Difference in Unique Interactome Gene List	355.0	156.994366	332.030791	-448.000000	759.000000	55733.000000
Difference in NonNeighborhood Gene List Size	355.0	161.484507	330.371588	-438.000000	785.000000	57327.000000
Difference in Neighbrohood Gene List Percentage	355.0	-0.015735	0.025755	-0.156875	0.111862	-5.585807
Difference in Unique Neighborhood Gene List Percentage	355.0	-0.018879	0.029960	-0.187500	0.097598	-6.702136
Difference in Neighborhood Gene List Percentage with Intersecting Interactome Gene List Size	355.0	-0.031667	0.054810	-0.333333	0.100000	-11.241614
Difference in Unique Neighborhood Gene List Percentage with Intersecting Interactome Gene List Size	355.0	-0.031667	0.054810	-0.333333	0.100000	-11.241614
Difference in Intersecting Neighborhood Gene List Percentage	355.0	0.003145	0.017389	-0.029221	0.063076	1.116329
Difference in Unique Interactome Gene List Percentage	355.0	0.014041	0.257770	-0.638644	0.345613	4.984605
Difference in intersecting Interactome Gene List Percentage	355.0	-0.014041	0.257770	-0.345613	0.638644	-4.984605
Difference in Drug Target Gene List Size	355.0	0.416901	4.114292	-14.000000	14.000000	148.000000
Difference in Unique Drug Target Gene List Size	355.0	0.416901	4.114292	-14.000000	14.000000	148.000000
Difference in intersecting Drug Target Gene List Percentage	355.0	0.010928	0.046905	-0.192857	0.283842	3.879475
Difference in Ratio of Drug Targets to Neighborhood Genes	355.0	0.387121	0.676623	-1.023810	6.666667	137.428045

Two Versions – Numerical Attributes Across Grouped by Adverse Events Statistics Table (Shortened due to visual constraints):

	Set Symmetric Difference in Neighborhood Gene List Size						Set Symmetric Difference in Interactome Gene List Size				...	Difference in intersecting Drug Target Gene List Percentage				Difference in Ratio of Drug Targets to Neighborhood Genes					
	count	mean	std	min	max	sum	count	mean	std	min	...	std	min	max	sum	count	mean	std	min	max	sum
Phenotype																					
Cerebral Infarction	5	5.200000	3.492850	3	11	26	5	117.0	0.0	117	...	0.001685	0.000716	0.004656	0.008955	5	0.582976	0.349928	0.287500	1.178571	2.914881
Deep Vein Thrombosis	2	1.000000	1.414214	0	2	2	2	87.0	0.0	87	...	0.000174	0.000737	0.000983	0.001720	2	0.333333	0.235702	0.166667	0.500000	0.666667
Delirium	11	2.909091	1.300350	1	5	32	11	19.0	0.0	19	...	0.058314	-0.192857	-0.023377	-1.145455	11	0.360390	0.900306	-1.023810	2.000000	3.964286
Edema	18	15.055556	5.895883	6	31	271	18	265.0	0.0	265	...	0.080945	0.015769	0.283842	2.254964	18	1.366778	0.812243	0.151515	3.035714	24.602007
Gastric ulcer	26	8.153846	3.585441	4	18	212	26	50.0	0.0	50	...	0.016324	0.002018	0.074652	0.548789	26	0.272938	0.479583	-0.542857	1.584416	7.096385
Hemorrhage	12	2.333333	1.154701	1	5	28	12	40.0	0.0	40	...	0.035898	0.014264	0.156907	0.656156	12	0.981944	0.697233	0.250000	2.500000	11.783333
Hyperlipidemia	19	11.894737	5.445638	2	22	226	19	168.0	0.0	168	...	0.008814	0.002748	0.026565	0.263818	19	0.572629	0.386547	0.000000	1.541667	10.879942
Myocardial Infarction	80	29.525000	14.287265	6	69	2362	80	1147.0	0.0	1147	...	0.003571	-0.015500	-0.000419	-0.322572	80	0.022592	0.142775	-0.750000	0.575163	1.807352
Pancreatitis	24	5.708333	3.406888	0	14	137	24	153.0	0.0	153	...	0.000959	-0.004203	-0.000114	-0.034081	24	0.510017	0.451404	-0.107407	1.833333	12.240396
Peripheral Neuropathy	1	12.000000	NaN	12	12	12	1	534.0	NaN	534	...	NaN	0.009620	0.009620	0.009620	1	0.563025	NaN	0.563025	0.563025	0.563025
Pneumonia	70	12.885714	5.647990	1	26	902	70	213.0	0.0	213	...	0.001014	-0.004259	-0.000115	-0.090016	70	0.121827	0.194791	-0.304511	1.000000	8.527875
Proteinuria	4	5.500000	3.696846	1	10	22	4	96.0	0.0	96	...	0.000783	0.000190	0.002095	0.004761	4	0.493538	0.434581	0.083333	1.083333	1.974153
Pulmonary Edema	3	2.666667	2.516611	0	5	8	3	32.0	0.0	32	...	0.000000	0.000000	0.000000	0.000000	3	0.692063	0.804612	-0.166667	1.428571	2.076190
Sepsis	33	14.393939	8.108404	1	30	475	33	354.0	0.0	354	...	0.001214	0.000157	0.005811	0.045231	33	0.175253	0.252252	-0.234524	1.000000	5.783364
Tardive Dyskinesia	24	7.250000	2.326595	2	12	174	24	26.0	0.0	26	...	0.043451	0.021875	0.161875	1.903125	24	1.370294	0.643492	0.366667	3.255639	32.887050
Thrombocytopenia	23	10.913043	8.039183	0	25	251	23	237.0	0.0	237	...	0.007843	-0.026296	-0.001011	-0.225541	23	0.420049	1.387911	-0.294118	6.666667	9.661138

16 rows × 114 columns