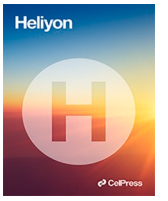
Heliyon 10 (2024) e34858

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Importance of health history analysis in Parkinson's disease

Vinayak Majhi [a], Sudip Paul [b], Goutam Saha [c], Ajaya Jang Kunwar [d], Manob Jyoti Saikia [e,*]

[a] Amity Innovation and Design Centre, Amity University Uttar Pradesh, Noida, 301213, Uttar Pradesh, India
[b] Department of Biomedical Engineering, North-Eastern Hill University, Shillong, 793022, Meghalaya, India
[c] Department of Information Technology, North-Eastern Hill University, Shillong, 793022, Meghalaya, India
[d] Kathmandu Center for Genomics and Research Laboratory, Lalitpur, Nepal
[e] Department of Electrical Engineering, University of North Florida, Jacksonville, FL, 32224, USA

ARTICLE INFO

Keywords:
Parkinson's disease
Health history
Surgical history
Clinical history

ABSTRACT

The objective of this research article is to investigate the impact of various health history factors on the risk of developing Parkinson's disease (PD). From the medical history we can identify PD Symptoms and this also help to detect the progression of PD symptoms. By conducting statistical analyses, the study seeks to identify independent risk and protective factors for Parkinson's disease (PD), considering variations in impact across genders and BMI categories.

Introduction: In the diagnosis of PD the analysis of previous health history is very rare in practice while the main diagnosis have been done through the different motor and non-motor symptoms taking in consideration besides the cardinal symptoms of PD for identification and determination the stages of PD. Here we have analyzed the impact of 56 different diseases, symptoms, and surgeries which a subject may have experienced in their life before PD, considered as a health history.

Methods: The behavioral impact for each types of health history have been analyzed statistically with 31,265 subjects including PD, and Control. In this analysis we have calculated the variation of impact for both the Male, and Female, as well as subjects BMI.

Results: 98.12 % PD patients, where 97.63 % Male PD, and 98.71 % Female PD were found with at least one health history record. Coronary heart disease odds ratio (OR) 2.15 (1.85–2.51), Colon Cancer OR 2.11 (1.45–3.05), Cranial brain surgery OR 6.21 (5.11–7.56) have the higher risks to PD. Having the history of Asthma OR 0.66 (0.6–0.72), Anemia OR 0.56 (0.51–0.63), Cirrhosis in Liver OR 0.7 (0.57–0.86), Cosmetic surgery OR 0.7 (0.64–0.77), and Gastritis OR 0.78 (0.71–0.87) have been found to be protective to PD. The risk of developing PD varies between male, and female including subjects BMI for each individual health history types. The diseases which reduce the oxygen saturation in blood like, anemia, asthma, and thalassemia act as protective to PD.

Conclusions: In this study we have analyzed fifty six diseases which include surgeries as a health history of PD patients. Study suggests that a thorough health history could greatly aid in understanding the onset and progression of Parkinson's disease (PD).

* Corresponding author.
E-mail address: manob.saikia@unf.edu (M.J. Saikia).

https://doi.org/10.1016/j.heliyon.2024.e34858
Received 22 July 2023; Received in revised form 4 July 2024; Accepted 17 July 2024
Available online 18 July 2024

Table 1
List of all considered participants health history related symptoms, diseases, and surgeries that need anesthesia with its frequency of occurrences and percentage of occurrences for each determined cohorts.

	All PD	All Control	PD Male	Control Male	PD Female	Control Female	All PD %	All Control %	PD Male %	Control Male %	PD Female %	Control Female %
Congestive heart failure	260	54	154	15	106	39	1.15	0.63	1.24	0.73	1.04	0.59
Valvular heart disease	204	44	150	18	54	26	0.90	0.51	1.21	0.88	0.53	0.40
Arrhythmia	748	262	421	70	327	192	3.30	3.04	3.39	3.42	3.19	2.92
Coronary heart disease	1091	198	907	107	184	91	4.82	2.30	7.31	5.23	1.80	1.39
Atrial fibrillation	1004	197	685	63	319	134	4.43	2.29	5.52	3.08	3.11	2.04
High blood pressure	5720	1597	3435	489	2285	1108	25.25	18.54	27.68	23.91	22.31	16.86
Asthma	1451	810	649	128	802	682	6.41	9.40	5.23	6.26	7.83	10.38
Emphysema	111	32	62	8	49	24	0.49	0.37	0.50	0.39	0.48	0.37
Chronic obstructive pulmonary disease (COPD)	461	138	250	41	211	97	2.04	1.60	2.01	2.00	2.06	1.48
Pneumonia	1204	464	593	69	611	395	5.32	5.39	4.78	3.37	5.97	6.01
Tuberculosis	64	24	35	8	29	16	0.28	0.28	0.28	0.39	0.28	0.24
Diabetes	1743	602	1085	174	658	428	7.70	6.99	8.74	8.51	6.43	6.51
Acid reflux (GERD)	6090	2221	3163	479	2927	1742	26.89	25.78	25.49	23.42	28.58	26.51
Gastritis	1242	595	485	90	757	505	5.48	6.91	3.91	4.40	7.39	7.69
Hiatal hernia	2029	657	1137	179	892	478	8.96	7.63	9.16	8.75	8.71	7.28
Ulcer	1032	432	520	77	512	355	4.56	5.01	4.19	3.77	5.00	5.40
Renal failure	148	51	87	15	61	36	0.65	0.59	0.70	0.73	0.60	0.55
Cysts	123	44	73	15	50	29	0.54	0.51	0.59	0.73	0.49	0.44
Kidney stones	1344	424	880	128	464	296	5.93	4.92	7.09	6.26	4.53	4.51
Cirrhosis	280	151	133	27	147	124	1.24	1.75	1.07	1.32	1.44	1.89
Chronic viral hepatitis (Hepatitis C or hep C)	133	49	73	16	60	33	0.59	0.57	0.59	0.78	0.59	0.50
Hepatitis A	219	72	134	10	85	62	0.97	0.84	1.08	0.49	0.83	0.94
Hepatitis B	123	25	83	6	40	19	0.54	0.29	0.67	0.29	0.39	0.29
Anemia	913	597	191	31	722	566	4.03	6.93	1.54	1.52	7.05	8.61
Thalassemia	46	19	19	2	27	17	0.20	0.22	0.15	0.10	0.26	0.26
Sickle cell disease	2	0	1	0	1	0	0.01	0.00	0.01	0.00	0.01	0.00
Cancer Bladder	136	36	112	16	24	20	0.60	0.42	0.90	0.78	0.23	0.30
Cancer Breast	723	330	9	3	714	327	3.19	3.83	0.07	0.15	6.97	4.98
Cancer Colon	182	33	104	13	78	20	0.80	0.38	0.84	0.64	0.76	0.30
Cancer Kidney (Renal cancer)	99	21	68	5	31	16	0.44	0.24	0.55	0.24	0.30	0.24
Cancer Leukemia	66	19	45	12	21	7	0.29	0.22	0.36	0.59	0.21	0.11
Cancer Liver (Hepatic cancer)	6	2	4	0	2	2	0.03	0.02	0.03	0.00	0.02	0.03
Cancer Lung	50	18	28	6	22	12	0.22	0.21	0.23	0.29	0.21	0.18
Cancer Lymphoma	132	39	90	12	42	27	0.58	0.45	0.73	0.59	0.41	0.41
Cancer Melanoma	2346	575	1413	155	933	420	10.36	6.67	11.39	7.58	9.11	6.39
Cancer Prostate	985	99	985	95	0	4	4.35	1.15	7.94	4.65	0.00	0.06
Cancer Thyroid	162	65	48	8	114	57	0.72	0.75	0.39	0.39	1.11	0.87
Cancer Skin (non-melanoma)	1807	459	1095	127	712	332	7.98	5.33	8.82	6.21	6.95	5.05
Cancer Uterine	120	42	0	0	120	42	0.53	0.49	0.00	0.00	1.17	0.64
Depression	10209	4049	4954	741	5255	3308	45.07	47.00	39.92	36.23	51.31	50.35
Osteoarthritis/degenerative arthritis	7854	2789	3632	497	4222	2292	34.68	32.37	29.27	24.30	41.23	34.89
Rheumatoid arthritis	761	303	365	57	396	246	3.36	3.52	2.94	2.79	3.87	3.74
Back Pain	11969	3502	6557	841	5412	2661	52.84	40.65	52.84	41.12	52.85	40.50
Anxiety	10210	3787	5025	728	5185	3059	45.08	43.96	40.49	35.60	50.63	46.56
Stroke	1134	260	627	78	507	182	5.01	3.02	5.05	3.81	4.95	2.77

(*continued on next page*)

Table 1 (*continued*)

	All PD	All Control	PD Male	Control Male	PD Female	Control Female	All PD %	All Control %	PD Male %	Control Male %	PD Female %	Control Female %
Traumatic brain injury (TBI)	1804	540	1070	141	734	399	7.96	6.27	8.62	6.89	7.17	6.07
Cardiac surgery	1523	247	1169	124	354	123	6.72	2.87	9.42	6.06	3.46	1.87
Orthopedic surgery	10219	3111	5660	783	4559	2328	45.12	36.11	45.61	38.29	44.52	35.43
Gastrointestinal surgery	5579	1998	2738	413	2841	1585	24.63	23.19	22.06	20.20	27.74	24.12
Cranial or brain surgery	1656	108	1016	40	640	68	7.31	1.25	8.19	1.96	6.25	1.04
Tumor removal	2965	1225	1058	139	1907	1086	13.09	14.22	8.53	6.80	18.62	16.53
Pulmonary (lung) surgery	168	45	89	14	79	31	0.74	0.52	0.72	0.68	0.77	0.47
ENT surgery	4306	1741	2159	336	2147	1405	19.01	20.21	17.40	16.43	20.96	21.39
Eye surgery	5495	1379	2860	297	2635	1082	24.26	16.01	23.05	14.52	25.73	16.47
Reproductive surgery	4553	2057	1509	191	3044	1866	20.10	23.88	12.16	9.34	29.72	28.40
Cosmetic surgery	1496	786	295	42	1201	744	6.60	9.12	2.38	2.05	11.73	11.32

The color distribution of the cells from lighter to darker gradient is representing the lower to higher occurrences, or percentages.

Table 2
List of Chi-square value and Odds ratios of all significant health history type.

	All study participants				Male				Female			
	$\chi^2(1)$	p value	Odds Ratio with 95 % CI	p value	$\chi^2(1)$	p value	Odds Ratio with 95 % CI	p value	$\chi^2(1)$	p value	Odds Ratio with 95 % CI	p value
Congestive heart failure	18.74	1.50E-05	1.84 (1.37–2.47)	4.80E-05	4.42	3.56E-02	1.7 (1–2.9)	5.05E-02	9.59	1.96E-03	1.75 (1.21–2.53)	2.87E-03
Valvular heart disease	13.18	2.83E-04	1.77 (1.28–2.45)	6.12E-04	.	.	.	.	.	.	.	.
Coronary heart disease	112.15	3.31E-26	2.15 (1.85–2.51)	1.34E-22	12.49	4.10E-04	1.43 (1.16–1.76)	6.97E-04	4.30	3.80E-02	1.3 (1.01–1.68)	4.06E-02
Atrial fibrillation	86.11	1.71E-20	1.98 (1.7–2.31)	4.63E-18	24.19	8.71E-07	1.84 (1.41–2.39)	5.00E-06	18.30	1.90E-05	1.54 (1.26–1.89)	3.00E-05
High blood pressure	162.51	3.20E-37	1.48 (1.4–1.58)	9.31E-36	12.90	3.29E-04	1.22 (1.09–1.36)	3.89E-04	75.04	4.62E-18	1.42 (1.31–1.53)	1.10E-17
Asthma	79.50	4.83E-19	0.66 (0.6–0.72)	1.04E-19	.	.	.	.	31.81	1.70E-08	0.73 (0.66–0.82)	1.45E-08
Chronic obstructive pulmonary disease (COPD)	6.47	1.10E-02	1.28 (1.05–1.55)	1.27E-02	.	.	.	.	7.80	5.23E-03	1.4 (1.1–1.79)	6.11E-03
Pneumonia	.	.	.	.	8.57	3.42E-03	1.44 (1.12–1.85)	5.10E-03	.	.	.	.
Diabetes	4.56	3.26E-02	1.11 (1.01–1.22)	3.39E-02	.	.	.	.	.	.	.	.
Acid reflux (GERD)	3.93	4.73E-02	1.06 (1–1.12)	4.78E-02	4.03	4.46E-02	1.12 (1–1.25)	4.62E-02	8.55	3.45E-03	1.11 (1.03–1.19)	3.52E-03
Gastritis	22.16	3.00E-06	0.78 (0.71–0.87)	2.00E-06	.	.	.	.	.	.	.	.
Hiatal hernia	14.42	1.47E-04	1.19 (1.09–1.31)	1.76E-04	.	.	.	.	11.15	8.41E-04	1.22 (1.08–1.37)	9.23E-04
Kidney stones	12.31	4.50E-04	1.22 (1.09–1.36)	5.47E-04	.	.	.	.	.	.	.	.
Cirrhosis in Liver	11.67	6.35E-04	0.7 (0.57–0.86)	4.97E-04	.	.	.	.	5.07	2.44E-02	0.76 (0.59–0.96)	2.36E-02
Hepatitis A	.	.	.	.	7.42	6.44E-03	2.22 (1.17–4.23)	1.52E-02	.	.	.	.
Hepatitis B	9.35	2.23E-03	1.88 (1.22–2.89)	4.19E-03	4.87	2.74E-02	2.29 (1–5.25)	5.06E-02	.	.	.	.
Anemia	106.50	5.72E-25	0.56 (0.51–0.63)	4.73E-26	.	.	.	.	13.69	2.16E-04	0.8 (0.8–0.9)	2.02E-04
Cancer Bladder	4.02	4.49E-02	1.44 (1–2.08)	5.25E-02	.	.	.	.	.	.	.	.
Cancer Breast	7.62	5.77E-03	0.83 (0.72–0.95)	5.24E-03	.	.	.	.	28.16	1.12E-07	1.43 (1.25–1.64)	1.86E-07
Cancer Colon	18.17	2.00E-05	2.11 (1.45–3.05)	8.50E-05	.	.	.	.	15.81	7.00E-05	2.51 (1.54–4.11)	2.43E-04
Cancer Kidney (Renal cancer)	6.69	9.70E-03	1.8 (1.12–2.88)	1.49E-02	.	.	.	.	.	.	.	.
Cancer Melanoma	106.38	6.08E-25	1.62 (1.47–1.78)	3.45E-23	28.63	8.76E-08	1.57 (1.32–1.86)	3.55E-07	41.06	1.48E-10	1.47 (1.3–1.65)	3.19E-10
Cancer Prostate	.	.	.	.	30.87	2.75E-08	1.77 (1.43–2.2)	2.21E-07	.	.	.	.

(*continued on next page*)

Table 2 (*continued*)

	All study participants				Male				Female			
	$\chi^2(1)$	p value	Odds Ratio with 95 % CI	p value	$\chi^2(1)$	p value	Odds Ratio with 95 % CI	p value	$\chi^2(1)$	p value	Odds Ratio with 95 % CI	p value
Cancer Skin (non-melanoma)	69.12	9.24E-17	1.54 (1.39–1.71)	1.05E-15	16.71	4.40E-05	1.46 (1.21–1.77)	9.00E-05	25.41	4.63E-07	1.4 (1.23–1.61)	7.11E-07
Cancer Uterine	.	.	.	.	.	.	.	.	12.57	3.93E-04	1.84 (1.3–2.62)	6.83E-04
Depression	9.33	2.26E-03	0.93 (0.88–0.97)	2.25E-03	10.09	1.49E-03	1.17 (1.06–1.29)	1.58E-03	.	.	.	.
Osteoarthritis/degenerative arthritis	14.80	1.19E-04	1.11 (1.05–1.17)	1.25E-04	21.80	3.00E-06	1.29 (1.16–1.44)	4.00E-06	68.16	1.51E-16	1.31 (1.23–1.4)	1.94E-16
Back Pain	372.98	4.20E-83	1.64 (1.56–1.72)	4.70E-82	96.76	7.83E-23	1.6 (1.46–1.76)	1.63E-22	245.40	2.61E-55	1.65 (1.55–1.75)	9.36E-55
Anxiety	.	.	.	.	17.76	2.50E-05	1.23 (1.12–1.36)	2.80E-05	26.54	2.58E-07	1.18 (1.11–1.25)	2.62E-07
Stroke	62.53	2.62E-15	1.69 (1.48–1.94)	5.02E-14	6.18	1.29E-02	1.34 (1.06–1.71)	1.63E-02	50.95	9.50E-13	1.83 (1.54–2.17)	6.71E-12
Traumatic brain injury (TBI)	26.80	2.26E-07	1.29 (1.17–1.43)	3.82E-07	7.17	7.43E-03	1.27 (1.06–1.53)	9.12E-03	7.72	5.47E-03	1.19 (1.05–1.35)	5.81E-03
Cardiac surgery	197.73	6.54E-45	2.44 (2.13–2.8)	1.81E-37	26.67	2.42E-07	1.61 (1.33–1.95)	1.00E-06	38.51	5.44E-10	1.88 (1.52–2.31)	2.76E-09
Orthopedic surgery	209.16	2.10E-47	1.45 (1.38–1.53)	1.07E-46	38.50	5.47E-10	1.35 (1.23–1.49)	7.34E-10	137.48	9.47E-32	1.46 (1.37–1.56)	1.98E-31
Gastrointestinal surgery	7.08	7.78E-03	1.08 (1.02–1.15)	7.98E-03	.	.	.	.	27.18	1.85E-07	1.21 (1.12–1.3)	2.08E-07
Cranial or brain surgery	557.47	2.98E-123	6.21 (5.11–7.56)	2.44E-74	133.33	7.65E-31	4.47 (3.25–6.15)	4.07E-20	325.62	8.67E-73	6.37 (4.95–8.2)	4.62E-47
Tumor removal	6.78	9.20E-03	0.91 (0.85–0.98)	8.87E-03	7.26	7.07E-03	1.28 (1.06–1.54)	8.72E-03	12.05	5.18E-04	1.16 (1.06–1.25)	5.46E-04
Pulmonary (lung) surgery	4.69	3.04E-02	1.42 (1.02–1.98)	3.61E-02	.	.	.	.	5.77	1.63E-02	1.64 (1.08–2.49)	2.00E-02
ENT surgery	5.70	1.70E-02	0.93 (0.87–0.99)	1.66E-02	.	.	.	.	.	.	.	.
Eye surgery	260.07	1.66E-58	1.68 (1.57–1.79)	4.52E-55	80.96	2.30E-19	1.76 (1.55–2.01)	1.21E-17	205.17	1.56E-46	1.76 (1.62–1.9)	1.24E-44
Reproductive surgery	52.40	4.53E-13	0.8 (0.76–0.85)	2.96E-13	14.21	1.63E-04	1.34 (1.15–1.57)	2.55E-04	.	.	.	.
Cosmetic surgery	56.12	6.83E-14	0.7 (0.64–0.77)	2.58E-14	.	.	.	.	.	.	.	.

(.) represents the insignificant results of logistics regression. The color gradient from white to blue indicating the OR greater than one, and white to red is less than one.

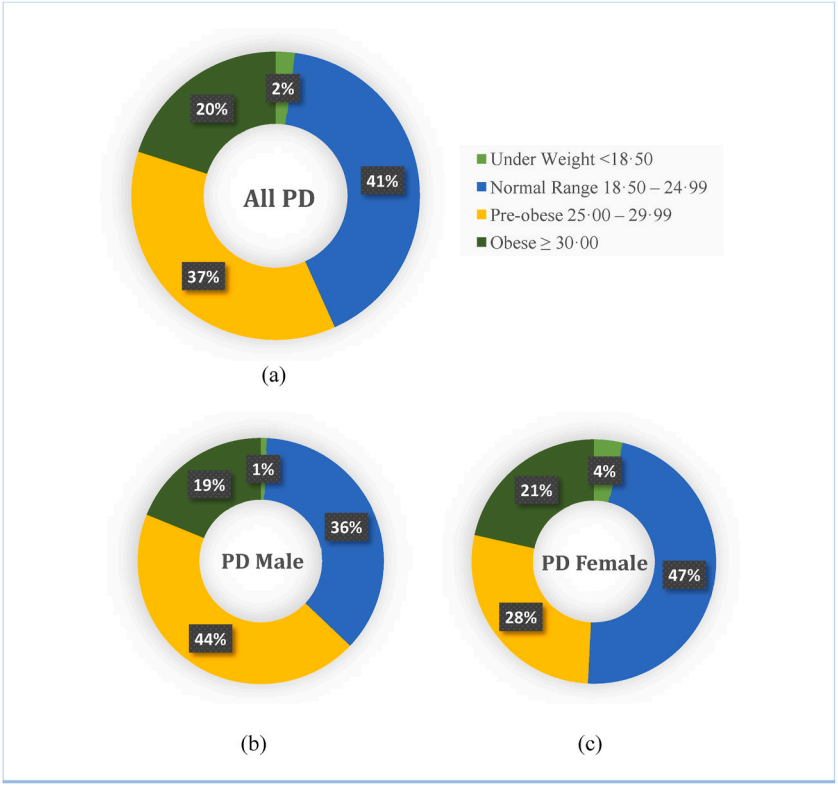


Fig. 1. The distribution of PD participants based on various BMI category represented for (a) among all PD participants; (b) among all male PD participants; and (c) among female PD participants.

1. Introduction

PD is a complex heterogeneous neuro degenerative disease, where the proper cause of the disease is still difficult to understand. In this scenario health history of a patient is always found to be an essential criteria for diagnosing a disease [1–4]. Sometimes, it plays a significant role to understand the pathways of disease progression and also helps to predict possible risk factors for a disease manifesting gradually [4]. PD is a complex neurodegenerative disease, where the proper cause, and cure are still challenging to understand [5]. PD identification is being made only when its' cardinal symptoms, Bradykinesia, Rigidity, Tremor, and Postural Instability, are predominant in a patient [6].

The health history serves a multifaceted purpose in PD diagnosis. First, it allows for a nuanced understanding of the patient's presenting symptoms. The temporal evolution of tremors, rigidity, bradykinesia (slowness of movement), and postural instability – all cardinal features of PD can be meticulously explored. This detailed information helps distinguish these symptoms from potential mimics like essential tremor, vascular Parkinsonism, or drug-induced movement disorders. This information can not only guide the diagnostic workup but also inform future discussions about potential disease progression and preventative measures.

The risk of developing PD mostly increases with growing age [7]. Most PD patients are found above the age of 60 years [7,8]. Through the journey up to this age, a subject may suffered with different diseases, may developed several symptoms, and might be gone through additional surgeries [9–12]. All of these facts are considered as the health history of a patient. Every disease, symptom, or surgery may have impact on PD [10,11], which will be our key findings in this research. During the progression of a particular disease, different symptoms may be developed in a patient. Throughout the process of its cure, or after cure, few symptoms may continue for a long time [13,14]. It may go on throughout the patient's whole life in some cases. Besides this, a few other diseases those are not curable, like PD, also remain with their own symptoms [15]. All of these facts may camouflage PD before its identification, which leads a health professional to misdiagnose it. An unidentified PD becomes more challenging to manage at its later stage and may a patient lead to permanent disability [16–18]. Several studies have looked into the various symptoms and diseases manifested with the PD [6, 10,11]. Many model based studies had also efficiently predicted the PD based on the different motor and non-motor symptoms [19,20]. Research into Parkinson's disease often focuses on singular risk or protective factors, yet few studies comprehensively analyze these factors concurrently. Identifying environmental factors, exploring their interactions, and assessing their predictive value for different subtypes of Parkinson's disease could lead to a more nuanced understanding of its etiology and progression, potentially informing personalized treatment approaches and preventive strategies [21]. However, there's been a lack of extensive research into how different health histories might impact PD risk among a large group of people. Our study posits that a thorough examination of health

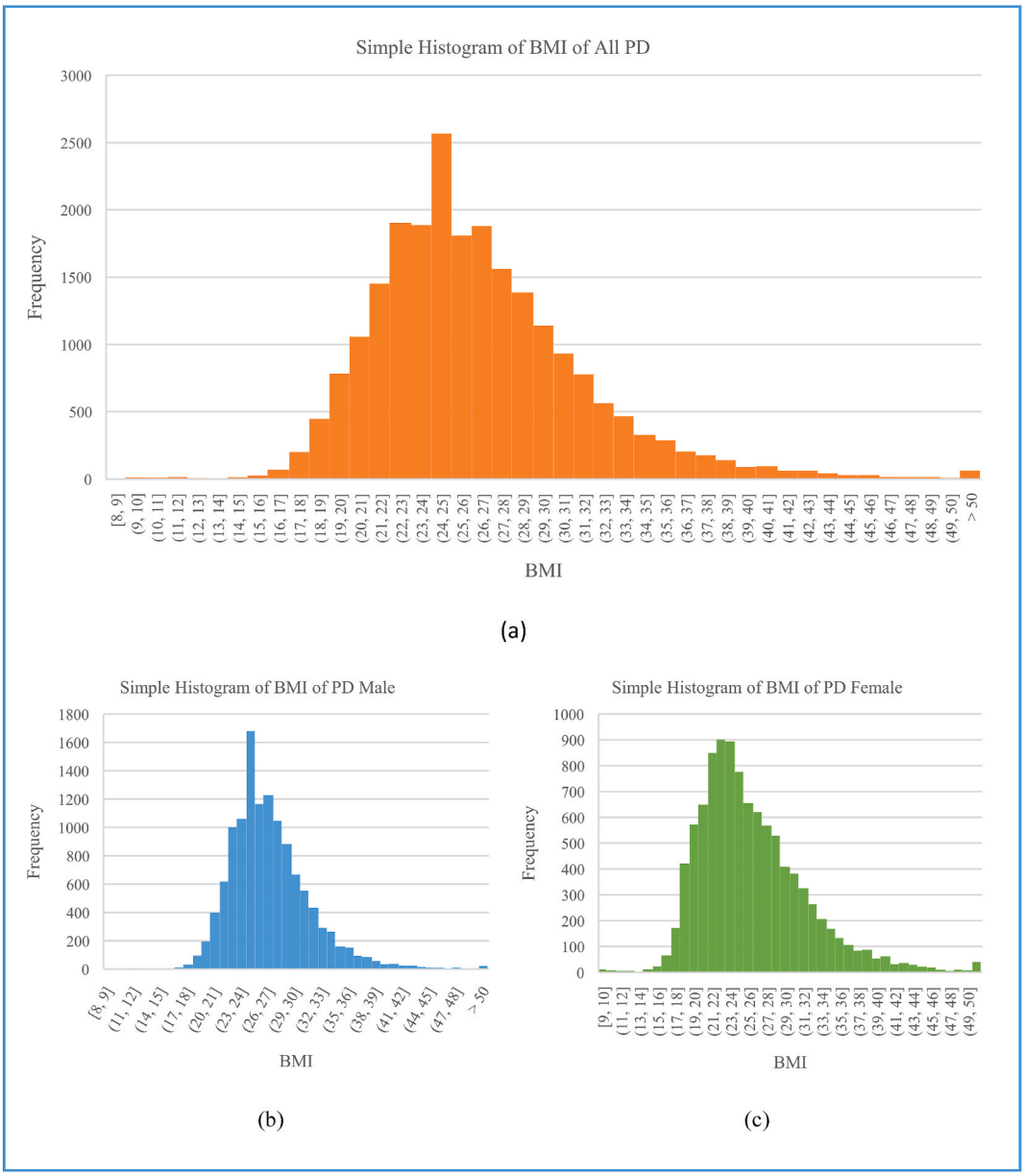


Fig. 2. Histograms of BMI value of (a) all PD participants; (b) male PD participants; and (c) female PD participants.

records may significantly contribute to elucidating the mechanisms underlying the onset and progression of Parkinson's disease. This study aims to fill that gap by analyzing the prevalence and risks associated with various health histories and their correlation with PD. Understanding these connections could potentially help to predict the onset of PD before it develops or identify individuals at risk [22, 23].

Besides health history analysis, BMI, which defines the physical constitution of a patient, can also be an essential feature for analyzing the effect of different health history types on PD. BMI vary upon the effect of different symptoms, disease, lifestyle, mental condition, and the environmental exposure to a patient has had [24,25]. In this hypothetical scenario, we are exploring the combined influence of past medical conditions, symptoms, and surgeries, along with current physical factors like BMI, on the likelihood of developing Parkinson's disease (PD). To test these hypotheses, we have identified a total of fifty-six symptoms, diseases, and surgeries as a health history marker (as shown in Table 1) and statistically analyze each symptom's impact individually on the PD for both Male and Female category. Besides that, we have also taken into consideration of BMI as a reference to find the changes of each symptom's, or disease's impact on PD.

Table 3
List of Chi-square value and odds ratios of all significant health history type grouped by BMI categories.

	General				Male				Female			
	$\chi2(1)$	p value	Odds Ratio	p value	$\chi2(1)$	p value	Odds Ratio	p value	$\chi2(1)$	p value	Odds Ratio	p value
Under Weight												
Atrial fibrillation	9.09	2.56E-03	8.75 (1.18–64.87)	3.38E-02	.	.	.	.	.	.	.	.
High blood pressure	5.30	2.13E-02	2.19 (1.06–4.53)	3.36E-02	.	.	.	.	.	.	.	.
Asthma	.	.	.	.	.	.	.	.	4.09	4.30E-02	0.44 (0.21–0.95)	3.71E-02
Cancer Melanoma	.	.	.	.	.	.	.	.	8.31	3.95E-03	2.9 (1.29–6.55)	1.01E-02
Cancer Skin (non-melanoma)	5.69	1.71E-02	2.77 (1.08–7.12)	3.45E-02	.	.	.	.	5.38	2.04E-02	2.74 (1.06–7.11)	3.82E-02
Stroke	8.62	3.32E-03	8.42 (1.13–62.5)	3.72E-02	.	.	.	.	.	.	.	.
Orthopedic surgery	10.12	1.47E-03	1.96 (1.28–3.01)	2.14E-03	.	.	.	.	8.75	3.10E-03	1.94 (1.23–3.05)	4.14E-03
Tumor removal	5.37	2.05E-02	1.99 (1.07–3.68)	2.94E-02	.	.	.	.	4.49	3.41E-02	1.91 (1.02–3.58)	4.46E-02
Normal Weight												
Congestive heart failure	18.47	1.70E-05	4.44 (1.92–10.23)	4.71E-04	.	.	.	.	9.99	1.57E-03	3.74 (1.45–9.63)	6.25E-03
Coronary heart disease	43.55	4.14E-11	2.31 (1.76–3.03)	1.64E-09	.	.	.	.	.	.	.	.
Atrial fibrillation	31.40	2.10E-08	2.01 (1.55–2.62)	1.81E-07	.	.	.	.	10.83	9.99E-04	1.73 (1.23–2.43)	1.53E-03
High blood pressure	95.34	1.60E-22	1.79 (1.58–2.02)	9.14E-21	6.31	1.20E-02	1.32 (1.06–1.66)	1.43E-02	58.26	2.30E-14	1.77 (1.52–2.06)	1.39E-13
Asthma	30.50	3.33E-08	0.64 (0.54–0.74)	1.75E-08	7.43	6.41E-03	0.62 (0.45–0.86)	4.38E-03	15.99	6.40E-05	0.68 (0.56–0.82)	5.70E-05
Diabetes	14.84	1.17E-04	1.56 (1.23–1.97)	2.15E-04	6.46	1.10E-02	1.73 (1.1–2.72)	1.79E-02	.	.	.	.
Acid reflux (GERD)	10.49	1.20E-03	1.17 (1.06–1.29)	1.31E-03	4.22	3.99E-02	1.24 (1.01–1.52)	4.33E-02	4.16	4.13E-02	1.13 (1–1.27)	4.21E-02
Gastritis	5.66	1.74E-02	0.82 (0.69–0.96)	1.61E-02	.	.	.	.	.	.	.	.
Hiatal hernia	22.19	2.00E-06	1.48 (1.25–1.75)	5.00E-06	.	.	.	.	5.18	2.29E-02	1.28 (1.03–1.58)	2.46E-02
Kidney stones	12.05	5.17E-04	1.4 (1.15–1.7)	7.33E-04	7.82	5.18E-03	1.7 (1.14–2.52)	8.90E-03	.	.	.	.
Hepatitis B	6.13	1.33E-02	2.36 (1.11–4.98)	2.48E-02	.	.	.	.	.	.	.	.
Anemia	42.23	8.12E-11	0.56 (0.47–0.67)	3.01E-11	.	.	.	.	8.37	3.81E-03	0.76 (0.64–0.92)	3.60E-03
Thalassemia	4.18	4.08E-02	0.42 (0.18–0.95)	3.67E-02	.	.	.	.	.	.	.	.
Cancer Bladder	9.76	1.78E-03	3.21 (1.37–7.49)	7.03E-03	.	.	.	.	.	.	.	.

(*continued on next page*)

Table 3 (*continued*)

	General				Male				Female			
	χ2(1)	p value	Odds Ratio	p value	χ2(1)	p value	Odds Ratio	p value	χ2(1)	p value	Odds Ratio	p value
Cancer Breast	.	.	.	.	.	.	.	.	20.38	6.00E-06	1.59 (1.29–1.95)	1.10E-05
Cancer Colon	10.47	1.22E-03	2.56 (1.36–4.83)	3.76E-03	.	.	.	.	7.33	6.79E-03	2.64 (1.22–5.7)	1.34E-02
Cancer Kidney (Renal cancer)	7.55	6.01E-03	3.97 (1.21–13)	2.26E-02	.	.	.	.	.	.	.	.
Cancer Melanoma	59.44	1.26E-14	1.74 (1.5–2.02)	1.85E-13	16.43	5.10E-05	1.8 (1.33–2.44)	1.53E-04	25.18	5.22E-07	1.56 (1.3–1.86)	9.77E-07
Cancer Prostate	.	.	.	.	18.05	2.10E-05	2.14 (1.45–3.14)	1.18E-04	.	.	.	.
Cancer Skin (non-melanoma)	40.28	2.20E-10	1.66 (1.41–1.95)	1.21E-09	12.38	4.33E-04	1.75 (1.25–2.45)	1.01E-03	13.78	2.05E-04	1.44 (1.18–1.76)	2.73E-04
Osteoarthritis/degenerative arthritis	24.00	9.62E-07	1.24 (1.14–1.35)	1.00E-06	12.89	3.30E-04	1.44 (1.17–1.76)	4.81E-04	52.80	3.69E-13	1.44 (1.31–1.6)	5.76E-13
Back Pain	201.74	8.69E-46	1.77 (1.64–1.92)	1.13E-44	27.13	1.90E-07	1.54 (1.31–1.82)	2.51E-07	150.34	1.46E-34	1.81 (1.64–1.99)	7.02E-34
Anxiety	.	.	.	.	7.93	4.85E-03	1.28 (1.08–1.52)	5.31E-03	13.06	3.01E-04	1.19 (1.08–1.3)	3.06E-04
Stroke	17.13	3.50E-05	1.62 (1.27–2.05)	7.50E-05	.	.	.	.	16.43	5.10E-05	1.8 (1.34–2.42)	1.01E-04
Traumatic brain injury (TBI)	11.08	8.72E-04	1.29 (1.11–1.51)	1.09E-03	4.56	3.27E-02	1.4 (1.02–1.94)	3.99E-02	.	.	.	.
Cardiac surgery	87.33	9.20E-21	2.79 (2.19–3.56)	8.59E-17	16.87	4.00E-05	2.07 (1.41–3.02)	1.84E-04	17.18	3.40E-05	1.95 (1.4–2.71)	8.00E-05
Orthopedic surgery	114.61	9.58E-27	1.56 (1.44–1.69)	4.49E-26	25.67	4.05E-07	1.55 (1.3–1.84)	6.58E-07	72.45	1.71E-17	1.53 (1.38–1.68)	3.26E-17
Gastrointestinal surgery	23.58	1.00E-06	1.28 (1.16–1.41)	2.00E-06	.	.	.	.	24.19	8.71E-07	1.35 (1.19–1.52)	1.00E-06
Cranial or brain surgery	220.14	8.43E-50	6.79 (4.86–9.48)	2.59E-29	34.03	5.44E-09	3.62 (2.14–6.11)	1.00E-06	153.70	2.69E-35	8.05 (5.21–12.46)	7.42E-21
Tumor removal	.	.	.	.	4.24	3.95E-02	1.41 (1–1.98)	4.79E-02	8.66	3.24E-03	1.21 (1.07–1.38)	3.46E-03
Eye surgery	142.72	6.76E-33	1.84 (1.66–2.04)	1.44E-30	33.59	6.79E-09	1.88 (1.5–2.37)	4.71E-08	112.92	2.25E-26	1.89 (1.68–2.14)	5.44E-25
Reproductive surgery	.	.	.	.	8.89	2.87E-03	1.53 (1.14–2.06)	4.49E-03	7.81	5.20E-03	1.16 (1.05–1.29)	5.37E-03
Cosmetic surgery	23.67	1.00E-06	0.72 (0.64–0.82)	7.86E-07	.	.	.	.	.	.	.	.
Pre-Obese												
Congestive heart failure	7.28	6.99E-03	1.94 (1.16–3.25)	1.22E-02	6.55	1.05E-02	3.04 (1.11–8.33)	3.08E-02	.	.	.	.
Valvular heart disease	12.51	4.05E-04	2.72 (1.45–5.1)	1.81E-03	.	.	.	.	.	.	.	.
Coronary heart disease	55.06	1.17E-13	2.4 (1.86–3.09)	1.73E-11	16.36	5.20E-05	1.91 (1.36–2.68)	1.92E-04	.	.	.	.
Atrial fibrillation	42.69	6.42E-11	2.36 (1.78–3.13)	3.09E-09	14.76	1.22E-04	2.14 (1.39–3.29)	5.27E-04	8.77	3.07E-03	1.8 (1.2–2.7)	4.34E-03

(*continued on next page*)

Table 3 (*continued*)

	General				Male				Female			
	$\chi2(1)$	p value	Odds Ratio	p value	$\chi2(1)$	p value	Odds Ratio	p value	$\chi2(1)$	p value	Odds Ratio	p value
High blood pressure	65.01	7.47E-16	1.52 (1.37–1.69)	3.36E-15	14.08	1.75E-04	1.37 (1.16–1.62)	2.38E-04	27.31	1.73E-07	1.46 (1.27–1.69)	2.38E-07
Asthma	14.08	1.75E-04	0.73 (0.62–0.86)	1.33E-04	.	.	.	.	7.15	7.51E-03	0.76 (0.62–0.93)	7.29E-03
Chronic obstructive pulmonary disease (COPD)	4.87	2.73E-02	1.46 (1.03–2.08)	3.34E-02	.	.	.	.	.	.	.	.
Pneumonia	.	.	.	.	6.87	8.76E-03	1.68 (1.11–2.56)	1.42E-02	.	.	.	.
Acid reflux (GERD)	.	.	.	.	3.90	4.83E-02	1.18 (1–1.4)	5.08E-02	5.08	2.43E-02	1.16 (1.02–1.31)	2.46E-02
Gastritis	6.09	1.36E-02	0.8 (0.67–0.95)	1.23E-02	.	.	.	.	.	.	.	.
Hiatal hernia	4.78	2.89E-02	1.19 (1.02–1.38)	3.10E-02	.	.	.	.	7.36	6.66E-03	1.32 (1.08–1.63)	7.28E-03
Ulcer	1.38	4.07E-02	0.81 (0.67–0.99)	3.80E-02	.	.	.	.	.	.	.	.
Kidney stones	4.44	3.51E-02	1.22 (1.01–1.47)	3.82E-02	.	.	.	.	.	.	.	.
Cirrhosis	3.92	4.78E-02	0.69 (0.48–0.99)	4.28E-02	.	.	.	.	.	.	.	.
Anemia	28.85	7.83E-08	0.57 (0.47–0.7)	3.39E-08	.	.	.	.	.	.	.	.
Cancer Breast	10.09	1.49E-03	0.67 (0.52–0.85)	1.15E-03	.	.	.	.	6.42	1.13E-02	1.37 (1.07–1.76)	1.25E-02
Cancer Colon	4.49	3.40E-02	1.91 (1–3.63)	4.85E-02	.	.	.	.	.	.	.	.
Cancer Melanoma	31.91	1.62E-08	1.57 (1.33–1.84)	5.45E-08	8.09	4.44E-03	1.42 (1.11–1.83)	6.15E-03	8.79	3.04E-03	1.4 (1.12–1.76)	3.49E-03
Cancer Prostate	.	.	.	.	8.55	3.46E-03	1.54 (1.14–2.08)	5.44E-03	.	.	.	.
Cancer Skin (non-melanoma)	21.88	3.00E-06	1.52 (1.27–1.83)	7.00E-06	.	.	.	.	9.04	2.65E-03	1.48 (1.14–1.92)	3.19E-03
Cancer Uterine	.	.	.	.	.	.	.	.	5.20	2.26E-02	2.26 (1.07–4.77)	3.23E-02
Depression	.	.	.	.	12.45	4.18E-04	1.31 (1.12–1.52)	4.78E-04	.	.	.	.
Osteoarthritis/degenerative arthritis	.	.	.	.	8.74	3.11E-03	1.27 (1.08–1.5)	3.56E-03	12.03	5.22E-04	1.23 (1.1–1.39)	5.38E-04
Back Pain	113.55	1.64E-26	1.6 (1.46–1.74)	3.21E-26	51.27	8.04E-13	1.69 (1.46–1.95)	1.37E-12	65.88	4.79E-16	1.62 (1.44–1.82)	6.35E-16
Anxiety	.	.	.	.	16.24	5.60E-05	1.36 (1.17–1.57)	6.80E-05	6.58	1.03E-02	1.16 (1.04–1.31)	1.04E-02
Stroke	17.93	2.30E-05	1.6 (1.27–2)	5.10E-05	.	.	.	.	11.18	8.29E-04	1.67 (1.23–2.27)	1.16E-03
Traumatic brain injury (TBI)	21.49	4.00E-06	1.51 (1.26–1.81)	8.00E-06	5.46	1.94E-02	1.39 (1.04–1.85)	2.43E-02	8.16	4.28E-03	1.43 (1.11–1.83)	4.95E-03
Cardiac surger	69.06	9.55E-17	2.33 (1.87–2.89)	3.25E-14	9.75	1.79E-03	1.53 (1.16–2.02)	2.92E-03	11.70	6.24E-04	1.88 (1.29–2.74)	1.02E-03

(*continued on next page*)

Table 3 (*continued*)

	General				Male				Female			
	$\chi2(1)$	p value	Odds Ratio	p value	$\chi2(1)$	p value	Odds Ratio	p value	$\chi2(1)$	p value	Odds Ratio	p value
Orthopedic surgery	55.68	8.51E-14	1.39 (1.28–1.52)	1.22E-13	12.84	3.39E-04	1.3 (1.13–1.5)	3.68E-04	40.43	2.04E-10	1.47 (1.3–1.65)	2.45E-10
Gastrointestinal surgery	.	.	.	.	5.38	2.04E-02	1.23 (1.03–1.48)	2.25E-02	15.89	6.70E-05	1.3 (1.14–1.48)	7.40E-05
Cranial or brain surgery	188.39	7.15E-43	5.8 (4.2–8)	1.18E-26	65.36	6.24E-16	5.22 (3.1–8.77)	4.58E-10	93.60	3.86E-22	5.54 (3.64–8.44)	1.45E-15
Tumor removal	6.54	1.05E-02	0.85 (0.75–0.96)	9.86E-03	4.59	3.21E-02	1.35 (1.02–1.8)	3.83E-02	.	.	.	.
Pulmonary (lung) surgery	.	.	.	.	.	.	.	.	11.42	7.27E-04	5.43 (1.63–18.07)	5.77E-03
Eye surgery	77.75	1.17E-18	1.63 (1.46–1.82)	1.23E-17	38.05	6.89E-10	1.81 (1.48–2.2)	5.03E-09	53.51	2.57E-13	1.7 (1.47–1.96)	6.81E-13
Reproductive surgery	31.87	1.65E-08	0.74 (0.67–0.82)	1.12E-08	4.89	2.70E-02	1.29 (1.02–1.64)	3.12E-02	.	.	.	.
Cosmetic surgery	36.07	1.90E-09	0.6 (0.51–0.71)	7.39E-10	.	.	.	.	.	.	.	.
Obese												
Coronary heart disease	14.72	1.25E-04	1.68 (1.28–2.22)	2.35E-04	.	.	.	.	.	.	.	.
Atrial fibrillation	12.78	3.50E-04	1.6 (1.23–2.09)	5.70E-04	6.72	9.56E-03	1.84 (1.12–3.03)	1.61E-02	.	.	.	.
High blood pressure	43.49	4.26E-11	1.44 (1.29–1.6)	6.20E-11	.	.	.	.	25.35	4.77E-07	1.42 (1.24–1.62)	5.48E-07
Asthma	23.47	1.00E-06	0.68 (0.58–0.79)	9.44E-07	.	.	.	.	.	.	.	.
Emphysema	8.61	3.35E-03	3.09 (1.31–7.31)	1.00E-02	.	.	.	.	9.60	1.95E-03	5 (1.48–16.85)	9.46E-03
Chronic obstructive pulmonary disease (COPD)	4.46	3.47E-02	1.4 (1.02–1.92)	3.94E-02	.	.	.	.	7.30	6.90E-03	1.74 (1.15–2.63)	8.73E-03
Pneumonia	.	.	.	.	7.13	7.60E-03	1.93 (1.14–3.27)	1.39E-02	.	.	.	.
Acid reflux (GERD)	.	.	.	.	.	.	.	.	9.65	1.89E-03	1.23 (1.08–1.4)	1.93E-03
Gastritis	9.26	2.34E-03	0.74 (0.61–0.9)	2.08E-03	.	.	.	.	.	.	.	.

(*continued on next page*)

Table 3 (*continued*)

	General				Male				Female			
	$\chi2(1)$	p value	Odds Ratio	p value	$\chi2(1)$	p value	Odds Ratio	p value	$\chi2(1)$	p value	Odds Ratio	p value
Hiatal hernia	.	.	.	.	.	.	.	.	5.39	2.02E-02	1.26 (1.04–1.53)	2.10E-02
Cirrhosis	4.40	3.60E-02	0.74 (0.57–0.98)	3.37E-02	.	.	.	.	.	.	.	.
Anemia	25.93	3.54E-07	0.6 (0.49–0.73)	2.37E-07	.	.	.	.	.	.	.	.
Cancer Breast	4.41	3.57E-02	0.76 (0.58–0.98)	3.35E-02	.	.	.	.	.	.	.	.
Cancer Melanoma	9.82	1.73E-03	1.37 (1.12–1.67)	2.14E-03	4.18	4.10E-02	1.47 (1–2.16)	4.97E-02	.	.	.	.
Cancer Prostate	.	.	.	.	4.27	3.88E-02	1.66 (1–2.78)	5.17E-02	.	.	.	.
Cancer Skin (non-melanoma)	4.85	2.76E-02	1.28 (1.02–1.59)	3.01E-02	.	.	.	.	.	.	.	.
Osteoarthritis/degenerative arthritis	.	.	.	.	4.10	4.30E-02	1.25 (1.01–1.54)	4.48E-02	24.05	9.39E-07	1.38 (1.21–1.57)	9.81E-07
Back Pain	86.31	1.54E-20	1.64 (1.48–1.82)	1.63E-20	23.63	1.00E-06	1.65 (1.35–2.02)	1.00E-06	59.70	1.10E-14	1.66 (1.46–1.9)	1.30E-14
Anxiety	.	.	.	.	.	.	.	.	10.58	1.14E-03	1.24 (1.09–1.41)	1.15E-03
Stroke	26.70	2.37E-07	1.89 (1.47–2.44)	9.75E-07	7.44	6.38E-03	2.04 (1.16–3.56)	1.27E-02	22.94	2.00E-06	2.01 (1.49–2.71)	4.00E-06
Cardiac surgery	39.91	2.65E-10	2.19 (1.69–2.84)	3.95E-09	.	.	.	.	9.05	2.62E-03	1.77 (1.21–2.61)	3.54E-03
Orthopedic surgery	51.88	5.91E-13	1.46 (1.32–1.63)	6.88E-13	6.28	1.22E-02	1.29 (1.06–1.58)	1.23E-02	37.51	9.09E-10	1.49 (1.31–1.7)	1.02E-09
Cranial or brain surgery	147.00	7.86E-34	6.47 (4.39–9.54)	3.94E-21	37.20	1.06E-09	5.14 (2.62–10.09)	2.00E-06	79.88	3.98E-19	6.05 (3.74–9.81)	2.56E-13
ENT surgery	4.04	4.45E-02	0.88 (0.78–1)	4.37E-02	.	.	.	.	.	.	.	.
Eye surgery	42.58	6.80E-11	1.53 (1.35–1.75)	1.64E-10	9.81	1.74E-03	1.51 (1.16–1.97)	2.44E-03	41.54	1.15E-10	1.67 (1.42–1.95)	2.14E-10
Reproductive surgery	25.28	4.97E-07	0.74 (0.66–0.83)	4.20E-07	.	.	.	.	.	.	.	.
Cosmetic surgery	3.87	4.92E-02	0.8 (0.64–1)	4.70E-02	.	.	.	.	.	.	.	.

(.) represents the insignificant results of logistics regression. The color gradient from white to blue indicating the OR greater than one, and white to red is less than one.

2. Methodology

2.1. Study design and participants

In this study we have used Fox Insight (FI) dataset updated up to 08th March 2022, sponsored by the Michael J. Fox Foundation. This study protocol and informed consent are reviewed by the New England IRB (IRB#: 120160179, Legacy IRB#: 14–236, Sponsor Protocol Number: 1, Study Title: Fox Insight).

FI dataset is consist of more than 6000 variables from which we have selected our study cohorts based upon the questionnaires tagged as 'Your Health History'. In this section there are 104 variables which describe different types of diseases, symptoms, and surgeries that a person ever had experienced in their life. It also describes about, whether the patient have received their treatments or not, and whether these diseases or symptoms have led into their daily life activity in limitation or not. From this 104 questionnaires we have considered only 56 questionnaires that describes, whether a person suffered from a particular disease or symptom or had experienced any particular surgery that used anesthesia. The list of our selected health history related attributes have been shown in Table 1. All the responses of these questionnaires are dichotomous, 'Yes', or 'No' which represented as 1, or 0 respectively. Besides these questionnaires we have also selected the data about participants' Sex 'Male' or 'Female' which is also a binary question, where 1 is represented as 'Male' and 2 is represented as 'Female'. Finally, we have taken the data about the participants' height in centimeter scale and weight in kilogram to obtain the BMI using the formula BMI = Weight (kilogram)/(height in centimeter/100)^2 [26]. Here, we have obtained total n = 53,471 participants for both suffering from PD as well as Control. To distinguish PD and Control group we have considered our key variable 'CurrPDDiag' where 1 is represented as 'PD' and 0 is represented as 'Control' group and finally we got n = 37,944 PD and n = 15,527 Control participants.

2.2. Data pre-processing

The obtained data initially loaded into the MySQL database for pre-processing through quires. The usage of database engine, like MySQL is beneficial for easy joining of multiple dataset, sorting, missing value management, implementing arithmetic, and logical operations [27,28]. Here, we got four different datasets carrying different information about each participant, which are 'general' data, containing information about whether a participant is suffering from PD, or not; next is 'health history' data containing participants' health history related information, third one is 'about' data which contains the information about participant height, weight, and sex, and finally 'users' data where the information about diagnosis age of a PD patient is defined. We have imported all these four datasets into our MySQL database as different tables and joined [29] them using the common field 'fox_insight_id', which is the identification number for each participants.

All records with missing values were discarded from our main study cohorts. In addition, we also discarded the recordings where the participants chose not to answer the questions. After removing all missing value finally we got total n = 31,265 participants, where n = 22,650 are PD and n = 8615 are Control participants. Next, we split our primary dataset into four different subsets based upon the participants' BMI. These are under weight (BMI <18.5), normal range (BMI 18.50–24.99), pre-obese (BMI 25.00–29.99), and obese (BMI ≥30.00) [26]. Then statistical analysis is done on the data set, which contain fifty-six participants' health history records.

2.3. Statistical analysis

The statistical analysis has been performed using IBM SPSS version 28.0.1.1. Initially frequencies for each health history types have been calculated. Next we have performed the Pearson correlation coefficient among all considered attributes. These helped us to understand the inter relationship, correlation, or dependencies between each considered attributes [30]. The binary logistics regression have been performed to get the Odds ratio (OR), and chi-square (χ^2) goodness-of-fit of the logistics regression for each considered health history types [31–33]. The logistics regression has been performed in two different ways. Initially it has been performed for each health history types for both uncategorized, and categorized by sex to obtain the pattern and variance for each category on PD [34]. In other way we have performed the analysis with our sub dataset divided by the BMI category. Here we have shown, how each history types found to be risk factor or protective factor to PD depending upon the variation of different BMI groups. These calculations have been done separately for both the sex Male and Female, as well as uncategorized by sex. The significant level for all the statistical analysis is set to $p < 0.05$.

3. Result

From the FI dataset of n = 53,471 participants after excluding n = 22,206 (41.53 %) finally we obtained n = 31,265 (58.47 %) participants for analysis, where n = 22,650 (72.45 %) PD and n = 8615 (27.55 %) Control. In the data set n = 12,409 (39.69 %) Male PD, n = 10,241 (32.76 %) Female PD, n = 2045 (6.54 %) Male Control, and n = 6570 (21.01 %) Female Control subjects. From this primary case-control cohort, we calculated the frequency of occurrence for each of our fifty-six individual health histories. Additionally, we calculated the percentage of occurrence relative to the entire study cohort, as outlined in Table 1. These findings support our decision to conduct binary logistic regression with 1 degree of freedom to obtain the odds ratio.

The Pearson's correlation coefficient, among all the health history related symptoms, diseases, surgeries, sex, and BMI, are shown in the table of appendix1. Here we have found that several symptoms, diseases, and surgeries are significantly correlated with each other. These findings ensure us to conduct the binary logistics regression with 1 degree of freedom to obtain odds ratio [35,36].

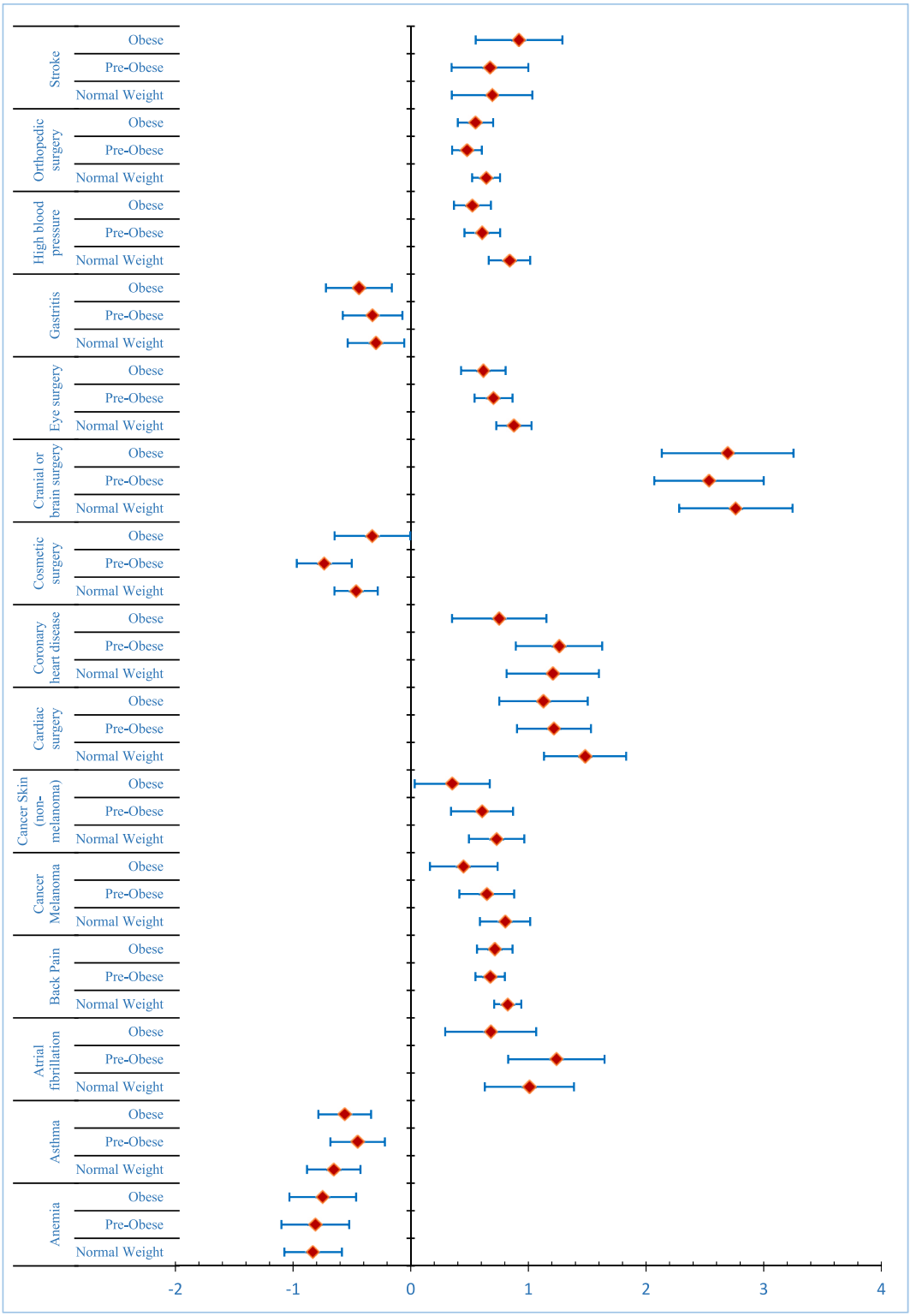


Fig. 3. Comparative forest plot for all participants log of odds ratios with base 2 with 95 % confidence interval for significant health history type comparing with normal weighted, pre-obese, and obese subjects.

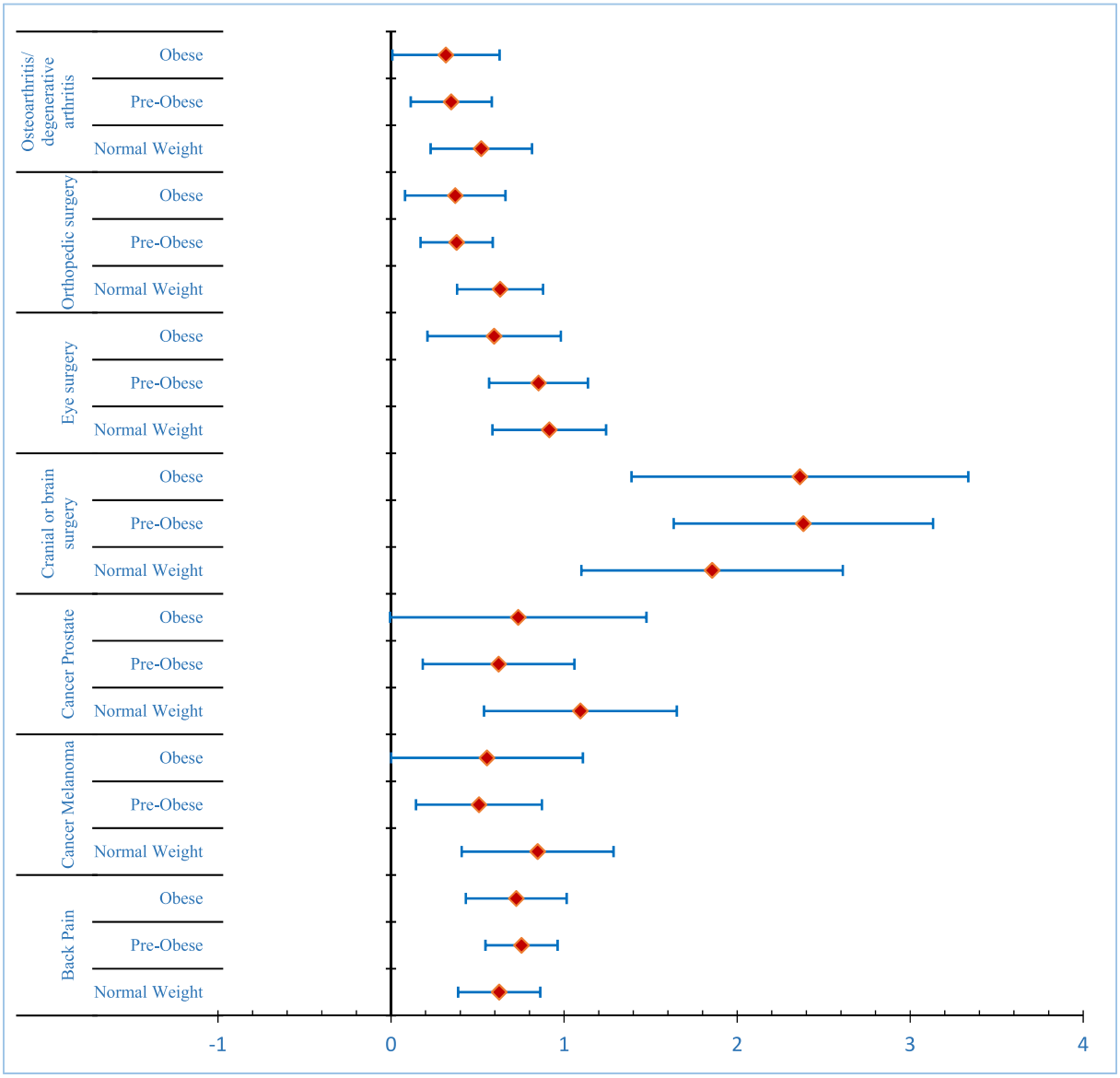


Fig. 4. Comparative forest plot for male participants' log of odds ratios with base 2 with 95 % confidence interval for significant health history type comparing with normal weighted, pre-obese, and obese subjects.

Initially the binary logistics regression has been performed with our primary cohort for both uncategorized by sex, and categorized by Male, and Female sex to find the impact of each health history type. Here we have found that Male subjects are more likely to exhibit PD [37] with OR 3.89 (3.68–4.12; $p < 0.00001$). We have also found most of the symptoms are likely to exhibit PD. The patients having the history of Coronary heart disease [38], Colon Cancer [39], cardiac surgery [40] are more than two times, and cranial brain surgery [41] are more than six times riskier to PD. For Male patients having the clinical history of Hepatitis A, Hepatitis B are more than two times, and cranial brain surgery are more than four times riskier to PD. Whereas Colon Cancer, and cranial brain surgery are the most riskier to PD for Female subjects. Besides these several other symptoms, diseases, and surgery like Asthma [42], Anemia [43], Cirrhosis of Liver, Cosmetic surgery are exhibiting less likely to PD. The following Table 2 contains the list of all significant health history related symptoms, diseases and surgeries with their Chi-square value and odds ratio's for both uncategorized by sex and categorized by Male and Female.

Our results show the prevalence of having PD is higher in percentage to those, who had history of cardiac disease (14.49 % of total PD, 17.65 % of Male PD, and 10.65 % of Female PD) or hypertension (25.25 % of total PD) or have cardiac surgery (6.72 % of total PD).

After dividing our primary dataset by participant's BMI, with four different categories, underweight, normal range, pre-obese, and obese, which individually sub categorized by sex Male, and Female, we have the following distribution as shown in Fig. 1. Through this figure we can easily identify that PD patients are mostly belongs to over weighted category. Whereas Female PD mostly belongs to the

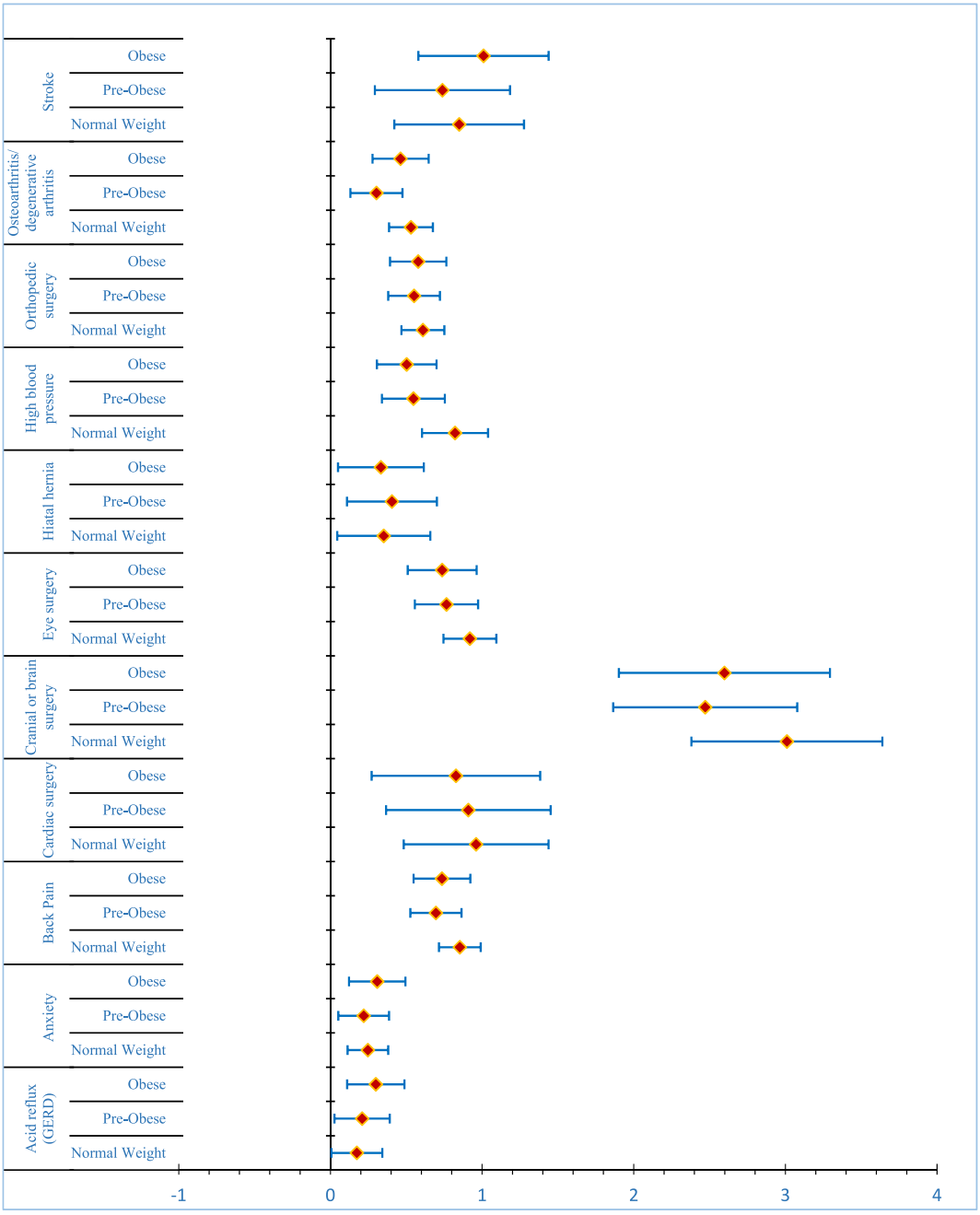


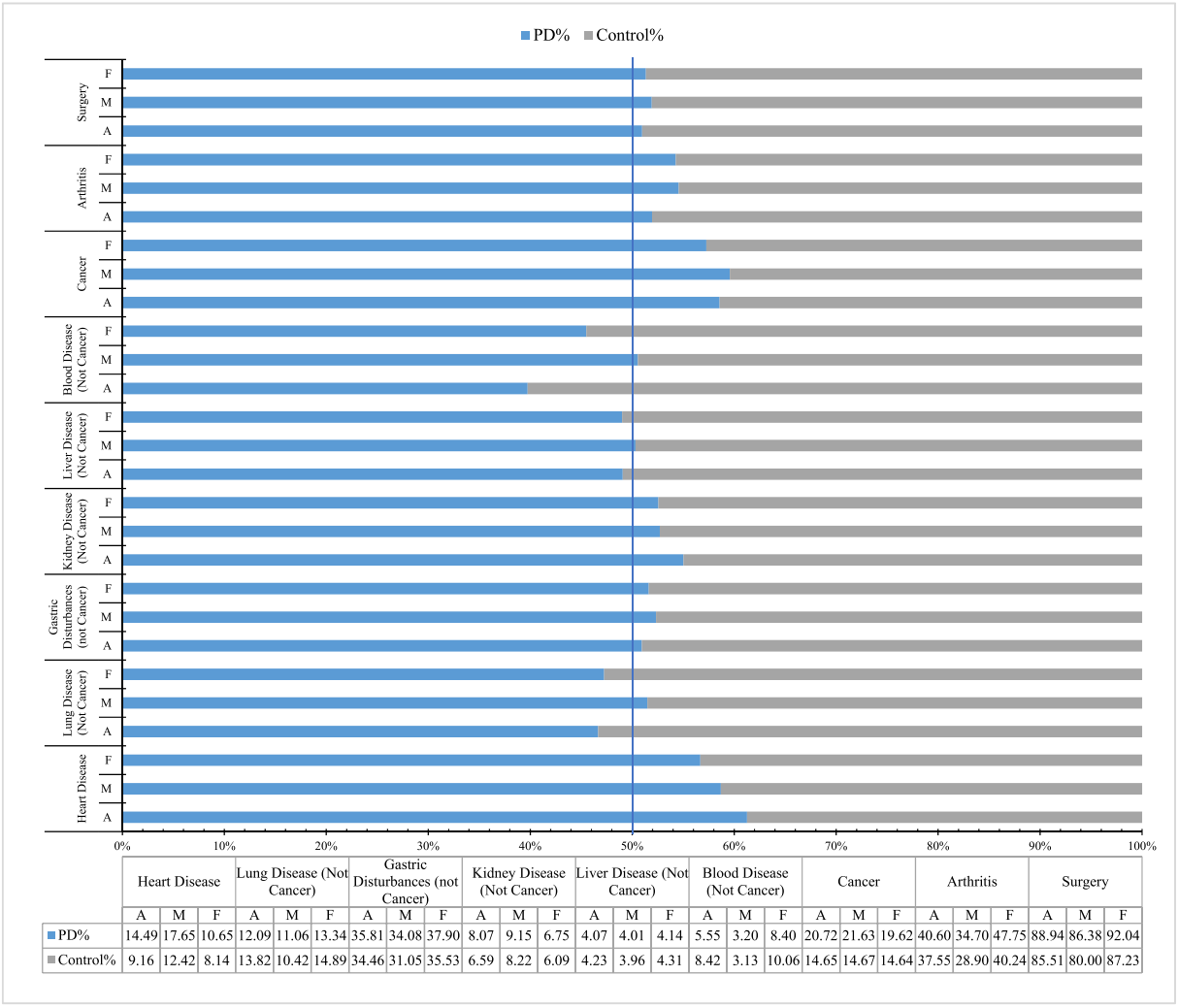
Fig. 5. Comparative forest plot for female participants' log of odds ratios with base 2 with 95 % confidence interval for significant health history type comparing with normal weighted, pre-obese, and obese subjects.

normal weighted category compared to Male PD.

We have also calculated the frequency of subjects, mean BMI with standard deviation (SD) and standard mean error for our all considered health history types as listed in appendix 2. The BMI distribution of the PD subjects uncategorized and categorized by Male, and Female is shown as histogram plot in Fig. 2. The BMI mean (SD) of PD participants are 26.49 (6.18), 26.78 (5.33) for Male PD, and 26.14 (7.06) for Female PD.

Calculated chi-square values and odds ratios by one-degree-of-freedom binary logistic regression between Control and PD for all of our listed health history types for four derived BMI categories are presented in Table 3.

In Table 3 we can easily spot few health history types like Atrial fibrillation, Stroke have the greater risk for PD if patients' is underweighted. Similarly Congestive heart failure, Cancer Bladder, Cancer Kidney (Renal cancer) are most likely to have with PD category for Normal weighted subjects. Cranial or brain surgery are found to be risk factor to PD for Normal weighted, Pre-obese, and



	Heart Disease			Lung Disease (Not Cancer)			Gastric Disturbances (not Cancer)			Kidney Disease (Not Cancer)			Liver Disease (Not Cancer)			Blood Disease (Not Cancer)			Cancer			Arthritis			Surgery		
	A	M	F	A	M	F	A	M	F	A	M	F	A	M	F	A	M	F	A	M	F	A	M	F	A	M	F
PD%	14.49	17.65	10.65	12.09	11.06	13.34	35.81	34.08	37.90	8.07	9.15	6.75	4.07	4.01	4.14	5.55	3.20	8.40	20.72	21.63	19.62	40.60	34.70	47.75	88.94	86.38	92.04
Control%	9.16	12.42	8.14	13.82	10.42	14.89	34.46	31.05	35.53	6.59	8.22	6.09	4.23	3.96	4.31	8.42	3.13	10.06	14.65	14.67	14.64	37.55	28.90	40.24	85.51	80.00	87.23

Fig. 6. 100 % stacked bar chart show the relative percentage distribution of major disease types based on their percentage of occurrence between PD and Control for All participants (A), Male participants (M), Female participants (F).

Obese subjects.

The impact of the presence of different diseases, symptoms, and surgical histories on PD may vary depending upon the subject's BMI and sex. In Figs. 3–5 we have shown the comparative forest plot of different significant health history types comparing with the variation depending upon the changes in BMI on PD. In this forest plot we have taken the log of odds ratio with 95 % confidence interval on base 2 to emphasize the impact scale. Through this plot we can easily compare that the presence of orthopedic surgery, high blood pressure, eye surgery, cranial or brain surgery, cardiac surgery, Skin cancer (non-melanoma), and Back pain as health history are more likely to exhibit PD for the subjects with normal weight compared to the pre-obese, and obese subjects. Where, history of stroke is more likely to exhibit PD for obese subjects, compare to pre-obese, and normal weighted subjects. Similarly, coronary heart disease, atrial fibrillation are mostly found with pre-obese PD subjects.

The symptoms, diseases or surgeries, which act as a protective toward PD also vary upon the BMI, like the presence of anemia, asthma in subjects health history may be more protective to PD for normal weighted subjects compared to pre-obese, and obese subjects. Likewise, gastritis, and cosmetic surgery are very less likely to occur in PD with obese and pre-obese subjects respectively. These patterns may also vary depending upon the subject's sex; like, cranial or brain surgery is found to be more likely with male pre-obese PD subjects where for female subjects it is more likely to normal weighted PD subjects. The following Figs. 3–5 are the comparative forest plot of logs of odds ratios with base 2 for all participants, male participants, and female participants respectively.

After categorizing according to different types of medical history, here we have a total of nine main categories. These are heart disease (congestive heart failure, Valvular heart disease, arrhythmia, coronary artery disease, atrial fibrillation), lung disease, non-cancer (asthma, emphysema, chronic obstructive pulmonary disease, pneumonia, tuberculosis), gastric disease, non-cancer (acid reflux (GERD), gastritis, hiatal hernia, ulcer), kidney disease, not cancer (kidney failure, cysts, kidney stones), liver disease, not cancer (cirrhosis, chronic viral hepatitis (hepatitis C or hep C), hepatitis A, hepatitis B), blood disease not cancer (anemia, thalassemia, sickle

cell anemia), cancer (bladder cancer, breast cancer, colon cancer, kidney cancer (kidney cancer), leukemia cancer, liver cancer (liver cancer), lung cancer, lymphoma cancer, melanoma cancer, prostate cancer, thyroid cancer, skin cancer (non-melanoma), uterine cancer), arthritis (osteoarthritis/degenerative arthritis, rheumatoid arthritis) and surgery who required anesthesia (cardiac surgery, orthopedic surgery, gastric D arm diseases). Surgery, cranial or brain surgery, tumor removal, lung (pulmonary) surgery, ENT surgery, eye surgery, reproductive surgery, cosmetic surgery). The subject distribution of these major disease categories with their respective percentage for each category is being shown in appendix 3 Table 1. The distribution of participants, based on different BMI categories with calculated BMI Mean and standard deviation, are shown in Table 2 of appendix 3. This distribution will help us to understand the disease prevalence by its category. The relative percentage distribution of these nine major distribution categories are shown in Fig. 6.

4. Discussion

The study aims to analyze the association between 56 different diseases, symptoms, and surgeries experienced by individuals before PD onset, and their subsequent diagnosis of PD. The research focuses on evaluating the behavioral impact of these health history factors among a large cohort of subjects, including both PD patients and controls. Through statistical analyses, this study endeavors to uncover the distinct risk and protective factors associated with Parkinson's disease, while also accounting for differences in impact across gender and BMI classifications. Additionally, the research aims to determine the specific diseases and conditions that may increase or decrease the risk of PD development, providing insights into potential preventive strategies and understanding underlying mechanisms.

A detailed health history plays a pivotal role in diagnosing PD. It empowers neurologists to comprehensively assess presenting symptoms, rule out alternative diagnoses, and understand individual risk factors. This information forms the bedrock for an accurate diagnosis and allows for the development of personalized treatment plans, ultimately improving patient outcomes. Our research sheds light on the often-overlooked role of health history in Parkinson's disease diagnosis, offering valuable insights into potential risk and protective factors. By identifying specific diseases and surgeries associated with PD, the study paves the way for more comprehensive screening and preventive measures, ultimately enhancing patient care and management.

This research encountered several limitations that warrant consideration. Firstly, due to the absence of PD stage scaling in the FI dataset, we couldn't categorize data based on different stages of PD, limiting our insights into disease progression. Additionally, the results for underweight participants often yielded insignificant values, likely due to the smaller number of participants in this category. Consequently, we couldn't compare the odds ratios of underweight individuals with those of normal-weight, pre-obese, and obese participants. Another limitation arose from the inability to analyze data based on race, as the dataset lacked sufficient representation from diverse racial backgrounds. Furthermore, due to the extensive nature of the dataset, we chose to focus only on statistically significant results with $p < 0.05$, discarding other findings deemed insignificant. Lastly, as our study adopted a cross-sectional design with a quantitative approach aimed at determining prevalence and risk factors, so it captures data at a single point in time, providing a snapshot of the relationship rather than a temporal sequence. Without longitudinal data, it's challenging to determine whether health history influences BMI or vice versa, as causality cannot be inferred from cross-sectional observations. Other factors also not accounted for in the study may confound the observed association, limiting the ability to establish a causal link. Additionally we refrained from delving into the underlying causes of observed statistical patterns. This aspect remains a potential avenue for exploration in future studies.

5. Conclusion

Our study has successfully demonstrated the association between various health history factors and Parkinson's disease (PD). Our findings support our hypothesis regarding the significant impact of different health history elements on PD, whether acting as risk factors or protective factors. While some diseases showed no significant impact on PD, we identified BMI as a crucial factor in health history analysis. Changes in BMI showed a notable association with different health history types, indicating that the influence of individual health factors varies with BMI changes in PD patients. Moreover, our analysis included considerations of sex differences, revealing variations in the impact of different health history factors between genders, both when uncategorized and when categorized by BMI type. These insights highlight the importance of accounting for health history when diagnosing PD. Overall this study proves itself as an essential by its derived results and features that open a new window to consider valuable insights that enhance our understanding of PD and its potential risk factors.

Data availability

Data used in the preparation of this manuscript were obtained from the Fox Insight database (https://foxinsight-info.michaeljfox.org/insight/explore/insight.jsp) on March 08, 2022. For up-to-date information on the study, visit https://foxinsight-info.michaeljfox.org/insight/explore/insight.jsp.

CRediT authorship contribution statement

Vinayak Majhi: Writing – original draft, Visualization, Validation, Software, Methodology, Formal analysis, Data curation, Conceptualization. **Sudip Paul:** Writing – original draft, Validation, Supervision, Software, Resources, Project administration, Methodology, Formal analysis, Conceptualization. **Goutam Saha:** Writing – original draft, Visualization, Validation, Software,

Resources, Methodology, Data curation. **Ajaya Jang Kunwar:** Writing – original draft, Validation, Methodology, Formal analysis, Data curation. **Manob Jyoti Saikia:** Writing – review & editing, Validation, Supervision, Resources, Project administration, Investigation, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank the Biomedical Sensors & Systems Lab, University of North Florida, Jacksonville, FL 32224, USA for supporting the article processing charges.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e34858.

References

[1] M.C. Peterson, J.H. Holbrook, D. Von Hales, N.L. Smith, L.V. Staker, Contributions of the history, physical examination, and laboratory investigation in making medical diagnoses, West. J. Med. 156 (2) (1992) 163–165, https://doi.org/10.1097/00006254-199210000-00013 (in eng).
[2] T. Tsukamoto, Y. Ohira, K. Noda, T. Takada, M. Ikusaka, The contribution of the medical history for the diagnosis of simulated cases by medical students, Int. J. Med. Educ. 3 (2012) 78–82, https://doi.org/10.5116/ijme.4f8a.e48c.
[3] G.O. Barnett, J.J. Cimino, J.A. Hupp, E.P. Hoffer, DXplain: an evolving diagnostic decision-support system, JAMA 258 (1) (1987) 67–74, https://doi.org/10.1001/jama.258.1.67.
[4] M.L. Graber, C. Byrne, D. Johnston, The impact of electronic health records on diagnosis, Diagnosis 4 (4) (2017) 211–223, https://doi.org/10.1515/dx-2017-0012.
[5] A.F.F. Ferreira, M.P. Singulani, H. Ulrich, Z.-P. Feng, H.-S. Sun, L.R. Britto, Inhibition of TRPM2 by AG490 is neuroprotective in a Parkinson's disease animal model, Mol. Neurobiol. 59 (3) (2022/03/01 2022) 1543–1559, https://doi.org/10.1007/s12035-022-02723-8.
[6] I. Cova, A. Priori, Diagnostic biomarkers for Parkinson's disease at a glance: where are we? J. Neural. Transm. 125 (10) (2018/10/01 2018) 1417–1432, https://doi.org/10.1007/s00702-018-1910-4.
[7] A. Reeve, E. Simcox, D. Turnbull, Ageing and Parkinson's disease: why is advancing age the biggest risk factor? Ageing Res. Rev. 14 (100) (2014) 19–30, https://doi.org/10.1016/j.arr.2014.01.004 (in eng).
[8] R. Cacabelos, Parkinson's disease: from pathogenesis to pharmacogenomics, Int. J. Mol. Sci. 18 (3) (2017) 551, https://doi.org/10.3390/ijms18030551 (in eng).
[9] A.U. Khan, M. Akram, M. Daniyal, R. Zainab, Awareness and current knowledge of Parkinson's disease: a neurodegenerative disorder, Int. J. Neurosci. 129 (1) (2019/01/02 2019) 55–93, https://doi.org/10.1080/00207454.2018.1486837.
[10] K. Seppi, et al., Update on treatments for nonmotor symptoms of Parkinson's disease-an evidence-based medicine review, Mov. Disord. 34 (2) (2019) 180–198, 10.1002%2Fmds.27602.
[11] J.A. Obeso, et al., Past, present, and future of Parkinson's disease: a special essay on the 200th Anniversary of the Shaking Palsy, Mov. Disord. 32 (9) (2017) 1264–1310, https://doi.org/10.1002/mds.27115 (in eng).
[12] G. Crosby, L. Vutskits, General anesthesia and progression of Parkinson disease: a shaky association, Anesth. Analg. 133 (5) (2021) 1138–1139, https://doi.org/10.1213/ane.0000000000005735.
[13] W.C. Summers, Virus infection, in: M. Schaechter (Ed.), Encyclopedia of Microbiology, third ed., Academic Press, Oxford, 2009, pp. 546–552.
[14] F. Salamanna, F. Veronesi, L. Martini, M.P. Landini, M. Fini, Post-COVID-19 syndrome: the persistent symptoms at the post-viral stage of the disease. A systematic review of the current data, Front. Med. 8 (2021) 653516, https://doi.org/10.3389/fmed.2021.653516 (in eng).
[15] E. Buzhor, et al., Cell-based therapy approaches: the hope for incurable diseases, Regen. Med. 9 (5) (2014) 649–672, https://doi.org/10.2217/rme.14.35 (in eng).
[16] M. Rodríguez-Violante, R. Zerón-Martínez, A. Cervantes-Arriaga, T. Corona, Who can diagnose Parkinson's disease first? Role of pre-motor symptoms, Arch. Med. Res. 48 (3) (Apr 2017) 221–227, https://doi.org/10.1016/j.arcmed.2017.08.005 (in eng).
[17] H.B. Tandeter, P. Shvartzman, Parkinson's disease camouflaging early signs of hypothyroidism, Postgrad. Med. 94 (5) (Oct 1993) 187–190, https://doi.org/10.1080/00325481.1993.11945738 (in eng).
[18] K.A. Severson, et al., Discovery of Parkinson's disease states and disease progression modelling: a longitudinal data study using machine learning, The Lancet Digital Health 3 (9) (2021) e555–e564, https://doi.org/10.1016/S2589-7500(21)00101-1.
[19] W. Wang, J. Lee, F. Harrou, Y. Sun, Early detection of Parkinson's disease using deep learning and machine learning, IEEE Access 8 (2020) 147635–147646, https://doi.org/10.1109/ACCESS.2020.3016062.
[20] S. Paul, et al., Bias investigation in artificial intelligence systems for early detection of Parkinson's disease: a narrative review, Diagnostics 12 (1) (2022) 166, https://doi.org/10.3390/diagnostics12010166.
[21] D. Belvisi, et al., Risk factors of Parkinson disease, Neurology 95 (18) (2020/11/03 2020) e2500–e2508, https://doi.org/10.1212/WNL.0000000000010813.
[22] J. Klucken, R. Krüger, P. Schmidt, B.R. Bloem, Management of Parkinson's disease 20 Years from now: towards digital health pathways, J. Parkinsons Dis. 8 (2018) S85–S94, 10.3233%2FJPD-181519.
[23] R. Prashanth, S. Dutta Roy, Early detection of Parkinson's disease through patient questionnaire and predictive modelling, Int. J. Med. Inf. 119 (2018/11/01/2018) 75–87, https://doi.org/10.1016/j.ijmedinf.2018.09.008.
[24] J. Chen, Z. Guan, L. Wang, G. Song, B. Ma, Y. Wang, Meta-analysis: overweight, obesity, and Parkinson's disease, International Journal of Endocrinology 2014 (2014/02/05 2014) 203930, 10.1155%2F2014%2F203930.
[25] N. Palacios, et al., Obesity, diabetes, and risk of Parkinson's disease, Mov. Disord. 26 (12) (2011) 2253–2259, https://doi.org/10.1002/mds.23855.
[26] W. G. I. Team, The SuRF Report 2. Surveillance of Chronic Disease Risk Factors: Country-Level Data and Comparable Estimates, WHO, Geneva, 2006.
[27] C. Ordonez, Data set preprocessing and transformation in a database system, Intell. Data Anal. 15 (2011) 613–631, https://doi.org/10.3233/IDA-2011-0485.
[28] R. Zach, S. Glawischnig, M. Hönisch, R. Appel, A. Mahdavi, MOST: An open-source, vendor and technology independent toolkit for building monitoring, data preprocessing, and visualization. eWork and eBusiness in ArchitectureEngineering and Construction, 2012, pp. 97–103.

[29] M. Norman, Database Design Manual: Using MySQL for Windows, 1 ed., Springer Science & Business Media, 2006.
[30] P. Schober, C. Boer, L.A. Schwarte, Correlation coefficients: appropriate use and interpretation, Anesth. Analg. 126 (5) (2018) 1763–1768, https://doi.org/10.1213/ane.0000000000002864.
[31] B. Zhang, A chi-squared goodness-of-fit test for logistic regression models based on case-control data, Biometrika 86 (3) (1999) 531–539, https://doi.org/10.1093/biomet/86.3.531.
[32] N. Balakrishnan, V. Voinov, M.S. Nikulin, Chi-squared Goodness of Fit Tests with Applications, Academic Press, 2013.
[33] F.E. Harrell, Binary logistic regression, in: J.F.E. Harrell (Ed.), Regression Modeling Strategies: with Applications to Linear Models, Logistic and Ordinal Regression, and Survival Analysis, Springer International Publishing, Cham, 2015, pp. 219–274.
[34] M.R. Cullen, A.R. Lemeshow, L.J. Russo, D.M. Barnes, Y. Ababio, A. Habtezion, Disease-specific health disparities: a targeted review focusing on race and ethnicity, Healthcare (Basel) 10 (4) (Mar 23 2022), https://doi.org/10.3390/healthcare10040603 (in eng).
[35] N. Senaviratna, T. Cooray, Diagnosing multicollinearity of logistic regression model, Asian Journal of Probability and Statistics 5 (2) (2019) 1–9, https://doi.org/10.9734/ajpas/2019/v5i230132.
[36] H. Midi, S.K. Sarkar, S. Rana, Collinearity diagnostics of binary logistic regression model, J. Interdiscipl. Math. 13 (3) (2010/06/01 2010) 253–267, https://doi.org/10.1080/09720502.2010.10700699.
[37] I.N. Miller, A. Cronin-Golomb, Gender differences in Parkinson's disease: clinical characteristics and cognition, Mov. Disord. 25 (16) (2010/12/15 2010) 2695–2703, https://doi.org/10.1002/mds.23388.
[38] S.K.K. Chua, S.E. Saffari, S.J.Y. Lee, E.-K. Tan, Association between Parkinson's disease and coronary artery disease: a systematic review and meta-analysis, J. Parkinsons Dis. 12 (2022) 1737–1748, https://doi.org/10.3233/JPD-223291.
[39] O. P. Trifonova et al., "Parkinson's disease: available clinical and promising omics tests for diagnostics, disease risk assessment, and pharmacotherapy personalization," Diagnostics, vol. 10, no. 5, doi: 10.3390/diagnostics10050339.
[40] T. Schroeter, et al., "Cardiac surgery in patients with Parkinson's disease: a retrospective analysis of a high-risk cohort," (in en), Thorac. Cardiovasc. Surg. 66 (8) (2018/12/12 2018) 629–636, https://doi.org/10.1055/s-0037-1603589.
[41] S. Padmakumar, P. Kulkarni, C.F. Ferris, B.S. Bleier, M.M. Amiji, Traumatic brain injury and the development of parkinsonism: understanding pathophysiology, animal models, and therapeutic targets, Biomed. Pharmacother. 149 (May 2022) 112812, https://doi.org/10.1016/j.biopha.2022.112812 (in eng).
[42] A. Barreiro-González, et al., Purtscher-like retinopathy associated with dermatomyositis and hemophagocytic lymphohistiocytosis, Arch. Soc. Esp. Oftalmol. 93 (4) (Apr 2018) 202–205, https://doi.org/10.1016/j.oftal.2017.03.009. Retinopatía Purtscher-like asociada a dermatomiositis y linfohistiocitosis hemofagocítica.
[43] I.Y. Cho, et al., Anemia and the risk of Parkinson's disease in Korean older adults: a nationwide population-based study, Sci. Rep. 10 (1) (2020/03/06 2020) 4268, https://doi.org/10.1038/s41598-020-61153-5.