

On Heredity Factors of Parkinson's Disease: A Parametric and Bayesian Analysis

Abolfazl Saghafi¹, Chris P. Tsokos², Rebecca D. Wooten³

¹Department of Mathematics, Physics and Statistics, University of the Sciences, Philadelphia, PA, USA

²Department of Mathematics and Statistics, University of South Florida, Tampa, FL, USA

³Department of Mathematics, Florida Southern College, Lakeland, FL, USA

Email: a.saghafi@uscience.edu, ctsokos@usf.edu, rwooten@usf.edu

How to cite this paper: Saghafi, A., Tsokos, C.P. and Wooten, R.D. (2018) On Heredity Factors of Parkinson's Disease: A Parametric and Bayesian Analysis. *Advances in Parkinson's Disease*, 7, *-*. <https://doi.org/10.4236/apd.2018.7.10001>

Received: **** *, **

Accepted: **** *, **

Published: **** *, **

Copyright © 2018 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Hereditary is one of the key risk factors of the Parkinson's disease (PD) and children of individuals with the Parkinson's carry a two-fold risk for the disease. In this article, chance of developing the Parkinson's disease is estimated for an individual in five types of families. That is, families with negative history of the PD (I), families with positive history where neither one of the parents (II), one of the parents (III-IV), or both parents (V) are diagnosed with the disease. After a sophisticated modeling, Maximum Likelihood and Bayesian Approach are used to estimate the chance of developing the Parkinson's in the five mentioned family types. It is extremely important knowing such probabilities as the individual can take precautionary measures to defy the odds. While many physicians have provided medical opinions on chance of developing the PD, our study is one of the first to provide statistical modeling and analysis with real data to support the conclusions.

Keywords

Parkinson's Disease, Heredity, Bayesian Estimation, Maximum Likelihood, Statistical Modelling

1. Introduction

Parkinson's disease (PD) is a chronic and progressive movement disorder, meaning that symptoms continue and worsen over time. Nearly one million Americans are living with Parkinson's disease and approximately 60,000 are diagnosed with PD each year. This number does not reflect thousands of cases that remain undetected. The cause for the PD is unknown, and although there is presently no cure, there are available treatments such as medication and surgery to manage its symptoms [1].

The diagnosis of PD depends upon the presence of one or more of the four most common motor symptoms of the disease. That is, tremor, bradykinesia, rigidity, and postural instability. In addition, there are other secondary and non-motor symptoms that affect many people and are increasingly recognized by doctors as important to diagnosing Parkinson's. These symptoms contribute to severe disability and impaired quality of life in advanced Parkinson's cases. Symptoms include anxiety, depression, cognitive mood swings, dementia, constipation, pain, genitourinary problems, sudden drop in blood pressure upon standing, excessive sweating, sleep disturbances, sense of smell, vision, memory, weight loss, psychosis, hallucinations and loss of energy, among others [2].

There are several research centers and foundations that study Parkinson's disease with the aim of providing education to the society about Parkinson's, providing facilities for people with Parkinson's, better understanding of the Parkinson's disease, reducing its effect in patients, and potentially finding a cure for the Parkinson's. Among them are National Parkinson Foundation, Parkinson's Disease Foundation, American Parkinson Disease Association, Davis Phinney Foundation, and Michael J. Fox Foundation for Parkinson's Research.

Through contact with The Michael J. Fox Foundation for Parkinson's Research, we were granted access to the vast database of Parkinson's Progression Markers Initiative (PPMI) [3] on different factors related to registered people with PD. Our aim is to study the heredity factors leading to Parkinson's by statistically modeling the existing data on healthy individuals and patients with Parkinson's disease. The total sample size in our study was 1258; 751 males and 507 females. However, more information was available through individual's relatives.

Figure 1 shows the outline of the available data to carry out this study. The available information included whether either one of the paternal/maternal grandparents had PD (0 for neither, 1 for either one, 2 for both), whether the biological father/mother had PD (0 for no, 1 for yes), number of paternal/maternal aunts/uncles with PD and in total, number of full/half siblings with PD and in total, and number of children so far diagnosed with PD. Note that the person himself/herself could be healthy or diagnosed with PD. The numbers in parenthesis shows the number of cases in each category. There was not enough information available on gender to perform gender related tests and comparisons.

2. Methods

2.1. Maximum Likelihood

The approach shown in **Figure 2** is followed which emphasizes discovering the hereditary importance of the PD. The data is first divided into two exclusive groups based on the heredity status; negative heredity ($H = 0$) and positive heredity ($H = 1$). Heredity is considered positive if at least one individual out of grandparents, parents, aunts/uncles, or full siblings carried the PD. Then, cases in positive heredity group categorized based on the disease status of parents. For

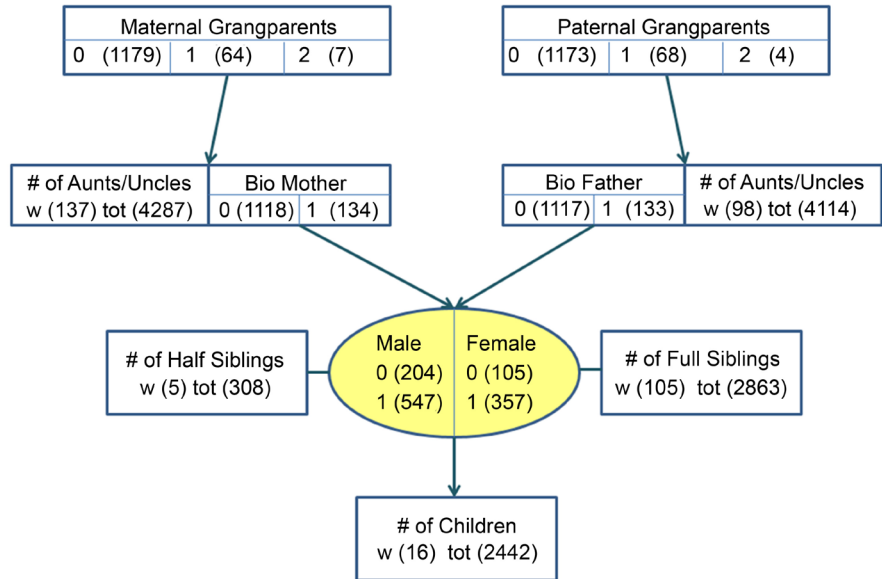


Figure 1. Schematic diagram of available data and the counts.

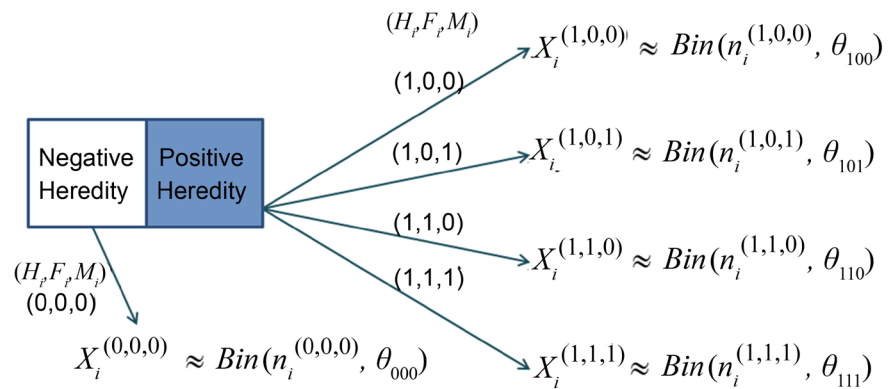


Figure 2. Flowchart of modeling approach.

case i , $(F_p, M_i) = (0, 0)$ when neither one of the parents carried Parkinson's, $(F_p, M_i) = (0, 1)$ when father was healthy, and mother was diagnosed with Parkinson's, etc. In this approach, the number of cases with Parkinson's in each one of the five categories follows a Binomial distribution with two parameters: total number of siblings in the family including the person himself/herself (n_i), and probability of developing Parkinson's (θ). Generally, for case i , one can write

$$X_i^{(j,k,l)} | (H_i = j, F_i = k, M_i = l) \sim \text{Bin}(n_i^{(j,k,l)}, \theta_{jkl}), j, k, l = 0, 1, \quad (1)$$

where $H_i = j$ with $j = 0, 1$ shows the negative/positive heredity group, $F_i = k$, $M_i = l$ with $k, l = 0, 1$ shows the Healthy/PD status of the parents, $n_i^{(j,k,l)}$ shows the total number of siblings in the family, and $0 \leq \theta_{jkl} \leq 1$ represents the probability of developing the PD. The likelihood function can then be written as

$$L(\theta_{jkl} | X^{(j,k,l)}) = \prod_{i=1}^{k_{jkl}} \binom{n_i^{(j,k,l)}}{x_i^{(j,k,l)}} \theta_{jkl}^{x_i^{(j,k,l)}} (1 - \theta_{jkl})^{n_i^{(j,k,l)} - x_i^{(j,k,l)}}, \quad (2)$$

where k_{jkl} is the number of cases in each of the five family types represented by $H_i = j, F_i = k, M_i = l$. Furthermore, it is easy to arrive at the following maximum likelihood estimator

$$\hat{\theta}_{jkl} = \frac{\sum_i X_i^{(j,k,l)}}{\sum_i n_i^{(j,k,l)}} \tag{3}$$

Table 1(a) provides maximum likelihood estimations for parameters θ_{jkl} in each of the five family types as well as the number of valid cases (k_{jkl}) in the dataset. The results show that the probability of developing the PD in families with negative heredity is 0.214. This estimation is based on 824 case subjects. As expected, this probability is higher in families with positive heredity. The prevalence of the PD for an offspring is 0.324 when neither one of the parents were diagnosed with the PD, increases to 0.274 when only the mother was diagnosed with the PD, and raises to 0.294 when only the father was diagnosed with the PD. The chance increases even more to 0.414 when both parents were diagnosed with Parkinson’s disease.

In deriving estimations of **Table 1(a)** only the information link between parents and the individual plus his/her siblings have been used. Using the information link between the person’s grandparents and parents leads to higher number of samples, thus more consistent estimations. The estimations in **Table 1(b)** use the combined likelihood, one from parents-children link and the other from grandparents-parents. The new estimations are significantly different in negative heredity group and where both parents carried the PD. This could trigger changing the prevalence through time. Moreover, since no information was provided on the gender of the grandparents with the PD, a combined probability has been estimated for the case of θ_{101} and θ_{110} . This combined probability shows a state where either one of the parents carried the PD.

The combined information suggests that the chance of developing the PD in families with positive PD history when neither one of the parents had the PD is five times more than that of with no history of the disease. It is about four times

Table 1. Maximum likelihood estimations for θ_{jkl} and the number of cases (a) using the parents-individuals link (b) using combined information with grandparents’ family.

| (a) | | | | |
|-----------------------------|----------------------------|---------------------------------|------------------------------|-------------------------|
| $\theta_{000}(k_{000})$ | $\theta_{100}(k_{100})$ | $\theta_{101}(k_{101})$ | $\theta_{110}(k_{110})$ | $\theta_{111}(k_{111})$ |
| $\frac{584}{2729} = 0.214$ | $\frac{189}{584} = 0.324$ | $\frac{104}{380} = 0.274$ | $\frac{113}{385} = 0.294$ | $\frac{12}{29} = 0.414$ |
| (825) | (165) | (125) | (124) | (9) |
| (b) | | | | |
| $\theta_{000}(k'_{000})$ | $\theta_{100}(k'_{100})$ | $\theta_{10/110}(k'_{10/110})$ | $\theta_{111}(k'_{111})$ | |
| $\frac{732}{11728} = 0.062$ | $\frac{376}{1196} = 0.314$ | $\frac{239}{887} = 0.269$ (253) | $\frac{19}{72} = 0.264$ (20) | |
| (2993) | (280) | | | |

more when one or both parents carry the disease. Surprisingly, the chances for developing the PD when neither one of the parents were diagnosed with the PD are significantly higher than the case where one or both parents are diagnosed with the disease (p-value = 0.00014 for Binomial test). This could suggest a dormant gene effect for the Parkinson's.

2.2. Bayesian Approach

The chance of passing the PD to next generations depends on many factors and could vary from one family to another. This random nature justifies using Bayesian approach for estimations. Moreover, one can use sets of hierarchical information as prior-likelihood and update prior information anytime new observations are added to the dataset.

To conduct a Bayesian approach, data in **Table 1(a)** that utilizes the information link between individuals plus full siblings and their biological parents is used as likelihood. There is available information on whether paternal/maternal aunts/uncles are diagnosed with the PD and whether grandparents had the disease. This information is utilized to derive Bayesian estimations for the model parameters θ_{jkl} following two approaches. In the first method, the frequency of the PD in each of the paternal and maternal grandparents' family is used as discrete prior. In the second method, this data is mixed with the information regarding the individual's family as likelihood and a uniform prior is utilized to derive estimations.

2.2.1. Discrete Prior

To select a prior for θ_{100} , cases with positive family history of PD were selected (decided based on the status of grandparents, aunts, and uncles) whose neither one of the paternal grandparents had PD ($H = 1, F = 0, M = 0$). Then, in each of such families, the chance of developing the Parkinson's disease is estimated by counting the number of cases with the PD divided by the total number of siblings. This estimator can be written as follows:

$$\frac{\text{Father's status} + \# \text{ of paternal aunts/uncles with PD}}{1 + \text{total} \# \text{ of paternal aunts/uncles}}. \quad (4)$$

Following the same procedure in the maternal family yields estimate of the chance of developing the PD using maternal family

$$\frac{\text{Mother's status} + \# \text{ of maternal aunts/uncles with PD}}{1 + \text{total} \# \text{ of maternal aunts/uncles}}. \quad (5)$$

These two separate estimations when computed for each case provide a frequency distribution that can be used as a priori information in estimating θ_{100} . Likewise, one can gather prior information for θ_{110} by frequency of disease in the paternal and maternal families with positive history where the grandfather did, and grandmother did not have the PD. However, the only information available in the grandparents' families is the sum of the PD status of grandmother/grandfather. In that case, the number of the PD diagnosed cases is

counted but the prior for θ_{101} and θ_{110} is set to be the same. Prior information for θ_{111} can be derived using the same technique but in different families with respect to grandparents' status. The same approach is used to derive prior for θ_{000} . **Table 2** provides the frequency distribution of the PD occurrences utilizing the above approach.

Table 2. Frequency of PD in the parents/aunts/uncles for (a) Negative Heredity group; (b) Positive Heredity group when ONE of the grandparents had PD; (c) Positive Heredity group when NEITHER ONE of the grandparents had PD; (d) Positive Heredity group when BOTH grandparents had PD.

| (a) | | |
|---------|-----------|---------|
| Support | Frequency | Percent |
| 0.000 | 2021 | 93.18% |
| 0.091 | 2 | 0.09% |
| 0.111 | 4 | 0.18% |
| 0.125 | 3 | 0.14% |
| 0.143 | 9 | 0.41% |
| 0.167 | 7 | 0.32% |
| 0.200 | 15 | 0.69% |
| 0.250 | 20 | 0.92% |
| 0.333 | 35 | 1.61% |
| 0.500 | 31 | 1.43% |
| 1.000 | 22 | 1.01% |
| Total | 2169 | 100% |

| (b) | | |
|---------|-----------|---------|
| Support | Frequency | Percent |
| 0.000 | 54 | 42.19% |
| 0.111 | 2 | 1.56% |
| 0.125 | 2 | 1.56% |
| 0.167 | 1 | 0.78% |
| 0.200 | 2 | 1.56% |
| 0.250 | 6 | 4.69% |
| 0.333 | 13 | 10.16% |
| 0.400 | 2 | 1.56% |
| 0.429 | 1 | 0.78% |
| 0.500 | 19 | 14.84% |
| 0.571 | 1 | 0.78% |
| 0.667 | 7 | 5.47% |
| 0.750 | 4 | 3.13% |
| 0.875 | 1 | 0.78% |
| 1.000 | 13 | 10.16% |
| Total | 128 | 100% |

(c)

| Support | Frequency | Percent |
|---------|-----------|---------|
| 0.091 | 2 | 1.74% |
| 0.100 | 2 | 1.74% |
| 0.111 | 2 | 1.74% |
| 0.125 | 5 | 4.35% |
| 0.143 | 10 | 8.70% |
| 0.167 | 11 | 9.57% |
| 0.182 | 1 | 0.87% |
| 0.200 | 6 | 5.22% |
| 0.250 | 10 | 8.70% |
| 0.286 | 5 | 4.35% |
| 0.333 | 21 | 18.26% |
| 0.400 | 6 | 5.22% |
| 0.500 | 16 | 13.91% |
| 0.600 | 3 | 2.61% |
| 0.667 | 6 | 5.22% |
| 0.750 | 4 | 3.48% |
| 0.833 | 2 | 1.74% |
| 1.000 | 3 | 2.61% |
| Total | 115 | 100% |

(d)

| Support | Frequency | Percent |
|---------|-----------|---------|
| 0.000 | 7 | 63.64% |
| 0.167 | 1 | 9.09% |
| 0.250 | 1 | 9.09% |
| 0.500 | 1 | 9.09% |
| 1.000 | 1 | 9.09% |
| Total | 11 | 100% |

To use these information as discrete priors, the set of $\{0.000, 0.001, 0.002, \dots, 0.999, 1\}$ with 101 values has been used as the distribution's support and a weight equal to frequencies in **Table 2** has been assigned to the respective values. Other values that had zero frequency have been given a weight of 0.001. Further, probabilities have been assigned to values in the support by dividing the frequencies by the total summation of the weights.

This approach does not change the mean of the priors significantly and provides a nonzero probability for other values in the support when mixed with likelihood. The prior then could be written as:

$$P\left(\theta_{jkl} = \frac{m}{100}\right) = p_m^{jkl} \quad m = 0, 1, \dots, 100, j, k, l = 0, 1, \quad (6)$$

where p_m^{jkl} is derived from **Table 2** after adding nonzero weights as described earlier. Combining the prior with the likelihood given in Equation (2) produces the following discrete posterior distribution for the five model parameters:

$$P\left(\theta_{jkl} = \frac{m}{100}\right) = \frac{p_m^{jkl} m^{\sum_{i=1}^{k_{jkl}} x_i^{(j,k,l)}} (100-m)^{\sum_{i=1}^{k_{jkl}} n_i^{(j,k,l)} - \sum_{i=1}^{k_{jkl}} x_i^{(j,k,l)}}}{\sum_{m=0}^{100} p_m^{jkl} m^{\sum_{i=1}^{k_{jkl}} x_i^{(j,k,l)}} (100-m)^{\sum_{i=1}^{k_{jkl}} n_i^{(j,k,l)} - \sum_{i=1}^{k_{jkl}} x_i^{(j,k,l)}}} \quad (7)$$

Table 3(a) provides parameters' estimate using posterior mean and the credible sets accompanied by their percent coverage. Estimation for θ_{000} is 0.200 whereas for θ_{100} it is equal to 0.3280. The relative risk of having the PD in positive heredity families whose neither one of the parents were diagnosed with the PD to families with negative heredity is $\frac{0.32801}{0.20012} = 1.64\%$. The estimation for

θ_{101} and θ_{110} are 0.2649 and 0.3148 respectively both with 99% credible set of [0.25, 0.33]. The chance of developing the PD increases to 0.4422 when both parents had PD which is 1.35% higher than the families where neither one of the parents were diagnosed with the PD. These estimations are close to the maximum likelihood estimations in **Table 1(a)**.

2.2.2. Uniform Prior

In this section, the available data from grandparents' family is considered as Binomial counts and is mixed with the data from the individual's family in the form of likelihood to derive Bayesian estimations by using non-informative uniform priors. In this case, the posterior distribution could be written as

$$f(\theta_{jkl}) = \frac{\theta_{jkl}^{\sum_{i=1}^{k'_{jkl}} x_i^{(j,k,l)}} (1-\theta_{jkl})^{\sum_{i=1}^{k'_{jkl}} n_i^{(j,k,l)} - \sum_{i=1}^{k'_{jkl}} x_i^{(j,k,l)}}}{\int_0^1 \theta_{jkl}^{\sum_{i=1}^{k'_{jkl}} x_i^{(j,k,l)}} (1-\theta_{jkl})^{\sum_{i=1}^{k'_{jkl}} n_i^{(j,k,l)} - \sum_{i=1}^{k'_{jkl}} x_i^{(j,k,l)}} d\theta_{jkl}}, \quad (8)$$

where k'_{jkl} accounts for the new sample cases in families when $H_i = j$, $F_i = k$, $M = l$ for fixed j, k, l . Since no information regarding the gender of the grandparents with the Parkinson's was available, the information from this link has been copied for both θ_{101} and θ_{110} . When combined with the primary likelihood, this provides distinct estimations for θ_{101} and θ_{110} .

The Bayesian computations in this section have been carried out using WinBUGS. Monte Carlo Simulations with three simultaneous chains have been utilized to arrive at stable estimations. A burn in of 110,000 with threads of 150,000 long has been used for this part of the analysis. **Table 3(b)** provides the results of the estimations.

The model parameter θ_{000} is estimated to be 0.0625 with 95% credible interval of (0.0582, 0.0669). For positive heredity group, θ_{100} through θ_{111} were estimated to be 0.3147, 0.2700, 0.2785, and 0.2702, respectively. As expected, all estimations are close to their respective maximum likelihood estimations provided in **Table 2(b)** since a non-informative uniform prior has been used. Looking at the relative risk of $\theta_{100}/\theta_{000} = 5.042$, the chance of developing the Parkinson's for an offspring in positive heredity family when neither one of the parents had the

Table 3. Bayesian estimations of the model parameters with (a) discrete prior (b) uniform prior.

| (a) | | | | |
|----------------|---------|----------|------------|------------------|
| Parameter | Mean | S.D. | % Coverage | Credible Set |
| θ_{000} | 0.20012 | 0.001657 | 99.353 | {0.20} |
| θ_{100} | 0.32801 | 0.008845 | 94.451 | {0.33} |
| θ_{101} | 0.26487 | 0.031059 | 98.649 | [0.25, 0.33] |
| θ_{110} | 0.31477 | 0.031213 | 98.779 | [0.25, 0.33] |
| θ_{111} | 0.44222 | 0.100040 | 90.687 | [0.17, 0.50] |
| (b) | | | | |
| Parameter | Mean | S.D. | % Coverage | Credible Set |
| θ_{000} | 0.0625 | 0.00223 | 95 | [0.0582, 0.0669] |
| θ_{100} | 0.3147 | 0.01341 | 95 | [0.2887, 0.3414] |
| θ_{101} | 0.2700 | 0.01490 | 95 | [0.2411, 0.2996] |
| θ_{110} | 0.2785 | 0.01500 | 95 | [0.2495, 0.3085] |
| θ_{111} | 0.2702 | 0.05119 | 95 | [0.1758, 0.3756] |

PD is about five times higher than an offspring in a family with negative heredity. Interestingly, children were less likely to have the PD when both parents had the PD than the condition where neither one of the parents were diagnosed with the PD. This might suggest the effect of dormant genes or lack of adequate data for case of positive PD status of both parents. This estimation is in accordance with some research studies [4] [5].

3. Results

The chance of developing the PD in families with negative heredity and in four family types with positive heredity has been estimated using four different approaches, two Maximum Likelihood and two Bayesian. **Table 4** presents all four estimations and their standard deviation. It is extremely important knowing such probabilities as the individual can take precautionary measures with respect to different therapies and physical exercises to defy the odds and preserving the quality of life for individuals with higher risk.

The information for grandparents and their families date respectively to two and one generation back thus might not be as reliable as it should be. There were registered cases having 18 and 21 aunts/uncles which might be due to registration error or might represent extreme cases that could affect the analysis to some degree. For this reason, the first and second-generation information of 47 cases that had more than 11 aunts/uncles has been excluded from the present study. It is more reasonable to use former less reliable information as prior knowledge and let the more recent and authentic information shape it to more reliable estimations. Thus, we opt to report the Bayesian estimations with discrete prior as the most reliable.

Table 4. Comparison of the estimations.

| Parameter | ML | | Bayes | |
|---------------------|---------------------------------------|--|-------------------|------------------|
| | Primary using 1 st link | Secondary using 2 nd links | Discrete Prior | Uniform Prior |
| θ_{000} (SD) | 0.214 (0.0079) | 0.016 (0.0022) | 0.20012 (0.0017) | 0.0625 (0.0022) |
| θ_{100} (SD) | 0.324 (0.0194) | 0.314 (0.0134) | 0.32801 (0.0088) | 0.3147 (0.0134) |
| θ_{101} (SD) | 0.274 (0.0229) | 0.269 (0.0149) | 0.26487 (0.0311) | 0.2700 (0.0149) |
| θ_{110} (SD) | 0.294 (0.0232) | 0.269 (0.0149) | 0.31477 (0.0312) | 0.2785 (0.0150) |
| θ_{111} (SD) | 0.414 (0.0915) | 0.264 (0.0519) | 0.44222 (0.1000) | 0.2702 (0.0514) |

For negative heredity group, estimations of θ_{000} vary from 0.016 to 0.214, both extreme estimations are ML estimations based on sample sizes of 2169 and 824. Increasing sample size should increase the consistency and efficiency of the ML estimations but one must consider the authenticity of information as well. This difference could also point out the change in prevalence of the Parkinson's through generations. The Bayesian method with discrete prior provides an estimation of 0.20012 meaning that a child in this family has a 20% chance of developing the Parkinson's disease.

Estimations for θ_{100} are less volatile among four different methods. In this case, Bayesian method with discrete prior estimates a chance of 33% for developing the Parkinson's for the children. When compared to θ_{000} , a relative risk of 1.59 is derived suggesting 1.59 times more chance of developing the PD if there is a positive Parkinson history in the family although neither one of the parents had the disease. This estimation is in accordance with findings of a community-based study in 1996 [6].

The chance of developing the PD in a family whose mother is diagnosed with the disease is estimated to be 0.26487 in comparison to 0.31477 when father had the Parkinson's; suggesting that the chance of passing the Parkinson's from father to children is slightly higher than passing it from mother to children [6]. Finally, there is 44% chance of developing the Parkinson's in a family whose both parents have the disease.

4. Conclusion and Discussion

Although a primary cause for Parkinson's disease is yet to be identified [7], several risk factors are known to be contributing to the disease. Among them are age [8], family history [3] [4] [8], sex [9] [10], environmental factors [3] [5] [11], and head trauma [12]. There is an overwhelming evidence for a role of heredity in susceptibility to Parkinson's disease [4] [8] [13]. While there have been some opinions on the chance of developing the PD based on family history in the news and the Internet, with no citation to any valid research article, there has not been a single statistical model to measure this effect reliably. This study that utilizes real data from the vast database of Parkinson's Progression Markers Initiative

(PPMI) [3] is one of the first to provide a sophisticated statistical model to support the conclusions. The provided Bayesian modeling that allows updating results when new information is added to the dataset, is very helpful in the ever-growing information age. Gender is thought to be one of the risk factors in developing the Parkinson's [9] [10]. Lack of enough gender related information in the available data prevented deriving separate estimations for men and women.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Parkinson's Disease Foundation. http://www.pdf.org/about_pd
- [2] Aarsland, D. and Kramberger, G.M. (2015) Neuropsychiatric Symptoms in Parkinson's Disease. *Journal of Parkinson's Disease*, **5**, 659-667. <https://doi.org/10.3233/JPD-150604>
- [3] Parkinson's Progression Markers Initiative (PPMI) Database. <https://www.ppmi-info.org/data>
- [4] Gwinn, K. (2009) Genetics and Parkinson's Disease: What Have We Learned? Winter 2009 Newsletter of Parkinson's Disease Foundation Inc.
- [5] Priyadarshia, A., Khudera, S.A., Schauba, E.A. and Priyadarshi, S.S. (2001) Environmental Risk Factors and Parkinson's Disease: A Meta-Analysis. *Environmental Research*, **86**, 122-127. <https://doi.org/10.1006/enrs.2001.4264>
- [6] Marder, K., Tang, M.X., Mejia, H., Alfaro, B., Cote, L., Louis, E., Groves, J. and Mayeux, R. (1996) Risk of Parkinson's Disease among first-Degree Relatives. *Neurology*, **47**, 155-160. <https://doi.org/10.1212/WNL.47.1.155>
- [7] De Lau, M.L. and Breteler, M. (2006) Epidemiology of Parkinson's Disease. *The Lancet Neurology*, **5**, 525-535. [https://doi.org/10.1016/S1474-4422\(06\)70471-9](https://doi.org/10.1016/S1474-4422(06)70471-9)
- [8] Gorell, J.M., Peterson, E.L., Rybicki, B.A. and Johnson, C.C. (2004) Multiple Risk Factors for Parkinson's Disease. *Journal of the Neurological Sciences*, **217**, 169-174. <https://doi.org/10.1016/j.jns.2003.09.014>
- [9] Cilia, R., Siri, C., Rusconi, D., Allegra, R., Ghiglietti, A., Sacilotto, G., Zini, M., Zecchinelli, A.L., Asselta, R., Duga, S., Paganoni, A.M., Pezzoli, G., Seia, M. and Goldwurm, S. (2014) LRRK2 Mutations in Parkinson's Disease: Confirmation of a Gender Effect in the Italian Population. *Parkinsonism & Related Disorders*, **20**, 911-914. <https://doi.org/10.1016/j.parkreldis.2014.04.016>
- [10] Saunders, R., Stanley, K., San Luciano, M., Barrett, M.J., Shanker, V., Raymond, D., Ozelius, L.J. and Bressman, S.B. (2012) Gender Differences in the Risk of Familial Parkinsonism: Beyond LRRK2. *Neuroscience Letters*, **496**, 125-128. <https://doi.org/10.1016/j.neulet.2011.03.098>
- [11] Sullivan, K.L., Mortimer, J.A., Wang, W., Zesiewicz, T.A., Brownlee, H.J. and Borstein, A.R. (2015) Occupational Characteristics and Patterns as Risk Factors for Parkinson's Disease: A Case Control Study. *Journal of Parkinson's Disease*, **5**, 813-820. <https://doi.org/10.3233/JPD-150635>
- [12] Taylor, C.A., SaintHilaire, M.H., Cupples, L.A., Thomas, C.A., Burchard, A.E.,

Feldman, R.G. and Myers, R.H. (1999) Environmental, Medical, and Family History Risk Factors for Parkinson's Disease: A New England-Based Case Control Study. *American Journal of Medical Genetics*, **88**, 742-749.

[https://doi.org/10.1002/\(SICI\)1096-8628\(19991215\)88:6<742::AID-AJMG29>3.0.CO;2-#](https://doi.org/10.1002/(SICI)1096-8628(19991215)88:6<742::AID-AJMG29>3.0.CO;2-#)

- [13] Michael, J. (2017) Fox Foundation for Parkinson's Research, Questions and Answers. <https://www.michaeljfox.org/understanding-parkinsons/i-have-got-what.php>