# Investigating the Contribution of Susceptibility Factors on Early Clinical Expression of Multiple Sclerosis

**Justin C. Yu, BS[1]; Jinghong Jiangyang, BS[1]; Erica Utigard, DO[2]; Douglas D. Gunzler, PhD[3]; Daniel D. Ontaneda, MD, MSc[2]; Farren B. S. Briggs, PhD, ScM[4]**

(1) Neuroimmunological Disorders Gene-Environment Epidemiology (NDGE) Laboratory, Department of Population and Quantitative Health Sciences, School of Medicine, Case Western Reserve University (2) The Mellen Center for MS, Cleveland Clinic (3) Department of Population and Quantitative Health Sciences, School of Medicine, Case Western Reserve University

## Hypothesis

There is evidence suggesting early clinical expression may predict long-term multiple sclerosis (MS) outcomes; however, little is known about factors influencing onset heterogeneity. **Given temporality, we hypothesize MS risk factors contribute to early clinical expression.**

## Project Overview

This study involved 1,524 persons with MS (PwMS) who participated in the Accelerated Cure Project for MS (ACP; www.acceleratedcure.org). ACP is an open-access repository of detailed epidemiologic data and blood-derived biospecimen samples of PwMS, other demyelinating diseases, and unaffected controls, recruited from 10 US neurology centers. Table 1 describes the MS study population.

We focused on **three** aspects of early clinical expression: **1. Age of onset** (**AOO**, age at 1st symptom); **2. Number of impaired functional domains** (**NIFDs,** PwMS completed 30 questions on symptoms at onset, which were categorized into 11 functional domains: motor, cerebellar, spasticity, optic nerve, sensory, cognitive, brainstem/bulbar, sexual, bladder/bowel, affect/mood, & fatigue); and **3. Early relapse activity** (**ERA**, or number of relapses in the 2 years after onset). For each outcome, we conducted variable selection using least absolute shrinkage and selection operator (LASSO) and all variables (**Table 1**). The resulting variables selected by LASSO and their respective outcomes were investigated using **linear (for AOO)** and **negative binomial (for NIFD and ERA) regressions** with robust standard errors. The analysis for ERA was restricted to those with relapsing remitting MS. *HLA-DRB1*15:01* was available in only 1,054 subjects. Considering the known differences in clinical presentation between genders and MS subtypes, we conducted stratified analyses and tested for equality of regression coefficients to determine if there were gender- or subtype-specific effects [1]. The test statistic for the equality of coefficients between the genders is:

$$Z = \beta_{Males} - \beta_{Females} \Big/ \sqrt{SE\beta^2_{Males} + SE\beta^2_{Females}}$$

## Table 1: Data Summary

| Characteristic | | All Mean(sd)/% | Relapse Remitting Mean(sd)/% | Primary Progressive Mean(sd)/% | $P_{RR\ vs\ PP}$ | Females Mean(sd)/% | Males Mean(sd)/% | $P_{F\ vs\ M}$ |
|---|---|---|---|---|---|---|---|---|
| **N** | | 1524 | 1414 | 110 | | 1185 | 339 | |
| **Birth Year** (median [IQR]) | | 1961.00 | 1962.00 | 1954.00 | **<0.001** | 1961.00 | 1961.00 | 0.56 |
| **AOO** | | 33.50 (9.87) | 32.87 (9.56) | 41.63 (10.29) | **<0.001** | 33.29 (9.94) | 34.25 (9.60) | 0.11 |
| **Males** | | 22.2% | 21.1% | 36.4% | **<0.001** | - | 100 | |
| **Progressive at onset** | | 7.2% | - | 100% | | 5.9% | 11.8% | **<0.001** |
| **Years of Education** | | 15.82 (2.92) | 15.83 (2.93) | 15.72 (2.86) | 0.71 | 15.72 (2.89) | 16.15 (3.01) | **0.02** |
| **Race** | **White** | 90.2% | 90.3% | 89.1% | | 89.3% | 93.5% | |
| | **Black** | 7.5% | 7.3% | 10.0% | 0.37 | 8.2% | 5.0% | 0.07 |
| | **Other** | 2.3% | 2.4% | 0.9% | | 2.5% | 1.5% | |
| **Smoker within 5 years of onset** | | 33.1% | 33.1% | 33.6% | 0.99 | 32.4% | 35.7% | 0.29 |
| **Had Mononucleosis before onset** | | 27.9% | 28.5% | 20.0% | 0.07 | 29.1% | 23.8% | 0.07 |
| **Diagnosis of Obesity before onset** | | 8.4% | 8.7% | 4.5% | 0.18 | 9.8% | 3.5% | **<0.001** |
| **NIFDs** (median [IQR]) | | 2.00 [1.00, 4.00] | 2.00 [1.00, 4.00] | 1.00 [1.00, 3.00] | **<0.001** | 2.00 [1.00, 4.00] | 2.00 [1.00, 3.00] | 0.11 |
| **NIFDs** | **1** | 32.8% | 33.0% | 30.9% | | 31.6% | 37.2% | |
| | **2-4** | 38.5% | 38.6% | 37.3% | **<0.001** | 39.2% | 36.0% | 0.36 |
| | **5-8** | 16.1 | 16.7% | 8.2% | | 16.6% | 14.2% | |
| | **9** | 2.7 | 2.8% | 0.9% | | 2.6% | 2.9% | |
| **ERA** | | | 2.09 (1.91) | | | 2.12 (2.00) | 1.98 (1.60) | 0.34 |
| ***HLA-DRB1*15:01*** (N) | | 1054 | 985 | 69 | | 821 | 233 | |
| **0 Risk Alleles** | | 53.3% | 53.3% | 53.6% | | 52.1% | 57.5% | 0.31 |
| **1 Risk Allele** | | 40.0% | 40.4% | 34.8% | 0.2 | 40.9% | 36.9% | |
| **2 Risk Alleles** | | 6.6% | 6.3% | 11.6% | | 6.9% | 5.6% | |

**Table 1:** Data summary of all individuals, including stratification by Relapse Remitting MS (RRMS) or Primary Progressive MS (PPMS) and females or males. Comparison were done using a one-way t-test, Chi-Square, or Kruskal Wallis Rank Sum test.

## Table 2A: AOO Stratified by Gender

| | All MS | p-value | All Males | p-value | All Females | p-value | $P_{F\ vs\ M}$ |
|---|---|---|---|---|---|---|---|
| **N** | **1525** | | **340** | | **1185** | | |
| **Gender (Males = 1)** | 0.033 (0.004, 0.061) | **0.025** | - | - | - | - | |
| **Subtype (PP = 1)** | 0.12 (0.073, 0.17) | **8.32x10-7** | 0.19 (0.12, 0.26) | **5.20x10-7** | 0.09 (0.028, 0.15) | **0.005** | **0.045** |
| **Years of Education** | -0.006 (-0.01, -0.001) | **0.01** | -0.006 (-0.014, 0.002) | 0.14 | -0.005 (-0.011, 0) | **0.043** | NS |
| **Birth Year** | -0.013 (-0.015, -0.012) | **2.47x10-86** | -0.012 (-0.014, -0.009) | **2.55x10-21** | -0.014 (-0.015, -0.012) | **8.16x10-67** | NS |
| **Smoker within 5 Years of Onset (Yes = 1)** | -0.081 (-0.11, -0.055) | **2.16x10-9** | -0.08 (-0.13, -0.027) | **3.32x10-3** | -0.084 (-0.12, -0.054) | **8.96x10-8** | NS |
| **Diagnosis of Obesity before Onset (Yes = 1)** | 0.077 (0.035, 0.12) | **3.72x10-4** | 0.17 (0.064, 0.27) | **0.002** | 0.069 (0.024, 0.11) | **0.003** | 0.087 |

**Table 2A:** Coefficients, 95% CIs, and p-values for **linear regression** with the **natural log of AOO** as the outcome. Variables were selected using LASSO which was performed using glmnet from the glmnet package on R version 3.3.2. Models were adjusted for recruitment site. Because natural log of AOO was used, the coefficients can be interpreted as a percent change using the following equation: $(e^{\beta_X} - 1) * 100$. **For example, for all MS, PwMS who were obese were 8% older at onset than non-obese PwMS. Males were 3% older than females at MS onset. Also, the AOO of PPMS was statistically different between genders, males with PP were 19% older and females were 9% older than their RRMS counterparts.**

## Table 3A: NIFDs Stratified by Gender

| | All MS | p-value | All Males | p-value | All Females | p-value | $P_{F\ vs\ M}$ |
|---|---|---|---|---|---|---|---|
| **N** | 942 | | 207 | | 735 | | |
| **Years of Education** | -0.021 (-0.036, 0.0056) | **0.007** | -0.021 (-0.068, 0.026) | 0.38 | -0.017 (-0.037, 0.0024) | 0.085 | NS |
| **Subtype (PP = 1)** | -0.17 (-0.33, -0.013) | **0.034** | -0.19 (-0.46, 0.083) | 0.17 | -0.099 (-0.4, 0.2) | 0.51 | NS |
| **ln(Age of Onset)** | 0.15 (0.01, 0.28) | **0.031** | 0.039 (-0.32, 0.4) | 0.83 | 0.27 (0.089, 0.45) | **0.0034** | NS |
| **Smoker within 5 years of Onset (Yes = 1)** | 0.079 (-0.007, 0.16) | 0.071 | 0.099 (-0.13, 0.33) | 0.39 | 0.13 (0.015, 0.24) | **0.027** | NS |
| **Diagnosis of Obesity before Onset (Yes = 1)** | 0.14 (0.003, 0.28) | **0.046** | 0.18 (-0.26, 0.63) | 0.42 | 0.15 (-0.034, 0.34) | 0.11 | NS |

**Table 3A:** Coefficients, 95% CIs, and p-values for **negative-binomial regression** with **NIFDs** as the outcome variable (1 or more impaired domain). Variables were selected using LASSO which was performed using cv.glmreg from the mpath package on R version 3.3.2. Models were adjusted for recruitment site. Coefficients can be interpreted as ratios. **For example, an PwMS who was obese at onset reported a 14% increase in the NIFDs at onset than non-obese PwMS. There were no differences by gender.**

## Table 4: ERA in RRMS Stratified by Gender

| | RRMS | p-value | Males | p-value | Females | p-value | $P_{F\ vs\ M}$ |
|---|---|---|---|---|---|---|---|
| **N** | **940** | | **200** | | **740** | | |
| **Years of Education** | -0.016 (-0.036, 0.004) | 0.11 | -0.013 (-0.05, 0.024) | 0.5 | -0.014 (-0.038, 0.0099) | 0.25 | NS |
| **ln(Age of Onset)** | -0.25 (-0.44, -0.059) | **0.01** | -0.047 (-0.41, 0.32) | 0.8 | -0.28 (-0.5, -0.068) | **0.0098** | NS |
| **Smoker within 5 years of Onset (Yes = 1)** | 0.11 (-0.013, 0.24) | 0.078 | 0.26 (0.044, 0.48) | **0.019** | 0.088 (-0.064, 0.24) | 0.26 | NS |
| **Diagnosis of Obesity before Onset (Yes = 1)** | 0.25 (0.04, 0.46) | **0.019** | 0.045 (-0.63, 0.71) | 0.9 | 0.26 (0.041, 0.48) | **0.02** | NS |

**Table 4:** Coefficients, 95% CIs, and p-values for **negative-binomial regression** with **ERA** as the outcome variable in RRMS alone. Variables were selected using LASSO which was performed using cv.glmreg from the mpath package on R version 3.3.2). Models were adjusted for recruitment site. Coefficients can be interpreted as ratios. **For example, if a person with RRMS were obese at onset, they reported a 25% increase in the number of relapses experiences during the first two years after onset than non-obese cases. There were no differences by gender.**

## Table 2B: AOO Stratified by Subtype

| | All MS | p-value | RR | p-value | PP | p-value | $P_{RR\ vs\ PP}$ |
|---|---|---|---|---|---|---|---|
| **N** | **1525** | | **1415** | | **110** | | |
| **Gender (Males = 1)** | 0.033 (0.004, 0.061) | **0.025** | 0.025 (-0.005, 0.056) | 0.1 | 0.11 (0.018, 0.2) | **0.02** | 0.091 |
| **Subtype (PP = 1)** | 0.12 (0.073, 0.17) | **8.32x10-7** | | | - | - | - |
| **Years of Education** | -0.006 (-0.01, -0.001) | **0.01** | -0.006 (-0.011, -0.002) | **0.0054** | 0.006 (-0.008, 0.021) | 0.39 | NS |
| **Birth Year** | -0.013 (-0.015, -0.012) | **2.47x10-86** | -0.013 (-0.015, -0.012) | **1.44x10-83** | -0.012 (-0.017, -0.007) | **2.19x10-5** | NS |
| **Smoker within 5 years of Onset (Yes = 1)** | -0.081 (-0.11, -0.055) | **2.16x10-9** | -0.075 (-0.1, -0.048) | **9.29x10-8** | -0.15 (-0.26, -0.038) | **0.01** | NS |
| **Diagnosis of Obesity before Onset (Yes = 1)** | 0.077 (0.035, 0.12) | **3.72x10-4** | 0.077 (0.033, 0.12) | **5.35x10-4** | 0.054 (-0.14, 0.25) | 0.58 | NS |

**Table 2B:** Coefficients, 95 percent CIs, and p-values for **linear regression** with the **natural log of AOO** as the outcome variable. Variables were selected using LASSO with robust standard errors (LASSO was performed using glmnet from the glmnet package on R version 3.3.2). Models were adjusted for recruitment site. Because natural log of AOO was used as the outcome variable, the coefficients can be interpreted as a percent change using the following equation: $(e^{\beta_X} - 1) * 100$. **For example, the AOO of PPMS is 12% later than for RRMS. There were no other differences by subtype.**

## Table 3B: NIFDs Stratified by Subtype

| | All MS | p-value | RR | p-value | PP | p-value | $P_{RR\ vs\ PP}$ |
|---|---|---|---|---|---|---|---|
| **N** | **942** | | **885** | | **57** | | |
| **Years of Education** | -0.021 (-0.036, 0.0056) | **0.007** | -0.021 (-0.038, -0.0044) | **0.014** | -0.027 (-0.12, 0.069) | 0.58 | NS |
| **Subtype (PP = 1)** | -0.17 (-0.33, -0.013) | **0.034** | | | | | |
| **ln(Age of Onset)** | 0.15 (0.01, 0.28) | **0.031** | 0.14 (-0.015, 0.29) | 0.077 | 0.036 (-0.67, 0.74) | 0.92 | NS |
| **Smoker within 5 years of Onset (Yes = 1)** | 0.079 (-0.007, 0.16) | 0.071 | 0.05 (-0.047, 0.15) | 0.31 | -0.36 (-0.78, 0.059) | 0.09 | 0.062 |
| **Diagnosis of Obesity before Onset (Yes = 1)** | 0.14 (0.003, 0.28) | **0.046** | 0.16 (0.0091, 0.31) | **0.038** | -0.56 (-1.65, 0.52) | 0.31 | NS |

**Table 3B:** Coefficients, 95 percent CIs, and p-values for **negative-binomial regression** with **NIFDs** as the outcome variable. Variables were selected using LASSO in cv.glmreg from the mpath package on R version 3.3.2. Models were adjusted for recruitment site. Coefficients can be interpreted as ratios. **For example, those with PPMS reported 17% less NIFDs at onset that those with RRMS at onset. There were no other differences by subtype.**

## Table 5: Effect of *HLA-DRB1*15:01*

| | All MS | p-value | Males | p-value | Females | p-value | $P_{F\ vs\ M}$ |
|---|---|---|---|---|---|---|---|
| **N** | **942** | | | | | | |
| **AOO** | -0.028 (-0.055, -0.001) | **0.041** | 0.004 (-0.057, 0.066) | 0.89 | -0.031 (-0.062, -0.001) | **0.041** | NS |
| **NIFDs** | -0.027 (-0.11, 0.053) | 0.51 | 0.093 (-0.089, 0.28) | 0.32 | -0.066 (-0.15, 0.021) | 0.13 | NS |
| **ERA** | -0.08 (-0.2, 0.041) | 0.19 | 0.13 (-0.082, 0.34) | 0.23 | -0.13 (-0.27, 0.007) | 0.063 | **0.044** |

**Table 5:** Coefficients, 95 percent CIs, and p-values for **regressions for AOO, NIFDs, and ERA** in models including ***HLA-DRB1*15:01*** (rs3135388). Variables were selection were done as previously described. **Each copy of the risk allele was associated with a XX% decrease in AOO, thus *HLA-DRB1*15:01* is associated with earlier onset.**

## Regression Results

- ***HLA-DRB1*15:01*** was significantly associated with earlier AOO, but was not associated with NIFDs or ERA.
- **Obesity** was significantly associated with all three outcomes. Obesity was associated with older AOO and had increases in NIFDs and ERAs,
- **Lower socioeconomic status** (SES; as measured by years of education) associated with earlier AOO and increases in NIFDs.
- **Smoking** was significantly associated with earlier AOO, and was suggestive of increased NIFDs and greater ERA.
- PwMS who were older at onset reported increases in NIFDs and lower ERA.
- Males with PPMS were much older than males with RRMS, and this difference was twice as much as the effect in females (Table 2A; 19% vs 9%).
- **These results demonstrate that MS risk factors do demonstrably influence early clinical expression of MS.**

**Reference** 1. Paternoster R, Brame R, Mazerolle P, Piquero A. Using the correct statistical test for the equality of regression coefficients. Criminology. 1998;36(4):859-66.