It is made available under a [CC-BY-NC 4.0 International license](http://creativecommons.org/licenses/by-nc/4.0/) .

Manuscript Information

Title: Pathogenesis of Multiple Sclerosis: Genetic, Environmental, and Random Mechanisms

Author: Douglas S. Goodin, MD

Department of Neurology, *University of California, San* Francisco & the *San Francisco VA Medical Center*, San Francisco, CA, USA

Address for Correspondence:

Douglas S. Goodin, MD

Department of Neurology; *University of California, San Francisco*

675 Nelson Rising Lane, Suite #221D

San Francisco, CA 94158

Phone: (415) 514 2464

Fax: (415) 514 2470

E mail: douglas.goodin@ucsf.edu

*Manuscript Subject Areas***:**

Multiple Sclerosis, pathogenesis, environmental factors, genetic factors, randomness, determinism

*Author Contributions:*Douglas S. Goodin:

Conceptualization; Formal analysis; Methodology; Software; Writing – original draft, review & editing

Funding: None

Conflicts of Interest: None

Acknowledgements:

I am especially indebted to John Petkau, PhD, Professor Emeritus, Department of Statistics, University of British Columbia, Canada, for enormous help with this project. He devoted many hours of his time to critically reviewing early versions of this analysis and contributed immensely both to the clarity and to the logical development of the mathematical and statistical arguments presented herein. I am also indebted to my mentor, Michael J. Aminoff, MD, Professor Emeritus, Department of Neurology, University of California, San Francisco, USA, for his invaluable help with this project. He critically, and thoughtfully, reviewed many drafts of this manuscript and contributed enormously to the logic and clarity of its presentation.

It is made available under a [CC-BY-NC 4.0 International license](http://creativecommons.org/licenses/by-nc/4.0/) .

Abstract

BACKGROUND: MS-pathogenesis requires both genetic factors and environmental events. The question remains, however, whether these factors and events *completely* describe the MS disease-process. This question was addressed using the Canadian MS-data, which includes 29,478 individuals, representing 65-83% of all Canadian MS-patients.

METHODS: The "*genetically-susceptible*" subset of the population, (G), includes *everyone* who has *any* of the population, (
er *some* environmen
ceptible" individual,
o cause MS in that p
ved in MS-pathogen
lue-ranges" for each
usible-ranges are the G), includes *everyone* who has *any*
tal-conditions. A "*sufficient*"
includes *every* set of environmenta
erson. This analysis incorporates
esis, only some of which are
n parameter. Those *parameter-value*
en determined. *non-zero* life-time chance of developing MS, under *some* environmental-conditions. A "*sufficient*" environmental-exposure, for *any* "*genetically-susceptible*" individual, includes *every* set of environmental conditions, each of which is *sufficient*, by itself, to *cause* MS in that person. This analysis incorporates several different *epidemiologic-parameters*, involved in MS-pathogenesis, only some of which are directly-observable, and establishes "*plausible*-*value-ranges"* for each parameter. Those *parameter-value* combinations (solutions) that fall within these *plausible-ranges* are then determined.

RESULTS: Only a fraction of the population can possibly be "*genetically-susceptible*". Thus, many individuals have *no possibility* of developing MS under *any* environmental conditions. Moreover, *some* "*genetically-susceptible*" individuals, despite their experiencing a "*sufficient*" environmental-exposure, *will never* develop disease.

CONCLUSIONS: This analysis *explicitly includes* all of those genetic factors and environmental events (including interactions), which are necessary for MS-pathogenesis*,* regardless of whether these are known, suspected, or as yet unrecognized. Nevertheless, in addition, "*true*" randomness seems to play a critical role in disease-pathogenesis. This observation provides empirical evidence that undermines the widely-held deterministic view of nature. Moreover, both sexes seem to have a similar genetic and environmental disease-basis. If so, this indicates that this random element is primarily responsible for the currently-observed differences in disease-expression between *susceptible-women* and *susceptible-men*.

It is made available under a [CC-BY-NC 4.0 International license](http://creativecommons.org/licenses/by-nc/4.0/) .

Introduction

 Multiple sclerosis (MS)-pathogenesis requires both environmental-events and genetic-factors [1- 4]. Considering genetics, familial-aggregation of MS-cases is well-established. MS-risk is increased ~30 fold in non-twin siblings and ~250-fold in monozygotic (*MZ*)-twins of an *MS-proband* [1,2,5]. Moreover, 233 MS-associated genetic-traits are now identified [6]. Nevertheless, MS-genetics is complex. The strongest MS-association is with the *Class-II* haplotype, HLA - $DRB1*15:01 \sim DQB1*06:02$, located at $(6p21)$, having an odd strongest MS-association is with the *Class-II* haplotype, *HLA-DRB1*15:01~DQB1*06:02,* located at 6*p*21), having an odds-ratio (*OR*) of (

MS-associations are quite weak [6] – (1

..414). Also, *DRB1**15:01~*DQB1**06.

aplotypes – the most frequent such hap

except *MZ*-twins) possesses a unique c

the observed *pro* ~3) in heterozygotes and of (
redian – OR = 1.158; inter
 \cdot 02 is highly "selected", accous
plotype – among European-de
combination of the 233 MS-ass
twin-concordance is: (0.11 –
disease-expression (*Table-4; l* ~6) in homozygotes [1,2,5,6]. Other

quartile – range = 1.080 –

unting for (~13%) of *DRB1~DQB1*-

cedents [1-8]. Moreover, *everyone*

sociated genetic-traits [3]. Finally,

0.46) and, consequently, genetics

Reference: MS-associations are quite weak $[6] - (median - OR = 1.158;$ interquartile - range = 1.080 - \therefore 02 is highly "selected", accounting for (~13%) of *DRB1~*.
plotype – among European-decedents [1-8]. Moreover, *ever*
combination of the 233 MS-associated genetic-traits [3]. Fin-
twin-concordance is: (0.11 – 0.46) 1.414). Also, *DRB1*15:01~DQB1*06:02* is highly "selected", accounting for (
haplotypes – the most frequent such haplotype – among European-decedents [1-
(except *MZ*-twins) possesses a unique combination of the 233 MS-as ~13%) of *DRB1~DQB1-*
8]. Moreover, *everyone*
netic-traits [3]. Finally,
, consequently, genetics
3*]*).
nnth-of-birth effect,
nmental-events in MShaplotypes – the most frequent such haplotype – among European-decedents [1-8]. Moreover, *everyone* (except *MZ-*twins) possesses a unique combination of the 233 MS-associated genetic-traits [3]. Finally, the observed *probability-range* for *MZ-*twin-concordance is: plays only a minor role in determining disease-expression (*Table-4; Reference:[3]*).

-linking MS-risk to the solar cycle, likely implicates intrauterine/perinatal environmental-events in MS--Because this gradient is also evident for *MZ-*twin-recurrence-rates (*Table-4; Reference:[3]*), 0.11 – 0.46) and, consequently, genetics

ble-4; Reference:[3]).

ll-documented month-of-birth effect,

e/perinatal environmental-events in MS-

recurrence-rate for dizygotic (DZ)-twins

trauterine/perinatal environmental MS is also linked to environmental-events. First, a well-documented *month-of-birth* effect, pathogenesis [2,9-11]. Second, given an *MS-proband*, the MS-recurrence-rate for dizygotic (*DZ*)-twins exceeds that for non-twin siblings [2,3,5] – also implicating intrauterine/perinatal environmental-events [2,3]. Third, MS becomes increasingly prevalent farther north or south from equatorial-regions [2,12]. environmental-factors are likely responsible. Fourth, a prior Epstein-Barr viral (*EBV*) infection is found in almost all ($> 99\%$) current MS patients [2,13,14]. If these rare *EBV-negative* patients represent *false*-> 99%) *current* MS patients [2,13,14]. If these rare *EBV-negative* patients represent *false-*
sts – either from inherent-errors when using *any* fixed antibody-titer "*cut-off*" to determine
vity, or from *only* deter *negative-tests* – either from inherent-errors when using *any* fixed antibody-titer "*cut-off*" to determine *EBV-positivity*, or from *only* determining antibody-responses to *some EBV-antigens* [2] – then one can conclude that *EBV-*infection is a necessary-factor in *every* causal-pathway, which led to MS in these individuals [2]. Regardless, an *EBV*-infection *must* somehow be involved in MS-pathogenesis [2,13,14]. Lastly, smoking and vitamin-D deficiency are implicated in MS-pathogenesis [2,15,16].

[4] in a relatively non-mathematical format to make its conclusions accessible. The terms and definitions This manuscript presents an analysis regarding genetic and environmental susceptibility to MS used for this analysis are presented in *Table-1*. For interested readers, the mathematical-development is presented in the *Supplemental-Material*. This analysis is based on the *Canadian-Collaborative-Projecton-Genetic-Susceptibility-to-Multiple-Sclerosis* (*CCPGSMS*)-dataset [5,8,9,17-23] – a summary of which is provided in the *Supplemental-Material*-*Sections:10a-b*. The *CCPGSMS*-dataset includes 29,478 MSpatients $(born: 1891 - 1993)$ – estimated to represent $(65 - 83%)$ of Canadian MS-patients [5,23,24]. $born: 1891 - 1993)$ – estimated to represent (
ort is assumed to represent a large random samp. 65 – 83%) of Canadian MS-patients [5,23,24].
le of the Canadian MS-population. Also, this This cohort is assumed to represent a large random sample of the Canadian MS-population. Also, this

It is made available under a [CC-BY-NC 4.0 International license](http://creativecommons.org/licenses/by-nc/4.0/) .

single population provides point-estimates and confidence-intervals for the MS-recurrence-rates in *MZ*twins, *DZ*-twins, and non-twin siblings (*S*), and for the time-dependent changes in the *female-to-male* (*F:M*)-*sex-ratio*.

Methods

1. Genetic-Susceptibility

A population (Z), consists of N individuals. The "genetically-susceptible" (G) -subset includes Z), consists of *N* individuals. The "*genetically-susceptible*" (
non-zero chance of developing MS under *some* environmenta
ls in the (G) -subset $(i = 1, 2, ..., m)$ has a unique-genotype [
aand, randomly-selected from (Z) G)-subset includes
1-conditions. Each G_i]. The probability
[$P(G) = m/N$].
any specific *Time-*
velop, MS. The *everyone* who has *any* non-zero chance of developing MS under *some* environmental-conditions. Each of the $(m \le N)$ individuals in the (G) -subset $(i = 1, 2, ..., m)$ has a unique-genotype $[G_i]$. The probability $m \leq N$) individuals in the (
e event that a *proband*, rand
obership in (G) is assumed in
od (E_T) – see legend Table-
The (*MS*)-subset include
ability of the event that a *pro*
se relevant-exposures occurr G)-subset (
omly-select
ndependent
2; *considera*
es *everyone*
oband, rand
red during ($i = 1, 2, ..., m$) has a unique-genotype $[G_i]$. The probability
red from (Z), is a (G)-subset member is: $[P(G) = m/N]$.
of the environmental-conditions during *any* specific *Time-
ing the definition of* (E_T).
who either has, or of the event that a *proband*, randomly-selected from (Z), is a (G)-subset member is: $[P(G) = m/N]$. Membership in (G) is assumed independent of the environmental-conditions during any specific Time-*Period* (E_T) – see legend Table-2; considering the definition of (E_T) .

Z), is a (
*i*ronmen
finition o
er has, or
cted fron
lled the *M*
obability G)-subset member is: $[P(\text{tal-conditions during any } f(E_T))$.
will subsequently develop
m (Z), is both an (MS)-sul
MS-penetrance for the pop of the event that a proban G = m/N].
specific *Time*
 ρ , MS. The
bset member and member and member and member and members of Z (E_T) , or $P(MS | G, E_T)$. Both *penetrance-values* depend upon the environmental-conditions during (E_T) . G) is assumed independent of the environmental-conditions during *any* specific *Time*-
 e legend Table-2; considering the definition of (E_T) .

(b)-subset includes *everyone* who either has, or will subsequently develo The ity set E_T free $P($) – *see legend Table-2; considering the definition of* (

e (*MS*)-subset includes *everyone* who either has, or wise of the event that a *proband*, randomly-selected from (

vant-exposures occurred during (E_T) , is call $\begin{align*}\n\text{all} \\
\text{in} \\
\text{$).
sul, i
ne
e e e
etr The (MS)-subset includes everyone who either has, or will subsequently develop, MS. The -*penetrance*-*values*, during *MS*)-subset includes *everyone* who either has, or will subsequently develop, MS. The the event that a *proband*, randomly-selected from (*Z*), is both an (*MS*)-subset membe nt-exposures occurred during (E_T), is calle probability of the event that a *proband*, randomly-selected from Z), is both an (
penetrance for
the event that a
enetrance for the
environmental
ceptible-men (l *MS*)-subset member and
the population (*Z*)
 ι *proband*, randomly-
le (*G*)-subset during
-conditions during (*E_T*).
M, *G*). Their *MS*whose relevant-exposures occurred during (E_T) , is called the *MS-penetrance* for the population (Z) E_T)
er,
 $\frac{1}{2}$
 $\frac{1}{2}$
 $\frac{1}{2}$
 $\frac{1}{2}$ %), is called the *MS-penetrance* for the population (
, the probability of the event that a *proband*, rando
is called the *MS-penetrance* for the (G) -subset due
s depend upon the environmental-conditions durine
nen) during $\frac{1}{2}$
 $\frac{1}{2}$), or $P($
 $\text{om } (G)$,
 $\langle MS | G \rangle$

ed are the values, is an $(M$
 E_T). B
ne subse
, during
Z
and: Z). Also, during (
 MS)-subset meml

oth *penetrance*-v

ts of *susceptible*
 (E_T) , are:
 $w = P(MS \mid G, H)$
 $m = P(MS \mid G, H)$ ber
validation
 3π
 3π), the probability of the event that a *proband*, randomly-

; is called the *MS-penetrance* for the (G) -subset during
 ues depend upon the environmental-conditions during (*E*
 pmen (*F*, *G*) and *susceptible-men* selected from *G*), is an (
 G, E_T). If
re the subset
ues, during
and:
MS-penetrial MS)-subset member, is called the *MS-penetrance* for the (

3oth *penetrance-values* depend upon the environmental-co

ets of *susceptible-women* (*F*, *G*) and *susceptible-men* (*M*, *C*
 $g(E_T)$, are:
 $Zw = P(MS | G, F, E_T)$ *G*)-subset during
nditions during (*l*
G). Their *MS*-
robabilities" for .
Als
en
use), or $P($
o define
etrance
The
reptible d are the s
values, du
and
ese *MS-per*
women and 9. Both *penetrance-values* depend upon the environmental-conditions during (

ubsets of *susceptible-women* (*F*, *G*) and *susceptible-men* (*M*, *G*). Their *MS*-

ring (*E*_{*T*}), are:
 $Zw = P(MS | G, F, E_T)$

1: $Zm = P(MS | G,$.
.
en).
in Also defined are the subsets of *susceptible-women*

$$
Zw = P(MS \mid G, F, E_T)
$$

and:
$$
Zm = P(MS \mid G, M, E_T).
$$

-These *MS-penetrance*-*values*, -*susceptible-women* and *susceptible-men* during (G) is independent of environmental-conditions, the $(F:M)$ -*sex-ratio* always reflects the ratio of these two F, G) and *susceptible-men* (
m), are also called the "*failu*
). During *any* (E_T), because the *F*:*M*)-*sex-ratio* always reflects *M*, *G*). Their *MS*-
re-probabilities"
the proportion of w
is the ratio of these -*failure-probabilities* (*see Supplemental-Material-Section:5d*). $y =$
 $x = c$
 ce
 ce
 p
 m), are:
 $P(MS)$
 $= P(MS)$
 $= MS$ $Zm = P($
rance-valu
usceptible
ronmenta
Suppleme $MS + G, M, E$
 $ues, (Zw)$ and
 e -men during
 1 -conditions, that
 R -Materia *cance-valu*
us*ceptible*
ronmental
Sus<mark>ceptib</mark>i *es*, (*Zw*) and
e-men during (
l-conditions, the
htal-Material
ility).
 $(Z E)$
ie $-Se$ Zw) and (ι during (E
ditions, the
Material-S
family of Zm), are also called the "*failure-probabilities*" for r_T). During *any* (E_T), because the proportion of *wom*
($F:M$)-*sex-ratio* always reflects the ratio of these two
ection:5*d*).
exposures is defined that includes r
e (
See
f e:). During *any* (
F:M)-sex-ratio
ction:5d).
xposures is def.
mt", by itself, t $\frac{1}{2}$
alw), because the proportion of *women* in
vays reflects the ratio of these two
d that includes *every* set of
ause MS in that person. Moreover,

2. Environmental-Susceptibility

For each (G) -subset member, a -environmental-exposures, each of which is "*sufficient*", by itself, to *cause* MS in that person. Moreover, G)-subset member, a family of exposures is defined that includes *every* set of obsures, each of which is "*sufficient*", by itself, to *cause* MS in that person. Mc individual to develop MS, that person must experience at for *any* susceptible-individual to develop MS, that person must experience at least one of the "*sufficientexposure-sets*" within *their* family. Individuals sharing the same family of *sufficient-exposures* – although possibly requiring different "*critical-exposure-intensities*" [4] – belong to the same "*exposure-group*".

It is made available under a [CC-BY-NC 4.0 International license](http://creativecommons.org/licenses/by-nc/4.0/) .

Certain environments may be *sufficient* to *cause* MS in *anyone* but are so improbable (e.g., intentional inoculation of someone with myelin proteins or other agents) that, effectively, they never occur spontaneously. Nevertheless, even individuals who can *only* develop MS under such extreme (or unlikely) conditions, are still (G) -subset members – i.e., they *can* develop MS under *some* environmentalconditions.

G)-subset members – i.e., they *can* develop MS under *some* environmental-
event (E) – i.e., that a randomly-selected member of (G) , during (E_T) ,
dificient to *cause* MS in them – is represented as: $P(E \mid G, E_T)$. A
The probability of the event (E) – i.e., that a randomly-selected member of (G) , during (E_T) , experiences an environment *sufficient* to *cause* MS in them – is represented as: $P(E \mid G, E_T)$. A mathematical definition for the (E) -event is provided in the Supplemental-Material-Section:1a.

-present or absent. The *only* requirement is that each *exposure-set*, taken together, is *sufficient*, by itself, to E) – i.e., that a randomly-selected member of (
to *cause* MS in them – is represented as: $P(E$ |
vent is provided in the *Supplemental-Material-S*
es is *completely* undefined and agnostic regardi
l; 2) when, during li G), during (

G, E_T). A

Section: la.

ng: 1) how 1

exposures n

es are; 5) when

exposures n 3-),
ny
d to
l to 1-*cause* MS in a specific susceptible-individual or in susceptible-individuals belonging to the same al-Section
al-Section
se expose
se expose
served
setting and the set of the set (a). A
in:1*a*.
b) how sures
5) w sures E)-event is provided in the *Supplemental-Material-Section:1a.*

bosures is completely undefined and agnostic regarding: 1) how

olved; 2) when, during life, and in what order, these exposures io

on of the required exposu Each set of *sufficient-exposures* is *completely* undefined and agnostic regarding: 1) how many environmental-exposures are involved; 2) when, during life, and in what order, these exposures need to occur; 3) the intensity and duration of the required exposures; 4) what these exposures are; 5) whether *any* of these exposures needs to interact with *any* genetic-factors; and 6) whether certain exposures need to be *exposure-group*.

3. MZ-Twins, DZ-Twins, and Siblings

The term (MZ) represents the event that a *proband*, randomly-selected from (Z), is an (MZ)- MZ) represents the event that a *proband*, randomly-selected from (equivalently, is an MZ -twin. This *proband*'s twin is called their "*cc proband* belongs to the (*MS*, *MZ*)-subset and their *co-twin* belongs ility *Z*), is an (
 $\frac{1}{2}$, is to (*MZ*)
 $\frac{1}{2}$ (*MS*, *MZ*)

eas (*MZ_M*, *S*) for *DZ* MZ)-
ne
is the
is the
s)
(cosubset member or, equivalently, is an *MZ-*twin. This *proband*'s twin is called their "*co-twin*". The probability that the *proband* belongs to the (MS, MZ)-subset and their *co-twin* belongs to (MZ) is the *MS*, *MZ*)-subset and their *co-twin* belongs to (*MS*)
ongs to (*MZ*) and their *co-twin* belongs to (*MS*
is subset (or event) for the *proband*, whereas (
co-twin. The analogous subsets (or events) for
 S_{MS}) are d MZ) is the
 (MZ) .
 MZ_{MS})
 DZ "co-

selected

estimated by same as the probability that the *proband* belongs to (MZ) and their *co-twin* belongs to (MS, MZ) . MZ) and their *co-twin* belongs to (
(or event) for the *proband*, wherear
The analogous subsets (or events)
defined similarly (*Table-1*).
life-time probability that a random
in belongs to (MZ_{MS}) – a probabili
ce-ra *MS*, *MZ*).
as (MZ_{MS})
for *DZ* "*c*
ly-selected
ty estimate Therefore, for clarity, (*MS, MZ*) indicates this subset (or event) for the *proband*, whereas (MZ_{MS}) indicates the same subset (or event) for their *co-twin*. The analogous subsets (or events) for *DZ "cotwins*" (DZ_{MS}) and non-twin "*co-siblings*" (S_{MS}) are defined similarly (*Table-1*).

proband belongs to (*MS*, *MZ*), given that the -the *proband-wise* (or *case-wise*) *MZ*-twin-concordance-rate [25]. *MS*, *MZ*) indicates this subset (or event) for the *proband*, whereas (
et (or event) for their *co-twin*. The analogous subsets (or events) for
-twin "*co-siblings*" (S_{MS}) are defined similarly (*Table-1*).
P(*MS* \cdot DZ " c .
selected
stimate Consequently, $P(MS \mid MZ_{MS})$ represents the life-time probability that a randomly-selected

The term $P(MS)$ -environments. This adjustment – made by multiplying the *proband-wise MZ*-twin-concordance-rate by DZ_{MS}) and non-twin "*co-siblings*" (
Consequently, $P(MS \mid MZ_{MS})$ represe
l belongs to (MS, MZ) , given that the
nand-wise (or *case-wise*) MZ -twin-cc
The term $P(MS \mid IG_{MS})$ represents then
is, in addition to sharing "i S_{MS}) are defined similarly (*Table-1*).
sents the life-time probability that a ra
eir *co-twin* belongs to (MZ_{MS}) – a pro
oncordance-rate [25].
his concordance-rate – i.e., $P(MS \mid MZ)$
genotypes (*IG*), also share int the $(S: DZ)$ concordance- $MS \mid MZ_{MS}$) represents the life-time probability that a randomly-selected MZ), given that their *co-twin* belongs to (MZ_{MS}) – a probability estimated $-wise)$ MZ -twin-concordance-rate [25].
 lG_{MS}) represents this conc -concordance-rates (*see Supplemental-Material-Section:2a*). *MS*, *MZ*), given that their *co-twin* belongs to (
case-wise) *MZ*-twin-concordance-rate [25].
MS $|IG_{MS})$ represents this concordance-rate –
1 to sharing "identical" genotypes (*IG*), also sh
djustment – made by mul MZ_{MS}) – a probability estimated by
i.e., $P(MS \mid MZ_{MS})$ – adjusted becau
are intrauterine and, probably, other
wise MZ-twin-concordance-rate by
n to the observed MZ-twin $MS \mid IG_{MS}$) represents this concordance-rate – i.e., $P($ n to sharing "identical" genotypes (*IG*), also share intindigustment – made by multiplying the *proband-wise* M ance-ratio [4] – isolates the genetic-contribution $MS + MZ_{MS}$) – adjusted because
rauterine and, probably, other
 dZ -twin-concordance-rate by
e observed MZ -twin
because twin *MZ*-twins, in addition to sharing "identical" genotypes (*IG*), also share intrauterine and, probably, other S: DZ) concordance-ratio [4] – isolates the genetic-contribution to the observed *MZ*-twin ordance-rates (*see Supplemental-Material-Section:2a*).

It is made available under a [CC-BY-NC 4.0 International license](http://creativecommons.org/licenses/by-nc/4.0/) .

4. **Estimating P**(
If the population
 H) and (*G*) are also

(*G*) are also

(*H*) such that:

If this ratio is (*H*) such this ratio is (*H*) $n($
 $id\epsilon$
 $P($
 $1)$
 i ht If the population (Z) and the (G)-subset are identical, then, during any (E_T) , the MS-penetrance of (Z) and (G) are also identical. Consequently, the ratio of these two *penetrance-values* [4] estimates

$$
P(G) = P(MS \mid E_T) / P(MS \mid G, E_T) \qquad \qquad \text{Equation-1}
$$

-indicating that only *some* members of *Z*) and the (

entical. Con
 $(G) = P(MS)$, then *everyc*

e *MS-penetr*

members of G)-subset are identical, then, during *any* (
sequently, the ratio of these two *penetrano*
 $S | E_T \rangle / P(MS | G, E_T)$
one in the population can develop MS und
ance of (G) exceeds that of (Z), then this
(Z) have *any* possibi e
ce-
er
rat
5, r), the *MS-penetrance*
values [4] estimates
Equation-1
<u>some</u> environmental-
io is less than (1),
egardless of *any* -(*z*) and (
 i) such the H

If the ditions. I

cating the osure the *G*) are also identical. Consequently, the ratio of these two *penetrance-values* [4] estimates
hat:
 $P(G) = P(MS \mid E_T)/P(MS \mid G, E_T)$ *Equation-1*
nis ratio is (1), then *everyone* in the population can develop MS under <u>some</u> en reaches 100% under *any* realistic conditions, if G) such that:
If this r
nditions. How
licating that o
posure they ei
ches 100% un then *eve*

MS-pen

nembers

and or count

alistic contracts *ryone* in
 etrance

of (*Z*) h

uld have

mditions

mstance the position of $(G$
ave a
b had.
b, if $(G$
wher the population
f (G) exceeds
we any possibi
had. Even if th
if (Z) and (G)
where: $[p = P($ *Equation-1*
can develop MS under <u>some</u> environmenta
that of (Z), then this ratio is less than (1),
lity of developing MS, regardless of *any*
e "*exposure-probability*", $P(E | G, E_T)$, nev
are the same, then this ratio is If this ratio is (1) , then *everyone* in the population can develop MS under *some* environmental-1), then *everyone* in the population can develop MS under <u>some</u> environmental-
the *MS-penetrance* of (*G*) exceeds that of (*Z*), then this ratio is less than (1),
ie members of (*Z*) have *any* possibility of develo conditions. However, if the *MS-penetrance* of (G) exceeds that of (Z) , then this ratio is less than (1) , G) exceeds that of (
 any possibility of de

d. Even if the "*expos*

(Z) and (G) are the

ere: $[p = P(F \mid G) \neq$ Z), then this ratio is less than (

eveloping MS, regardless of an

sure-probability", $P(E \mid G, E_T)$,

same, then this ratio is (1) dur
 $P(F)$, it must be that: ($P(G)$) 1),
 y , nev
ing
 $) <$ Z) have *any* possibility of developing MS, regardless of *any*
have had. Even if the "*exposure-probability*", $P(E | G, E_T)$, i
tions, if (Z) and (G) are the same, then this ratio is (1) durin
nnce where: $[p = P(F | G) \neq P(F)]$, i exposure they either have had or could have had. Even if the "*exposure-probability*", $P(E \mid G, E_T)$, never is (1) du
hat: $(P($
hat: $(P($
 $cal-$ Mode
 $[3, 4, 23]$), never
 $\text{array}\n\text{array}\n\text{if } \text{array}\n\text{if } \text{array}\n\text{if } \text{limits}\n\text{if } \text{limits}\n$ *every* (E_T) . Moreover, in *any* circumstance where: $[p = P(F | G) \neq P(F)]$, it must be that: $(P(G) < 1)$.

5. Data-Analysis

-*Cross-sectional-Models use* data from the "*current*" 1-data regarding changes in MS-epidemiology, which have occurred over the last century [3,4,23] – *see* -*also Supplemental-Material; Figure-S1*. *Cross-sectional-Models* make the two common assumptions that: Z) and (
re: $[p =$
ee "*curre*
ich have
ectional G) are the same, then this ratio is (
 $P(F \mid G) \neq P(F)$, it must be that:
 $int''(E_T) - Table-2$. *Longitudinal*-

occurred over the last century [3,4]
 Models make the two common ass 1) during
 $(P(G) <$

Models us

.23] – see

umptions 1) *MZ-twinning* is independent of genotype and: 2) *MS-penetrance* is independent of (*MZ*)-subs $\begin{array}{c} \nD \ C \n\text{gar} \n\end{array}$). Moreover, in *any* circumstance where: $[p = P($
 ata-Analysis
 ross-sectional-Models use data from the "*current*"

ding changes in MS-epidemiology, which have oo
 lemental-Material; Figure-S1. Cross-sectional-M
 $r(E_T) - Ta$
ccurred over
odels make
etrance is in $2D$, it must be that: χ
ble-2. *Longitudinal-M*
the last century [3,4,2
the two common assure els use $- see$
tions that et -(solutions) that fall within the "*plausible*-*value-ranges*" for each *parameter*. For each *Model*, r
|
|els
| can
| al-
| ve) – *Table-2. Longitudinal-Models* use
ed over the last century [3,4,23] – *see*
make the two common assumptions th
ce is independent of (*MZ*)-subset
Models make neither assumption.
d" and "non-observed" *parameter-v* MZ)-subset
assumption.
" *parameter-
combination*
Model, (~10
), is (5.7)-fo membership (*Supplemental-Material-Section:4a*). *Longitudinal-Models* make neither assumption. Initially, "*plausible*-*value-ranges*" are defined for both "observed" and "non-observed" *parameter-values* (*Table-2*). Subsequently, a "substitution-analysis" determined those *parameter-value*-*combinations* possible *parameter-value-combinations* were systematically-interrogated.

-greater than *MS-penetrance* for comparable *male-probands* (*Table-2*). Moreover, *currently*, both the \sim 10¹¹)
')-fold
oth the
ly, the
al-Mater Currently, MS-penetrance for female-probands, whose *co-twin* belongs to (MZ_{MS}) , is (5.7) -fold $(F:M)$ -*sex-ratio* and $P(MS)$ are increasing [2-4,23]. Under such circumstances, almost certainly, the *MS*) are increasing [2-4,23]. Under such circumstances, almost certainly, the
in *susceptible-women* exceeds that in *susceptible-men* (*see Supplemental-Mate*
fore, it is assumed that, *currently*:
 $Zw = P(MS \mid F, G) > P(MS \mid M,$ *current MS-penetrance* in *susceptible-women* exceeds that in *susceptible-men* (*see Supplemental-Material-Sections:3a&7g*). Therefore, it is assumed that, *currently*:

e regardin
| *Models*
| g the circumstar
:
: $\frac{4}{5}$ and $\frac{4}{5}$ for $\frac{4}{5}$ f No assumptions are made regarding the circumstances of other *Time-Periods*

1-*6. Cross-sectional Models:*

It is made available under a [CC-BY-NC 4.0 International license](http://creativecommons.org/licenses/by-nc/4.0/) .

For notational simplicity, *parameter-abbreviations* are used: *MS-penetrance* for the *i th* susceptible individual is: { $x_i = P(MS | G_i)$ }; the set (X) consists of *MS-penetrance* values for all susceptible-
individuals – i.e., (X) = ($x_1, x_2, ..., x_m$); the variance of (X) is: σ_X^2 ; *MS-penetrance* for the (G)
is: $x = P(MS | G)$; th individuals – i.e., $(X) = (x_1, x_2, ..., x_m)$; the variance of (X) is: σ_X^2 ; *MS-penetrance* for the G)-subset
mental-
netranceis: $x = P(MS | G)$; the adjusted MZ-twin concordance-rate is: $x' = P(MS | IG_{MS})$.

 ν ; the a
 ν (E_T),
 a , durin

tly, dur

4). $x_1, x_2, ..., x_m$); the variance of (
djusted *MZ*-twin concordance-rat
the *MS-penetrance* for (*Z*) is *P*(*l*
ng *any* (*E_T*), the *MS-penetrance* f
 $x = (x'/2) \pm \sqrt{(x'/2)^2 - 1}$
ing *any* (*E_T*), the probability *P*(*G*) X) is: σ_X^2

ie is: $x' =$

MS). As

is: σ_Y^2

is: estimation Dun
erial-Se
Cor
es (Met $MS | G)$; the adjusted *MZ*-twin concordance-rate is: $x' = P(G)$
ring *any* (E_T) , the *MS-penetrance* for (Z) is $P(MS)$. As den
x = $(x'/2) \pm \sqrt{(x'/2)^2 - \sigma_X^2}$
asequently, during *any* (E_T) , the probability $P(G)$ is estimate $MS \mid IG_{MS}$).
nonstrated in
by the ratio During *any* (E_T) , the *MS-penetrance* for (Z) is $P(MS)$. As demonstrated in the *Supplemental*-), the *MS-penetrance* for (

ring *any* (*E_T*), the *MS-pen*
 $x = (x'/2) \pm \sqrt{(x)}$

uring *any* (*E_T*), the probabi
 Models: Z) is $P($
etrance
'/2)² –
lity $P(G)$ *MS*). As demonstrated in the *Supplemental*-
for (*G*) is:
 $\frac{2}{\sigma_X^2}$
) is estimated by the ratio of these *penetranc Material*-*Section:4a*, during *any*

$$
x = (x'/2) \pm \sqrt{(x'/2)^2 - \sigma_X^2}
$$

-*values* (*Methods #4*). $\frac{1}{4}$
 $\frac{1}{4}$
 $\frac{1}{4}$ $\frac{1}{2}$ = (E)), the *MS-penetrance* for (
 $(x'/2) \pm \sqrt{(x'/2)^2 - \sigma_X^2}$

r), the probability $P(G)$ is expressed.

General Consideration *G*) is:
stimat
ons Consequently, during any (E_T) , the probability $P(G)$ is estimated by the ratio of these *penetrance*-

-*7. Longitudinal Models:*

General Considerations

event (E) occurs for a randomly-selected member of the *G*) is estimated by the ratio of these *penetrance*-

<u>lerations</u>

he exposure (*u*) is defined as the *odds* that the

(*G*)-subset during *any Time-Period* (*see* Using standard survival-analysis methods [26], the exposure (*u*) is defined as the *odds* that the E) occurs for a randomly-selected member of the (
mental-Material-Sections:5a-c). Hazard-functions is
idard manner [26] and, if these unknown hazard-fu:
 $(R > 0)$ is defined such that:
 $k(u) = R * h(u)$
The exposure-level ($u = a$) *G*)-subset during *any Time-Period* (*see n men, h*(*u*), and *women, k*(*u*), are defineding are proportional, a proportionali
notions are proportional, a proportionali
).
riod, is then converted into "*cumulative* ite *Supplemental-Material-Sections: 5a-c*). Hazard-functions in *men*, $h(u)$, and *women*, $k(u)$, are defined in u), are defined in
oportionality-
'cumulative-
ied hazardthe standard manner [26] and, if these unknown hazard-functions are proportional, a proportionalityfactor $(R > 0)$ is defined such that:

$$
k(u) = R * h(u).
$$

functions from an *exposure-level* of: $(u = 0)$ to an *exposure-level* of *u*), and *women*, *k*(
e proportional, a pr
en converted into '
ls of these unspecif
f: $(u = a)$. $R > 0$) is defined such that:

The *exposure-level* $(u = a)$
 functions", $H(a)$ and $K(a)$,

ns from an *exposure-level* of:
 {NB: Cumulative-hazard me
 $\frac{dK}{dt}$ *S MS The manning of* $(u = a)$ some Time-
spresent defi
(a) to an expo
exposure, no
(a) or $K(a)$,
ivalent. Hov u).
Pert
inite
sure
ot fa
if p The *exposure-level* $(u = a)$, during *some Time-Period*, is then converted into "*cumulativehazard-functions*", $H(a)$ and $K(a)$, which represent definite integrals of these unspecified hazard-

-*develops MS. The mapping of* Therefore, both exposure-measures are equivalent. However, the failure-probabilities (Zw and Zm) are exponentially related to cumulative-hazard and, therefore, are mathematically-tractable, despite the $u = a$), during *some Time-Period*, is then converted into "*cumulative-*
 $K(a)$, which represent definite integrals of these unspecified hazard-
 evel of: $(u = 0)$ to an *exposure-level* of: $(u = a)$.
 zard measures expos underlying hazard-functions being unspecified. Moreover, <u>any</u> two points on <u>any</u> exponential curve *a*) and *K*
(*osure-leve.
tive-hazare*
pping of (1
ure-measu
to cumulat
ctions bei *a*), which represent definite integrals of these unspecified hazard-
 l of: $(u = 0)$ to an *exposure-level* of: $(u = a)$.
 l measures exposure, not failure. Failure is the event that the prob
 $u = a$) to $H(a)$ or $K(a)$, u = 0) to an *exposure-level* of: (

sures exposure, not failure. Failu

) to H(a) or K(a), if proportiona

re equivalent. However, the failu

uzard and, therefore, are mathem

specified. Moreover, <u>any</u> two poin $u = a$).
 ire is the
 il, is "*on*
 ire-proba
 atically-
 its on <u>an</u> {NB: Cumulative-hazard measures exposure, not failure. Failure is the event that the proband u = a) to *H*

ures are equiv

tive-hazard a

ing unspecifie

r MS, the fail

e Supplemen a) or *K*(a), if proportional, is "one-to-one and onto" [4].
 alent. However, the failure-probabilities (Zw and Zm) are
 md, therefore, are mathematically-tractable, despite the
 d. Moreover, <u>any</u> two points on <u>any</u> defines the entire curve.}

-*probability* approaches unity (*see Supplemental-Material*-*Sections:5b-e*). Moreover, the limiting-value for Unlike true-survival, for MS, the *failure-probability* may not approach 100% as the *exposure*this *failure-probability* in *susceptible-men* (c) and *susceptible-women* (d) may not be the same. Also, *c*) and *susceptible-women* (
e *parameter-values* of (Zw) *d*) may not be the same. Also,
, (Zm) , $P(MS)$, and the $(F:M)$ *c*) and (*d*) are constants, estimated from the *parameter-values* of (ZW), (Zm), $P($ *MS*), and the $(F:M)$ -

It is made available under a [CC-BY-NC 4.0 International license](http://creativecommons.org/licenses/by-nc/4.0/) .

sex-ratio "observed" during *any* two *Time-Periods*.

The *exposure-level* at which MS becomes possible (i.e., the threshold) must be zero for *susceptible-women* or *susceptible-men* or both. The difference between the threshold in *women* (λ_w) and that in *men* (λ_m) is defined as:
 $(\lambda = \lambda_w - \lambda_m)$.

And, therefore: if: $(\lambda_w > \lambda_m)$; then (λ) is positive and

$$
(\lambda = \lambda_w - \lambda_m).
$$

if:
$$
(\lambda_w < \lambda_m)
$$
; then (λ) is negative and $(\lambda_w = 0)$

if:
$$
(\lambda_w = \lambda_m)
$$
; then: $(\lambda = \lambda_w = \lambda_m = 0)$

that in *men* (λ_m) is defined as:

And, therefore: if: $(\lambda_w$

if: $(\lambda_w$

if: $(\lambda_w$

As demonstrated in the
 $[H(a) \ge \lambda)]$, the *cumulative-has* And, therefore: if: $(\lambda_w > \lambda_m)$; then (λ) is positive and $(\lambda_m = 0)$
if: $(\lambda_w < \lambda_m)$; then (λ) is negative and $(\lambda_w = 0)$
if: $(\lambda_w = \lambda_m)$; then: $(\lambda = \lambda_w = \lambda_m = 0)$
As demonstrated in the *Supplemental-Material-Section:7a*, i And, therefore: if: $(\lambda_w > \lambda_m)$; then (
if: $(\lambda_w < \lambda_m)$; then (
if: $(\lambda_w = \lambda_m)$; then:
As demonstrated in the *Supplemental-i*
[$H(a) \ge \lambda$], the *cumulative-hazards* in *men* and
 $K(a)$
Moreover, any *causal-chain* leading to $λ$) is positive and $(λ_m = 0)$
 $λ$) is negative and $(λ_w = 0)$
 $(λ = λ_w = λ_m = 0)$
 Material-Section:7a, if haza

and *women* are related such there are related such there are related such there are re As demonstrated in the *Supplemental-Material*-*Section:7a*, if hazards are proportional and if:

$$
K(a) = R * [H(a) - \lambda]
$$

if: $(\lambda_w < \lambda_m)$; then (λ) is negative and $(\lambda_w = 0)$
if: $(\lambda_w = \lambda_m)$; then: $(\lambda = \lambda_w = \lambda_m = 0)$
l in the *Supplemental-Material-Section:7a*, if hazar
utive-hazards in *men* and *women* are related such th
 $K(a) = R * [H(a) - \lambda]$
 -*sufficient* to *cause* MS in them, and if, in this circumstance, this person's probability of developing $\lambda_w = \lambda_m$); then: $(\lambda = \lambda_w = \lambda_m = 0)$
the *Supplemental-Material-Section:7a*,
-hazards in *men* and *women* are related
 $K(a) = R * [H(a) - \lambda]$
al-chain leading to disease can *only* incl
ny interactions). Therefore, if *any* Consequently, if randomness plays *no role* in MS-pathogenesis, then: $K(a) \ge \lambda$)], the *cumulative-hazards* in *men* and *women* are related such that:
 $K(a) = R * [H(a) - \lambda]$

Moreover, any *causal-chain* leading to disease can *only* include geneti

nts, or both (including *any* interactions). $\frac{1}{2}$ is disease can Therefore, if
is circumstance part, must be in MS-pathog *any* disparity any many in any mer

e, this p

due to a

enesis, thetween Moreover, any *causal-chain* leading to disease can *only* include genetic-factors, environmentalevents, or both (including *any* interactions). Therefore, if *any* member of G) experiences an exposure
robability of developing (*M*.
andom mechanism.
 $= d = 1$) – *see Discussion*.
and *men* in their likelihood o
een (*c*) and (*d*), between
ient-exposure, or between be is less than (100%) ; then their outcome, in part, must be due to a "truly" random mechanism.

M-Also, regardless of proportionality, *any* disparity between *women* and *men* in their likelihood of $\frac{1}{2}$ of both -100%); then their outcome, in part, must be due to a "*truly*" random mechanism.
y, if randomness plays *no role* in MS-pathogenesis, then: $(c = d = 1) - see Disc$, regardless of proportionality, *any* disparity between *women* and $c = d = 1$) – *see Discussion.*
 en and *men* in their likelihood

etween (*c*) and (*d*), between
 fficient-exposure, or between \leq 1), also assumes that any

exclusively, to a difference in t developing MS, during *any Time-Period*, must be due to a difference between (c) and (d) , between c) and (
exposure
so assun
ely, to a d), between

i, or between

ines that any

difference in *susceptible*-*men* and *women* in the likelihood of their experiencing a *sufficient-exposure*, or between both (*Supplemental-Material-Section:5d*). Therefore, assuming that: $c = d \le 1$), also assumes that any
due, exclusively, to a difference in
alausible-parameter-value" ranges s
as take during different *Time-Period* difference between *men* and *women* in their *failure-probability* is due, exclusively, to a difference in their likelihood of experiencing a *sufficient*-*exposure*.

Non-proportional Hazard

If hazards in *women* and *men* are not proportional, the "*plausible*-*parameter-value"* ranges still limit possible solutions. However, any difference that these *values* take during different *Time-Pe*r*iods* could be attributed, both potentially and plausibly, to the different environmental-circumstances of different times and places.

Proportional Hazard

It is made available under a [CC-BY-NC 4.0 International license](http://creativecommons.org/licenses/by-nc/4.0/) .

The "*apparent*" value of (R) , or (R^{app}) , is defined as *the* value of (R) when: $(c = d \le 1)$. As demonstrated in *Supplemental-Material*-*Sections:7c&7g*, for proportional-hazards with proportionalityfactor (R) , three conditions must hold:

- 1) if: $R \le 1$; or, if: $R < R^{app}$; or, if: $\lambda \le 0$; then: $c < d$
- 2) if: $c = d \le 1$; then, both: $R > 1$ and: $\lambda > 0$
- 3) if: $R > 1$; then: $\lambda > 0$

Condition #1 excludes *any* possibility that: $(c = d \le 1)$ – *see Figures-1&2 and Results.*

-must come a point where only *susceptible*-*men* can develop MS. This implies that, at (or below) this *sufficient-R*), three conditions must hold:

1) if: $R \le 1$; or,

2) if: $c = d \le 1$

3) if: $R > 1$; the

Condition #1 excludes any possi

Conditions #2&3 (where: $\lambda > 0$)

me a point where only *susceptible Conditions #2&3* (where: $\lambda > 0$), require that, as the *odds* of a *sufficient-exposure* decreases, there *exposure-level*, $(R = 0)$. Consequently, the additional requirement that: $(R > 1)$ poses a potential paradox.

There are two obvious ways to avoid this paradox (*see Supplemental-Material-Sections:7d-h*). First, if the hazards are non-proportional, although this creates other problems. For example, *women* and *men* in the same *exposure-group*, necessarily, have proportional-hazards (*Supplemental-Material-Section:7h*). Therefore, if *women* and *men* are never in the same *exposure-group*, each sex *must* develop MS in response to distinct $\{E_i\}$ families, in which case *female-MS* and *male-MS* would represent different diseases.

Second, *Condition #1* is compatible with any (λ) so that, if: $(\lambda > 0)$ and $(R \le 1)$, then, at every $sufficient-exposure-level (u = a)$, the probability that a *susceptible-man*, randomly-selected, will experience a *sufficient-exposure* is as great, or greater, than this probability for a *susceptible*-*woman*.

Results

1. Cross-sectional Models

For all *Cross-sectional* analyses [4], the supported-range for $P(G)$ is:

From *Equation-1*, assuming: $(x \ge x)$

2. Longitudinal Models

 G) is:

for $P($ *I*, assuming:
0.003 $\leq P($
Iodels
roportional o $|G| < 0.83$
 $\therefore (x \ge x'/2)$
 $|G| < 0.55$
 $\therefore (x \ge 0.52)$ $x \geq x'$
 > 0.5

propor
 ≥ 0.5
 $(x - d)$ /2); the supported-range for $P($
i5
tional-hazards and, if proportion
i2
1-1), the supported ranges for the i is:
aal, ang
as three *dodels*
roportional
0.001 < *P*(
-hazards and
en and *susce* $G) < 0.55$
or proportic $G) \leq 0.52$
d: $(c = d =$ For either non-proportional or proportional-hazards and, if proportional, any (R) , the supportedrange for $P(G)$ is:

$$
0.001 < P(G) \leq 0.52
$$

For proportional-hazards and: $(c = d = 1)$, the supported ranges for the threshold-difference *R*), the supported-
Id-difference
 R^{app}); and the *G*) is:
portio
ible-wo
i for re hazards and
 en and *susce*

ving a *suffic* G) \leq 0.52
 \pm : $(c = d =$

ptible-men (

ient-exposure between *susceptible-women* and *susceptible-men* (λ); for the proportionality-factor ($R =$
probability-ratio for receiving a *sufficient*-exposure – i.e., $P(E \mid F, G)/P(E \mid M, G)$ – are
probability-ratio for receiving a between *susceptible-women* and *susceptible-men* (λ); for the proportionality-factor ($R = R^{app}$); and the *probability-ratio* for receiving a *sufficient*-exposure – i.e., $P(E \mid F, G) / P(E \mid M, G)$ – are:
 \therefore *n* = are:

It is made available under a [CC-BY-NC 4.0 International license](http://creativecommons.org/licenses/by-nc/4.0/) .

 $0.0005 \le \lambda \le 0.13$
 $1.3 \le R = R^{app} \le 1$
 $1.2 \le P(E \mid F, G)/P$

-hazards and both (*R*

veloping MS in *susce*
 $0.002 < \lambda < 2.4$
 $0.002 \le c \le 0.786$

 $1.3 \le R = R^{app} \le 1177$
 $1.2 \le P(E \mid F, G)/P(E \mid M)$

-hazards and both $(R = 1)$

veloping MS in *susceptible*
 $0.002 < \lambda < 2.4$
 $0.002 \le c \le 0.786$ -hazards

veloping $\geq 0.002 <$

0.002 \leq

1.00001 and both (*R*
MS in *suscep*.
 $\lambda < 2.4$
 $c \le 0.786$
sions from th $E \mid M, G$) ≤ 32
= 1) & (d = 1
tible-men (c) are
tible-men (c) are
is analysis are: For proportional-hazards and both $(R = 1) \& (d = 1)$, the supported-ranges for (λ) and for the $R = 1$) & (
eptible-men
b
this analysis
tion is susce $d = 1$), the supported-ranges for (λ) and for the

(*c*) are:

s are: 1) the *penetrance* of (G) is greater than the

entihle to developing MS and: 2) at *maximum* limiting probability of developing MS in *susceptible-men* (c) are:

$$
0.002 < \lambda < 2.4
$$
\n
$$
0.002 \leq c \leq 0.786
$$

Discussion

of (Z) and, thus, not everyone in the popular $0.002 < \lambda < 2.4$
 $0.002 \le c \le 0.78$

al conclusions from

ting probability for

These two conclu $0.002 \leq c \leq 0.786$
al conclusions from t
ryone in the populati
ting probability for d
These two conclusio The two principal conclusions from this analysis are: 1) the *penetrance* of (G) is greater than that G) is greater than that
1: 2) at *maximum*
is less than that for
ported *MZ*-twin-(*z*) and, thus, *not everyone* in the population is susceptible to developing MS and: 2) at *maximum osure-levels*, the limiting probability for developing MS in *susceptible-men* (*c*) is less than that for *ceptibleexposure-levels*, the limiting probability for developing MS in *susceptible*-*men susceptible-women* (*d*). These two conclusions, stated explicitly, are:

$$
1. \ P(G) \leq 0.52
$$

and:
$$
2 \cdot c < d \le 1
$$

-*Conclusion #1* seems inescapable (*see Results*). Indeed, given *any* of the reported *MZ-*twin--concordance-rates, the notion that *MS-penetrance* for 4; Reference: [3]). Therefore, a large proportion of the population (Z) must be in c) is less than that for
reported *MZ*-twin-
(*Z*) is untenable (*Tabi*
mpervious to developin -MS, regardless of *any* environmental-events they either have experienced or could have experienced. *d*). These two conclusions, stated explicitly, are:

1. $P(G) \le 0.52$

2. $c < d \le 1$

#*I* seems inescapable (*see Results*). Indeed, given

ne notion that *MS-penetrance* for (*G*) is the same

herefore, a large proportio G) \leq 0.52
 \leq **d** \leq 1

nescapable

hat *MS-pen*

large proportional-ever
 $\#2$, however Example 1.1 and MS -
fore, a large provided a large provided as
well as l and l a

-*chromosome* risk-variant is present in both sexes [6] – *see also Supplemental-Material*-*Section:7f*. Moreover, G) is the same as that for (
population (Z) must be improved in the set of could
exist where: $(c = d \le 1)$ r
onal-hazards, which requir
ven the genetic and environ
3 MS-associated loci are *au* Z) is untenable (*Table*-
pervious to developing
1 have experienced.
might be possible.
es *female-MS* and *male*
nmental evidence, this
tosomal, and the single X -the MS-association with different *HLA-*haplotypes is the same for both sexes – *see Tables-3&4;* Z) must be impervious to developing

ienced or could have experienced.
 $(c = d \le 1)$ might be possible.

s, which requires *female-MS* and *male*

etic and environmental evidence, this

atted loci are *autosomal*, and the s Regarding *Conclusion #2*, however, scenarios exist where: $(c = d \le 1)$ might be possible. -5,8,22,27,28]. Thus, both twin and non-twin siblings (*male* or *female*) of an *MS-proband* have increased MS $c = d \le 1$) might be possible.
which requires *female-MS* and
ic and environmental evidence,
ed loci are *autosomal*, and the sir
ental-Material-Section:7f. More
oth sexes – *see Tables-3&4*;
MS in *women* and *men* [2-
 Principal among these is the possibility of non-proportional-hazards, which requires *female-MS* and *male-MS* to be different diseases (*Methods #7*). However, given the genetic and environmental evidence, this too seems untenable. For example, all but one of the 233 MS-associated loci are *autosomal*, and the single *X-Reference:[4]*. Family studies also suggest a common genetic-basis for MS in *women* and *men* [2 risk, regardless of *proband* sex [5,27,28]. Similarly, both *sons* and *daughters* of conjugal couples have markedly increased MS-risk [8,27]. Also, *male* and *female* full- or half-siblings with an MS-*proband* parent (*mother* or *father*) have increased MS-risk [2,8,22,27]. Each of these observations, supports a similar (if not the same) genetic-basis for MS in both sexes.

Also, contrary to those circumstances *required* whenever the proportionality-factor (R) is greater R) is greater
IS-associated
ffect is than (1) – i.e., when $(R > 1)$ – women don't seem more likely than men to experience the MS-associated 1) – i.e., when $($
nmental-events $($ $R > 1$) – *women* don't seem more likely than *men* to experience the MS-associated either known or suspected). Thus, for both sexes, the *month-of-birth* effect is environmental-events (either known or suspected). Thus, for both sexes, the *month-of-birth* effect is

It is made available under a [CC-BY-NC 4.0 International license](http://creativecommons.org/licenses/by-nc/4.0/) .

equally evident $[2,4,9-11]$; the latitude gradient is the same $[2,4,12]$; the impact of intrauterine/perinatal environments is similar (*Supplemental-Material-Section:2c*); *EBV* infection is equally common [2,4,13,14]; vitamin-D levels are the same [2,4,15,16]; and smoking tobacco is actually less common among *women* [2,4]. Collectively, these observations suggest that, *currently*, each sex experiences the relevant environmental-events in an approximately equivalent manner. Taken together, this genetic and environmental evidence implies that *female-MS* and *male-MS* represent the same underlying diseaseprocess and, therefore, that the hazards are proportional (*Methods #7*).

Also, several lines of evidence indicate that the proportional-hazard condition of: $(c = d \le 1)$ is -steeply ascending and present only a very narrow exposure-window to explain the (*F:M*)*-sex-ratio* data $c = d \le 1$) is
 ≈ 1), which is

thenever

ird, the

the tare ways to

whenever:
 $R > 1$) are unlikely. First, the environmental-observations (*described above*) suggest that: $(R^{app} > R \approx 1)$, which is -[23] – *see Figures-2A&2B*. Moreover, following this narrow-window, the (*F:M*)*-sex-ratio* decreases with $R^{app} > R \approx 1$), which is
Figure-1, whenever
ds #7). Third, the
lough there are ways to
s required *whenever*:
 $I \le 1$) & ($R > 1$) are
ne ($F:M$)-sex-ratio data impossible whenever: $(c = d \le 1)$ – see Methods #7 & Results. Second, as in Figure-1, whenever -increasing exposure. By contrast, the Canadian MS-data documents a steadily-progressive rise in the $c = d \le 1$) – *see Methods #7 & Results*. Second, as in *Figure-1*, whenever $R \le 1$), the condition that: $(c < d)$ is established (*Methods #7*). Third, the & $(\lambda > 0)$ creates a potential paradox (*Methods #7*). Although th -(*F:M*)*-sex-ratio* over the past century [4,23] – *see also Supplemental-Material; Figure-S1*. $\lambda \leq 0$) or whenever (
lternative of: $(R > 1)$
ationalize this potentia
 $c < d \leq 1$) are far les
teeply ascending and
23] – *see Figures*-2Ac
ncreasing exposure. B $R \le 1$), the condition that: (
 $\& (\lambda > 0)$ creates a potential

al paradox with: ($c = d \le 1$

ss extreme [4]. Finally, the r

present only a very narrow e
 $\& 2B$. Moreover, following th

y contrast, the Canadian MS $c < d$) is established (*Methods #7*). Third, the
ul paradox (*Methods #7*). Although there are wa
), in every case, the conditions required *whenev*
esponse-curves when: $(c = d \le 1) \& (R > 1)$
xposure-window to explain the (alternative of: $(R > 1)$ & $(\lambda > 0)$ creates a potential paradox (*Methods #7*). Although there are ways to potential pa

e far less ex

ng and pres

res-2A&2B

ssure. By co

o over the p $\lambda > 0$) creates a potential paradox (*Methods #7*). Although there are ways to aradox with: $(c = d \le 1)$, in every case, the conditions required *whenever*:
streme [4]. Finally, the response-curves when: $(c = d \le 1)$ & $(R >$ rationalize this potential paradox with: $(c = d \le 1)$, in every case, the conditions required whenever: $c = d \le 1$), in every case, the conditions required *whenever*:

nally, the response-curves when: $(c = d \le 1) \& (R > 1)$ are

y narrow exposure-window to explain the $(F:M)$ -sex-ratio dat

ollowing this narrow-window, the $(F$ $c < d \le 1$) are far less extreme [4]. Finally, the response-curves when: (
teeply ascending and present only a very narrow exposure-window to exp
23] – *see Figures-2A&2B*. Moreover, following this narrow-window, the there $c = d \le 1$) & (
lain the $(F:M)$ -s.
 $(F:M)$ -sex-ratio
illy-progressive
rial; Figure-S1.
develop MS, ev
lian MS-data [5, $R > 1$) are
ex-ratio dat.
decreases w
rise in the
en when a
8,9,17-23]

-susceptible-genotype co-occurs with a *sufficient-exposure*. Thus, the Canadian MS-data [5,8,9,17-23] -seems to indicate that MS-pathogenesis involves a "*truly*" random element. This cannot be attributed to Nevertheless, whenever $(c < d)$, some *susceptible-men* will *never* develop MS, even when a -This also cannot be attributed to the possibility that *some* individuals can *only* develop MS under $c < d$), some *susceptible-men* will *never* develop MS, even when a
ith a *sufficient-exposure*. Thus, the Canadian MS-data [5,8,9,17-23]
enesis involves a "*truly*" random element. This cannot be attributed to
l-factors other, unidentified, environmental-factors (e.g., other infections, diseases, nutritional deficiencies, toxicexposures, etc.) because each set of environmental-exposures is defined to be *sufficient*, by itself, to *cause* MS in a specific susceptible-individual. If other conditions were necessary for this individual to develop MS, then one (or more) of their *sufficient-exposure sets* would include these conditions (*Methods #2*). improbable conditions. Thus, the estimates for (c) and (d) are based solely upon observable *parameter*c) and (
ibuted tc
omen co
ts in dise
s that the
nisms.
ces whe *d*) are based solely upon observable *parameter*-

b mild or asymptomatic-disease because this

mpared to the *current* (*F*:*M*)-*sex-ratio* in MS

ease-expression requires replication.

e behavior of some complex physi *values* (*Methods #7*). Finally, this cannot be attributed to mild or asymptomatic-disease because this *disease-type* occurs disproportionately among *women* compared to the *current* (*F:M*)-*sex-ratio* in MS [4,23]. Naturally, invoking "*truly*" random-events in disease-expression requires replication. Nevertheless, *any* finding that: $(c < d)$ indicates that the behavior of some complex physical-systems $c < d$) indicates that the behavior of some complex physical-systems
random-mechanisms.
hose circumstances where: $(R = 1)$ & $(d = 1)$ and, also, considering and also, considering (e.g., organisms) involve "*truly*" random-mechanisms.

Moreover, considering those circumstances where: $(R = 1)$ & $(d = 1)$ and, also, considering a $d = 1$) and, also, considering a

It is made available under a CC-BY-NC 4.0 International license.

man, randomly-selected from the (M, G) -subset, who also experiences a *sufficient-environment*, the -random-processes – the only difference is that, for the biased-coin, the two possible outcomes are not *M*, *G*)-subset, who also experiences a *sufficient-environment*, the is $(21 - 99\%)$ – *see Results*. Consequently, both the genetic and the conclusion that: $(R \approx 1)$ – *see above* – also, support the lement" of diseasechance that he *will not* develop MS is $(21 – 99%) - see Results$. Consequently, both the genetic and -equally likely. In the context of MS-pathogenesis, the characteristics of "*female-ness*" and "*male-ness*" 21 – 99%) – *see Results*. Consequently, both the genetic and conclusion that: $(R \approx 1)$ – *see above* – also, support the ent" of disease-pathogenesis, which is primarily responsible fo tly-observed between *susceptible-w* environmental data, which support the conclusion that: $(R \approx 1)$ – see above – also, support the would each simply be envisioned as biasing the coin differently (whatever characteristics are implied by $R \approx 1$) – *see above* – also, support the
thogenesis, which is primarily responsi
veen *susceptible-women* and *men*. Impo
men over *men* does <u>not</u> imply that the p
ed-coin compared to a fair-coin – if bot
ased-coin, conclusion that it is this "random-element" of disease-pathogenesis, which is primarily responsible for the difference in disease-expression currently-observed between *susceptible*-*women* and *men*. Importantly, the fact that a process favors disease-development in *women* over *men* does *not* imply that the process *must be* non-random. For example, when flipping a biased-coin compared to a fair-coin – if both are these two terms).

Other authors, modeling immune-system function, also invoke random-events in MS-diseaseexpression [4]. In these cases, however, randomness is incorporated into their *Models* to reproduce the MS-disease-process more faithfully. However, the fact that including randomness improves a *model's* performance doesn't constitute a *test* of whether "*true*" randomness ever occurs. For example, the outcome of a dice-roll may be most accurately *modeled* by treating this outcome as a random-variable with a well-defined probability-distribution. Nevertheless, the question remains whether this probability distribution represents a *complete* description of the process, or whether this distribution is merely a convenience, compensating for our ignorance regarding the initial orientation of the dice and the direction, location, and magnitude of the forces that act on the dice during the roll [4,29,30].

It is hard to imagine that the outcomes of complex-biological-processes such as evolution and immune-function are pre-determined events, especially considering the fact that both processes are so remarkably adaptive to contemporary external-events [4,30]. Nevertheless, *proving* that any macroscopicprocess is "*truly*" random is difficult. This requires an experiment (i.e., a *test*), in which the outcome predicted by determinism differs from that predicted by non-determinism.

The Canadian MS-data presents an opportunity to apply just such a *test*. Thus, the widely-held deterministic-hypothesis [4,30] *requires* that: $(c = d = 1)$. Any observation that either: $(c < d = 1)$ or: -randomness also seems critical to disease-pathogenesis. Moreover, both sexes seem to have the same $c = d = 1$). *Any* observation that either: (
is a component of disease-development and
S-data [5,8,9,17-23], which strongly implie
on-deterministic hypothesis. Importantly, t
environmental-events (including interactic
whet $c < d = 1$) or:

I undermines the

es that: $(c < d)$,

his analysis

ons), which are

ns are known,

es, "*true*" $(c \le d < 1)$ indicates that "*true*" randomness is a component of disease-development and undermines the $c \le d < 1$) indicates that "*true*" randomness is a component of disease-development and undermines the leterministic-view. Therefore, the Canadian MS-data [5,8,9,17-23], which strongly implies that: $(c < d)$, provides empir deterministic-view. Therefore, the Canadian MS-data [5,8,9,17-23], which strongly implies that: $(c < d)$, $c < d$),
ysis
ch are
nown,
" provides empirical evidence in support of the non-deterministic hypothesis. Importantly, this analysis *explicitly includes* all those genetic-factors and environmental-events (including interactions), which are necessary for MS-pathogenesis*,* regardless of whether these factors, events, and interactions are known, suspected, or as yet unrecognized. Nevertheless, in addition to these necessary prerequisites, "*true*"

It is made available under a [CC-BY-NC 4.0 International license](http://creativecommons.org/licenses/by-nc/4.0/) .

underlying-disease. Thus, both seem to have a similar genetic-basis and, also, a similar response to the same environmental disease-determinants (*see above*). These observations suggest both that the hazards are proportional (*Methods #7*) and that $(R \approx 1)$. If correct, this indicates that it is this "*truly*" random- $R \approx 1$). If correct, this indicates that it is this "*truly*" random-
is primarily responsible for the currently-observed differences
women and *susceptible-men*. element in disease-pathogenesis, which is primarily responsible for the currently-observed differences in disease-expression between *susceptible*-*women* and *susceptible*-*men*.

References

- 1. Compston A, Confavreux C, Lassman H, McDonald I, Miller D, Noseworthy J, Smith K, Wekerly H. McAlpine 's Multiple Sclerosis. Churchill Livingstone, Elsevier. 2006: pp.287-346.
- 2. Goodin DS. The Epidemiology of multiple sclerosis: Insights to disease pathogenesis. In. *Handbook of Clinical Neurol*ogy Aminoff MJ, Boller F, Swaab DF (eds). Elsevier, London. 2014;122:231-266.
- 3. Goodin DS, Khankhanian P, Gourraud PA, Vince N. The nature of genetic and environmental susceptibility to multiple sclerosis. *PLoS One.* 2021;16(3): e0246157.
- 4. Goodin DS, Khankhanian P, Gourraud PA, Vince N. Multiple Sclerosis: Exploring the limits and implications of genetic and environmental susceptibility. *PLoS. One.* 2023;18(6): e0285599.
- 5. Willer CJ, Dyment DA, Rusch NJ, et al. and the Canadian Collaborative Study Group. Twin concordance and sibling recurrence rates in multiple sclerosis. *Proc Natl Acad Sci U S A.* 2003;100:12877–82.
- 6. International Multiple Sclerosis Genetics Consortium. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science.* 2019;365 (6460): doi:10.1126/science.aav7188.
- 7. Goodin DS, Khankhanian P, Gourraud PA, Vince N. Genetic susceptibility to multiple sclerosis: interactions between conserved extended haplotypes of the MHC and other susceptibility regions. *BMC Med Genomics.* 2021;14:183
- 8. Sadovnick AD, Ebers GC, Dyment DA, Risch NJ. The Canadian Collaborative Study Group. Evidence for genetic basis of multiple sclerosis. *Lancet*. 1996;347(9017):1728-1730.
- 9. Willer CJ, Dyment DA, Sadovnick AD, et al. Timing of birth and risk of multiple sclerosis: population based study. *Br Med J.* 2005;330:120-124.
- 10. Staples J, Ponsonby AL, Lim L. Low maternal exposure to ultraviolet radiation in pregnancy, month of birth, and risk of multiple sclerosis in offspring: longitudinal analysis. *Br Med J.* 2010;340:c1640.
- 11. Pantavou KG, Bagos PG. Season of birth and multiple sclerosis: a systematic review and multivariate meta-analysis *J Neurol.* 2020;267:2815-2822.
- 12. Sabel CE, Pearson JF, Mason DF, et al. The latitude gradient for multiple sclerosis prevalence is established in the early life course. *Brain.* 2021;144:2038-2046.
- 13. Kreft KL, VanNierop, GP, Scherbeijn SMJ, et al.Elevated EBNA-1 IgG in MS is associated with genetic MS risk variants. *Neurol Neuroimmunol Neuroinflamm* 2017;4:e406
- 14. Bjornevik K, Cortese M, Healy BC, et al. Longitudinal analysis reveals high prevalence of Epstein-Barr virus associated with multiple sclerosis. *Science.* 2022;375:296-301.
- 15. Munger KL, Levin LI, Hollis BW, et al. Serum 25-hydroxyvitamin D levels and risk of multiple sclerosis. *JAMA*. 2006;296:2832-2838
- 16. Sowah D, Fan X, Dennett L, et al. Vitamin D levels and deficiency with different occupations: a systematic review. *BMC Public Health.* 2017;22;17:519.
- 17. Sadovnick AD, Ebers GC, Dyment DA, Risch NJ, and the Canadian Collaborative Study Group. Evidence for genetic basis of multiple sclerosis. *Lancet.* 1996;347:1728–1730
- 18. Sadovnick AD, Risch NJ, Ebers GC. Canadian collaborative project on genetic susceptibility to MS, phase 2: Rationale and method. Canadian Collaborative Study Group. *Can J Neurol Sci*. 1998;25:216-21.
- 19. Ebers GC, Yee IML, Sadovnick AD, Duquette P, and the Canadian Collaborative Study Group. Conjugal multiple sclerosis: Population based prevalence and recurrence risks in offspring. *Ann Neurol.* 2000;48:927–931.
- 20. Ebers GC, Sadovnick AD, Dyment DA, et al. Parent-of-origin effect in multiple sclerosis: observations in half-siblings. *Lancet.* 2004;363:1773–1774.
- 21. Sadovnick AD, Yee IML, Ebers GC, and the Canadian Collaborative Study Group. Multiple sclerosis and birth order: A longitudinal cohort study. *Lancet Neurol.* 2005;4:611–617.
- 22. Dyment DA, Yee IML, Ebers GC, Sadovnick AD, and the Canadian Collaborative Study Group. Multiple sclerosis in step siblings: Recurrence risk and ascertainment. *J Neurol Neurosurg Psychiatry.* 2006;77:258–259
- 23. Orton SM, Herrera BM, Yee IM, et al., and the Canadian Collaborative Study Group. Sex-ratio of multiple sclerosis in Canada: A longitudinal study. *Lancet Neurol*. 2006 ;5 :932–936.
- 24. Canadian Census. 2010. Available from: https://www150.statcan.gc.ca/n1/en/pub/89-503 $x/2010001/article/11475-eng.pdf?st=WVL9Ggm (last accessed 15 May 2023)$
- 25. Witte JS, Carlin JB, Hopper JL. Likelihood-based approach to estimating twin concordance for dichotomous traits. *Genetic Epidemiol*. 1999;16:290–304.
- 26. Fisher LD, van Belle G. Biostatistics: A Methodology for the Health Sciences, John Wiley & Sons, New York, 1993; pp. 369-373 & 786-829.
- 27. Robertson NP, O'Riordan JI, Chataway J, et al. Offspring recurrence rates and clinical characteristics of conjugal multiple sclerosis. *Lancet*. 1997;349:1587-1590.
- 28. Sadovnick AD, Ebers GC, Dyment DA, Risch NJ. The Canadian Collaborative Study Group. Evidence for genetic basis of multiple sclerosis. *Lancet*. 1996;347:1728-1730.
- 29. Green B. Until the end of time. Alfred A Knopf, Penguin Random House, New York, USA. (2020).
- 30. Layzer D. Why we are free: Consciousness, free will and creativity in a unified scientific worldview. Information Publisher. 2021;ISBN-10 0983580251.

It is made available under a [CC-BY-NC 4.0 International license](http://creativecommons.org/licenses/by-nc/4.0/) .

Figure Legends

Figure 1. Response-curves representing the likelihood of developing MS in genetically-susceptible *women* (black lines) and *men* (red lines) with an increasing probability of a *sufficient* environmental exposure – *see Methods #2 & Supplemental-Material*-*Section:1a*. The curves depicted in *Panels A* and *B* are proportional, with a proportionality-factor (R) , although the environmental threshold is greater for *men* than for *women* – i.e., under conditions in which: $(\lambda < 0)$ – *see Text*. The curves depicted in *Panels C* and *D* are "strictly" proportional, meaning that the environmental threshold is the same for both *men* and *women* – i.e., under conditions in which: $(\lambda = \lambda_w = \lambda_m = 0)$ – *see Text*. The blue lines represent the change in the $(F:M)$ -sex*ratio* (plotted at various scales; indicated in each *Panel*) with increasing exposure. The thin grey vertical lines represent the portion of the response-curve that covers the change in the (*F:M*)*-sex-ratio* from 2.2 to 3.2 (i.e., the actual change observed in Canada [23] between *Time-Periods #1 & #2*). The grey lines are omitted in *Panel C* because the observed (*F:M*)-*sex-ratio* change is not possible under these conditions. In *Panel A*, although the $(F:M)$ -*sex-ratio* change is possible, the condition $(Zw > Zm)$ is never possible throughout the entire response-curve. Response-curves A, B, and D reflect conditions in which $(R < 1)$; whereas curve C reflects conditions in which $(R > 1)$. If $(R = 1)$, the blue line in *Panel C* would be flat (*see Supplemental*-*Material-Sections;7c-f*). Response-curves *A* and *C* reflect conditions in which $(c = d = 1)$; whereas curves *B* and *D* reflect those conditions in which $(c < d = 1)$.

Figure 2. Response-curves for the likelihood of developing MS in genetically *susceptible*-*women* (black lines) and *men* (red lines) with an increasing probability of a *sufficient* environmental exposure – *see Methods #2 & Supplemental-Material*-*Section:1a*. The curves depicted are proportional, with a proportionality-factor , although the environmental threshold is greater for *women* than that it is in *men* – i.e., these are conditions in which: $(\lambda > 0)$. Also, all of these response-curves represent actual solutions. The blue lines represent the change in the (*F:M*)*-sex-ratio* (plotted at various scales; indicated in each *Panel*) with increasing exposure. *Panels A* and *B* are for conditions where: $(c = d = 1)$. The value of (R) , specific for this condition, is termed (R^{app}) . Indeed, for *every* condition in which: $(c = d \le 1)$, both: $(R = R^{app})$ and the response curves for *men* and *women* have the same relationship to each other (*see Supplemental-Material-Sections:7c-f*). By contrast, *Panels C* and *D* represent conditions where: $(c < d \le 1)$ and, in these circumstances: $(R < R^{app})$. To account for the observed increase in the (*F:M*)*-sex-ratio*, the response-curves in Panels *A* and *B require* that the Canadian observations $[23]$ were made within a very narrow window – i.e., for most of these responsecurves, the (*F:M*)*-sex-ratio* is actually decreasing. By contrast, the response-curves in *Panels C* and *D* demonstrate an increasing (*F:M*)*-sex-ratio* for *every* two-point interval of exposure along the entire responsecurves for *women* and *men*. The thin grey vertical lines represent the portion of these response-curves (for the depicted solution), which represents the actual change in the (*F:M*)*-sex-ratio* for specific "solutions" between *Time-Periods #1 & #2*.

It is made available under a [CC-BY-NC 4.0 International license](http://creativecommons.org/licenses/by-nc/4.0/) .

Table 1. Definition of Terms

It is made available under a [CC-BY-NC 4.0 International license](http://creativecommons.org/licenses/by-nc/4.0/) .

Table 2. *Parameter-values* – Point Estimates and Plausible Ranges *

* Estimated values and "*plausible*" ranges for observed and non-observed *parameters* [4] – *Supplemental Material* (*Sections 10a-b*)*.* Because the MS-status of individuals born during *Time-Period #2* (1976–1980), cannot be determined until 25- 35 years later, all parameter estimates – except $P(F \mid MS)_1$ – are for the "*current*" *Time-Period* (2001–2015). Estimates for all observed *parameter-values* – except $P(MS)$ and $P(F)$ – are exclusively from the *CCPGSMS*-dataset [5,8,9,17-23]. The estimate for $P(MS)$ is based upon three measures: 1) the population-prevalence of MS; 2) the age-specific prevalence of MS in the ageband of 45-54 years; and 3) the proportion of death certificates mentioning MS [3]. The parameter $P(F)$ is taken from the 2010 Canadian census [24]. Also, $P(MS)$ has been increasing in many regions around the world – especially among *women* [3,4]. In Canada, based on the point-estimates provided in this *Table*, it has increased by $(≥ 32%)$ between the two *Time-Periods – see Supplemental-Material*-*Section:8a*). If all of the environmental events, relevant to *MS-*pathogenesis, take place prior to the age of 30 years, then, for an individual born in 1975, (E_T) would extend from 1975 to 2005 whereas, for a person born in 1980, (E_T) would extend from 1980 to 2010. If the relevant age-window is different than 30 years, then the definition of (E_T) would change accordingly.

† Ranges represent the 95% confidence intervals [4]. To include a broader range of possible solutions, the range for $P(MS)$ was expanded beyond the range of: $\{0.0025 \le P(MS) \le 0.0046\}$, which was supported by the three *above* methods [3]. The range for $P(F \mid MS)$ was similarly expanded [4], as was the range for the $(S: DZ)$ concordance-ratio, considering the

It is made available under a [CC-BY-NC 4.0 International license](http://creativecommons.org/licenses/by-nc/4.0/) .

theoretical constraint [4] that: $[(S: DZ) \le 1]$. Because $P(F)$ is taken from a census of the entire "*current*" Canadian population at the time (2010), there is no estimated range [24].

†† Ranges represent the "*plausible*" *parameter-value* range for each parameter. For example, because, *currently*, both *men* and *women* can (and do) develop MS, $P(G)$ cannot be (0) and $P(F|G)$ cannot be either (0) or (1) . Also, the theoretical *upper-limit* for the value of the ratio (C) is 0.9 [2-4]. In addition, a greater than 4-fold increase in the prevalence of (MS) over the last 35 years seems implausible based upon the available world-wide evidence [2-4]; including the evidence for MS in Canada – see *Supplemental Material* (*Sections 8a & 10a-b*); *see also Rosati G, Neurol Sci 2001;22:117-39*.

`

It is made available under a [CC-BY-NC 4.0 International license](http://creativecommons.org/licenses/by-nc/4.0/) .

It is made available under a [CC-BY-NC 4.0 International license](http://creativecommons.org/licenses/by-nc/4.0/) .

