

# Chapter 1

## General Introduction

## 1.1 | MULTIPLE SCLEROSIS

Multiple Sclerosis (MS) is a chronic inflammatory and neurodegenerative disease of the central nervous system that is characterized by an unpredictable clinical course, progression of disability over time, and an onset in young adults between 20 and 40 years old [1]. Worldwide, approximately 2.5 million people are estimated to be affected by MS with a prevalence of 1 case in 1000 in Northern Europe.

In general, MS is considered to be an auto immune disease and both environmental as well as genetic factors are presumed to contribute to its complex, multifactorial etiology. Evidence supporting genetic factors include racial susceptibility, with MS being particularly prevalent in people of Northern European extraction and virtually absent in native black Africans, and the presence of familial aggregation in family studies which showed the risk of MS in relatives of affected individuals to be increased, and to be greater in first-degree than in second-degree relatives [2]. Among environmental factors, there is evidence for infections (e.g. Epstein Barr virus) and chemo-physical factors (sunlight, vitamin D3 status, dietary factors) [3] playing a role, but none have been formally confirmed.

Histopathologically, MS is characterized by temporary focal breakdowns of the blood-brain-barrier causing influx of white blood cells into the surrounding brain parenchyma. The subsequent inflammation in those areas damage the insulating myelin surrounding axons and finally the axons itself, resulting in multiple sclerotic plaques which disrupt the conduction of nerve pulses to and from the brain. Depending on the location of the plaques, this can lead to a variety of neurological symptoms, including visual disturbances, muscle weakness, sensory deficit, movements disorders, brainstem dysfunction, urinary problems, cognitive decline, tremor and fatigue.

While the clinical course of MS is unpredictable, in approximately 80% of patients the disease starts in the relapsing remitting (RR) form, in which clearly defined periods of neurological symptoms (relapse) are alternated with periods of remission where partial to full recovery of the relapse occurs [4]. After a variable number of years, the disease course progresses into the Secondary Progressive (SP) form, where the relapses and remissions are replaced by a pattern of gradual worsening due to the accumulation of irreversible neurological damage [5]. About 15% of the cases, MS presents itself in the so called Primary Progressive form (PP), where patients show progressive neurological deterioration from onset and do not suffer from relapses and remissions [6].

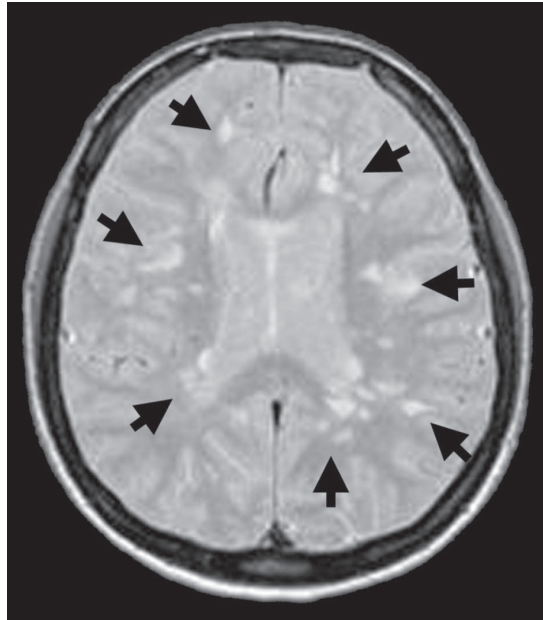
At present, therapeutic options are limited and a curative treatment does not exist for MS. Much research has focussed on the inflammatory components of the disease, and currently approved treatments that alter the immune response of MS (e.g. interferons, glatiramer acetate and natalizumab [7,8,9]), have shown to reduce the number of relapses, delay clinical disease progression, and reduce the accumulation of brain abnormalities on magnetic resonance imaging (MRI). In the past decade however, evidence has emerged that neurodegeneration is another key pathological feature of MS [10], even in the absence of inflammation [11,12]. Therefore, novel therapeutic strategies have increasingly turned their attention toward neuroprotection (slowing degeneration of neural tissue) and neurorepair (restoring neural tissue integrity and function) [13] and various potential candidates have appeared, including sodium channel blockers [14,15], neuroprotective compounds licensed for diseases other than MS such as riluzole (which is currently used for patients with amyotrophic lateral sclerosis) and stem cell transplantations [16]. Unfortunately, the majority of such agents have only shown their neuroprotective potential *in vitro* or in animal models, and developments are slowly progressing into clinical stages.

## 1.2 | MRI IMAGING IN MS

MRI has emerged as an important paraclinical tool to visualize the pathology of MS *in vivo*, and has greatly impacted our thinking about the disease process. MRI exhibits greater sensitivity to disease activity, is more closely associated with histopathology findings compared to clinical outcomes, and highly reproducible [17]. Depending on the MRI sequence applied, various pathological aspects of MS can be visualized, from which those used throughout this thesis will be described below.

### T2 lesions

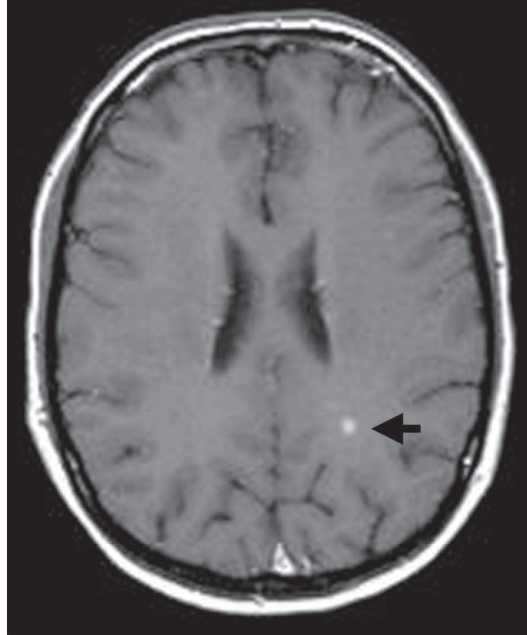
T2 weighted MR images will depict any alteration in the brain as bright, hyperintense spots compared to the surrounding brain tissue, which implies that inflammation, oedema, demyelination, axonal loss, gliosis or even remyelination [18,19] all have the same appearance on this MRI sequence. T2 weighted signal changes therefore, are very sensitive for lesional disease activity of MS, but are pathologically a-specific. This feature of T2 weighted imaging is known to be partly responsible for the poor relationship between the number of T2 lesions and clinical disability, a finding known as the clinoradiological paradox.



**Figure 1** | Appearance of MS lesions on T2 proton density weighted MRI images. Shown by the black arrows are multiple lesions in the deep white matter and surrounding the cortex and ventricles.

### Gadolinium enhancing lesions

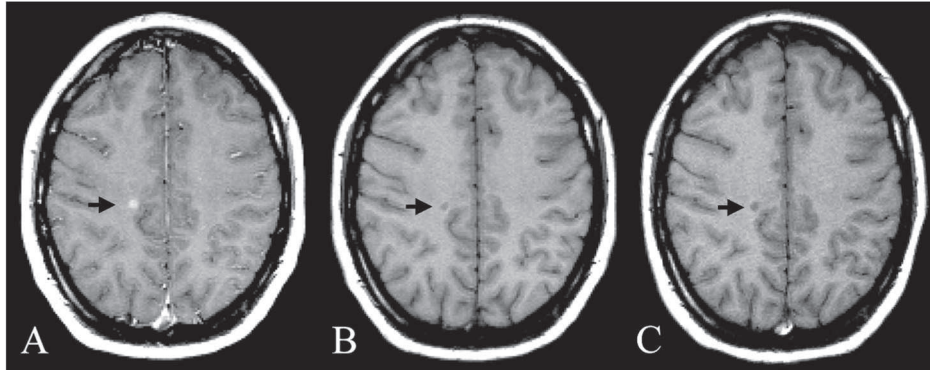
T1 weighted MR images are another way of depicting MS lesions, and are used with and without the administration of intravenously injected contrast. The introduction of the contrast-enhancing agent gadolinium (Gd) has had a profound impact on the clinical and research value of MRI in MS. Disturbances in the blood-brain barrier (BBB) cause leakage of Gd into the surrounding brain tissue which appears as hyperintense abnormalities on T1 weighted images. In MS, opening of the BBB due to the focal inflammatory responses by auto-reactive immune cells, is one of the earliest events in the development of an MS lesion, and Gd enhancing lesions thus represent new active lesions. Compared to the accompanying clinical activity within MS patients, the number and extent of Gd enhancing lesions are much higher [20], resulting in serial monthly enhanced MRI scans to be 10 times more sensitive in detecting MS activity than clinical relapses and disability measures [21]. Gd enhanced MRI therefore, has become a valuable outcome measure in clinical trials monitoring the effect of anti-inflammatory agents.



**Figure 2** | Appearance of newly formed MS lesions on Gd enhanced T1 weighted MRI images. Shown by the black arrow is a clearly visible enhancing lesion located in the deep white matter.

### **Persistent Black Holes**

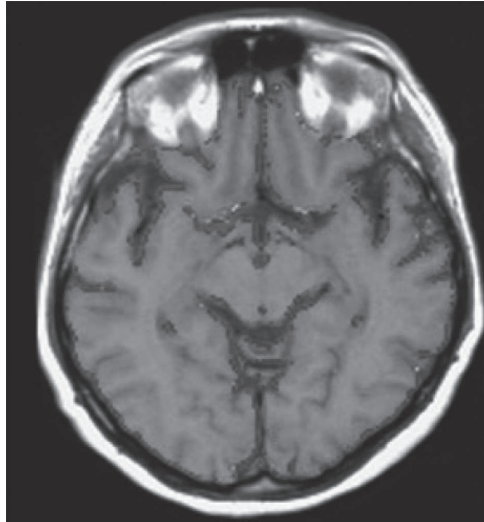
Black Holes (BH) are hypointense (darker than the surrounding brain tissue, therefore called “black”) lesions on T1-weighted images. A majority of the newly formed Gd enhancing lesions appear as hypointense lesions on the corresponding unenhanced T1-weighted image, caused by a decreased signal intensity owing to both increased extracellular water due to active inflammation as well as due to demyelination. Once contrast enhancement subsides, approximately 45% of the initial “acute” hypointense lesions gradually return to isointensity, reflecting both remyelination and the loss of extracellular edema, and around 35% of the acute hypointense lesions will remain hypointense [22]. Histopathological studies have shown that these persistent BHs (PBH) specifically reflect axonal loss and loss of myelin, with the volume of PBHs correlating well with disability [23,24]. The number of PBHs therefore, is a potential outcome measure for permanent tissue destruction in MS, and applicable for monitoring the effect of neuroprotective agents.



**Figure 3** | T1 weighted MR images of the evolution of a contrast enhancing lesion at month 0 (A), being hypointense on the corresponding unenhanced T1 weighted scan (B), evolving into a Persisting T1 hypointense lesion at month 3 (C). Lesions are marked by black arrows [22].

### Cerebral atrophy

Brain tissue loss is a prominent feature in the pathology of MS and occurs at a significantly higher rate in patients with MS compared to the normal aging brain [23,24]. Neuroaxonal loss is considered to be an important driving mechanism, although other pathological and physiological factors are also known to alter brain volume, *e.g.* oedema, inflammation, gliosis, remyelination and demyelination. Not only does neurodegeneration occur focally in MS lesions, but also diffusely in the normal appearing brain tissue where widespread abnormalities (not visualized by MRI) are known to occur. Therefore, cerebral atrophy is recognized as a global marker of the neurodegenerative components of MS, and a conceivable outcome measure for clinical trials measuring the efficacy of neuroprotective agents [27]. MRI detects the rate of cerebral atrophy *in vivo* in a sensitive and reproducible manner. Although there are various techniques available, the automated registration based method SIENA (Structural Image Evaluation Using Normalization of Atrophy) is one of the most promising in terms of sensitivity and study power [28].



**Figure 4** | T1 weighted MR image showing the result of the semi-automated brain volume measurement technique SIENA to determine the percentage brain volume change (PBVC) over a given follow-up period. Coloured areas represent the positive (red/yellow) and negative (blue / light blue) brain edge displacement.

### **Magnetisation transfer imaging**

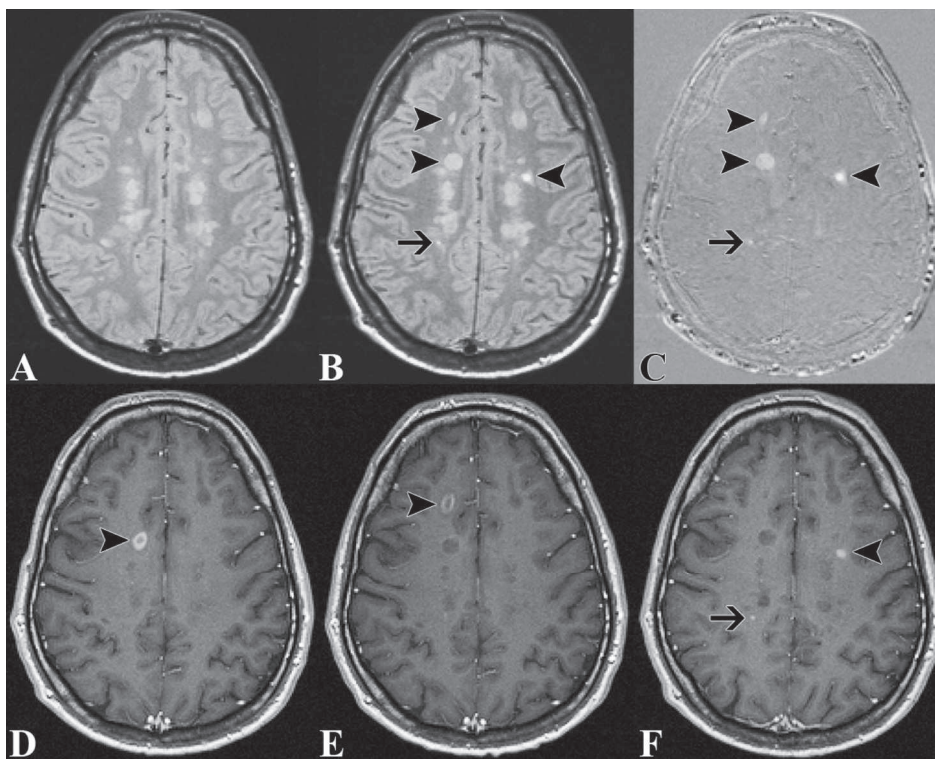
Magnetisation transfer (MT) imaging is a slightly more complex MRI technique based on the exchange of magnetization between protons in a restricted environment and those where motion is relatively free. In brain tissue, protons in a restricted environment are those bound to macromolecules in predominantly myelin, and protons in a free environment are those bound to solvent molecules or free water. In undisrupted white matter, the magnetic transfer ratio (MTR) is high due to the bound protons within myelin, whereas in MS lesions, demyelination causes an increase of the unbound protons and a significant decrease in MTR [29]. Since recent studies have shown that the MTR is mainly driven by the integrity of myelin given the axons remained intact [30] monitoring the MTR of lesions is a promising outcome measure for clinical trials assessing the efficacy of remyelinating agents.

### **T2 subtraction imaging**

T2 subtraction imaging is a post processing method to efficiently determine the residual lesional disease activity of MS. By performing a T2 weighted scan at the start and end of a study, and digitally subtracting the first scan from the last scan after



both scans are aligned (registered), the resulting T2 subtraction scan shows only those lesions which changed and originated during follow-up. Compared to the detection of active (new and enlarged) T2 lesions using serial conventional MR imaging which is complicated by repositioning errors and a background of unaltered non-active lesions, this technique has demonstrated increased sensitivity and higher interobserver agreement in the detection of active T2 lesions [31,32].



**Figure 5** | Examples of Gd-enhancing T1w lesions visible as positive activity on the T2w subtraction image. MS activity on A, halfway registered month one T2w MR image; B, halfway-registered month nine T2w image; C, T2w subtraction image; D, E, F, month three, month five, and month nine non-registered Gd-enhanced T1w images respectively. **Arrowheads** = Gd-enhancing T1w lesions on various timepoints, also visible on the follow-up registered T2w image, and easily identified on the T2w subtraction image. **Arrows** = very small lesion not readily identified on the T2w subtraction image which did correspond to a Gd-enhancing T1w lesion [32].



### 1.3 | STATISTICAL MODELLING OF MRI DATA

A general aim in medical sciences is to make statements about a characteristic in the population being studied, for example a statement about the effect of a novel drug on the number of enhancing lesions in patients with MS. Ideally, one would conduct the study in a very large population of MS patients, so there is little uncertainty about the authenticity of the detected effect one would find. In practice however, research is commonly conducted on a selected sample of MS patients acting as a representative of a larger population of MS patients one wants to make statements about. When several samples of MS patients are taken from the total MS population, some unknown mechanism is thought to have distributed the individuals or variable of interest among the chosen subgroup(s). One can think of the number of enhancing lesions an MS patient experiences to follow some type of fixed pattern. With a probability distribution this mechanism or pattern is formalized into a mathematical formula as an approximation to reality. The most well known distribution in this regard is the Normal distribution (box 1).

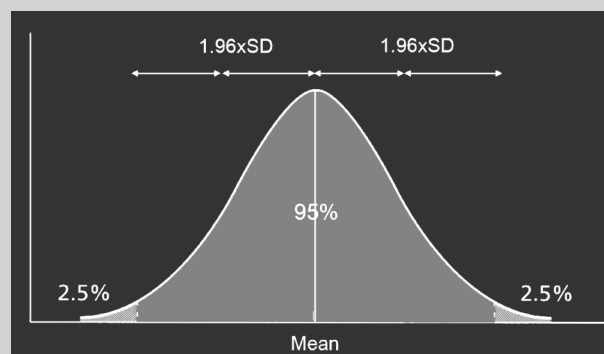
In statistics, probability distributions are used as a tool for dealing with uncertainty. It is important that the applied distribution describes the unknown data generating mechanism in the population as adequately as possible, and for this purpose we use our collected data (e.g. the number of enhancing lesions) within the sample (or more formally: within the population) which gives an impression of the underlying mechanism.

Then, the number of enhancing lesions is assumed to follow the chosen distribution and is, for example, used to generate and simulate data to perform advanced sample-size calculations (see also paragraph 1.4). If an inappropriate distribution is selected, e.g. one that doesn't describe the data generating mechanism and the empirical data well, the subsequent calculations will be incorrect, thereby resulting in erroneous estimations on the required sample size. Therefore, it should be determined whether a probability distribution fits the data and hence the underlying mechanism appropriately. Customary methods are visual assessment of a frequency histogram of the data, as is frequently applied for normally distributed data, obtaining a Q-Q probability plot, or performing so called goodness of fit (GOF) tests. The latter determines the compatibility of a random sample with a statistical distribution by measuring the distance between the data and the intended distribution, and comparing that distance to some threshold value. If the distance is less than the threshold value, the fit is considered good.

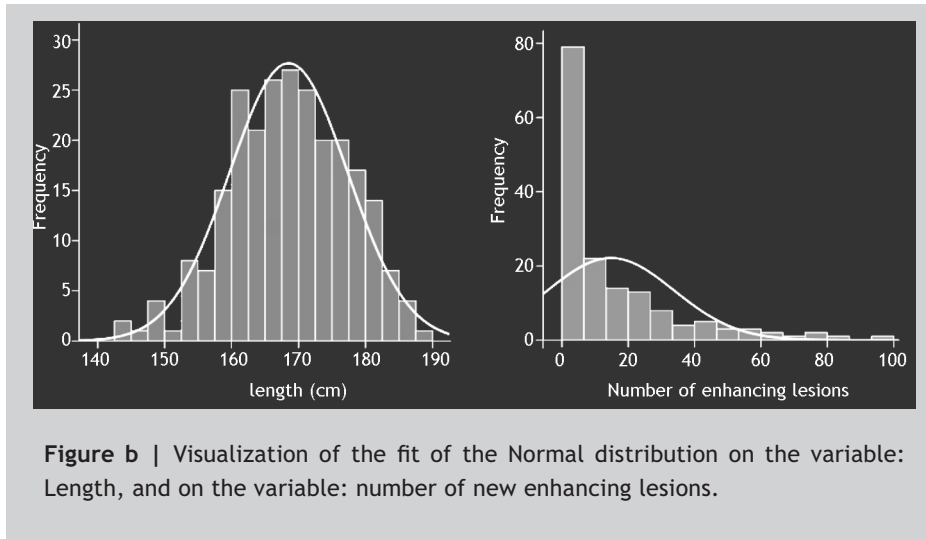
## BOX 1 | A FAMILIAR STATISTICAL DISTRIBUTION: THE NORMAL DISTRIBUTION

The Normal distribution has been developed more than 250 years ago and, because of its convenient mathematical properties and simplicity, is one of the most frequently used distributions known. Most of the parametrical (*e.g.* based on a statistical distribution: in contrast to non-parametrical) techniques used in statistics today assume the collected data to be normally distributed. Its form has the characteristic shape of a bell (figure A).

As all statistical distributions, the normal distribution is described by so-called “parameters”, which are numerical characteristics of a population. The normal distribution consists of two parameters, namely the “mean” and “standard deviation”. The mean defines where the “centre” of the distribution is located, and the standard deviation how wide or how narrow the bell shape is (figure A). If both parameters are specified, the shape of the distribution is known exactly. The Normal distribution lends itself well for describing common variables such as a person’s length (figure B). For other variables, like the number of enhancing lesions within MS patients for example, the Normal distribution is not a good fit since the majority of patients will experience around 0-5 lesions within an average trial whilst only a few experience more, resulting in data with a long right-sided tail (figure B). As will be addressed in chapter 2.1, alternative statistical distributions are needed to describe this outcome measure.



**Figure a** | Visualization of the Normal distribution, characteristically bell shaped, and described by two parameters: the mean, and the standard deviation (SD). By definition, 95% of all the observations are known to lie within two times the SD below and above the mean, leaving the remaining 5% of the observations on either side of the 95%.



#### 1.4 | CLINICAL TRIALS, STATISTICAL POWER AND SAMPLE SIZE

To demonstrate the efficacy of a new experimental drug, the scientific standard of excellence is to perform a randomized “clinical trial”. Compared to an observational study, where investigators simply observe what happens to patients who for various reasons do or do not receive the novel drug, the conditions of a clinical trial (selection of patients, type of treatment, measurement of the outcome) are highly controlled for the purpose of creating two groups of patients without systematic differences, other than the treatment itself, that could lead to misleading conclusions about the treatment effect. In this way, any differences found in the measured outcome between the two groups can be solely ascribed to the effect of the treatment. Ideally, the group receiving the active treatment is compared with a group receiving a placebo, an intervention intended to be indistinguishable from the active treatment whether in physical appearance, color, taste or smell.

When a clinical trial has been performed, the next question is whether there is a genuine difference in the outcome measure between the placebo group and the treatment group. Since the data have been collected in two samples of patients instead of the population as a whole, some uncertainty exists about the true values in the population and in theory, the difference found between the groups may have been caused by chance alone.

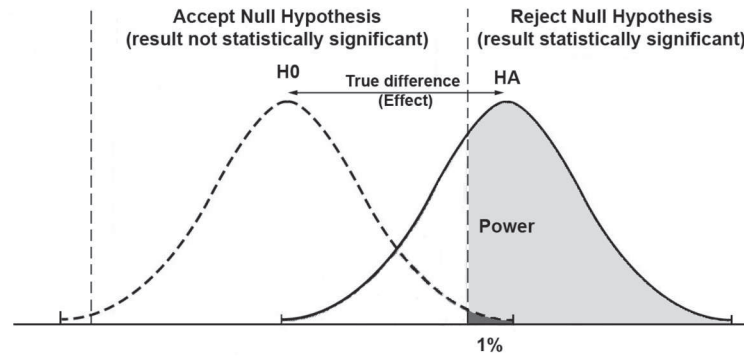
The traditional way of assessing the role of chance herein is so-called hypothesis testing, which determines whether a difference between the placebo and treatment group is present or not by using statistical tests to examine the hypothesis that there is no difference. The result of such a test is expressed by the familiar p-value, which is the chance of obtaining the observed difference between the groups when in fact there is no effect. A small p-value leads to the conclusion that the observed difference is “statistically significant” (i.e. unlikely to be caused by chance).

The question is when a p-value is deemed small enough for this conclusion. In science and statistics it has become customary to mark p-values falling below the cutoff value 0.05 (e.g. a chance of 1 in 20) “significant”: a threshold value called “alpha” or “significance level”. Assuming the situation that there is no difference between the groups being studied in a trial, a p-value lower than 0.05 indicates that the chance of finding a result/effect/difference between the groups in the trial is so small, that this assumption should be rejected in favor of the alternative hypothesis that there truly is a difference between the groups. In case of a clinical trial studying the effect of an experimental treatment, a significant p-value thus indicates that there is sufficient evidence to assume that the difference found derived from the administered treatment, thereby rejecting the hypothesis that the treatment has no effect.

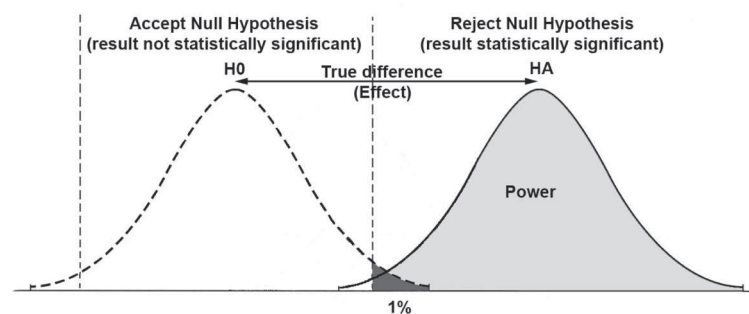
The statistical power of a test reflects the probability that a statistical test will reject the hypothesis that there is no difference or effect (e.g. the chance of finding a p-value equal or below 0.05; a significant result), for the alternative hypothesis (that there is an effect) in case that in fact there is no effect. Another way of formulating this is “the ability of a statistical test to detect an effect, given that the effect actually exists.” Although there are no formal standards, a power of 80% is generally regarded as acceptable.

With the aforementioned Normal distribution, the concept of statistical power can be visualized (figure 6). Three factors greatly determine the statistical power of a trial: the magnitude of the postulated effect, the specified level of significance (alpha), and the number of patients participating in the trial, often referred to as sample size. The larger the postulated effect (figure 7) or the more liberal the level of significance (figure 8), the more likely the test and the trial will yield a significant p-value. Since the observed variation in the samples depend directly on the number of patients participating in the trial, a larger sample size also improves statistical power (figure 9).

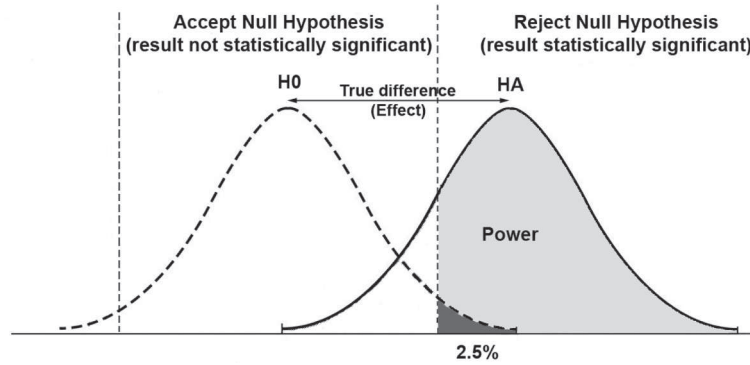
A power analysis is an important aspect in the design of a clinical trial. Power analysis is used to anticipate the likelihood that a study will yield a significant effect, and mainly focuses on determining the required sample size for a given treatment effect (as assessed by a predefined outcome measure) and a predefined alpha. When the data is assumed to follow the Normal distribution, standardized formulas are available to calculate the sample size. However, when the data is differently distributed, like for example the number of enhancing lesions, alternative methods need to be applied. One approach is to apply non-parametric techniques, for example based on statistical resampling methods such as bootstrapping, and assume no distribution at all, whereas another option, among others applied in this thesis, is to fit an alternative distribution to the data and implement this distribution in a simulation procedure. By drawing data from the new distribution, placebo groups and treatment groups, after artificially simulating a treatment effect, can be generated and the required number of patients to obtain sufficient statistical power can be calculated. If a power analysis is not performed prior to clinical trials, the applied sample size might either be too small with the trial thereby lacking the precision to demonstrate the anticipated treatment effect or, the sample size might be too large which would unnecessarily expose patients to potentially harmful side-effects and be a waste of value time and resources, often for a minimal gain.



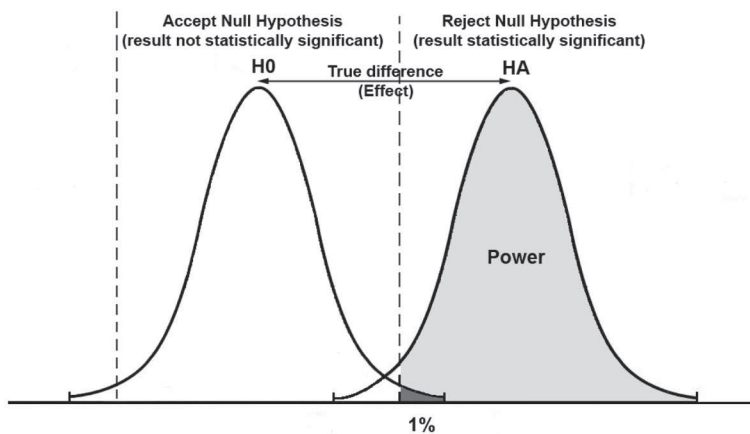
**Figure 6 | A.** Visualisation of the statistical power within a statistical test. The curves represent the distribution of the mean difference in the number of lesions between untreated and treated patients (e.g. the test statistic, in this simplified example for the purpose of explanation assumed to follow the normal distribution with a known variance) wherein the null hypothesis ( $H_0$ ) represents the assumption that there is no difference in the number of lesions between untreated and treated patients in the population ( $H_0=0$ ), and the alternative hypothesis ( $H_A$ ) represents the assumption that there is a difference in the number of lesions between untreated patients and treated patients in the population. When the mean difference in lesion number between treated and untreated patients exceeds the predefined alpha (here: 1%), the difference is deemed “statistically significant”. The surface beneath the  $H_A$  represents the statistical power of the test.



**B.** A larger measurable effect (a larger difference between the alternative hypothesis and the null hypothesis) yields larger statistical power.



C. Increase of the significance level (here from 1% to 2.5% compared to figure a) causes an increase of the statistical power, assuming the same effect size.



D. Increase of the number of participating patients causes the standard error (SE) to decrease in size, thereby changing the shape of the distribution (bell shape becomes smaller). Subsequently, the statistical power of the test increases, assuming the true difference (effect) between the alternative hypothesis and the null hypothesis to remain equal.



## 1.5 | AIMS AND OUTLINES OF THIS THESIS

With the fast development of new therapeutic agents for MS and the widespread use of approved therapies altering the practice of trials in MS, use of more sensitive and powerful outcome measures to maximize the ability of detecting treatment effects, and the need for more pathological specific outcome measures is becoming increasingly important. The previous paragraphs have illustrated the importance of MRI outcome measures in clinical trials and underlined the role of statistical power in efficient trial design.

The aim of this thesis was to explore the statistical power of conventional and non-conventional MRI measures and investigate their feasibility as primary outcome in clinical trials of MS. To do so, the statistical distribution of the applied MRI measures is investigated and parametric power analyses are performed to determine the required sample size for placebo controlled clinical trials.

In the first part of this thesis, conventional and non-conventional measures for monitoring inflammatory activity will be explored. First, in **Chapter 2.1**, we will assess the statistical distribution of the number of enhancing lesions across MS patients in multiple datasets, and evaluate whether the currently proposed statistical distribution for this measure is valid. Next, in **Chapter 2.2**, we will compare the sensitivity and statistical power of the number of enhancing lesions as primary outcome measure in clinical trials with an alternative approach of monitoring the number of newly formed lesions: T2 subtraction imaging. In **chapter 2.3**, the statistical modelling of enhancing lesion volumes in MS patients is studied, and its statistical power compared with the number of enhancing lesions as outcome measure.

of MRI outcome measures in clinical trials capable of monitoring treatments inducing neuroprotection and repair are described. In **Chapter 3.1** we determine the required sample size for clinical trials using the number of PBHs as primary outcome measure, in **chapter 3.2** the required number of patients in short interval clinical trials for the rate of cerebral atrophy as primary outcome measure is determined and in **chapter 3.3**, we determine the required sample size for clinical trials using the recovery of average lesion MTR as primary outcome measure.

Finally in **chapter 4**, the results of the previous chapters will be summarized and discussed, and directions for future research will be presented.

## REFERENCE LIST

1. **Pugliatti M, Rosati G, Carton H, Riise T, Drulovic J, Vécsei L, Milanov I.** The epidemiology of multiple sclerosis in Europe. *Eur J Neurol.* 2006;13(7): 700-22.
2. **Dyment DA, Ebers GC, Sadovnick AD.** Genetics of multiple sclerosis. *Lancet Neurol.* 2004;3(2):104-10.
3. **Giovannoni G, Ebers G.** Multiple sclerosis: the environment and causation. *Curr Opin Neurol.* 2007; 20(3): 261-268.
4. **Weinshenker BG.** The natural history of multiple sclerosis. *Neurol Clin.* 1995; 13(1): 119-46.
5. **Lublin FD, Reingold SC.** Defining the clinical course of multiple sclerosis: results of an international survey. National Multiple Sclerosis Society (USA) Advisory Committee on Clinical Trials of New Agents in Multiple Sclerosis. *Neurology.* 1996; 46(4): 907-11.
6. **Miller DH, Leary SM.** Primary-progressive multiple sclerosis. *Lancet Neurol.* 2007; 6(10): 903-12.
7. **PRISMS Study Group.** Randomised, double-blind, placebo-controlled study of interferon beta-1a in relapsing remitting multiple sclerosis. *Lancet* 1998; 7;352 (9139): 1498-1504.
8. **Comi G, Filippi M, Wolinsky JS.** European Canadian multicenter, double-blind, randomized, placebo-controlled study of the effects of glatiramer acetate on magnetic resonance imaging-measured disease activity and burden in patients with relapsing multiple sclerosis. European Canadian Glatiramer Acetate Study Group. *Ann Neurol.* 2001; 49(3): 290-297.
9. **Polman CH, O'Conner PW, Havrdova E et al.** A randomized, placebo-controlled trial of natalizumab for relapsing multiple sclerosis. *N Eng J Med.* 2006; 2; 354(9):899-910.
10. **van Waesberghe JH, Kamphorst W, De Groot CJ, van Walderveen MA, Castelijns JA, Ravid R, Nijeholt GJ, van d,V, Polman CH, Thompson AJ, Barkhof F (1999)** Axonal loss in multiple sclerosis lesions: magnetic resonance imaging insights into substrates of disability. *Ann Neurol* 1999; 46:747-754.
11. **Inglese M, Mancardi GL, Pagani E, et al.** Brain tissue loss occurs after suppression of enhancement in patients with multiple sclerosis treated with autologous haematopoietic stem cell transplantation. *J Neurol Neurosurg Psychiatry* 2004;4:643-644.
12. **Coles AJ, Wing MG, Molyneux P, et al.** Monoclonal antibody treatment exposes three mechanisms underlying the clinical course of multiple sclerosis. *Ann Neurol* 1999; 3: 296-304.
13. **Lopez-Diego RS, Weiner, HL.** Novel therapeutic strategies for multiple sclerosis - a multifaceted adversary. *Nat. Rev. Drug Discov.* 2008: 909-925.
14. **Kapoor R, Davies M, Blaker PA et al.** Blockers of sodium and calcium entry protect axons from nitric oxid-mediated degeneration. *Ann Neurol.* 2003; 53: 174-180.
15. **Bechtold DA, Kapoor R, Smith KJ.** Axonal protection using flecainide in experimental autoimmune encephalomyelitis. *Ann Neurol.* 2004; 55: 607-616.
16. **Schippling S, Heesen C, Zander A, Martin R.** Stem cell transplantation in multiple sclerosis. *J Neurol.* 2008; 255: S6: 43-47.

17. Rovira A, León A. MR in the diagnosis and monitoring of multiple sclerosis: an overview. *Eur J Radiol.* 2008;67(3): 409-14.
18. Miller DH, Grossman RI, Reingold SC, McFarland HF. The role of magnetic resonance techniques in understanding and managing multiple sclerosis *Brain.* 1998;121 ( Pt 1): 3-24.
19. Barkhof F, Bruck W, De Groot CJ, Bergers E, Hulshof S, Geurts J, Polman CH, van der Valk P. Remyelinated lesions in multiple sclerosis: magnetic resonance image appearance. *Arch Neurol.* 2003;60(8): 1073-1081.
20. Miller DH et al. Serial gadolinium-enhanced magnetic resonance imaging in multiple sclerosis. *Brain* 1988; 111: 927- 939.
21. Filippi M et al. Guidelines for using quantitative measures of brain magnetic resonance imaging abnormalities in monitoring the treatment of multiple sclerosis. *Ann Neurol.* 1998; 43: 499-506.
22. van Waesberghe, JH, van Walderveen, MA, Castelijns, JA, Scheltens, P, Lycklama, Nijeholt, GJ, Polman, CH, et al. Patterns of lesion development in multiple sclerosis: longitudinal observations with T1-weighted spin echo and magnetization transfer MR. *Am J Neuroradiol.* 1998; 19: 675-683.
23. van Walderveen MA, Kamphorst W, Scheltens P, van Waesberghe JH, Ravid R, Valk J, Polman CH, Barkhof F. Histopathologic correlate of hypointense lesions on T1-weighted spinecho MRI in multiple sclerosis. *Neurology* 1998; 50:1282-1288.
24. Bitsch A, Kuhlmann T, Stadelmann C, Lassmann H, Lucchinetti C, Bruck W. A Longitudinal MRI study of histopathologically defined hypointense multiple sclerosis lesions. *Annals of Neurology* 2001; 49: 793-796.
25. Fox NC, Jenkins R, Leary SM, Stevenson VL, Losseff NA, Crum WR, Harvey RJ, Rossor MN, Miller DH, Thompson AJ. Progressive cerebral atrophy in MS. *Neurology* 2000; 54: 807-812.
26. Kalkers NF, Ameziane N, Bot JC, Minneboo A, Polman CH, Barkhof F. Longitudinal brain volume measurement in multiple sclerosis: rate of brain atrophy is independent of the disease subtype. *Arch Neurol.* 2002; 59: 1572-1576.
27. Barkhof F, Calabresi PA, Miller DH, Reingold SC. Imaging outcomes for neuroprotection and repair in multiple sclerosis trials. *Nat Rev Neurol.* 2009; 5(5): 256-266.
28. Anderson VM, Bartlett JW, Fox NC, Fisniku L, Miller DH. Detecting treatment effects on brain atrophy in relapsing remitting multiple sclerosis: Sample size estimates. *J Neurol* 2007; 254:1588-1594.
29. Schmierer K, Scaravilli F, Altmann DR, Barker GJ, Miller DH. Magnetization transfer ratio and myelin in postmortem multiple sclerosis brain. *Ann Neurol.* 2004; 56(3): 407-415.
30. Chen JT, Collins DL, Atkins HL, Freedman MS, Arnold DL. Magnetization transfer ratio evolution with demyelination and remyelination in multiple sclerosis lesions. *Ann Neurol.* 2008;63(2): 254-262.
31. Tan IL, van Schijndel RA, Fazekas F, et al. Image registration and subtraction to detect active T2 lesions in MS: an interobserver study. *J Neurol* 2002; 249: 767-73.
32. Moraal B, Meier DS, Poppe PA, et al. Subtraction MR Images in a Multiple Sclerosis Multicenter Clinical Trial Setting. *Radiology* 2009; 250:506-14.