Basic Neuroscience

Rotarod motor performance and advanced spinal cord lesion image analysis refine assessment of neurodegeneration in experimental autoimmune encephalomyelitis

Robert van den Berg [a,b], Jon D. Laman [c], Marjan van Meurs [d], Rogier Q. Hintzen [b], Casper C. Hoogenraad [a,*]

[a] Cell Biology, Utrecht University, Utrecht, The Netherlands
[b] Department of Neurology, Erasmus MC, Rotterdam, The Netherlands
[c] Department of Neuroscience, University Groningen, University Medical Center Groningen, The Netherlands
[d] Department of Immunology, Erasmus MC, Rotterdam, The Netherlands

HIGHLIGHTS

- The rotarod provides a quantitative and reproducible way of measuring motor performance in EAE.
- Clinical score and rotarod performance are highly correlated.
- Rotarod performance is correlated with demyelination in the motor systems of the spinal cord.

ARTICLE INFO

ABSTRACT

Background: Experimental autoimmune encephalomyelitis (EAE) is a commonly used experimental model for multiple sclerosis (MS). Experience with this model mainly comes from the field of immunology, while data on its use in studying the neurodegenerative aspects of MS is scarce.
New method: The aim of this study is to improve and refine methods to assess neurodegeneration and function in EAE. Using the rotarod, a tool used in neuroscience to monitor motor performance, we evaluated the correlation between motor performance, disease severity as measured using a clinical scale and area covered by inflammatory lesions.
Results: The included parameters are highly correlated in a non-linear manner, with motor performance rapidly decreasing in the intermediate values of the clinical scale. The relation between motor performance and histopathological damage is exclusively determined by lesions in the ventral and lateral columns, based on a new method of analysis of the entire spinal cord. Using a set of definitions for distinct disease milestones, we quantified disease duration as well as severity.
Comparison with existing methods: The rotarod measures motor performance in a more objective and quantitative manner compared to using a clinical score. The outcome showes a strong correlation to the surface area of inflammatory lesions in the motor systems of the spinal cord.
Conclusions: These results provide an improved workflow for interpreting the outcome of EAE from a neurological point of view, with the eventual goal of dissecting neurodegeneration and evaluating neuroprotective drugs in EAE for application in MS.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Multiple sclerosis (MS) is the most common cause of neurological disability in young people. It affects 0.1–0.2% of the population with onset of disease at a mean age of 28 years (Goodin, 2014). As of yet, no curative treatment has been developed, despite very good progress in reduction of the inflammatory component, mostly by biological anti-inflammatory drugs (Ransohoff et al., 2015). MS is a complex disease entity, with characteristics of both an immunological and a neurodegenerative disorder (Stys et al., 2012). Research into the causes of MS is difficult, partially because a large discrepancy exists between biological and clinical onset with patients

* Corresponding author. Tel.: +31 30 2534585; fax: +31 30 2532837.
E-mail address: c.hoogenraad@uu.nl (C.C. Hoogenraad).

developing symptoms after years of subclinical disease. As more time passes, epidemiological or histological traces of any causative factor will be harder to detect. Also, the processes leading to disability in MS occur in the brain and spinal cord, areas of the central nervous system (CNS) that are hard to probe or study at the microscopic level from the outside. The majority of MS patients are initially diagnosed with relapsing-remitting MS, a form of the disease in which short periods of attacks are followed by timeframes in which disease activity is nearly absent. Although almost all individuals eventually show progression to a continuously active form of the disease as damage accumulates (Ebers, 2001; Tremlett et al., 2008). The variation seen in MS patients supports the growing conviction that MS should be considered a collection of several disorders with different causes and pathways involved, but with a relatively similar set of symptoms as result (Lassmann, 2005).

All these challenges stress the importance of an accurate experimental model for MS. Ever since it was first described in the 1920s, as a chance finding during vaccination trials (Koritschoner and Schweinburg, 1925), Experimental Autoimmune Encephalomyelitis (EAE) has been the most widely used model system for MS (Mix et al., 2010; Baker and Amor, 2015). In EAE animals are inoculated with myelin components, triggering an immune response against the injected compound which also targets the native myelin sheath. The resulting disease depends on the species and strain of animals selected, as well as the molecule or cell used to evoke an immune response and any adjuvants used to stimulate cell-mediated immunity (Gold et al., 2006). Although EAE has greatly contributed to our understanding of MS (Baxter, 2007; Steinman and Zamvil, 2006), several disadvantages limit its use (Steinman and Zamvil, 2005; Ransohoff, 2012). The first and foremost of these is that in many protocols, EAE is an inflammatory disease of the CNS, only partially mimicking the neurodegenerative process seen in MS (Sriram and Steiner, 2005). Another concern involves the type of immune response, which in rodent EAE is mainly $CD4^+$ T-cell mediated, while active MS lesions in human patients tend to show a higher involvement of $CD8^+$ cells (Friese and Fugger, 2009; 't Hart et al., 2011).

EAE is not a trivial experimental procedure, and the consistency of measurements can vary depending on the experience of the person who implements the complex protocol (Stromnes and Goverman, 2006). Even seemingly minor decisions -such as the adjuvant used for immunization (Smith et al., 2011) and the dose and timing of pertussis toxin- will influence the disease course. Similar attention should be paid to the acquisition and analysis of results (Baker et al., 2011). It is common for EAE experiments to measure disease severity using a clinical score running from 0 to 5, with 0 indicating no symptoms and 5 death (e.g. Kassiotis et al., 1999). This scale is observer-dependent, non-linear and produces a categorical variable, implying that any statistical analysis involving a mean score has to be interpreted with caution (Baker et al., 2014). Two animals with a mean score of 2.5 can both be moderately sick or one animal is dead while the other shows no symptoms. Representing the results of an EAE experiment accurately requires a thorough statistical analysis (Fleming et al., 2005). Failure to take the proper precautions may contribute to publication bias (Tsilidis et al., 2013).

In the field of neuroscience, it has been common for several decades to measure motor performance of rodents using a rotarod, a bar rotating at a slowly increasing speed. The animal is forced to increase its walking speed to keep up with the bar, until it can no longer maintain its balance and drops on a switch, connected to a clock. The latency time to fall can then be used as an objective and non-invasive measurement of motor performance (Jones and Roberts, 1968). As common as this method is in neuroscience, relatively few studies have explored its application in EAE (Jones et al., 2008; Al-Izki et al., 2012; Moore et al., 2014). In this paper,

we demonstrate the use of a rotarod to evaluate the results of EAE induction in mice. We compare the rotarod measurements to a clinical score as well as to the number of inflammatory lesions found in the spinal cord of these animals. Using a set of definitions for crucial disease events based on clinical score and rotarod, we also provide a quantifications of EAE disease course over time. The clinical impact of any CNS lesion is determined by its location as well as its size, where a minor lesion in a critical system can lead to severe clinical disability. Furthermore, we develop a novel image analysis approach in which for each spinal cord lesion the information on anatomical context is preserved, allowing for a better understanding of the three-dimensional distribution of inflammatory lesions and their relation to clinically observed disease severity.

2. Materials and methods

2.1. Ethics statement

All animal experiments were performed in compliance with the guidelines for the welfare of experimental animals issued by the Government of The Netherlands. All animal experiments were approved by the Animal Ethical Review Committee (DEC) of the Erasmus Medical Centre under protocol number 128-12-01. All experiments were performed in the animal facilities of the Erasmus Medical Centre.

2.2. EAE induction and assessment

C57BL/6Tac animals were originally obtained from Xenogen Biosciences (currently: Taconic Farms, Inc. NJ, USA). At the start of experiments, the strain had been in our facility for four generations. Animals were housed in specified pathogen-free conditions. Both male and female animals were used for the experiments. The mice were 12 weeks old at the start of the experiment, all animals were born within 3 days of eachother. EAE was induced by subcutaneous immunisation at 4 sites in the flanks with 200 μg MOG_{35-55} peptide in total, emulsified in CFA (Difco/Voigt, Lawrence, KS, USA) on day 0 as described previously (Stromnes and Goverman, 2006). 500 μg heat-killed *Mycobacterium tuberculosis* H37RA (Difco/Voigt, Lawrence, KS, USA) was added per 100 μl CFA to evoke a better immune response and more extensive neuropathology. At day 0 and 2 mice were injected i.p. with 150 ng pertussis toxin (PTX, Sigma-Aldrich, MO, USA). After EAE induction, mice were monitored daily, and weight and clinical status were recorded. A clinical score was used as defined in Table 1. If an animal was scored higher than 2, gel packs and soft food pellets were provided. In addition,

Table 1
Scale for Clinical Score. During this experiment, all animals with a score higher than 2.0 were provided with gel packs and soft food pellets. If an animal reached a score of 4.5, it was euthanized. Scores of 4 and above are rare, only animal reached 4.5 and was euthanized.

Score	Symptoms
0.0	None
0.5	Weakness of the tail.
1.0	Paralysis of the tail without weakness of the limbs.
1.5	Weakness of the limbs without paralysis of the tail.
2.0	Weakness of the limbs with weakness or paralysis of the tail.
2.5	Complete paralysis of one of the hind limbs, or weakness of three or four limbs while retaining the ability to walk.
3.0	Complete paralysis of front or hind limbs, or weakness of three or four limbs with the loss of walking ability.
3.5	Complete paralysis of the hind body, animal cannot turn its body.
4.0	Complete paralysis of front and hind body.
4.5	Complete paralysis with inability to eat and drink.
5.0	Death due to experiment.

from day 6 onward motor performance was tested daily using a rotarod device. On day 28, animals were humanely euthanized and perfused with 4% paraformaldehyde.

2.3. Rotarod test

Animals were trained to walk on the accelerating rotarod (Ugo Basile 47600, Milan, Italy). The rotarod consists of a cylinder with a diameter of 3 cm on which 5 animals can run simultaneously, separated by panels of sufficient size to prevent the animals from detecting each other visually. The speed of the rod was linearly increased from 4 rpm to 40 rpm during 300 s, after which the animals were returned to their cages. When the animal was not capable of maintaining its balance and fell of the device, it triggered a sensor and the time was recorded. The first week of the experiment, before the onset of symptoms, was used to train the animals in the use of the device and to obtain baseline values. Animals were subjected to the rotarod on a daily basis, unless they were unable to move due to paralysis of the limbs (EAE score higher than 3), in which case their latency time was recorded as 1 s.

2.4. Histology

After perfusion with 4% paraformaldehyde, the brain and spinal cord were extracted and embedded in paraffin using an automated Microm HMP110 paraffin tissue processor. Paraffin material was cut into 5 μm thick coronal (brain) or transverse (spinal cord) sections. Prior to staining, sections were deparaffinised in xylene and rehydrated in ethanol of decreasing percentage (from 100% to 70%) followed by rinsing with distilled water. Histochemical staining was performed according to Klüver and Barrera (1953). In short, sections were stained overnight with Luxol Fast Blue solution (Sigma–Aldrich) at 60° and rinsed in 95% ethanol and distilled water. To wash away non-specific labelling, sections were rinsed with a 0.05% lithium carbonate solution (Sigma–Aldrich) and washed with 70% ethanol. Sections were then stained with a 0.5% Eosin-Y solution (Sigma–Aldrich), washed with distilled water and stained with a 0.25% Chresyl Echt Violet solution (Chresyl violet acetate, Sigma–Aldrich). After these stainings, slides are rinsed with distilled water, dehydrated using an ethanol series with increasing percentage (80–100%), followed by two washes with xylene.

2.5. Computational analysis of spinal cord sections

Slices were imaged on a Zeiss Axioskop upright microscope with 2.5× and 10× objectives. Processing was performed using the tools available in the open source image analysis package FIJI (Schindelin et al., 2012). All images were aligned to a line from the gracile fasciculus through the central canal to the ventral fasciculus with the dorsal side facing up. The contours of the white matter, the grey matter and any lesions present in the slice were manually tracked in FIJI using the multipoint selection tool, resulting in a table of coordinates. These coordinates were loaded by a custom script written in R (R Core Team, 2014) and converted into areas for which both the absolute surface and for each lesion its area relative to the white matter was calculated. All images from a single animal were aligned according to their position in the paraffin block and matched by eye with the corresponding element from the reference series to determine the spinal cord segment.

2.6. Statistical analysis

All statistical analysis was performed in R (R Core Team, 2014). Variables following a normal distribution were compared using Welch two sample t-test. Count or proportional data, such as the disease categories and time points were compared using Fisher's exact test for count data. Linear correlations were tested using Pearson's product-moment correlation, except for the relation between initial weight and clinical score/rotarod performance (Fig. 4D and E), where a linear model was fit to the data predicting either maximal clinical score or minimal rotarod performance based on gender and weight. An analysis of variance (ANOVA) was then performed to determine the significance of the gender and weight terms. To determine the best description for the relation between clinical score and rotarod performance (Fig. 1G), a two-step approach was used. By eye, the shape of the curve suggested a logistic decay function. In the first round, a brute force algorithm (Grothendieck, 2013) was used to get an initial estimate of the fit parameters. In the second round, these estimated values were used as starting points for a nonlinear least-squares estimate of the optimal parameters. The optimal fit was achieved when assuming a limit of 286 s rotarod performance, an EC_{50} midway point at a clinical score of 1.79, and slope coefficients of −0.08 and −5.21.

3. Results

3.1. Clinical score and rotarod performance show excellent correlation in a non-linear fashion

For the methodological purposes of this study and guided by the IACUC, a single large EAE experiment was designed using both male and female mice, with a modified induction protocol to enhance neurological damage. The same experiment was co-opted to assess whether knocking out the molecular motor protein Kif1b affected EAE (Van den Berg et al, manuscript in preparation). During the course of EAE, all animals were scored daily by the same individual. To avoid an observer-expectancy effect, a coding system was used to keep the observer unaware of animal identity and score of the previous day. A classic and validated clinical EAE scale from 0 to 5 was used, as described in Table 1. At the same time each day, all animals were placed on the rotarod and latency time to fall was measured. Approximately 7 days after induction, the first animals started to develop symptoms (Fig. 1A). During the first few days, animals need to get used to the rotarod, leading to a predicted initial increase in performance. All animals reached maximum performance before onset of symptoms. Ten days after induction, an increase in clinical score is mirrored in a decrease in rotarod latency time. (Fig. 1B).

It should be noted that a significant fraction of the animals in which EAE had been induced did not develop symptoms. The choice of whether or not to include these animals in the analysis significantly influences the outcome of the experiment, as illustrated by the two curves in Fig. 1A and B. To circumvent this issue, the score and rotarod results for each day where converted to a categorical variable according to the criteria shown in Table 2. The number of animals assigned to each category as a fraction of the total population can then be used as a more informative readout of disease severity compared to only mean values. Although the patterns for the categories 'asymptomatic' (Fig. 1C) and 'severe' (Fig. 1E) are quite similar when defined according to clinical score or rotarod performance, only a few animals are considered moderately sick according to rotarod performance (Fig. 1D). This becomes even more apparent when calculating the predictive value of

Table 2
Disease severity criteria. Definitions used to divide clinical score and rotarod performance into categories.

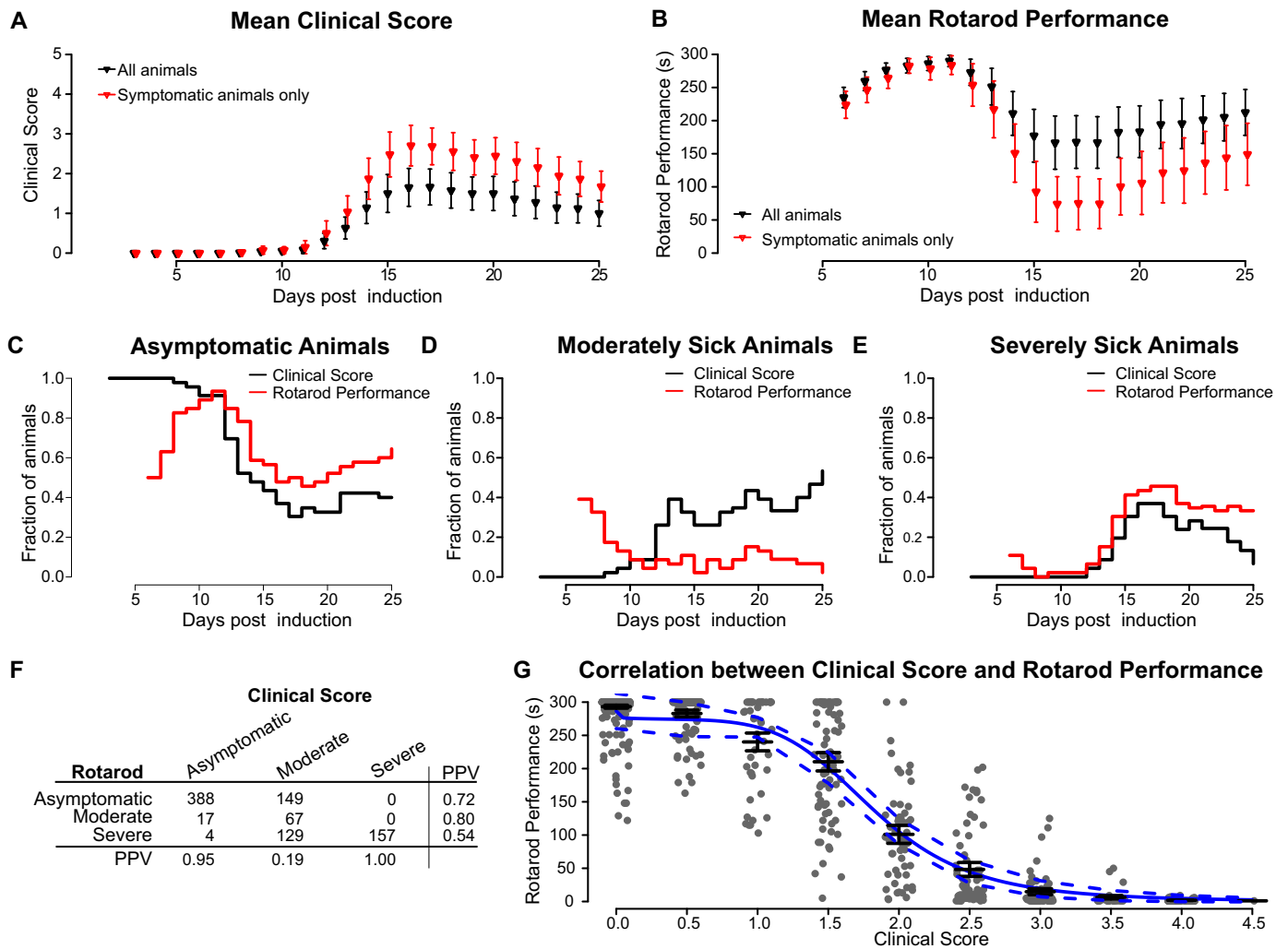| Category	Clinical Score	Rotarod Performance
Asymptomatic	0.0	>250 s
Moderate	0.5–2.5	150–250 s
Severe	>2.5	<150 s

Fig. 1. Rotarod motor performance is highly correlated to clinical score. (A) Average clinical score per day for all animals (black) and only the animals that develop symptoms (red). (B) Average rotarod performance in seconds for all animals (black) and only the animals that develop symptoms (red). The increase in clinical score around day 12 corresponds to a decrease in rotarod performance. (C–E) To get a more meaningful estimate of the variation in disease intensity over time, animals were classified into three categories. The plots show the fraction of the total population which shows no symptoms at all (C), moderate (D) or severe symptoms (E) according to either the clinical score (black) or the rotarod performance (red). Rotarod performance and clinical score show a similar pattern for asymptomatic and severely sick animals, but differ in their classification of moderately sick animals. As predicted, the fraction of asymptomatic animals according to rotarod performance is not 1 at the start of the experiment, due to the learning curve the animals are experiencing while first using the device. (F) Table of concordance between clinical score and rotarod. On each day the animals were observed and classified as 'asymptomatic', 'moderately sick' or 'severely sick' according to the definitions described in the text. There is a high level of agreement between clinical score and rotarod performance, with the notable exception of the clinical score category 'moderately sick'. PPV indicates positive predictive value, the number of true positives divided by the total number of observations, in this situation using the other test (clinical score or rotarod) as gold standard. (G) There is a non-linear correlation between clinical score and rotarod performance, best fitted using a logistic decay function. Rotarod performance is not affected in the lower clinical scores, but rapidly decreases at scores higher than 1. Solid blue line indicates optimal nonlinear least squares fit, dashed lines indicate 95% confidence interval. Error bars in A–B and G indicate 95% confidence interval.

clinical score for rotarod performance. When an animal is categorised as 'asymptomatic' according to the clinical score, there is a 0.95 probability it will be similarly categorised according to the rotarod performance. In contrast, if the clinical score of an animal results in a classification of moderately sick, the probability that its rotarod performance falls in the same category is only 0.19 (Fig. 1F). This discrepancy can be explained by the distinctively non-linear correlation between clinical score and rotarod performance, which can best be described using a logistic decay function (Fig. 1G). Over the clinical score interval of 1.5–2.5 there is a rapid decrease in motor performance as measured by rotarod. While animals with a clinical score lower than 1.5 usually retain maximum rotarod performance and animals with clinical score over 2.5 are barely able to walk the rod, all loss of motor function takes place in the small interval between score 1.5 and 2.5.

3.2. Analysing the natural history of EAE

The use of categories provides a more accurate representation of disease severity when compared to solely presenting a mean value. A similar approach can be used to analyse the chronological progression of disease. We defined four pivotal events in the natural history of EAE, according to both clinical score and rotarod (Table 3). These definitions were applied to the disease course of the animals included in the experiment, resulting in the corresponding dates for each event (applied to an individual animal in Fig. 2A and B, to the entire population in Fig. 2C and D). Clinical score and rotarod are in agreement on the time point of most of the events, with the largest discrepancy in the median observed for recovery (Fig. 2E, Table 4). When comparing the date for each individual event according to rotarod performance and clinical score, a strong linear correlation

Table 3
Criteria used to define disease events.

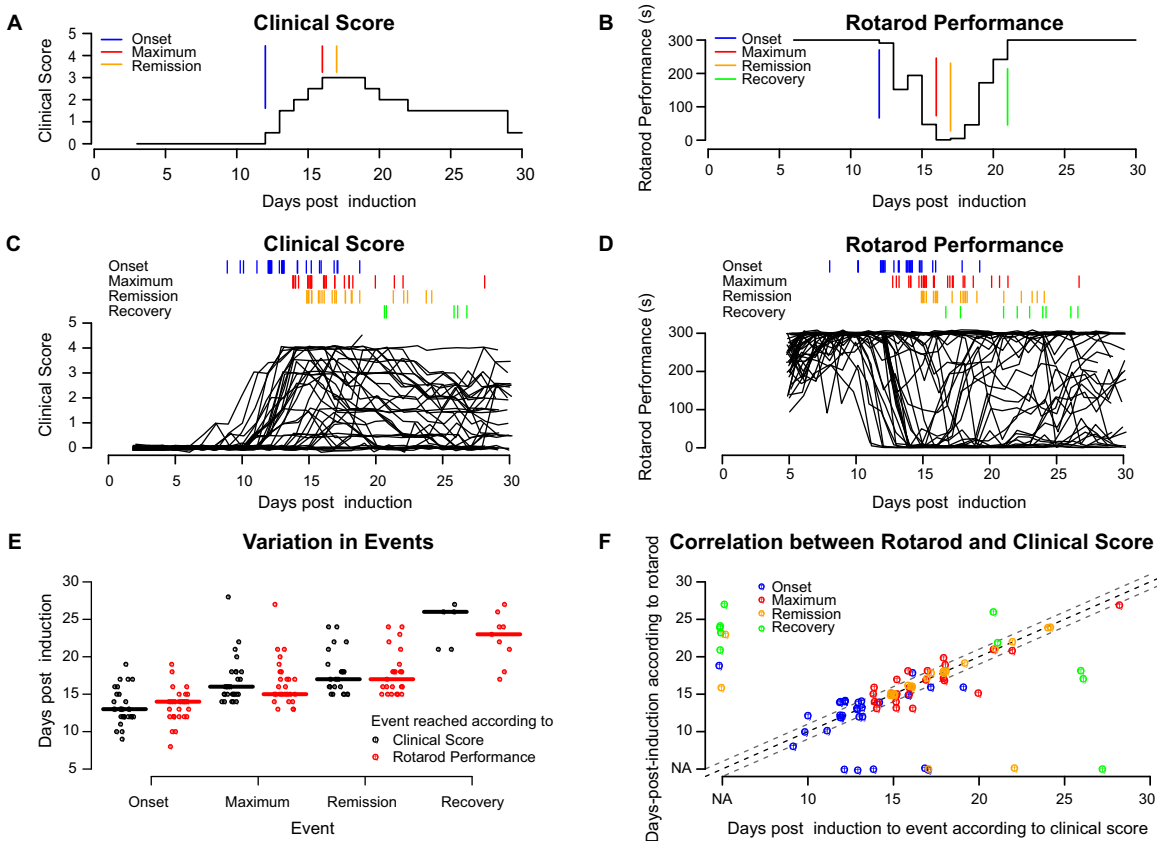| Event	Clinical Score	Rotarod Performance
Onset	Day 1 in a period of at least 3 days with clinical score >0.	Day 1 in a period of at least 3 days with rotarod performance lower than the maximum score during the first week of testing.
Maximum	Day on which the maximum clinical score is recorded.	Day at which the minimal rotarod performance is observed.
Remission	A period of at least 3 consecutive days after the disease maximum in which the animal shows a decreasing or equal score, with at least two of the days showing improvement.	A period of at least 3 consecutive days after the disease maximum in which the animals shows an increasing or equal rotarod performance, with at least two of the days showing improvement.
Recovery	A period of at least 3 consecutive days after the disease maximum in which the animals scores 0.	A period of at least 3 consecutive days after the disease maximum in which the animal achieves a rotarod score of at least 90% of the maximal performance before disease onset.

Fig. 2. Major disease events according to rotarod and clinical score. (A) and (B) Disease course for one typical animal illustrating the different disease events as defined in the text, according to clinical score (A) and rotarod performance (B). Although this animal shows clear signs of recovery, it never fulfils the criteria for complete remission according to clinical score. (C) and (D) Disease course for all animals with time points for disease events shown on top. Although the patterns for clinical score (C) and rotarod performance (C) are highly similar, complete remission is much more rare according to clinical score. (E) Median time to disease event according to clinical score (black) and rotarod performance (red), horizontal bars indicating the median. All event definitions are reached either on the same day or with one day difference, with the exception of recovery. (F) Correlation between date according to clinical score and rotarod performance for each individual event. If positioned on the middle grey line, a disease event takes place on the exact same day according to both clinical score and rotarod performance, with both parallel lines indicating plus or minus one day. All events show an excellent correlation, with the notable exception of recovery. The NA mark on the axes indicates that the criteria for this event was fulfilled according to one but not both measurements.

Table 4
Days post-induction to disease events. Date values indicate median and range in days post-induction. *P*-values indicate the probability assigned by Fishers' exact test for count data to the hypothesis that the date of the event is independent of the measurement method chosen, i.e. clinical score or rotarod.

Event	According to Score		According to Rotarod		P-value
	Median	Range	Median	Range	
Onset	13	(9–19)	14	(8–19)	0.21
Maximum	16	(14–28)	15	(13–27)	0.63
Remission	17	(15–24)	17	(15–24)	0.99
Recovery	26	(21–27)	23	(17–27)	0.89

is observed (Fig. 2F, $\rho = 0.87$, $P < 0.001$, Pearson's product-moment correlation). Of all events ($N = 79$), 51.9% is assigned to the same date according to rotarod performance and clinical score, with an additional 25.3% showing a difference of a single day. In 9.2% of the cases, a disease milestone – mainly recovery – is reached according to rotarod performance, but not according to clinical score. The opposite is true in another 9.2%, where clinical score predicts an event – onset of disease – that does not take place according to rotarod performance. In these cases, an animal experiences a temporary weakness or paralysis of the tail, without a drop in motor performance.
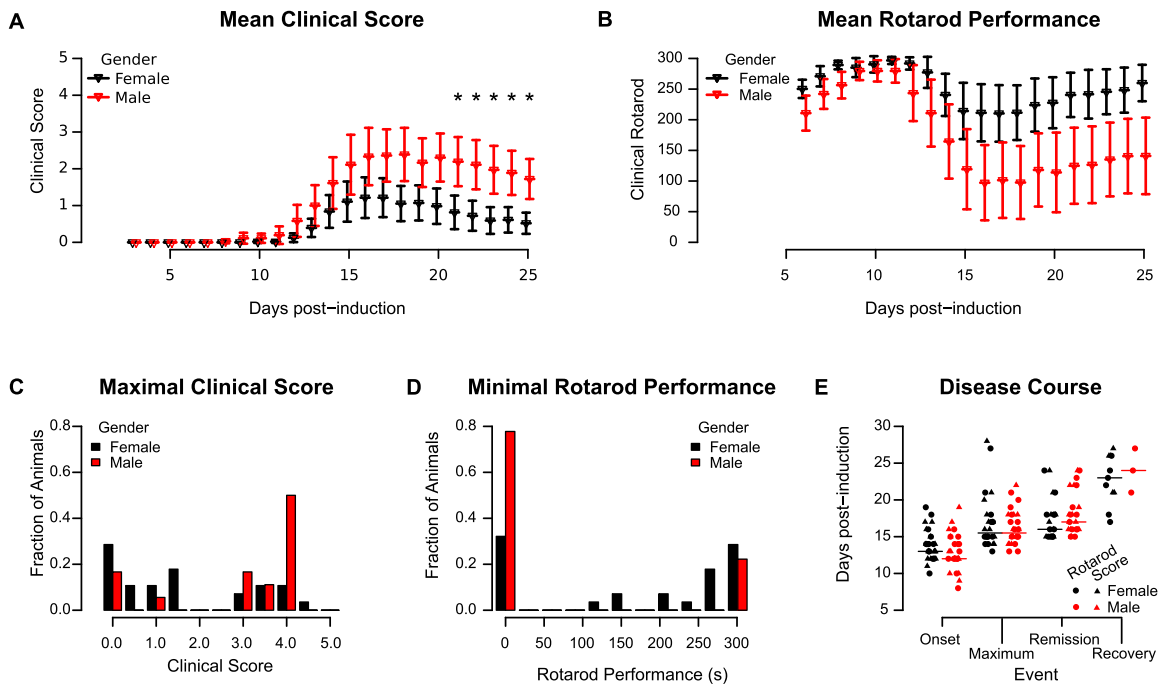
Fig. 3. Gender strongly influences severity of EAE. Male animals experience a more severe form of EAE compared to females, both according to clinical score (A) and as measured by rotarod performance (B). A significantly larger fraction of male animals reaches a clinical score of 4.0 compared to female animals (C) and are no longer able to walk the rotarod (D). This difference is only present when considering disease severity, not when analysing duration or speed of onset. All disease events are reached on the same day or within one day of each other by both male and female animals (E). Error bars in (A) and (B) indicate 95% confidence interval, asterisks indicate significance after correcting for multiple comparison using Bonferroni's method: $^{*} = p< 0.05$. Horizontal lines in (E) indicate the median.

3.3. EAE is more severe in males, but follows the same time pattern

The susceptibility to MS as well as the disease course differ substantially between men and women (Schwendimann and Alekseeva, 2007; Bove and Chitnis, 2014). A similar disparity has been observed for male and female animals subjected to EAE (Massella et al., 2012). In humans, although women have an approximately two- to threefold higher risk of developing MS, the disease tends to be more rapidly progressive in men (Bove and Chitnis, 2013). In our animals, we observed a similar pattern, with a more severe disease course in males compared to females according to both clinical score (Fig. 3A) and rotarod performance (Fig. 3B). A significantly larger fraction of the male animals reached a maximum clinical score of 4.0 compared to the female animals (Fig. 3C, $P = 0.03$), and only 9 out of 28 (32.1%) female animals reached a minimum rotarod latency time of less than 25 s, compared to 14 out of 18 (77.8%) male animals (Fig. 3D, $P = 0.04$). This gender disparity is limited to the severity of disease, there is no difference in time to reach any of the disease milestones (Fig. 3E). No difference in susceptibility to EAE was observed. In both gender groups some animals did not develop EAE, with no significant differences between the groups.

3.4. Weight at disease onset does not influence EAE outcome

Another commonly used indicator of disease is weight (Fleming et al., 2005), which has the advantage that it can be measured objectively with minimal stress to the animal. A disadvantage can be that weight at the start of the experiment varies greatly between individual animals and can differ between experimental groups. Also, since weight loss depends on the metabolic response of the animal to the stress associated with EAE, the magnitude of the weight loss is only indirectly related to disease intensity. We observed a highly significant linear correlation between weight and clinical score (Fig. 4A, $\rho = -0.11$, $P = 1.33 \times 10^{-7}$) and rotarod performance (Fig. 4B, $\rho = 0.12$, $P = 2.3 \times 10^{-4}$), independent of gender. As expected, male animals are significantly heavier at the start of the experiment with a mean weight of 25.1 g versus 19.4 g for the female animals (Fig. 4C, $P = 2.8 \times 10^{-12}$). This might raise the question if this weight difference plays a role in the gender disparity in disease intensity. However, no correlation exists between weight before the start of the experiment and the maximum disease intensity, both according to clinical score (Fig. 4D, $P = 0.82$) and rotarod performance (Fig. 4E, $P = 0.90$). Although limited by its high variance, weight can be used as a marker for disease progression. The weight of an animal at the start of an EAE experiment does not predict disease severity.

3.5. Both rotarod performance and clinical score correlate with the area of inflammatory lesions

Although non-invasive measurements can teach us a great deal about the severity and time course of EAE, the gold standard remains histological analysis. However, little research has been performed correlating clinical score to histological evidence of CNS damage. The available evidence suggests that the correlation with atrophy in the spinal cord might only be weak (Källen and Nilsson, 1986), although a stronger correlation exists with brain atrophy (Paz Soldán et al., 2015). Interestingly, a similar disparity between clinical signs and damage visible on MRI exists in MS patients (clinic-radiological paradox), possibly related to the inability of MRI to reliably and quantitatively visualise axonal damage (Rocca et al., 2013). To determine the correlation between our clinical measurements and the damage to the nervous system, the spinal cords from a representative subset of animals ($N = 12$) were embedded in paraffin and used to generate a series of microscopy slides covering the entire extent of the spinal cord. These slides were stained using the
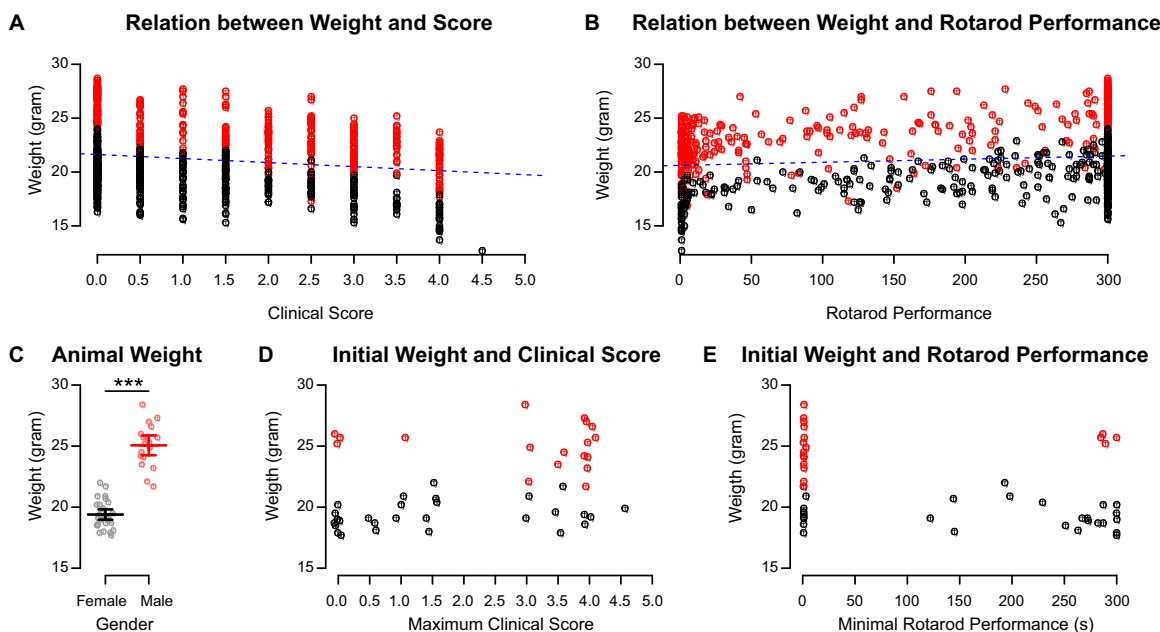
Fig. 4. Weight difference between genders does not explain disease severity. When measuring the animals weight over time, there is a small but significant negative correlation with clinical score (A) and a similar sized and significant positive correlation with rotarod performance (B). This effect is caused by the weight loss associated with severe disease. Although males are significantly heavier at the start of the experiment (C), the initial weight shows no significant correlation with either the maximum clinical score achieved by the animal (D) or the minimal rotarod performance (E). In all panels, data points from male animals are indicated in red, and females in black. A total of 966 weight measurements were included in A and B, a subset of 46 measurements (one for each animal on the first day) were used for panels C–E. Error bars in C indicate 95% confidence interval.

protocol provided by Klüver and Barrera (1953), a method deploying luxol fast blue and eosin Y to chemically stain myelin, neuropil and neurons. Loss of luxol fast blue staining does not represent pure demyelination, but is more often the result of the combination of demyelination, axonal loss and immune cell infiltration (Baker et al., 2011). The correlation between loss of luxol fast blue and axonal degeneration has been confirmed using in vivo imaging studies (Kim et al., 2006; Budde et al., 2009).

The resulting stained slides were imaged in a systematic manner and manually traced, resulting in a list of coordinates for the outline of the white matter, the border between white and grey matter and the boundary of inflammatory lesions, if present. These resulting outlines were matched to an anatomical atlas of the spinal cord (Watson et al., 2009) to determine the exact segment of the spinal cord the slide was located in, as well as the anatomical structure which each of the inflammatory lesions occupied (Fig. 5A$_{1-3}$). Using a custom written script in the programming language Python, the coordinate list was loaded into the open source animation suite Blender (Blender Online Community, 2015) to create a data-driven visualisation of the three-dimensional organisation of the lesions (Fig. 5B and C).

In agreement with previous reports, we find a large variation in the amount of histological damage, both between animals but also along the spinal cord of individual animals. Within the same animal, parts of the spinal cord can be covered with lesions (Fig. 5B), while a different section located only a few segments away in caudal direction can show virtually no lesions at all (Fig. 5C). Still, a significant correlation can be found between the mean percentage of white matter containing lesions and both the clinical score (Fig. 6A, $\rho = 0.70$, $P = 0.01$) and the rotarod performance (Fig. 6B, $\rho = -0.61$, $P = 0.04$). However, it should be observed that these numbers are based on the assumption of a linear correlation, an assumption that is hard to confirm given the high variability. Although the amount of damage within one individual can vary from 0 to 50% of the white matter containing inflammatory lesions in the span of a few spinal cord segments, it appears there are some 'hot spots'

in the spinal cord with a somewhat higher probability of containing inflammatory lesions. The most easily visible of these is located from approximately L3 to S2, the region of the spinal cord innervating the hind limbs (Fig. 6C).

3.6. Inflammatory lesions are located in every section of the white matter, close to the pia mater

A total of 444 lesions were detected in 12 animals. For each lesion, the anatomical structure it occupied and the spinal cord segment in which it was located was determined. If the lesion crossed the border between two anatomical structures, it was excluded from further analysis ($N = 43/9.7\%$, all of these lesions crossed the boundary between anterior and lateral funiculus). Since many anatomical structures only contained one or two lesions, the counts were aggregated into three funiculi. The posterior funiculus contained counts from the gracile fasciculus, the dorsal corticospinal tract and the cuneate fasciculus. The lateral funiculus included the counts from the rubrospinal tract, the lateral corticospinal tract, the lateral spinal nucleus and the lateral cervical nucleus, while the anterior funiculus contained the anterior spinothalamic, olivospinal and vestibulospinal tracts. The absolute number of lesions detected was approximately equal for the lateral and anterior funiculus and nearly halve this amount for the posterior funiculus (Fig. 6D). To get an estimate of the relative surface area of each of the anatomical subdivisions, the anatomical atlas used as a reference (Watson et al., 2009) was traced in a similar manner as the microscopy slides. When correcting the lesion counts for the relative area of each of the funiculi, a minor preferential localisation in the posterior funiculus appears, which is significant compared to the anterior funiculus ($P = 0.006$), but not compared to the lateral funiculus ($P = 0.07$). Interestingly, the tracts located in the posterior funiculus are involved in the transmission of information on fine touch, pressure and vibration, properties that are not tested in the clinical score and only mildly relevant for rotarod performance. It is therefore not surprising that when only the lesions in
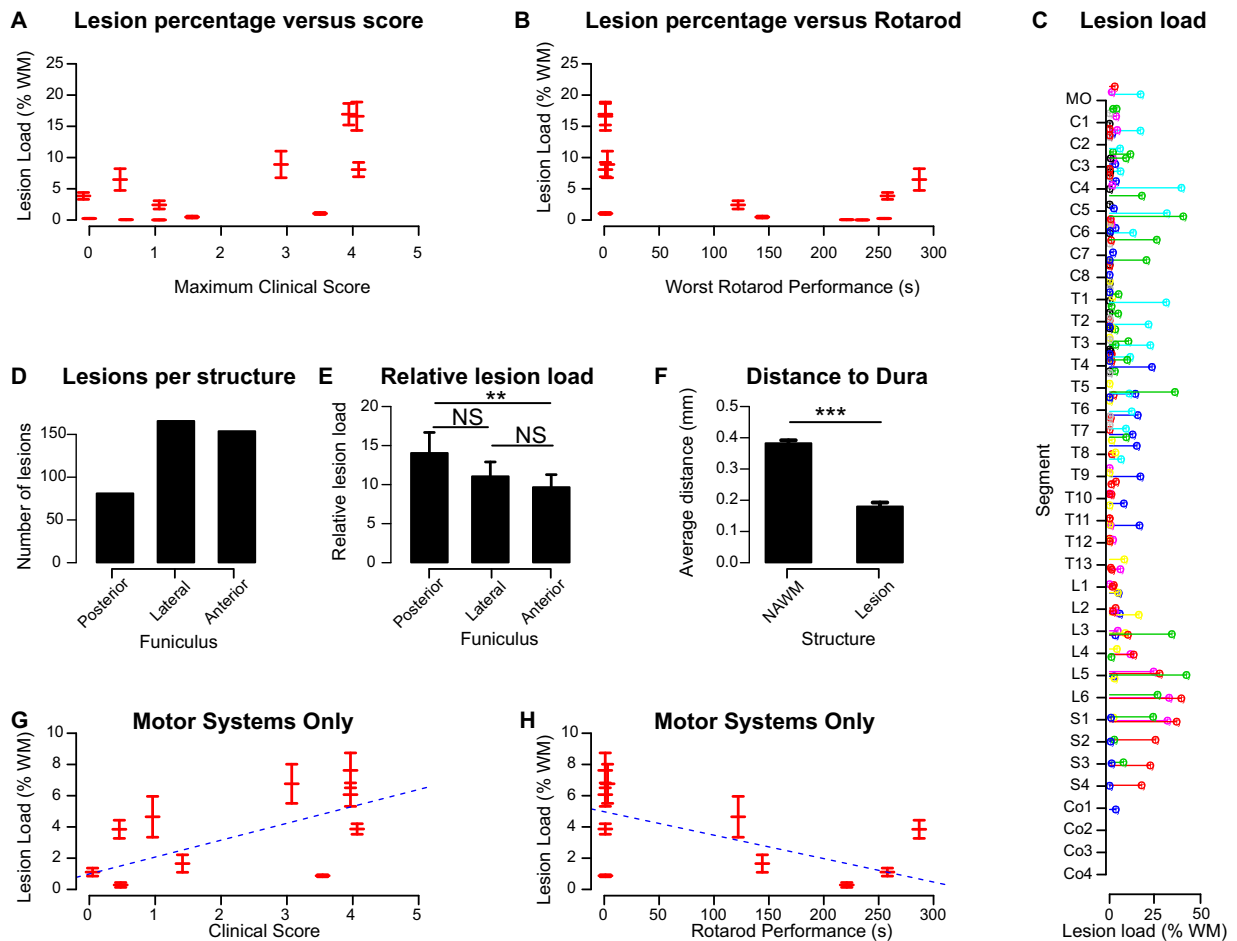
Fig. 5. Computational methods for calculating lesion load and distance to pia mater. A standard series of Klüver-Barrera stained paraffin sections was prepared for each animal, covering the entire available spinal cord at regular intervals (A_1). The outline of the spinal cord, the border between grey and white matter and all inflammatory lesions were traced. Within the area bounded by the traces, 250 points were sampled at random to function as starting points for distance calculation, as indicated with 'x' marks in (A_2). For area calculations, the point clouds obtained by tracing were converted to polygons (A_3). These point clouds also form the basis of the 3D reconstructions shown in (B) and (C), which show two spinal cord segments from the same animal. In the low lumbar and high sacral region, inflammatory lesions cover most of the lateral and anterior funiculus, while at low thoracic level, only a few lesions are visible.

the motor systems are considered, the correlation with both clinical score and rotarod remains the same, while the quality of the fit improves (Fig. 6H and I). No significant correlation exists between the amount of damage in the posterior funiculus and clinical score ($\rho = 0.35$, $P = 0.26$) or rotarod performance ($\rho = -0.23$, $P = 0.46$, plots not shown).

To determine the relation between the position of the inflammatory lesions and the distance to the outer border of the spinal cord formed by the pia mater, a number of random points were generated within the area of the lesions and the minimal distance to the pia mater was calculated for each point. A similar sampling was performed on the white matter (Fig. 5A_2). This method was chosen to take the irregular shape of the lesions into account, which would make a distance calculation based on centroid highly inaccurate. The average point within the inflammatory lesions is located 0.18 mm away from the pia mater, 2 times closer than the white matter as a whole (Fig. 6F, $P < 0.001$). This proximity to the pia mater is reflected by an increased distance to the boundary between white and grey matter, increased from 0.34 mm for the white matter as a whole to 0.57 mm for the lesions (Fig. 6G, $P < 0.001$). This subpial preferential localisation of lesions could suggest that meningeal inflammation plays a role in the pathogenesis of EAE, as has been suggested previously based on findings in the cerebral cortex (Errede et al., 2012; Kramann et al., 2015).

4. Discussion

Driven by the need for developing and evaluating neuroprotective therapies to ameliorate disease progression in MS, multiple attempts have been made to improve the assessment of neurological impairment seen in EAE (e.g. Buddeberg et al., 2004; Jones et al., 2008). The rotarod performance test provides an objective measurement of motor skills with a high level of inter-rater reliability. While the clinical score results in an ordinal variable, the rotarod provides a discrete numerical outcome (length of time till fall). In this study, we validated the usefulness of the rotarod to evaluate the course of EAE by comparing it to a commonly used clinical score and comparing both to the amount of damage observed in the spinal cord of the mice. Although there is an excellent negative correlation between neurological disability as determined using a clinical scale and the ability of the animal to maintain its balance on the rotarod, this relationship is decisively non-linear. It closely resembles a Gompertz curve, a commonly used mathematical model for a time series with slow growth at start and end of an interval, but fast growth in the middle (Gompertz, 1825; Winsor, 1932), used for example in Chatterjee et al. (2015). When provided with a negative exponent, this curve can be used to model a logistic decay process. Using this model, we find that motor performance barely decreases when neurological disability is scored below 1.5,
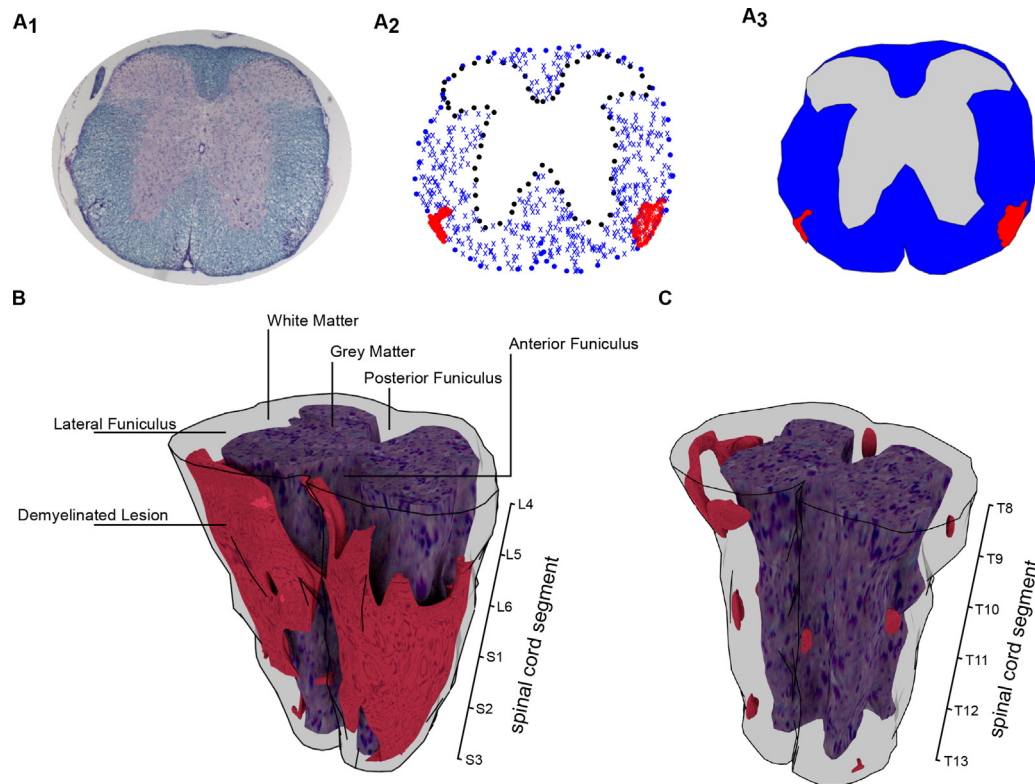
Fig. 6. Lesion load shows a high inter- and intraindividual variability. Animals with a higher percentage of the white matter covered with inflammatory lesions tend to have reached higher clinical scores (A) and showed a shorter rotarod latency time (B). However, there is a large variability in the relative percentage of lesions between animals as well as within the spinal cord of the same animal. (C) Distribution of relative lesion load along the spinal cord, each colour indicating one individual animal. MO = medulla oblongata. (D) Although every part of the white matter can be affected by inflammatory lesions, most lesions are present in the anterior and lateral funiculus, with a smaller number located in the posterior funiculus. (E) When corrected for the relative size of these anatomical structures, there appears to be a significantly higher lesion load in the posterior funiculus. Inflammatory lesions are preferentially located in the periphery of the white matter, in close proximity to the pia mater (F). G–H Recalculating the correlation between lesion load and clinical score or rotarod performance based solely on the damage in the ventral and lateral columns, containing the motor tracts, provides a similar correlation as shown in panels (A) and (B) with a better quality of fit. NAWM = normal appearing white matter. Error bars in (A) and (B) and (E)–(H) indicate standard error of the mean, asterisks in (E)–(F) indicate significance, $^{**} = p< 0.01$, $^{***} = p< 0.001$.

reflecting the fact that the criteria for these scores are mainly aimed at tail dysfunction. In contrast, the score interval from 1.5 to 2.5 shows a rapid decay of motor skills as measured using the rotarod. This part of the clinical disability scale seems to reflect a transitional state, since only a very small number of animals reach a maximum score in this interval (Fig. 3C). Animals either progress to a higher disability score or improve. If they progress and reach a clinical score of 3, the paralysis of the front or hind limbs will minimise their ability to use the rotarod. Therefore, while the rotarod provides an objective and quantitative measure of disability, a well-chosen clinical scale still has an added value in the intervals where the ability of the animal to walk the rotarod is either unaffected or, on the high end of the scale, very limited. However, we have also found that the relation between clinical score and the motor skills as measured using rotarod is not linear. A one-unit increase in clinical score is not reflected by a similar sized decrease in rotarod performance, strongly suggesting that the clinical score is an unequally spaced ordinal variable. Comparing mean clinical score between groups, as is commonly used in the EAE field, is based on the assumption that the measured categories reflect an underlying continuous variable. Since this is assumption is apparently not valid, analysis of clinical score using statistical tools developed for numerical variables can lead to artefacts (Bollen and Barb, 1981).

The majority of studies published on EAE focuses on disease intensity only. Clearly defined disease milestones (see Table 3) allow quantifying the time course of EAE as well as its severity. Both a clinical scale and the rotarod performance are useful to this end. The clinical score is more sensitive to mild cases of EAE not affecting motor performance, where the animal's ability to walk the rotarod is not affected. The rotarod on the other hand is better suitable to monitor the slow recovery experienced by most animals. This method of analysing disease course is not unlike the common practice of measuring MS progression over time using Kurtzke's Expanded Disability Status Scale (EDSS) (Kurtzke, 1983). Interestingly, several recent studies have suggested that a clinical test with a high similarity to the rotarod, gait analysis, could provide a prognosis of disability progression in MS (Hamilton et al., 2009; Flegel et al., 2012; Socie et al., 2013).

As in MS patients, we observed a large gender disparity in EAE severity in our mice. This gap only affects the magnitude of disability, not its progression over time. A similar effect was recently reported in a rat EAE model (Nacka-Aleksić et al., 2015). Although gender might be an unwanted confounder in most EAE experiments, the fact that a gender gap exists in EAE could be used to study the still unexplained female predominance seen in MS. It could also help to study the phenomenon that the reported female to male ratio in MS patients has been constantly increasing over time (Dunn and Steinman, 2013). The available research suggest that at least part of the difference is caused by hormonal influences, with female animals being more severely affected depending on the phase of their cycle (Rahn et al., 2014). Other studies have pointed at the role of genes on the sex chromosomes, both in the central nervous system and in the immune system (Du et al., 2014). Although including both genders in EAE experiments has the disadvantage of increased variability within groups, it is a necessary step to enhance our knowledge on sex differences in MS.

Furthermore, in some countries regulatory bodies object against using only a single gender, necessitating sacrifice without experiment of the second gender.

The second major finding of this study provides a method for quantifying the amount of CNS damage in EAE. To this end, inflammatory lesions are measured while taking their three-dimensional surroundings into account, preserving information on all spinal cord segments and the and anatomical structures these lesions occupy. In most EAE studies, the number or relative area of inflammatory lesions is used as a measure for CNS pathology. However, a small lesion located in a critical system like the corticospinal tract will contribute disproportionally to disability. On the other hand a large lesion in a system with more redundancy such as the spinocerebellar tracts might not result in any clinical signs.

We now demonstrate a correlation between the magnitude of disease as measured using a clinical scale or rotarod performance and damage in the spinal cord. Importantly, this correlation is entirely based on damage to the lateral and ventral spinal cord. These parts contain amongst others the fibres forming the pyramidal and extrapyramidal tracts, responsible for controlling motor functions. Lesions in the posterior system, mainly involved in conveying sensory information, did not contribute to the correlation between neurological disability and histological damage. We suggest that adding a measure for sensory disability to the clinical test battery will provide a better estimate of CNS pathology. Several studies have been performed on neuropathic pain in EAE (Tian et al., 2013) and various sensory deficits have been reported to occur (Thibault et al., 2011). However, such measurements cause considerable stress to the animals. There is no method currently available for measuring vibration sense in rodents, although a recently published quantitative measurement using paw retraction in response to cold (Duraku et al., 2014) could be used. However, such a measurement would presumably be subjected to the same limitation as the rotarod. Animals which are too sick to walk the rod will also be unable to retract their paws in response to stimuli, prohibiting use in severely affected animals.

As previously suggested by other authors (Jones et al., 2008), the dorsal column of the spinal cord sustains a relatively large amount of damage compared to its size. Various explanations can be brought forward and should be explored, such as proximity to blood vessels, organisation of the meninges, and a possible preferential localisation in ascending versus descending fibre tracts. However, this relatively small preference might also be related to a higher axon density in the dorsal column compared to the ventral and lateral columns, as observed using quantitative MRI (Cohen-Adad and Wheeler-Kingshott, 2014). Inflammatory lesions are preferentially located near the edge of the white matter, close to the pia mater, supporting the hypothesis that the meninges can function as an entrance point for immune cell infiltration into the central nervous system (Polfliet et al., 2002; Walker-Caulfield et al., 2015). The peripheral accumulation of damage might also explain the frequent observation that disability starts in the hind limbs of animals. Several key systems in the spinal cord, including the pyramidal tract, share a somatotopical organisation in which the fibres from the sacral spinal cord segments are located on the outside and fibres from the cervical cord on the inside. The subpial localisation of lesions suggests that the function of the axons associated with the sacral and lumbar spinal cord segments is impaired first, while the thoracic and cervical fibres are preserved.

Overall, the workflow outlined in this paper provides a systematic method of quantifying neurological disability in EAE, both in magnitude and in terms of disease course. The severity of disease can be correlated to the area of inflammatory lesions found in the motor systems of the spinal cord. Our findings also emphasize the importance of gender as an influential factor, determining disease severity in EAE as it does in MS. Finally, we have shown that when studying damage to the CNS, lesion size and location is vital in order to link clinical signs to histopathology. In this manner, the vast immunological experience acquired using EAE can be combined with knowledge and techniques from the fields of neuroscience and experimental neurology, further upgrading EAE into a more accurate model for studying the neurodegenerative aspects of MS.

Acknowledgements

References

Al-Izki S, Pryce G, O'Neill JK, Butter C, Giovannoni G, Amor S, Baker D. Practical guide to the induction of relapsing progressive experimental autoimmune encephalomyelitis in the Biozzi ABH mouse. Mult Scler Relat Disord 2012;1(1):29–38, http://dx.doi.org/10.1016/j.msard.2011.09.001.

Baker D, Amor S. Mouse models of multiple sclerosis: lost in translation? Curr Pharm Des 2015;21(18):2440–52.

Baker D, Gerritsen W, Rundle J, Amor S. Critical appraisal of animal models of multiple sclerosis. Mult Scler 2011;17(6):647–57.

Baker D, Lidster K, Sottomayor A, Amor S. Two years later: journals are not yet enforcing the arrive guidelines on reporting standards for pre-clinical animal studies. PLoS Biol 2014;12(1):e1001756.

Baxter AG. The origin and application of experimental autoimmune encephalomyelitis. Nat Rev Immunol 2007;7(11):904–12.

Blender Online Community. Blender – a 3D modelling and rendering package. Amsterdam: Blender Foundation, Blender Institute; 2015 http://www.blender.org.

Bollen K, Barb K. Pearson's r and coarsely categorized measures. Am Sociol Rev 1981;46(2):232–9.

Bove R, Chitnis T. Sexual disparities in the incidence and course of MS. Clin Immunol 2013;149(2):201–10.

Bove R, Chitnis T. The role of gender and sex hormones in determining the onset and outcome of multiple sclerosis. Mult Scler 2014;20(5):520–6.

Budde MD, Xie M, Cross AH, Song SK. Axial diffusivity is the primary correlate of axonal injury in the experimental autoimmune encephalomyelitis spinal cord: a quantitative pixelwise analysis. J Neurosci 2009;29(9):2805–13.

Buddeberg BS, Kerschensteiner M, Merkler D, Stadelmann C, Schwab ME. Behavioral testing strategies in a localized animal model of multiple sclerosis. J Neuroimmunol 2004;153(1–2):158–70.

Chatterjee T, Chatterjee BK, Majumdar D, Chakrabarti P. Antibacterial effect of silver nanoparticles and the modeling of bacterial growth kinetics using a modified Gompertz model. Biochim Biophys Acta 2015;1850(2):299–306.

Cohen-Adad J, Wheeler-Kingshott C. Quantitative MRI of the spinal cord. Elsevier Science; 2014 https://books.google.nl/books?id=VjBnAgAAQBAJ.

Du S, Itoh N, Askarinam S, Hill H, Arnold AP, Voskuhl RR. Xy sex chromosome complement, compared with xx, in the CNS confers greater neurodegeneration during experimental autoimmune encephalomyelitis. Proc Natl Acad Sci U S A 2014;111(7):2806–11.

Dunn SE, Steinman L. The gender gap in multiple sclerosis: intersection of science and society. JAMA Neurol 2013;70(5):634–5.

Duraku LS, Niehof SP, Misirli Y, Everaers M, Hoendervangers S, Holstege J, Boele HJ, Koekkoek SK, Smits ES, Selles RW, Walbeehm ET. Rotterdam advanced multiple plate: a novel method to measure cold hyperalgesia and allodynia in freely behaving rodents. J Neurosci Methods 2014;224:1–12.

Ebers GC. Natural history of multiple sclerosis. J Neurol Neurosurg Psychiatry 2001;71(Suppl 2), ii16–9.

Errede M, Girolamo F, Ferrara G, Strippoli M, Morando S, Boldrin V, Rizzi M, Uccelli A, Perris R, Bendotti C, Salmona M, Roncali L, Virgintino D. Blood-brain barrier alterations in the cerebral cortex in experimental autoimmune encephalomyelitis. J Neuropathol Exp Neurol 2012;71(10):840–54.

Flegel M, Knox K, Nickel D. Step-length variability in minimally disabled women with multiple sclerosis or clinically isolated syndrome. Int J MS Care 2012;14(1):26–30.

Fleming KK, Bovaird JA, Mosier MC, Emerson MR, LeVine SM, Marquis JG. Statistical analysis of data from studies on experimental autoimmune encephalomyelitis. J Neuroimmunol 2005;170(1-2):71–84.

Friese MA, Fugger L. Pathogenic CD8(+) t cells in multiple sclerosis. Ann Neurol 2009;66(2):132–41.

Gold R, Linington C, Lassmann H. Understanding pathogenesis and therapy of multiple sclerosis via animal models: 70 years of merits and culprits in experimental autoimmune encephalomyelitis research. Brain 2006;129(Pt 8):1953–71.

Gompertz B. On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. Philos Trans R Soc London 1825;115:513–85.

Goodin DS. The epidemiology of multiple sclerosis: insights to disease pathogenesis. Handb Clin Neurol 2014;122:231–66.

Grothendieck G. nls2: Non-linear regression with brute force, r package version 0.2; 2013 http://CRAN.R-project.org/package=nls2.

Hamilton F, Rochester L, Paul L, Rafferty D, O'Leary CP, Evans JJ. Walking and talking: an investigation of cognitive-motor dual tasking in multiple sclerosis. Mult Scler 2009;15(10):1215–27.

Jones BJ, Roberts DJ. The quantitative measurement of motor in co-ordination in naive mice using an accelerating rotarod. J Pharm Pharmacol 1968;20(4):302–4.

Jones MV, Nguyen TT, Deboy CA, Griffin JW, Whartenby KA, Kerr DA, Calabresi PA. Behavioral and pathological outcomes in MOG 35–55 experimental autoimmune encephalomyelitis. J Neuroimmunol 2008;199(1-2):83–93.

Källen B, Nilsson O. Dissociation between histological and clinical signs of experimental auto-immune encephalomyelitis. Acta Pathol Microbiol Immunol Scand A 1986;94(2):159–64.

Kassiotis G, Pasparakis M, Kollias G, Probert L. Tnf accelerates the onset but does not alter the incidence and severity of myelin basic protein-induced experimental autoimmune encephalomyelitis. Eur J Immunol 1999;29(3):774–80.

Kim JH, Budde MD, Liang H-F, Klein RS, Russell JH, Cross AH, Song S-K. Detecting axon damage in spinal cord from a mouse model of multiple sclerosis. Neurobiol Dis 2006;21(3):626–32, http://dx.doi.org/10.1016/j.nbd.2005.09.009 http://www.sciencedirect.com/science/article/pii/S0969996105002500.

Klüver H, Barrera E. A method for the combined staining of cells and fibers in the nervous system. J Neuropathol Exp Neurol 1953;12(4):400–3.

Koritschoner R, Schweinburg F. Induktion von paralyse und rückenmarksentzündung durch immunisierung von kaninchen mit menschlichem rückenmarksgewebe. Z Immunitäts Exp Therapie 1925;42:217–83.

Kramann N, Neid K, Menken L, Schlumbohm C, Stadelmann C, Fuchs E, et al. Increased meningeal t and plasma cell infiltration is associated with early subpial cortical demyelination in common marmosets with experimental autoimmune encephalomyelitis. Brain Pathol 2015;25(3):276–86.

Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (edss). Neurology 1983;33(11):1444–52.

Lassmann H. Heterogeneity of multiple sclerosis: implications for therapy targeting regeneration. Ernst Schering Res Found Workshop 2005;53:11–22.

Massella A, D'Intino G, Fernández M, Sivilia S, Lorenzini L, Giatti S, Melcangi RC, Calzà L, Giardino L. Gender effect on neurodegeneration and myelin markers in an animal model for multiple sclerosis. BMC Neurosci 2012;13:12.

Mix E, Meyer-Rienecker H, Hartung HP, Zettl UK. Animal models of multiple sclerosis-potentials and limitations. Prog Neurobiol 2010;92(3):386–404.

Moore SM, Khalaj AJ, Kumar S, Winchester Z, Yoon J, Yoo T, Martinez-Torres L, Yasui N, Katzenellenbogen JA, Tiwari-Woodruff SK. Multiple functional therapeutic effects of the estrogen receptor lγ agonist indazole-Cl in a mouse model of multiple sclerosis. Proc Natl Acad Sci U S A 2014;111(50):18061–6, http://dx.doi.org/10.1073/pnas.1411294111.

Nacka-Aleksić M, Djikić J, Pilipović I, Stojić-Vukanić Z, Kosec D, Bufan B, et al. Male rats develop more severe experimental autoimmune encephalomyelitis than female rats: sexual dimorphism and diergism at the spinal cord level. Brain Behav Immun 2015;9:101–18.

Paz Soldán MM, Raman MR, Gamez JD, Lohrey AK, Chen Y, Pirko I, et al. Correlation of brain atrophy, disability, and spinal cord atrophy in a murine model of multiple sclerosis. J Neuroimaging 2015;25(4):595–9.

Polfliet MM, Van de Veerdonk F, Döpp EA, Van Kesteren-Hendrikx EM, Van Rooijen N, Dijkstra CD, Van den Berg TK. The role of perivascular and meningeal macrophages in experimental allergic encephalomyelitis. J Neuroimmunol 2002;122(1-2):1–8.

R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2014 http://www.R-project.org/.

Rahn EJ, Iannitti T, Donahue RR, Taylor BK. Sex differences in a mouse model of multiple sclerosis: neuropathic pain behavior in females but not males and protection from neurological deficits during proestrus. Biol Sex Differ 2014;5(1):4.

Ransohoff RM, Hafler DA, Lucchinetti CF. Multiple sclerosis – a quiet revolution. Nat Rev Neurol 2015;11(5):246.

Ransohoff RM. Animal models of multiple sclerosis: the good, the bad and the bottom line. Nat Neurosci 2012;15(8):1074–7.

Rocca MA, Messina R, Filippi M. Multiple sclerosis imaging: recent advances. J Neurol 2013;260(3):929–35, http://dx.doi.org/10.1007/s00415-012-6788-8.

Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez J-Y, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A. Fiji: an open-source platform for biological-image analysis. Nat Methods 2012;9(7):676–82, http://dx.doi.org/10.1038/nmeth.2019.

Schwendimann RN, Alekseeva N. Gender issues in multiple sclerosis. Int Rev Neurobiol 2007;79:377–92.

Smith AJ, Liu Y, Peng H, Beers R, Racke MK, Lovett-Racke AE. Comparison of a classical th1 bacteria versus a th17 bacteria as adjuvant in the induction of experimental autoimmune encephalomyelitis. J Neuroimmunol 2011;237(1-2):33–8.

Socie MJ, Motl RW, Pula JH, Sandroff BM, Sosnoff JJ. Gait variability and disability in multiple sclerosis. Gait Posture 2013;38(1):51–5.

Sriram S, Steiner I. Experimental allergic encephalomyelitis: a misleading model of multiple sclerosis. Ann Neurol 2005;58(6):939–45.

Steinman L, Zamvil SS. Virtues and pitfalls of EAE for the development of therapies for multiple sclerosis. Trends Immunol 2005;26(11):565–71.

Steinman L, Zamvil SS. How to successfully apply animal studies in experimental allergic encephalomyelitis to research on multiple sclerosis. Ann Neurol 2006;60(1):12–21.

Stromnes IM, Goverman JM. Active induction of experimental allergic encephalomyelitis. Nat Protoc 2006;1(4):1810–9.

Stys PK, Zamponi GW, Van Minnen J, Geurts JJ. Will the real multiple sclerosis please stand up? Nat Rev Neurosci 2012;13(7):507–14.

't Hart BA, Gran B, Weissert R. EAE: imperfect but useful models of multiple sclerosis. Trends Mol Med 2011;17(3):119–25.

Thibault K, Calvino B, Pezet S. Characterisation of sensory abnormalities observed in an animal model of multiple sclerosis: a behavioural and pharmacological study. Eur J Pain 2011;15(3), 231.e1-16.

Tian DH, Perera CJ, Moalem-Taylor G. Neuropathic pain in animal models of nervous system autoimmune diseases. Mediat Inflamm 2013;2013:298326.

Tremlett H, Yinshan Z, Devonshire V. Natural history of secondary-progressive multiple sclerosis. Mult Scler 2008;14(3):314–24.

Tsilidis KK, Panagiotou OA, Sena ES, Aretouli E, Evangelou E, Howells DW, Al-Shahi Salman R, Macleod MR, Ioannidis JP. Evaluation of excess significance bias in animal studies of neurological diseases. PLoS Biol 2013;11(7):e1001609.

Walker-Caulfield ME, Hatfield JK, Brown MA. Dynamic changes in meningeal inflammation correspond to clinical exacerbations in a murine model of relapsing-remitting multiple sclerosis. J Neuroimmunol 2015;278:112–22.

Watson C, Paxinos G, Kayalioglu G, editors. The spinal cord. San Diego: Academic Press; 2009.

Winsor CP. The gompertz curve as a growth curve. Proc Natl Acad Sci U S A 1932;18(1):1–8.