



University of Thessaly
School of Health Sciences - Faculty of Medicine
Master of Science Thesis

**“Assessment of the reporting quality of observational
Studies about response to treatment with interferons in patients with
RRMS published from 2015 to 2018
using the STROBE statement”**

by

Gkotzamanis Viktor

**«Αξιολόγηση της ποιότητας αναφοράς μελετών παρατήρησης σχετικά
με την ανταπόκριση στη θεραπεία με ιντερφερόνες σε ασθενείς με
RRMS δημοσιευμένων από το 2015 έως το 2018 με τη χρήση της
δήλωσης STROBE»**

Γκοτζαμάνης Βίκτωρας , Ιατρός

Τριμελής Επιτροπή:

**Στεφανίδης Ιωάννης, Καθηγητής Πανεπιστημίου Θεσσαλίας, Τμήμα
Ιατρικής (επιβλέπων)**

**Δοξάνη Χρυσούλα, Επιστημονικός Συνεργάτης Πανεπιστημίου
Θεσσαλίας**

**Ζιντζαράς Ηλίας, Καθηγητής Πανεπιστημίου Θεσσαλίας, Τμήμα
Ιατρικής**

MSc Program:

**“Research Methodology in Biomedicine,
Biostatistics & Clinical Bioinformatics”**

Larissa, September 2018

Contents

| | |
|-------------------------------|------------------|
| 1. <u>Summary</u> | <u>3</u> |
| 2. <u>Introduction</u> | <u>4</u> |
| 3. <u>Methods</u> | <u>10</u> |
| 4. <u>Results</u> | <u>11</u> |
| 5. <u>Discussion</u> | <u>15</u> |
| 6. <u>Conclusion</u> | <u>16</u> |
| 7. <u>References</u> | <u>16</u> |

Περίληψη

ΣΤΟΧΟΣ: Να γίνει ποιοτική αξιολόγηση μελετών παρατήρησης σχετικά με την ανταπόκριση στη θεραπεία με ιντερφερόνες σε ασθενείς πάσχοντες από RRMS, από το 2015 ως το 2018, χρησιμοποιώντας το STROBE checklist

Μέθοδοι: Πραγματοποιήθηκε έρευνα στην πλατφόρμα PubMed και επιλέχθηκαν 9 έρευνες σχετικές με το θέμα μας δημοσιευμένες μεταξύ 2015 και 2018. Η συμμόρφωση των μελετών με τις οδηγίες ποσοτικοποιήθηκε χρησιμοποιώντας ένα σύστημα βαθμολόγησης που έδινε 1 βαθμό για κάθε αντικείμενο του ερωτηματολογίου που έπαιρνε την απάντηση «ΝΑΙ» και 0 για κάθε «ΟΧΙ». Στη συνέχεια τα άρθρα χωρίστηκαν σε 2 ομάδες ανάλογα με το impact factor του περιοδικού όπου δημοσιεύτηκαν και οι μέσες τιμές των βαθμολογιών συμμόρφωσης συγκρίθηκαν με ένα t-test για ανεξάρτητα δείγματα. Τέλος τα αποτελέσματα συγκρίθηκαν με αυτά μιας παρόμοιας μελέτης που αξιολογούσε την ποιότητα μελετών δημοσιευμένων στο Journal of Hand Surgery χρησιμοποιώντας το STROBE checklist

Αποτελέσματα: Η συνολική συμμόρφωση ήταν 64,6% (95% CI 50%-79%). Οι βαθμοί που χάθηκαν ήταν κυρίως αποτέλεσμα της μη αναφοράς στην περίληψη του τύπου της μελέτης (55%), μη επισήμανσης πιθανού συστηματικού σφάλματος που μπορεί να υπεισέρχεται στη μελέτη (22%), παράλειψης αναφοράς του τρόπου με τον οποίο αποφασίστηκε το μέγεθος του δείγματος (0%), και μη παρουσίασης των πηγών χρηματοδότησης. Οι μελέτες που είχαν δημοσιευτεί σε περιοδικά με impact factor πάνω από 3 είχαν κατά 6,14 βαθμούς υψηλότερη μέση βαθμολογία συμμόρφωσης (p value=0,04, 95%CI 2,8-9,4). Η συμμόρφωση που παρατηρήθηκε δεν διέφερε στατιστικά σημαντικά από αυτήν που αναφερόταν στη μελέτη του Journal of Hand Surgery για την περίοδο του 2016.

Συμπέρασμα: Η συμμόρφωση με τις κατευθυντήριες οδηγίες STROBE φάνηκε να είναι σε καλό επίπεδο, δείχνοντας μια ανοδική τάση στην πάροδο του χρόνου. Η καλύτερη συμμόρφωση φαίνεται να σχετίζεται με δημοσίευση σε περιοδικό με καλύτερο impact factor. Ωστόσο, φαίνεται να υπάρχει ακόμα χώρος για βελτίωση ειδικά σε ό,τι έχει να κάνει με την επισήμανση πιθανού συστηματικού σφάλματος καθώς και με την παρουσίαση του τρόπου που αποφασίζεται το μέγεθος του δείγματος

Abstract

PURPOSE: To use the Strengthening The Reporting of Observational studies in Epidemiology (STROBE) checklist to evaluate the quality of observational studies regarding the response to treatment with interferons in patients with RRMS from 2015 to 2018

METHODS: PubMed research was conducted and 9 reports of observational studies regarding our topic published between 2015 and 2018 were selected. The compliance of the reports was quantified using a point system that awarded 1 point for every “YES” in the STROBE checklist and 0 for every “NO”. The mean compliance scores were then compared, after grouping the reports based on the impact factor of the journal they were published in, using independent sample t-test. Ultimately the results were compared to those of a similar study assessing the quality of reports published in the Journal of Hand Surgery

RESULTS: Overall compliance was 64,6% (95% CI 50%-79%). The missing points were a result mostly of not indicating the study design in the abstract (55%), not addressing potential bias (22%), not mentioning how the size of the sample was arrived at (0%) and not presenting the sources of their funding (0%). Reports published in journals with impact factor higher than 3 were found to have 6,14 points higher compliance score (p value=0,04, 95% CI 2,8-9,4). Compliance was found not to differ statistically significantly from that of the 2011 period in the Journal of Hand Surgery

Conclusion: Compliance with the STROBE guidelines seems quite good, showing a rising trend through the years. It is evident that better compliance is associated with a publication in a journal with higher impact factor. However, there is still room for improvement especially as regards addressing sources of potential bias and describing how the size of the study was decided

Introduction

Much of our knowledge about various medical conditions, risk factors and associations between exposures and clinical outcomes comes from observational studies. In this sense, it is obvious how critical it is to ensure the high quality of observational studies since they provide the basis for understanding the pathophysiological mechanisms underlying medical conditions and practice of evidence based medicine. The Strengthening the Reporting of Observation Studies in Epidemiology (STROBE) statement was produced in 2007 to “improve the quality of observational study reporting, improve transparency in reporting, and allow for critical assessment by others of the strengths and weaknesses in study design, conduct, and analysis. A team of 23 editors, epidemiologists, methodologists, statisticians, and practitioners from Europe and North America created this statement. The STROBE statement provides a 22-item checklist of items for inclusion in the reporting of all

observational studies including components of the study design, interventions, data collection, analytic techniques, and potential bias. Available checklists are specific for each of the 3 observational study designs (i.e., cohort, case-control, and cross-sectional). Use of these checklists is intended to improve the reader's ability to assess, interpret, and generalize study findings. (1)

MS is a neurological disease affecting an estimated number of 2,500,000 people worldwide (2) with important impact on the quality of the life of the patients. RRMS is the most common and well-studied form of MS with a lot of research still ongoing. There is already a number of approved disease-modifying treatments applied to many patients. Interferons are first disease modifying therapy approved by the FDA in 1993. They were shown to decrease the frequency of relapses, the MRI lesion burden as well as the disability caused by the disease and set the standard to which the later treatments compared. (3) However, despite their proven efficacy and safety, the response of the patients is still quite heterogeneous for reasons that remain unknown.

Evaluation of the response to treatment with the means available today in the everyday clinical practice such as MRI imaging of new lesions and clinical evaluation (EDSS score), is quite accurate, however the time needed for the discrimination of the optimal and suboptimal respondents is 1 to 2 years. This time interval is quite critical for the progression as initial disease course is regarded as an important prognostic factor, so that non-respondents have increased risk of facing more relapses and more severe disease course. Ideally, we would like to have biomarkers that even before the initiation of the treatment would indicate those patients that are expected to respond optimally and those would have to seek alternative treatment in order to avoid losing any precious time. That is why studies aiming in this direction, investigating the response to the treatment with interferons, were chosen to be evaluated for the purposes of this thesis.

STROBE Statement—checklist of items that should be included in reports of observational studies

| | Item No | Recommendation |
|---------------------------|----------------|---|
| Title and abstract | 1 | (a) Indicate the study's design with a commonly used term in the title or the abstract |
| | | (b) Provide in the abstract an informative and balanced summary of what was done and what was found |
| Introduction | | |

| | | |
|----------------------|---|--|
| Background/rationale | 2 | Explain the scientific background and rationale for the investigation being reported |
| Objectives | 3 | State specific objectives, including any prespecified hypotheses |
| Methods | | |
| Study design | 4 | Present key elements of study design early in the paper |
| Setting | 5 | Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection |
| Participants | 6 | <p><i>(a) Cohort study</i>—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up</p> <p><i>Case-control study</i>—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls</p> <p><i>Cross-sectional study</i>—Give the eligibility criteria, and the sources and methods of selection of participants</p> <hr/> <p><i>(b) Cohort study</i>—For matched studies, give matching criteria and number of exposed and unexposed</p> <p><i>Case-control study</i>—For matched studies, give matching criteria and the number of controls per case</p> |
| Variables | 7 | Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable |

| | | |
|---------------------------|-----|---|
| Data sources/ measurement | 8* | For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group |
| Bias | 9 | Describe any efforts to address potential sources of bias |
| Study size | 10 | Explain how the study size was arrived at |
| Quantitative variables | 11 | Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why |
| Statistical methods | 12 | <p>(a) Describe all statistical methods, including those used to control for confounding</p> <p>(b) Describe any methods used to examine subgroups and interactions</p> <p>(c) Explain how missing data were addressed</p> <p>(d) <i>Cohort study</i>—If applicable, explain how loss to follow-up was addressed</p> <p><i>Case-control study</i>—If applicable, explain how matching of cases and controls was addressed</p> <p><i>Cross-sectional study</i>—If applicable, describe analytical methods taking account of sampling strategy</p> <p>(e) Describe any sensitivity analyses</p> |
| Results | | |
| Participants | 13* | (a) Report numbers of individuals at each stage of study—eg numbers potentially |

| | | |
|------------------|-----|---|
| | | <p>eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed</p> <hr/> <p>(b) Give reasons for non-participation at each stage</p> <hr/> <p>(c) Consider use of a flow diagram</p> |
| Descriptive data | 14* | <p>(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders</p> <hr/> <p>(b) Indicate number of participants with missing data for each variable of interest</p> <hr/> <p>(c) <i>Cohort study</i>—Summarise follow-up time (eg, average and total amount)</p> |
| Outcome data | 15* | <p><i>Cohort study</i>—Report numbers of outcome events or summary measures over time</p> <hr/> <p><i>Case-control study</i>—Report numbers in each exposure category, or summary measures of exposure</p> <hr/> <p><i>Cross-sectional study</i>—Report numbers of outcome events or summary measures</p> |
| Main results | 16 | <p>(a) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included</p> <hr/> <p>(b) Report category boundaries when continuous variables were categorized</p> <hr/> <p>(c) If relevant, consider</p> |

| | | |
|--------------------------|----|--|
| | | translating estimates of relative risk into absolute risk for a meaningful time period |
| Other analyses | 17 | Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses |
| Discussion | | |
| Key results | 18 | Summarise key results with reference to study objectives |
| Limitations | 19 | Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias |
| Interpretation | 20 | Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence |
| Generalisability | 21 | Discuss the generalisability (external validity) of the study results |
| Other information | | |
| Funding | 22 | Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based |

Methods

PubMed research was conducted with keywords “rrms”, “interferons” and “response”. The publication date filter was set from 2015 to the present day. From the 29 given results 9 were reports of observational studies and truly relevant to our topic. These 9 articles were evaluated using the STROBE checklist. For each item of the checklist answered with “YES” 1 point was given to the article so that an article that was fully compliant with the STROBE guidelines would have a total of 22 points. Some items consisted of more than one bullets with each bullet requiring more than one pieces of information. For items like this, the approach used was that the point was awarded to the articles that provided most of the required information with no significant omissions. This was decided so that no item would gain extra impact on the overall score, as would happen if extra points were given for every sub-item. The scores were then transformed to compliance rates using the SPSS software.

The data was passed to Excel sheet so that total score per item and per article could be calculated and presented (table 1)

Then the articles were split into 2 groups based on the impact factor of the journal they were published in so that we could investigate if there is an association between that and the compliance with the STROBE guidelines. The cutoff point was set to 3 so that the first group consisted of articles published in journals with impact factor <3 and the second >3

Ultimately, the results were compared with those of a similar report of quality evaluation of observational trials in the Journal of Hand Surgery.

Statistical Analysis

The statistical analysis was conducted using the SPSS software. For the first part, the mean values of their score in the strobe checklist of the two groups were compared, using independent samples t-test. Normality of the distribution of the data was checked with Shapiro-Wilk and Kolmogorov-Smirnov tests. The level of significance was set to 95% (p value 5%)

Comparison with the results of the report in the Journal of Hand Surgery were based on the 95% CI of the rates of the compliance scores.

| article published in | 1 neuroim | 2 neuroim | 3 neuroim | 4 neur sci | 5 eur neur | 6 neur sci | 7 j neurol | 8 neuroim | 9 neur sci | total per item | |
|----------------------|--------------|--------------|--------------|---------------|---------------|---------------|---------------|--------------|---------------|----------------|---|
| impact factor | 2.959 | 2.959 | 2.959 | 2.353 | 3.692 | 2.353 | 3.470 | 2.959 | 2.353 | | |
| checklist number | | | | | | | | | | | |
| abstract | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 1 | 1 | 5 |
| Introduction | 2 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 9 |
| | 3 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 1 | 8 |
| methods | 4 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 5 |
| | 5 | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 6 |
| | 6 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 8 |
| | 7 | 1 | 0 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 7 |
| | 8 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 8 |
| | 9 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 2 |
| | 10 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | 11 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 6 |
| | 12 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 9 |
| | results | 13 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 14 | | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 7 |
| 15 | | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 4 |
| 16 | | 1 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 | 4 |
| 17 | | 1 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 3 |
| Discussion | 18 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 8 |
| | 19 | 0 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 4 |
| | 20 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 8 |
| | 21 | 0 | 1 | 0 | 0 | 1 | 1 | 1 | 1 | 1 | 6 |
| | 22 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| total per article | 15 | 10 | 11 | 7 | 19 | 17 | 19 | 13 | 15 | | |

Table 1

Results

Overall compliance of the studies evaluated was 64,6% (128 out of 198 max points), 95% CI (50%-79%) with range from 31% to 86%. Compliance rate per section can be seen in table 2

Abstract

The overall compliance of the studies with the STROBE guidelines as regards to the abstract was 55%. All the articles presented an informative abstract that summarized the purposes, methods, results and conclusions of the study, however only three of them indicated the study design with a commonly used term. Given that the first item of the bullet consisted of two bullets the one point was awarded to the articles that either covered both bullets or had a very well balanced abstract of what was to follow.

Introduction

Introduction was the area with the best compliance with an overall 94%(17 points out of 18 total items). The one missing point was due to an article not stating beforehand the hypotheses that were to be explored.

Methods

This area had intermediate compliance with an overall compliance of 61% (51 points out of 81 total items). This section includes items that require more than one pieces of specific information such as item number 5 about Setting, that requires setting, location, relevant dates, periods of recruitment, exposure, follow up and data collection. Answering an item like this with “yes” or “no” is not that simple as it cannot reflect the amount of the required information that was really provided. For items like this the one point was awarded when most of the requirements were met and no important detail was left unmentioned.

The item-by-item compliance was very heterogeneous ranging from 100% for the statistical analysis to 0% for the study size. Another item with very little compliance was number 9 regarding bias as only 22%(2 out of 9) made an attempt to address potential sources of bias due to the methods selected.

Results

The overall compliance of the results section was 60% (27 points out of 45 items) ranging from 100% (9 out 9 for item number 13) to 33% (3 out 9 for item number 27). Most studies presented enough data regarding the participants (demographics, numbers at each stage, follow up, etc.) scoring 16 out of 18 in the first two items. However, they did not report with the same consistency outcome data and main results omitting often to present unadjusted estimates and other analyses like analyses of subgroups etc.

Discussion

Discussion section had an overall compliance of 57% (26 points out of 45 items). Ranking of this section had a clear pattern as most articles summarized key results and presented a cautious approach in the interpretation of the results often mentioning the need for further research, scoring a total of 16 out 18 in these two items. However, only 4 of them included

in the discussion the existence of potential bias and its sources. Moreover, it is worth mentioning that none presented the sources of its funding.

| published in | neuroim | neuroim | neuroim | neur sci | eur neur | neur sci | j neurol | neuroim | neur sci | Total per Section |
|-------------------|---------|---------|---------|----------|----------|----------|----------|---------|----------|-------------------|
| impact factor | 2.959 | 2.959 | 2.959 | 2.353 | 3.692 | 2.353 | 3.470 | 2.959 | 2.353 | |
| Abstract | 0% | 0% | 0% | 100% | 100% | 100% | 0% | 100% | 100% | 55% |
| Introduction | 100% | 100% | 100% | 100% | 100% | 100% | 100% | 50% | 100% | 94% |
| Methods | 66% | 44% | 44% | 22% | 89% | 77% | 89% | 66% | 66% | 62% |
| Results | 100% | 20% | 40% | 40% | 80% | 80% | 100% | 40% | 40% | 60% |
| Discussion | 40% | 60% | 60% | 0% | 80% | 60% | 80% | 60% | 80% | 57% |
| Total per article | 68% | 45% | 50% | 31% | 86% | 77% | 86% | 59% | 68% | |

Table 2

Comparison Based on Impact factor

The heterogeneity of the results as it is reflected by the range of the compliance ratios (31%-86%) creates the question if there is a factor that can be associated with better compliance. It was assumed that publication in a journal with higher impact factor will be associated with better compliance ratio and vice versa. In order to test this hypothesis the articles were split in two groups based on the impact factor of the journal they were published in. The cutoff point was set to 3. The first group consisted of 7 articles published in journals with lower than 3 impact factor and the second group had 2 articles published in journals with higher than 3 impact factor.

The comparison was made using the SPSS software. Normality of data was tested with Shapiro-Wilk and Kolmogorov-Smirnov tests both showing that it was distributed normally (table 3). Then the mean values of the compliance scores of the two groups were compared using independent sample t test. Levene's test does not give statistical significance so we continue assuming equal variances. Despite the very limited sample, it was shown that there is a statistically significant mean difference in compliance score of the two groups of 6.46 points, p value=0,04, 95%CI (0,36-12,49). (table 4) These results confirm our initial hypothesis

Tests of Normality^a

| | ImpactFactor | Kolmogorov-Smirnov ^b | | | Shapiro-Wilk | | |
|-------|--------------|---------------------------------|----|-------------------|--------------|----|------|
| | | Statistic | df | Sig. | Statistic | df | Sig. |
| Score | 2,00 | ,187 | 7 | ,200 [*] | ,963 | 7 | ,845 |

*. This is a lower bound of the true significance.

a. Score is constant when impactFactor = 1,00. It has been omitted.

b. Lilliefors Significance Correction

Table 3

Independent Samples Test

| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | | | | | |
|-------|-----------------------------|---|------|------------------------------|-------|-----------------|-----------------|-----------------------|---|----------|
| | | F | Sig. | t | df | Sig. (2-tailed) | Mean Difference | Std. Error Difference | 95% Confidence Interval of the Difference | |
| | | | | | | | | | Lower | Upper |
| Score | Equal variances assumed | 4,715 | ,066 | 2,505 | 7 | ,041 | 6,42857 | 2,56632 | ,36019 | 12,49696 |
| | Equal variances not assumed | | | 4,920 | 6,000 | ,003 | 6,42857 | 1,30671 | 3,23117 | 9,62597 |

Table 4

Comparison with similar study

Ultimately seeking external validity to the results we came up with, with the results of another STROBE evaluation report regarding studies published in the journal of Hand Surgery. This report compared compliance of articles published in two separate 6-month periods, one in 2005 and one in 2011. The compliance rates had been found to be 38% (95% CI 35%-42%, range 10%-50%) for 2005 and 58% (95% CI 55%-60%, range 39%-85%)

The 64,6% compliance of our study seems to be in line with the 2011 results as the overlapping 95% CI (50%-79%). However, this broad 95% CI that reflects the effect that this small sample has on the study, would make it difficult for any statistically significant difference to be indicated even if it really exists.

Discussion

This study aimed to evaluate the quality of observational reports. The 64,6% rate of compliance that was found seems suboptimal. Although there are items that are consistently and thoroughly reported in the majority of the articles (scientific background, purposes of the study), there is also an important number that are poorly reported (address potential sources of bias) or even totally neglected (funding).

This heterogeneity in the reporting of different items even in the same article reflects the different challenges that are met when dealing with each item. Some of them being totally theoretical and not creating any conflict seem easy to report. Such an example set the two items regarding the introduction which were found to have excellent compliance. However, others require much more effort throughout the conduction of the entire study, not just the writing the report, in means of data collection and careful information classification, such as reporting of events at each stage, describing patient follow up etc. This is reflected on the worse compliance rates of the sections methods and results when compared to those of introduction, as in these sections the items are much more challenging. Moreover, items like number 9 (address bias) may create a conflict of interests as at some point they are equivalent to indicate the weakness of one's own study. So, in addition to the author not being aware of any potential bias in their study this would be an extra reason to explain the low compliance rates found in this item.

Lastly, a special mention has to be made to the items that were found to have 0 compliance, meaning that not a single article included the information required in those items. Firstly, it has to be stated that the rates found in this study are validated by the report in the Journal of Hand Surgery, which presented the exact same rates. These two items were about the size of the sample and the funding of the study. As regards to the funding, although it may illuminate any conflicts of interest that are not apparent and it helps a great deal in the transparency of whole process of research, it is not as vital for the report of a study. However, the procedure of deciding the sample size has a scientific background based on the magnitude of the association that is being investigated and the desirable power of the study. What might really be the case here, besides the authors ignoring or neglecting to mention how the sample size was decided, is that these studies aimed at a small population of patients with a certain disease, under a certain treatment located in certain facilities, so they just included all the available patients the could include.

What is more this study also indicated an association between better compliance rates and publication in a journal with higher impact factor. This was a hypothesis assumed by many but the fact that it shown in terms of statistical significance despite the very limited size of the sample is a potential indicator of size of the magnitude of this association.

As regards the interpretation of the results, it should be done with great caution. The main source of potential bias is the inevitable subjectivity that lies in the scoring of the checklist. As mentioned above, some items of the checklist consisted of multiple bullets, each requiring multiple pieces of information. Scoring these items was a challenging process that

aimed to reward those that provided most information without important omissions. It is clear that different reviewers might give different scores, so in the absence of previous consensus about it any scoring is subjective. Although the checklist structure facilitates its objective scoring and differences between individuals are not expected to be big, the presence of subjectivity is not eliminated and must be taken into account as potential source of bias. Another problem that must be addressed is the limited size of this study and the resulting skewness of the data derived from it. However, the tests of normality in the compliance score and statistical method of indicating the differences assure us that it is safe to make some conclusions although it would be best if these were validated by further research.

Conclusion

The compliance rates found in this study are good but there is certainly room for improvement. Taking into account previous studies that show an upward trend through the years and the fact that high status journals show better compliance we can conclude that reporting of observational studies is heading to the right direction and the quality of future studies is expected to be better and better.

References

1) Using the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement to Assess Reporting of Observational Trials in Hand Surgery

Amelia A. Sorensen, MD, Robert D. Wojahn, BS, Mary Claire Manske, MD, Ryan P. Calfee, MD, MSc

2) Multiple Sclerosis Trust webpage

3) Bradley's Neurology in Clinical Practice, 7th Edition

Robert B. Daroff, Joseph Jankovic, John C. Mazziotta, Scott L. Pomeroy

4) Higher expression of IL-12R β 2 is associated with lower risk of relapse in relapsing–remitting multiple sclerosis patients on interferon- β 1b therapy during 3-year follow-up, *J Neuroimmunol*

Emina Milosevic, Irena Dujmovic, Milos Markovic, Sarlota Mesaros, Goran Rakocevic, Jelena Drulovic, Marija Mostarica Stojkovic, Dusan Popadic

5) Toll-like receptor (TLR)7 and TLR9 agonists enhance interferon (IFN) beta-1a's immunoregulatory effects on B cells in patients with relapsing-remitting multiple sclerosis (RRMS), *j neuroimm*

Yazhong Tao , Xin Zhang , Silva Markovic-Plese

6) Chitinase 3-like 1 is associated with the response to interferon-beta treatment in multiple sclerosis, *j neuroimm*

Clara Matute-Blanch , Jordi R o , Luisa M Villar , Luciana Midaglia , Sunny Malhotra , Jos  C  lvarez-Cerme o , Angela Vidal-Jordana , Xavier Montalban, Manuel Comabella

7) Blood lymphocyte subsets identify optimal responders to IFN-beta in MS, *j neurol*

Raquel Alenda · Lucienne Costa-Frossard · Roberto Alvarez-Lafuente · Carmen Espejo, Eulalia Rodriguez-Martin · Susana Sainz de la Maza · Noelia Villarrubia · Jordi Rvo · Marva I. Dominguez-Mozo · Xavier Montalban · Jos  C. Alvarez-Cermeno · Luisa M. Villar

8) Different clinical response to interferon beta and glatiramer acetate related to the presence of oligoclonal IgM bands in CSF in multiple sclerosis patients

Bonaventura Casanova & Laura Lacruz¹ & Mar a Luisa Villar² & Jos  Andr s Dom nguez & Mar a Carcel n Gadea & Francisco Gasc n³ & Javier Mallada⁵ & David Herv s & Mar a Sim -Castell  & Jos  Carlos  lvarez-Cerme o & Carmen Calles⁸ & Javier Olascoaga & Llu s Rami -Torrent  & Carmen Alcal  & Angeles Cervell ⁴ & Isabel Bosc  & Francisco Carlos P rez-Mirallles & Francisco Coret, *Neurol Sci*

9) Validation of 1-year predictive score of long-term response to interferon-b in everyday clinical practice multiple sclerosis patients, *j eur neurol*.

M. Romeoa, V. Martinellia, M. Rodeghera, E. Peregoa, S. Maida, M. P. Sormanib, G. Comia, c and San Raffaele Multiple Sclerosis Clinical Group¹

10) miR-145 and miR20a-5p Potentially Mediate Pleiotropic Effects of Interferon-Beta Through Mitogen-Activated Protein Kinase Signaling Pathway in Multiple Sclerosis Patients, *jneurol sci*

Naeim Ehtesham & Fariborz Khorvash & Majid Kheirollahi

11) Down-regulation of TYK2, CBLB and LMP7 genes expression in relapsing-remitting multiple sclerosis patients treated with interferon-beta, *j neuroimm*

Mehrdokht Mazdeha, Najmeh Moradib, Elnaz Khoshroob, Zahra Shayestehb, Mohammad Taheric, Arezou Sayadc, Mir Davood Omranic, Mehrdad Hajilooib, Ghodratollah Roshanaeid, Ghasem Solgib

12) HLA-DRB1 does not have a role in clinical response to interferon-beta among Iranian multiple sclerosis patients, *jneurol sci*

Sara Samadzadeh , Elnaz Tabibian , Tayebeh Sabokbar , Abbas Shakoory , Shahram Rahimi Dehghan , Saeed Azad Armaki, Bahram Aslanbeigi, Roya Abolfazli

13) Papathanasiou AA, Zintzaras E. Assessing the quality of reporting of observational studies in cancer. *Ann Epidemiol.* 2010

14) The STROBE group

Coordinator

Myriam Cevallos Christen

Research Fellow, Institute of Social and Preventive Medicine, University of Bern, Finkenhubelweg 11, CH-3012 Bern, Switzerland

STROBE initiative group

Douglas G. Altman

Professor of Medical Statistics, Centre for Statistics in Medicine, Wolfson College Annexe, Linton Road, Oxford OX2 6UD, UK.

Matthias Egger

Professor of Epidemiology and Public Health, Institute of Social and Preventive Medicine, University of Bern, Finkenhubelweg 11, CH-3012 Bern, Switzerland

Peter C. Gøtzsche

Director, Nordic Cochrane Centre, Rigshospitalet, Dept. 3343, Blegdamsvej 9, DK-2100 Copenhagen, Denmark

Stuart J. Pocock

Professor of Medical Statistics, Medical Statistics Unit, London School of Hygiene and Tropical Medicine, Keppel Street, London WC1E 7HT, UK

Jan P. Vandenbroucke

Professor of Clinical Epidemiology, Dept. of Clinical Epidemiology, Leiden University Hospital, CO-P, PO Box 9600, NL-2300 RC Leiden, Netherlands

Erik von Elm

Codirector Cochrane Switzerland, Institute of Social and Preventive Medicine (IUMSP), Lausanne University Hospital, Corniche 10, CH-1010 Lausanne, Switzerland

2010 Revision group

Douglas G. Altman

Professor of Medical Statistics, Centre for Statistics in Medicine, Wolfson College Annexe, Linton Road, Oxford OX2 6UD, UK.

Myriam Cevallos Christen

Research Fellow, Institute of Social and Preventive Medicine, University of Bern, Finkenhubelweg 11, CH-3012 Bern, Switzerland

Bruno da Costa

PhD Candidate, Institute of Social and Preventive Medicine, University of Bern,
Finkenhubelweg 11, CH-3012 Bern, Switzerland

Matthias Egger

Professor of Epidemiology and Public Health, Institute of Social and Preventive Medicine,
University of Bern, Finkenhubelweg 11, CH-3012 Bern, Switzerland

Patrizia Frei

Cancer epidemiologist, Swiss Tropical and Public Health Institute, Department of
Epidemiology and Public Health, Socinstrasse 57, 4051 Base

Peter C. Gøtzsche

Director, Nordic Cochrane Centre, Rigshospitalet, Dept. 3343, Blegdamsvej 9, DK-2100
Copenhagen, Denmark

Cindy Mulrow

Senior Deputy Editor, Annals of Internal Medicine, Professor of Medicine, University of Texas
Health Science Centre, San Antonio, USA

Landon Myer

Associate Professor of School of Public Health and Family Medicine, Infectious Diseases
Epidemiology Unit, University of Cape Town, Cape Town, South Africa

Stuart J. Pocock

Professor of Medical Statistics, Medical Statistics Unit, London School of Hygiene and
Tropical Medicine, Keppel Street, London WC1E 7HT, UK

Charlie Poole

Associate Professor of Epidemiology, Department of Epidemiology, UNC Gillings School of
Global Public Health, University of North Carolina, North Carolina, USA

Andrew Renehan

Senior Lecturer in Cancer Studies and Surgery, University of Manchester, Christie Hospital NHS
Trust, Manchester, United Kingdom

Jan P. Vandenbroucke

Professor of Clinical Epidemiology, Dept. of Clinical Epidemiology, Leiden University Hospital,
CO-P, PO Box 9600, NL-2300 RC Leiden, Netherlands

Erik von Elm

Codirector Cochrane Switzerland, Institute of Social and Preventive Medicine (IUMSP),
Lausanne University Hospital, Corniche 10, CH-1010 Lausanne, Switzerland