DR. SIZHENG ZHAO (Orcid ID : 0000-0002-3558-7353)

**The impact of smoking on response to TNF inhibitors in axial spondyloarthritis: methodological considerations for longitudinal observational studies**

Sizheng Zhao MD[1,2,3], Kazuki Yoshida MD, PhD[3,4], Gareth T Jones PhD[5,6], David M Hughes PhD[7], Sara K Tedeschi MD[3], Houchen Lyu MD[8,9], Robert J Moots MD, PhD[1,2], Daniel H Solomon MD[3,10], Nicola J Goodson MD, PhD[1,2]

1 Musculoskeletal biology I, Institute of Ageing and Chronic Disease, University of Liverpool, Liverpool, UK

2 Department of Academic Rheumatology, Aintree University Hospital, Liverpool, UK

3 Division of Rheumatology, Immunology and Allergy, Brigham and Women's Hospital, Boston, Massachusetts, United States

4 Departments of Epidemiology and Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, United States

5 Epidemiology Group, School of Medicine, Medical Sciences and Nutrition, University of Aberdeen, Aberdeen, UK

6 Aberdeen Centre for Arthritis and Musculoskeletal Health, University of Aberdeen, Aberdeen, UK

7 Department of Biostatistics, Institute of Translational Medicine, University of Liverpool, Liverpool, UK

8 Department of Medicine, Harvard Medical School, Boston, Massachusetts, United States

9 Department of Orthopaedics, General Hospital of Chinese PLA, Beijing, China

10 Division of Pharmacoepidemiology and Pharmacoeconomics, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, United States

Correspondence to:

Dr Nicola J Goodson

Department of Academic Rheumatology

Aintree University Hospital

Liverpool

L9 7AL

UK

ngoodson@liverpool.ac.uk

## Abstract

**Objective**. Observational data facilitate examination of treatment-effect heterogeneity, but the risk of bias is substantial. We highlight methodological considerations through an analysis of whether smoking affects response to TNF inhibitors (TNFi) in axial spondyloarthritis (axSpA).

**Methods.** We used longitudinal data from the British Society for Rheumatology Biologics Register for Ankylosing Spondylitis. Participants fulfilling the ASAS criteria for axSpA, who started their first TNFi were eligible for analysis. To compare the impact of smoking status, weighted generalised estimating equations were used to examine changes in several continuous outcome measures, including BASDAI and ASDAS. Inverse-probability weights were used to account for differences in baseline covariates and excluded participants. We separately assessed response in the first 3 months to account for non-random dropout.

**Results**. Of 840 participants that started on TNFi, 1,641 assessments from 627 individuals were analysed (69% male, mean age 46 years). 33% were current smokers and 30% ex-smokers. Ex- and current smokers had worse disease than never smokers at baseline. Accounting for these differences, response did not differ according to smoking status. Compared against never smokers, ex-smokers ($\beta$=-0.6; 95%CI -1.4, 0.3) and current smokers ($\beta$=-0.4; 95%CI -1.1, 0.4) had similar response in BASDAI, and ASDAS (ex: $\beta$=-0.1; 95%CI -0.5, 0.3; current: $\beta$=-0.01; 95%CI -0.4, 0.4), at 3 months.

**Conclusions**. TNFi response did not differ according to baseline smoking status in this UK cohort. Conflicting results from previous studies were likely due to methodological differences. This analysis highlights potential sources of bias that should be addressed in future studies.

**Keywords:** axial spondyloarthritis, ankylosing spondylitis, registry, treatment response, inverse probability weights

**Key messages:**

- Observational studies are susceptible to bias and potential sources should be identified and addressed.
- Conflicting results from prior studies are likely explained by methodological differences.
- Response to TNF inhibitors did not differ according to smoking status in this large cohort.

**Introduction**

Disease registries are important resources for observational research. They provide high quality data for large numbers of patients generalisable to clinical practice, making them invaluable for comparing effectiveness of treatments. However, observational data are susceptible to bias, requiring rigorous methodological approaches when attempting to infer causation. Consider the effect of smoking on response to TNF inhibitors (TNFi) in axial spondyloarthritis (axSpA): an important clinical question since only half of patients respond (1, 2) and alternative treatment options are relatively limited. While some smaller studies reported that smoking did not have an important impact on TNFi response (3-5), others found apparently dramatic effect sizes with current smokers having half the odds of response compared to non-smokers (6, 7).

Several methodological differences may explain these conflicting results. First, many patient characteristics differ according to smoking status (3, 6, 7). Baseline disease severity is known to predict response to TNFi and should be adequately accounted for. Second, longitudinal studies apply selection criteria that typically exclude a proportion of participants from the analysis set, for example by requiring at least one follow-up visit. Such sample-restriction introduces selection bias if included and excluded participants differ in smoking status and other baseline characteristics. Third, continuation of TNFi typically requires demonstration of response, with variable enforcement and response definitions in different healthcare systems. Ongoing treatment in the UK is only funded if response is demonstrated at 12 weeks (allowing for some flexibility) as measured by the Bath AS Disease Activity Index (BASDAI) and spinal pain (8, 9). Dropout due to inefficacy or adverse events are examples of non-random censoring (ie. these events are associated with participant characteristics) that can present problems for observational analyses (10).

The aim of the current analyses was to examine the impact of smoking on response to TNFi in axSpA participants, while exploring common methodological issues and their solutions. We synthesise our discussion alongside a qualitative review of methods and results from existing studies.

**Methods**

Study design and population

The British Society for Rheumatology Biologics Register for Ankylosing Spondylitis (BSRBR-AS) is a UK-wide prospective cohort study of participants fulfilling the ASAS criteria for axSpA. The study protocol has been previously published (11). The current analysis focused on those who started their first TNFi, from December 2012 to June 2017. These participants were followed up at baseline, 3, 6 and 12 months and annually thereafter, at study visits and using postal questionnaires. Start and stop dates were recorded for each TNFi, along with the reason for stopping (adverse events, inefficacy or other reasons). The cohort was defined at baseline, therefore only those with a valid baseline questionnaire (within 1 year before and 7 days after starting TNFi) and had smoking status were eligible for analysis. Where there was more than one questionnaire in each time period, the nearest one to the per-protocol follow-up time was chosen. Ethical approval was obtained from the National Research Ethics Committee (reference 11/NE/0374) and informed consent was obtained from all participants.

Exposure and outcomes

Smoking status was self-reported as current, ex- or never in each questionnaire. Baseline smoking status was used to define the exposure; where this was missing, the earliest reported instance was used. Ex-smokers were defined as those who have not smoked for the past 3 months.

Outcomes included patient-reported disease activity (Bath AS Disease Activity Index (BASDAI), spinal pain), functional impairment (BASFI) and other aspects of disease severity such as the Bath AS Global Score (BASG), Chalder Fatigue Scale (12), Jenkins Sleep Evaluation Questionnaire (13) and the Hospital Anxiety and Depression Scale (HADS) (14). The AS Disease Activity Score (ASDAS) was calculated using CRP or, if unavailable, ESR.

We compared changes in outcomes (measured by the above scales) over time to their pre-treatment baseline, between difference categories of smokers. To allow comparison with existing studies, we also compared proportions meeting a binary response, BASDAI50/2 (50% or 2-unit reduction in BASDAI) at 3, 6 and 12 months.

Covariates

The following covariates were recorded at baseline and chosen *a priori* for their known or theoretical associations with TNFi response (1, 2, 15-17): age, gender, symptom duration, education, elevated baseline CRP (above upper normal limit), classification as AS (modified New York criteria (18)), HLA-B27 status, body mass index (BMI), index of multiple deprivation (in quintiles (19-21)) as a measure of socioeconomic status, alcohol status (as current, ex- or never) and comorbidity (categorised as 0, 1 or ≥2 from 13 conditions (11)). Time was categorised by per-protocol follow-up.

Statistical analysis

Baseline participant characteristics were summarised by smoking status. For each outcome variable, we compared its change over time according to smoking status using generalised estimating equations (GEE) (22). This was achieved using interaction terms between smoking status and the time variable: their coefficients are interpreted as the difference in response compared to the reference group (never smokers). Model predictions were plotted to visualise results. These models were weighted with weights constructed as follows.

We balanced differences in baseline characteristics between smoking exposure categories using inverse probability of "treatment" weights (IPTW) (23). This adjustment approach has an advantage over inclusion of the baseline characteristics in the outcome model (the theoretical basis is given in supplementary materials). A multinomial logistic model was used to construct IPTW for each smoking category. Independent variables for the weight model included all baseline covariates specified above as well as all baseline outcome measures (as a collective representation of disease severity). Studying the causal effect of baseline smoking status has conceptual difficulty: we cannot randomly assign an individual to "having smoked for 20 years" at the onset of a hypothetical trial (24). However, propensity score related methods are still useful for "unconfounded descriptive comparisons" (25, 26).

Including participants with a baseline questionnaire assumes this selected subset is representative of the initial cohort. We improved upon this approach by weighting individuals in such a way that baseline characteristics of the analysis set resembles the original eligible cohort. This is a form of inverse probability of censoring weights (IPCW) for censoring at the baseline. IPCWs were constructed from predicted values of logistic models using inclusion/exclusion status as the dependent variable, and smoking status and available baseline covariates as independent variables.

To address informative censoring after the baseline, we first limited the above analysis to response within 3 months (analysis 1), during which time dropout due to inefficacy should be minimal. Missing 3-month responses were modelled using time-varying IPCWs as described above with "missingness" as the dependent variable. This makes missingness random with respect to baseline characteristics. We then repeated the analysis for the subset of participants that remained on treatment from 6 months onwards (analysis 2) using baseline IPCWs to account for the excluded, as described above, but without additional use of time-varying IPCWs.

Lastly, BASDAI50/2 was used as the outcome in weighted logistic models. Dropout due to inefficacy was defined as non-response; other missing responses were modelled using IPCWs as described above. All weights were "stabilised" to have a mean of 1, allowing the overall sample size to remain unchanged (27). Missing covariates were imputed using chained equations (see supplement for details) (28). Analyses were performed in Stata version 13.

**Results**

Among a total of 2,420 participants in the BSRBR-AS, 840 commenced their first TNFi within the study period and provided smoking status. 213 participants were excluded because they did not have a valid baseline assessment. 627 participants were included in analyses, providing 1,641

questionnaire assessments. Excluded participants had shorter symptom duration and showed trends for having lower deprivation and higher educational attainment (differences shown in supplementary table 1).

Analysis 1: Comparing response at 3 months according to smoking status

Baseline characteristics of the analysis cohort are shown in table 1. Covariate were well balanced after IP weighting (supplementary figure 1). A third of participants were current smokers, 30% ex-smokers and 37% never smokers. Current smokers were younger, more frequently male and showed trends for having higher deprivation and lower educational attainment. Baseline values of most outcome variables were worse in ex- and current smokers. 218 participants had missing outcome measures at 3 months and were modelled using IPCWs; the number of participants who stopped their TNFi during this period was too small (n=1) to model separately.

All outcomes improved significantly after commencing TNFi. These changes were highly similar for each smoking status (BASDAI and BASFI shown in figure 1, remainder shown in supplementary figure 2). Although not statistically significant, interaction term coefficients suggest that BASDAI reduction at 3 months was greater for ex- (β= -0.58; 95%CI -1.41 to 0.25) and current smokers (β= -0.38; 95%CI-1.12 to 0.36) compared to never smokers (table 2). Results were similar for most other outcome measures, except ex-smokers had significantly greater improvement in fatigue.

Analysis 2: Comparing response after 6 months in those who remained on treatment

During the study period, 136 participants discontinued treatment: adverse event was labelled as the reasons for 49, inefficacy for 32 and other for 55. Proportions were not significantly different according to smoking status (table 1). Baseline characteristics for participants in analysis 2 are shown in supplementary table 2.

Progression of outcome measures after 6 months were similar between each smoking status (table 3). Compared to never smokers, BASDAI increased by 0.07 units more (95%CI -0.11 to 0.24) for ex-smokers and 0.04 units more (95%CI -0.13 to 0.22) for current smokers, per 6-month period. The only statistically significant differences were for fatigue and ASDAS. The slope coefficients suggest that current smokers may have poorer sustained treatment response after 6 months than never smokers (figure 2 and supplementary figure 3).

BASDAI50/2 response at 3 months was not significantly different for ex-smokers (OR 1.11; 95%CI 0.76 to 1.61) or current smokers (OR 0.97; 95%CI 0.66 to 1.44), compared with never smokers. Results were similar at 6 months (ex-smokers: OR 1.01, 95%CI 0.67 to 1.50; current smokers: OR 0.85, 95%CI 0.56 to 1.27). At 12 months, ex-smokers had higher odds of response (OR 1.65, 95%CI 1.11 to 2.45), but not current smokers (OR 1.12, 95%CI 0.74 to 1.77). Distributions of all IP weights are described in supplementary table 3.

**Discussion**

In this large UK cohort of axSpA participants, baseline smoking status was associated with significantly worse disease severity at baseline across all measures. However, it was not associated with response to the first TNFi. This applied to all outcome measures including disease activity, functional impairment, quality of life, fatigue, sleep and mental health. We demonstrated the importance of several methodological considerations for future studies of non-interventional exposures on treatment response, and offer inverse-probability weighting as a solution to reduce potential bias.

The main strength of this study is the quality of data. Several outcomes were measured that provide a holistic representation of disease severity and impact. The rich BSRBR-AS dataset also allowed us to adjust for a large number of confounders, minimising the impact of unmeasured confounding. Participants were recruited from both specialist and non-specialist secondary care centres, thus providing a relatively unselected population representative of UK clinical practice.

Our data did have some limitations. There was an unusually low proportion of discontinuation due to inefficacy, likely reflecting limitations in the way discontinuation reasons were labelled. However, the start and stop dates for each TNFi were diligently recorded, providing clear information about duration of use. The BSRBR-AS did not record exercise or other lifestyle factors that are potential confounders. We did adjust for socioeconomic status, and included alcohol-use as another representation of health-related behaviour; exercise did not contribute significant confounding in a previous study (6). We did not examine BASMI, ESR or CRP as these variables required a clinic visit, which made them distinct from questionnaire-derived variables that would require separate modelling. ESR/CRP were also different in that they were measured only when clinically indicated. Subtle differences between smoking status for some outcomes were not clinically important and should not be over-interpreted. It was interesting that current and ex-smokers had non-significantly greater improvement in analysis 1, which may be explained by regression towards the mean. Higher

odds of 12-month response in ex-smokers may be explained by the fact that people who give up smoking might also make other healthy decisions.

Results from three similar studies are summarised in table 4. The first, by Ciurea et al., used mixed models to show a statistically significant, but not clinically important, effect of smoking only among a subgroup of those with elevated baseline CRP (6). Their "step model" of initial response was analogous to analysis 1 in our study. The main difference was that patients needed to have at least one follow-up to be eligible. Conditioning on follow-up attendance is likely to introduce bias (29).

The study by Kydd et al. used linear mixed models to include individuals with only one data-point: the maximum number of patients at any assessment period was 252 despite a total sample size of 422. Only 99 patients had their outcomes of interest assessed before initiating TNFi, therefore making it difficult to adjust for baseline differences. In that study, the interaction terms, representing the difference in outcomes between smoking status, were not significant.

Previous studies that used binary definitions of response reported dramatic effect sizes related to current smoking (6, 7). The study by Ciurea et al., which found no clinically meaningful overall difference in continuous BASDAI according to smoking status, reported 46% reduced odds for achieving BASDAI50 response (OR 0.54; 95%CI 0.31 to 0.95) at 1 year (±6 months) (6). This was reproduced by Glintborg et al. using BASDAI50/2 at 3 and 6 months (7). Small changes in continuous outcomes should not translate to significant differences in the proportion of responders when they are dichotomised. There are two potential explanations. If all patients had identical improvements in BASDAI, smokers would still have poorer BASDAI50 response because of their higher baseline BASDAI. Hypothetically, if each participant in our cohort improved by an identical 3 units, their BASDAI50 responses at 6 months would be significantly different (40%, 35% and 28% for never, ex- and current smokers, respectively). Second, patients who discontinued treatment were labelled as non-responders. Whether smoking has a biological effect on response is a different question to whether it increases treatment discontinuation, yet this distinction is crucial for causal inference. Smokers may discontinue treatment for reasons other than inefficacy that may be confounded by, for instance, attitudes to health. While binary responses can be helpful in including dropout due to inefficacy, other reasons for dropout should be modelled separately. In the current study, smoking did not affect BASDAI50/2 when we separately accounted for subjects who were censored for reasons other than inefficacy. These negative results were supported by a post hoc analysis of the ABILITY-1 randomised clinical trial, where smoking status did not affect binary response or time to response (5).

Conclusion

We used this analysis to highlight methodological considerations for future observational studies aiming to explore causal effects of exposures on treatment response. In this large UK cohort of axSpA participants, response to the first TNFi did not differ significantly according to baseline smoking status. Prescribers should dispel any subconscious bias that smokers may not respond as well to treatment. Nevertheless, we emphasise the importance of smoking cessation, particularly given the high burden of cardiovascular disease in rheumatic patients (30, 31). Smoking is associated with more severe disease (activity, functional impairment and radiographic progression (32)); it is not known whether cessation leads to improvement in disease outcome.

**References**

1.       Lord PA, Farragher TM, Lunt M, Watson KD, Symmons DP, Hyrich KL, et al. Predictors of response to anti-TNF therapy in ankylosing spondylitis: results from the British Society for Rheumatology Biologics Register. Rheumatology. 2010;49(3):563-70.

2.       Glintborg B, Ostergaard M, Krogh NS, Dreyer L, Kristensen HL, Hetland ML. Predictors of treatment response and drug continuation in 842 patients with ankylosing spondylitis treated with anti-tumour necrosis factor: results from 8 years' surveillance in the Danish nationwide DANBIO registry. Annals of the rheumatic diseases. 2010;69(11):2002-8.

3.       Kydd AS, Chen JS, Makovey J, Chand V, Henderson L, Buchbinder R, et al. Smoking did not modify the effects of anti-TNF treatment on health-related quality of life among Australian ankylosing spondylitis patients. Rheumatology. 2015;54(2):310-7.

4.       Dellyes A, Lafforgue P, Pradel V, Pham T. Smoking is not associated with response to TNF blockers in patients with axial spondyloarthritis. ACR/ARHP Annual Meeting. 2014;Abstract no. 2584.

5.       Sieper J, Poddubnyy D, Pangan AL, Rathman SS, Anderson J. Smoking was not associated with response to adalimumab therapy in patients with nonradiographic axial spondyloarthritis. Annals of the rheumatic diseases. 2013;72:s3.

6.       Ciurea A, Scherer A, Weber U, Exer P, Bernhard J, Tamborrini G, et al. Impaired response to treatment with tumour necrosis factor alpha inhibitors in smokers with axial spondyloarthritis. Annals of the rheumatic diseases. 2016;75(3):532-9.

7.       Glintborg B, Hojgaard P, Lund Hetland M, Steen Krogh N, Kollerup G, Jensen J, et al. Impact of tobacco smoking on response to tumour necrosis factor-alpha inhibitor treatment in patients with ankylosing spondylitis: results from the Danish nationwide DANBIO registry. Rheumatology. 2016;55(4):659-68.

8.       National Institute for Health and Care Excellence. TNF-alpha inhibitors for ankylosing spondylitis and non-radiographic axial spondyloarthritis.  2016  [cited June 2018; Available from: https://www.nice.org.uk/guidance/ta383

9.       McAllister K, Goodson N, Warburton L, Rogers G. Spondyloarthritis: diagnosis and management: summary of NICE guidance. Bmj. 2017;356:j839.

10.      Kurland BF, Johnson LL, Egleston BL, Diehr PH. Longitudinal Data with Follow-up Truncated by Death: Match the Analysis Method to Research Aims. Statistical science : a review journal of the Institute of Mathematical Statistics. 2009;24(2):211.

11.      Macfarlane GJ, Barnish MS, Jones EA, Kay L, Keat A, Meldrum KT, et al. The British Society for Rheumatology Biologics Registers in Ankylosing Spondylitis (BSRBR-AS) study: Protocol for a prospective cohort study of the long-term safety and quality of life outcomes of biologic treatment. BMC musculoskeletal disorders. 2015;16:347.

12.      Jackson C. The Chalder Fatigue Scale (CFQ 11). Occup Med (Lond). 2015;65(1):86.

13.      Jenkins CD, Stanton BA, Niemcryk SJ, Rose RM. A scale for the estimation of sleep problems in clinical research. J Clin Epidemiol. 1988;41(4):313-21.

14.      Stern AF. The hospital anxiety and depression scale. Occup Med (Lond). 2014;64(5):393-4.

15.      Vastesaeger N, van der Heijde D, Inman RD, Wang Y, Deodhar A, Hsu B, et al. Predicting the outcome of ankylosing spondylitis therapy. Annals of the rheumatic diseases. 2011;70(6):973-81.

16.      Rudwaleit M, Listing J, Brandt J, Braun J, Sieper J. Prediction of a major clinical response (BASDAI 50) to tumour necrosis factor alpha blockers in ankylosing spondylitis. Annals of the rheumatic diseases. 2004;63(6):665-70.

17.      Brookhart MA, Schneeweiss S, Rothman KJ, Glynn RJ, Avorn J, Sturmer T. Variable selection for propensity score models. Am J Epidemiol. 2006;163(12):1149-56.

18.      van der Linden S, Valkenburg HA, Cats A. Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis and rheumatism. 1984;27(4):361-8.

19.    The Scottish Government. Scottish Index of Multiple Deprivation.   [cited Nov 2018; Available from: http://www.gov.scot/Topics/Statistics/SIMD

20.    Department for Communities and Local Government. English indices of deprivation 2015. 2015  [cited Nov 2018; Available from: www.gov.uk/government/statistics/english-indices-of-deprivation-2015

21.    The Welsh Government. Welsh Index of Multiple Deprivation.  2015  [cited Nov 2018; Available from: https://gov.wales/statistics-and-research/welsh-index-multiple-deprivation/

22.    Hernan MA, Brumback BA, Robins JM. Estimating the causal effect of zidovudine on CD4 count with a marginal structural model for repeated measures. Statistics in medicine. 2002;21(12):1689-709.

23.    Robins JM, Hernan MA, Brumback B. Marginal structural models and causal inference in epidemiology. Epidemiology. 2000;11(5):550-60.

24.    Hernan MA, Robins JM. Using Big Data to Emulate a Target Trial When a Randomized Trial Is Not Available. Am J Epidemiol. 2016;183(8):758-64.

25.    Li F, Morgan KL, Zaslavsky AM. Balancing Covariates via Propensity ScoreWeighting. Journal of the American Statistical Association. 2018;113(521):390-400.

26.    Li F, Zaslavsky AM, Landrum MB. Propensity score weighting with multilevel data. Statistics in medicine. 2013;32(19):3373-87.

27.    Cole SR, Hernan MA. Constructing inverse probability weights for marginal structural models. Am J Epidemiol. 2008;168(6):656-64.

28.    STATA MULTIPLE-IMPUTATION REFERENCE MANUAL (RELEASE 15).  2017  [cited July 2018; Available from: https://www.stata.com/manuals/mi.pdf

29.    Hernán MA, Robins JM. Causal Inference: Boca Raton: Chapman & Hall/CRC, forthcoming; 2018.

30.    Han C, Robinson DW, Jr., Hackett MV, Paramore LC, Fraeman KH, Bala MV. Cardiovascular disease and risk factors in patients with rheumatoid arthritis, psoriatic arthritis, and ankylosing spondylitis. The Journal of rheumatology. 2006;33(11):2167-72.

31.    Solomon DH, Goodson NJ, Katz JN, Weinblatt ME, Avorn J, Setoguchi S, et al. Patterns of cardiovascular risk in rheumatoid arthritis. Annals of the rheumatic diseases. 2006;65(12):1608-12.

32.    Poddubnyy D, Haibel H, Listing J, Marker-Hermann E, Zeidler H, Braun J, et al. Baseline radiographic damage, elevated acute-phase reactant levels, and cigarette smoking status predict spinal radiographic progression in early axial spondylarthritis. Arthritis and rheumatism. 2012;64(5):1388-98.

Figure 1. No statistically significant difference in BASDAI and BASFI response to TNF inhibitors at 3 months according to smoking status. Plots show predicted values from weighted generalised estimating equations. Responses using the remaining eight outcome measures were similar and are shown in supplementary figure 2.

Figure 2. No significant difference in response to TNF inhibitors after 6 months according to smoking status. Plots show predicted values from weighted generalised estimating equations. Responses using the remaining eight outcome measures were similar and are shown in supplementary figure 3.

| Table 1. Baseline participant and disease characteristics according to smoking status. | | Never smoker (n=234) | Ex-smoker (n=187) | Current smoker (n=206) | P-value |
|---|---|---|---|---|---|
| Age, mean (SD) years | | 45.0 (14.9) | 50.0 (13.2) | 42.3 (12.2) | <0.001 |
| Male | | 153 (65%) | 121 (65%) | 156 (76%) | 0.027 |
| Meets mNY criteria for AS | | 143 (61%) | 123 (66%) | 123 (60%) | 0.440 |
| HLA-B27 positive[+] | | 127 (71%) | 102 (77%) | 125 (81%) | 0.122 |
| Elevated CRP[*] | | 133 (58%) | 108 (61%) | 129 (66%) | 0.221 |
| Symptom duration, median (IQR) years | | 15.0 (5.4 to 29.3) | 20.8 (10.5 to 32.8) | 13.1 (5.6 to 23.6) | <0.001 |
| BMI, mean (SD) | | 27.7 (5.8) | 28.9 (5.0) | 27.5 (5.7) | 0.038 |
| Quintiles of Index of Multiple Deprivation | 1, most deprived | 36 (15%) | 29 (16%) | 67 (33%) | <0.001** |
| | 2 | 48 (21%) | 23 (12%) | 35 (17%) | |
| | 3 | 39 (17%) | 39 (21%) | 40 (19%) | |
| | 4 | 64 (27%) | 48 (26%) | 37 (18%) | |
| | 5, most affluent | 47 (20%) | 48 (26%) | 27 (13%) | |
| Highest level of education | Secondary school | 70 (30%) | 60 (32%) | 95 (47%) | <0.001 |
| | Apprenticeship | 16 (7%) | 22 (12%) | 24 (12%) | |
| | Further education college | 68 (29%) | 68 (37%) | 55 (27%) | |
| | University degree | 60 (26%) | 30 (16%) | 22 (11%) | |
| | Further degree | 18 (8%) | 5 (3%) | 6 (3%) | |
| Alcohol status | Current | 179 (77%) | 147 (79%) | 130 (63%) | 0.001 |
| | Ex | 32 (14%) | 31 (17%) | 50 (24%) | |
| | Never | 22 (9%) | 9 (5%) | 26 (13%) | |
| Number of comorbidities | 0 | 140 (60%) | 94 (51%) | 105 (51%) | 0.055** |
| | 1 | 63 (27%) | 53 (29%) | 67 (33%) | |
| | ≥2 | 29 (13%) | 38 (21%) | 34 (17%) | |
| Disease activity, median (IQR) | BASDAI | 6.4 (5.1 to 7.4) | 6.8 (5.5 to 8.1) | 7.2 (5.9 to 7.9) | 0.004 |
| | ASDAS[+] | 2.9 (2.4 to 3.4) | 3.0 (2.4 to 3.5) | 3.0 (2.6 to 3.6) | 0.042 |
| | Spinal pain | 7.0 (5.0 to 8.0) | 7.0 (5.0 to 8.0) | 7.0 (6.0 to 8.0) | 0.028 |
| BASFI, median (IQR) | | 6.0 (4.1 to 7.7) | 6.7 (5.0 to 8.3) | 7.1 (5.5 to 8.5) | <0.001 |
| ASQoL, median (IQR) | | 11.0 (8.0 to 14.0) | 13.0 (9.0 to 16.0) | 15.0 (11.0 to 17.0) | <0.001 |
| BASG[+], median (IQR) | | 7.5 (6.0 to 8.0) | 7.5 (6.0 to 8.5) | 7.5 (6.3 to 8.5) | 0.170 |
| Fatigue, median (IQR) | | 17.0 (14.0 to 21.0) | 17.0 (13.0 to 21.0) | 18.0 (14.0 to 22.0) | 0.390 |
| Sleep, median (IQR) | | 13.0 (8.0 to 17.0) | 15.0 (9.0 to 18.0) | 15.0 (11.0 to 19.0) | 0.004 |
| HADS, median (IQR) | Anxiety | 8.0 (5.0 to 11.0) | 8.5 (6.0 to 11.0) | 11.0 (8.0 to 14.0) | <0.001 |
| | Depression | 6.0 (3.0 to 9.0) | 7.5 (5.0 to 10.0) | 9.0 (6.0 to 12.0) | <0.001 |
| Remained on treatment | | 188 (80%) | 149 (80%) | 154 (75%) | 0.230 |
| Stopped treatment | Adverse events | 20 (9%) | 16 (9%) | 13 (6%) | |
| | Inefficacy | 11 (5%) | 6 (3%) | 15 (7%) | |
| | Other | 15 (6%) | 16 (9%) | 24 (12%) | |
| Data presented as mean (standard deviation), median (interquartile range), number (percentage). Comparisons used ANOVA or Kruskal–Wallis test for continuous variables, Chi-squared test for | | | | | |

categorical variables.

[+] Not all variables had complete data, HLA-B27 status was available for 468 participants, ASDAS for 539.

*Above upper normal limit.

**Non-parametric test for trend across ordered groups.

SD, standard deviation; IQR, interquartile range; mNY, modified New York criteria for Ankylosing Spondylitis; BMI, body mass index; BASDAI, Bath AS disease activity index; ASDAS, AS disease activity score; BASFI, Bath AS functional index; ASQoL, AS quality of life questionnaire; BASG, Bath AS Global Score; HADS, Hospital Anxiety and Depression Scale.

| Table 2. Coefficients of interaction terms between smoking status and time, showing the difference in 3-month response compared to never smokers (analysis 1). | | Never smoker | Ex-smoker | Current smoker |
|---|---|---|---|---|
| Disease activity | BASDAI | reference | -0.58 (-1.41 to 0.25) | -0.38 (-1.12 to 0.36) |
| | ASDAS | reference | -0.07 (-0.47 to 0.32) | -0.01 (-0.42 to 0.40) |
| | Spinal pain | reference | -0.67 (-1.61 to 0.26) | -0.36 (-1.32 to 0.60) |
| | BASFI | reference | -0.59 (-1.40 to 0.22) | 0.21 (-0.61 to 1.03) |
| | ASQoL | reference | -1.56 (-3.20 to 0.09) | -0.34 (-1.94 to 1.26) |
| | BASG | reference | -0.61 (-1.29 to 0.08) | -0.13 (-0.84 to 0.58) |
| | Fatigue | reference | -2.29 (-4.29 to -0.28) | -0.64 (-2.73 to 1.44) |
| | Sleep | reference | 0.22 (-1.82 to 2.25) | 0.67 (-1.29 to 2.63) |
| HADS | Anxiety | reference | -0.38 (-1.58 to 0.82) | -0.37 (-1.87 to 1.14) |
| | Depression | reference | -0.90 (-2.14 to 0.34) | -0.41 (-1.76 to 0.94) |
| Example interpretation of coefficients: ex-smokers had an additional 0.58-unit reduction in BASDAI compared to never smokers at 3 months. BASDAI, Bath AS disease activity index; ASDAS, AS disease activity score; BASFI, Bath AS functional index; ASQoL, AS quality of life questionnaire; BASG, Bath AS Global Score; HADS, Hospital Anxiety and Depression Scale. | | | | |

| | | Never smoker | Ex-smoker | Current smoker |
|---|---|---|---|---|
| Table 3. Coefficients of interaction terms between smoking status and time, showing the difference responses after 6 months, compared to never smokers (analysis 2). | | | | |
| Disease activity | BASDAI | reference | 0.07 (-0.11 to 0.24) | 0.04 (-0.13 to 0.22) |
| | ASDAS | reference | 0.02 (-0.10 to 0.13) | 0.10 (0.002 to 0.20) |
| | Spinal pain | reference | -0.01 (-0.28 to 0.26) | 0.17 (-0.05 to 0.38) |
| | BASFI | reference | 0.03 (-0.18 to 0.23) | 0.02 (-0.19 to 0.23) |
| | ASQoL | reference | 0.28 (-0.11 to 0.67) | 0.27 (-0.12 to 0.66) |
| | BASG | reference | 0.17 (-0.08 to 0.42) | 0.17 (-0.08 to 0.42) |
| | Fatigue | reference | 0.49 (0.05 to 0.93) | 0.46 (0.03 to 0.90) |
| | Sleep | reference | 0.26 (-0.17 to 0.70) | 0.26 (-0.17 to 0.69) |
| HADS | Anxiety | reference | 0.14 (-0.11 to 0.39) | -0.10 (-0.42 to 0.23) |
| | Depression | reference | 0.08 (-0.15 to 0.31) | 0.16 (-0.17 to 0.50) |

Coefficients shown for time in units of 6 months.

Example interpretation of coefficients: BASDAI in ex-smokers worsened by 0.07 units more than never smokers per 6-month period.

BASDAI, Bath AS disease activity index; ASDAS, AS disease activity score; BASFI, Bath AS functional index; ASQoL, AS quality of life questionnaire; BASG, Bath AS Global Score; HADS, Hospital Anxiety and Depression Scale.

| | Ciurea 2015 | Glintborg 2015 | Kydd 2015 | This study |
|---|---|---|---|---|
| | Table 4. Comparing published studies of the effect of smoking on response to TNFi. | | | |
| Cohort and analysis set | Swiss Clinical Quality Management Cohort (n=2973)

1880 fulfilled ASAS criteria with available smoking status.
698 eligible for analysis with at least one follow-up assessment. | DANBIO (n=1775)

1576 were eligible for inclusion.
1425 eligible for analysis with known smoking status. | Australian Rheumatology Association Database (n=561)

422 eligible with at least one assessment within 27 months of starting TNFi. | BSRBR-AS (n=2420)

947 exposed to first TNFi.
628 eligible for analysis with baseline assessment (excluded participants accounted for using IPCW). |
| Baseline characteristics | Mean BASDAI 5.5 (SD 1.9)

Significant differences according to smoking status: BASFI, BASMI, EQ5D, SF-12, elevated CRP; age, gender, education. | Median BASDAI 5.6 to 6.1

Significant differences according to smoking status: CRP, BASDAI, BASFI, BASMI, pain and physician/patient global; age, disease duration, gender. | Mean BASDAI 7.3 (SD 1.5)

Significant differences according to smoking status: SF-36, AQoL; age, gender, disease duration, education. | Mean BASDAI 6.3 (SD 1.8), median 6.6

Significant differences according to smoking status: BASDAI, ASDAS, BASFI, BASMI, ASQoL, JSEQ, HADS; age, gender, symptom duration, BMI, deprivation, education, alcohol |
| Exposure | 38% current, 24% ex, 38% never | 43% current, 16% ex ,41% never | 19% current, 33% ex, 49% never | 29% current, 32% ex, 39% never |
| Outcomes | 1. Continuous BASDAI/ASDAS
2. Binary BASDAI50, ASDAS-MI, ASAS40 at 1yr±6m | Binary BASDAI50/2 at 3 and 6m* | Continuous SF-36, AQoL, HAQ-S | 1. Continuous BASDAI, ASDAS, spinal pain, BASFI, ASQoL, BASG, fatigue, sleep, HADS |

| | | | | 2. Binary BASDAI50/2 |
|---|---|---|---|---|
| Methods | 1. Linear mixed model stratified by elevated baseline CRP<br><br>2. Logistic models | Logistic models stratified by gender and TNFi | Linear mixed model | 1. Weighted GEE<br><br>2. Weighted logistic models |
| Covariates | Age, symptom duration, sex, education, exercise, HLA-B27, classification as AS, BMI, baseline BASDAI (or ASDAS) and BASFI.<br><br>Logistic models additionally included elevated baseline CRP. | Age (quartile), gender, disease duration (tertiles), year starting TNFi (tertiles). Categorising continuous variables reduces control for confounding.<br><br>Sensitivity analysis additionally adjusted for baseline BASDAI, BASFI, BASMI, disease duration, physician global. | Age, gender, education, employment, comorbidity, use of DMARDs, NSAID and analgesic drugs.<br><br>Separate analysis additionally adjusted for baseline BASDAI. | None in the outcome model. IPTW included: age, gender, symptom duration, education, elevated baseline CRP, classification as AS, deprivation, BMI, comorbidities, HLA-B27, alcohol status; baseline BASDAI, ASDAS, spinal pain, BASFI, ASQoL, BASG, fatigue, sleep and HADS subscores. |
| Results | Among those with elevated baseline CRP, current smokers had poorer BASDAI (0.75 units, p=0.005) and ASDAS responses (0.69 units, p=0.001) than non-smokers.<br><br>Difference not significant in the subgroup without elevated CRP. | Current (OR 0.48, P<0.001) and ex-smokers (OR 0.53, p=0.002) were both less likely to achieve BASDAI50/2 compared with non-smokers at 3 months.<br><br>Results were similar at 6 months and according to TNFi types and | Coefficients for interaction terms were not reported, except that they were P>0.36 | Response to TNFi did not differ according to smoking status.<br><br>Smoking status did not affect odds of achieving BASDAI50/2 response. |

| | | gender. | | |
|---|---|---|---|---|
| | Current smokers had reduced odds of achieving BASDAI50 (OR 0.54, p=0.03), ASDAS-MI (OR 0.43, p=0.01) and ASAS40 (OR 0.43, p=0.004) compared with non-smokers at 1 year. Previous smoking did not influence response. | | | |

*Glintborg et al also studied time to discontinuation.

BASDAI50, 50% reduction in BASDAI; BASDAI50/2, 50% or 2-unit reduction in BASDAI; ASDAS-MI, ASDAS major improvement; ASAS40, 40% improvement in ASAS core set; BASDAI, Bath AS disease activity index; ASDAS, AS disease activity score; BASFI, Bath AS functional index; AQoL, assessment of quality of life; ASQoL, AS quality of life questionnaire; BASG, Bath AS Global Score; JSEQ, Jenkins Sleep Evaluation Questionnaire; HADS, Hospital Anxiety and Depression Scale; SF-12/36, 12- or 36-item short form health survey; HAQ-S, health assessment questionnaire for spondylitis; IPTW, inverse probability of treatment weight; IPCW, IP censoring weight.