

## PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (<http://bmjopen.bmj.com/site/about/resources/checklist.pdf>) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

## ARTICLE DETAILS

<b>TITLE (PROVISIONAL)</b>	Parkinson's disease, temporomandibular disorder pain, and bruxism and its clinical consequences. A protocol of a single-centre observational outpatient study
<b>AUTHORS</b>	Verhoeff, Merel; Koutris, Michail; Berendse, Henk; Dijk van, Karin; Lobbezoo, F.

## VERSION 1 – REVIEW

<b>REVIEWER</b>	Karen Raphael New York University, Oral & Maxillofacial, Radiology and Medicine
<b>REVIEW RETURNED</b>	30-Jun-2021

<b>GENERAL COMMENTS</b>	<p>Reviewer: K.G. Raphael, PhD – kgr234@nyu.edu Dept. of Oral &amp; Maxillofacial Pathology, Radiology and Medicine, New York University College of Dentistry</p> <p>Importance of the proposal. Although I remained concerned about the clinical or scientific impact of the proposed protocol, I have been convinced that the protocol is worthy of publication. At the very least, when viewed in combination my comments, it should provoke important discussion.</p> <p>In part, my willingness to reexamine this protocol is motivated by a large study utilizing claims data from the National Health Insurance program in Taiwan 1. The study reported more than a two-fold increased risk of temporomandibular muscle and joint disorders (TMD) in individualized diagnosed with Parkinson's Disease (PD) compared to those without the PD diagnosis, but the ICD codes used for a "TMD" diagnosis reference joint but not muscle-based TMD. Diagnoses were not standardized and were purely clinical. A review of patient and community studies using standardized research diagnostic criteria 2 concluded that muscle-based "myalgia" or myofascial pain TMD is more prevalent than diagnoses involving the temporomandibular joint. The authors of the first study 1 comment that research using empirically standardized diagnostic coding system for TMD, identifying which specific TMD diagnoses may be found at elevated rates among people with PD, would be important. The proposed protocol does satisfy this need. Moreover, a pilot study 3 conducted by a set of authors partially overlapping the set of authors of the current protocol argues for potentially elevated rates of TMD and bruxism in people with PD, but assessments in the pilot were based on self-report.</p> <p>In addition, the authors have been responsive to my earlier concerns about organization of primary versus secondary aims, incorporation of results from brain imaging when available to confirm PD diagnosis, and acknowledgment of difficulties making</p>
-------------------------	--

distinctions between oral dyskinesias or oromandibular dystonia when assessing bruxism using instrumental methods. They have only been partially responsive to my concerns about incorrectly defining bruxism as a disorder rather than a behavior, stating in the introduction “The most common oral movement disorder in the dental office is bruxism.” Additionally, despite multiple claims in the protocol that bruxism is a risk factor for TMD pain, a recent scoping review coauthored by protocol coauthor Dr. Lobbezoo concludes that studies using unbiased instrumental bruxism assessment methods do not support this relationship 4.

Importance and interpretation of TMD pain in people living with PD. Pain is among the most common nonmotor symptoms reported by people living with PD 5. Especially in early/moderate PD, pain exerts greater influence on quality of life than motor symptoms 6 Certainly, oral health problems may affect quality of life among people living with PD. Xerostomia, drooling, and dysphagia have been noted among German patients with PD 7. The extent to which orofacial pain due to TMD affects quality of life among people living with PD is unknown. Of note, when comparing the prevalence of domains of pain among a large sample of people with PD versus controls without PD,8 orofacial pain was one of two areas in which the patients and controls did not differ significantly; strikingly, only 13% of PD patients reported orofacial pain, for which only a subset is likely to represent a TMD. Some may have been reporting odontogenic pain. In contrast, approximately 78% of both patients and controls reported musculoskeletal pain. Given these data and the high prevalence of other types of pain relative to low prevalence of orofacial pain, it seems difficult to argue that a focus on TMD assessment or treatment is likely to affect quality of life for most people living with PD. Unfortunately, despite a multitude of proposed measurements, no assessment of quality of life is proposed in the protocol. Although the PDQ-8 9 is an 8-item questionnaire measuring PD-specific symptoms and is labeled as a “quality of life” measure in the Appendix, its items address specific aspects of PD health rather than quality of life, asking about mobility, muscle cramping, concentration problems, and other symptoms.

Given the vast array of motor, cognitive and autonomic dysfunctions with which individuals living with PD must cope, treatment of a diagnosable TMD may well be low on the health priority list for people living with PD. The speculation that TMD treatment would lead to long and burdensome visits at the dental office has not been addressed by any cited data, even if conduct of the proposed protocol would be able to establish comorbidity. Other orofacial problems such as drooling may be more distressing. Moreover, since pain thresholds are clearly altered in people living with PD 10, it seems likely that at least some false positive diagnoses will occur when conducting mandatory palpation of the masticatory muscles and TM joint, part of the standardized DC/TMD examination 11.

Allow me to borrow from the title of a recent editorial written by a physician who lives with PD 12: She quotes a sociologist who wrote a half-century ago 13, “Not everything that can be counted, counts.” This is of particular concern when proposing to assess a sample of individuals with PD for TMDs according to a validated set of diagnostic criteria 11 which, nevertheless, have never tested for validity in diverse samples of individuals who have other

serious health conditions. We 14 first proposed the application of Jerome Wakefield's "harmful dysfunction" analysis 15 16 to improve conceptual clarity in understanding whether bruxism was a disorder; it was adopted by international consensus 17 and subsequently applied to aid evolution of the International Classification of Sleep Disorders 18. For a person living with a chronic neurodegenerative condition that impairs mobility, autonomic function and cognition as it advances, does meeting diagnostic criteria for TMD represent a problem worthy of pursuing treatment? Does meeting diagnostic criteria for a TMD mean that a person living with PD experiences harm or dysfunction in the orofacial region? If a DC/TMD diagnosis occurs in a person living with PD and who is not seeking TMD treatment, it may not even represent a disorder from a "harmful dysfunction" analysis.

In their response to my initial questions about the clinical impact of the proposed protocol, the authors' cover letter states "Knowledge of the factors that can influence... TMD pain in patients with PD will help dentists and other oral health care providers to provide individualised (sic) care to ... alleviate ... TMD pain...in this vulnerable group of patients." What type of individualized care would a dentist provide? The best treatment may often be the treatment recommended for other types of pain in PD: dopaminergic medication 19 20 A dentist or orofacial pain specialist without expertise in PD may well make inappropriate treatment recommendations.

Importance and interpretation of bruxism in people living with PD. Bruxism is currently considered an oral behavior that may potentially represent a risk factor for oral health problems. This new view is based on a widely cited international consensus statement on assessment of bruxism 17. I as well as many of the protocol's authors participated in the meeting from which the consensus statement arose. That meeting was motivated in part by an earlier paper 14 in which I and several colleagues, including protocol coauthor Dr. Lobbezoo, challenged the quality of evidence used to consider bruxism as anything more than a behavior. Yes, using very weak self-reports or clinical reports of bruxism, one may find an association between sleep bruxism and TMDs, but this relationship disappears when using polysomnographic and other instrumental measures. Most recently, protocol coauthor Dr. Lobbezoo 4 reaffirmed the strongly held position of this reviewer that bruxism must be evaluated in a continuous manner from polysomnographic or ambulatory instrumental methods, given problems with non-instrumental methods. These non-instrumental methods can produce only a categorical rating (possible, probable) based on clinical or patient self-reports. Nevertheless, the protocol states that one of the main parameters is the categorical "presence of bruxism (sleep and/or awake)." This seems scientifically counter to Dr. Lobbezoo's own published scientific recommendations 4. If protocol participants were able to use the GrindCare or BruxApp for multiple days to obtain a continuously scored measure of sleep or awake bruxism, why would the continuous data be arbitrarily dichotomized to be combined with judgments of bruxism from inferior sources? No comparable data from a demographically similar group of controls without PD will even be available to provide a hint of how to select a particular point on the continuum to dichotomize. To quote the earlier cited publication, 4 "bruxism...must be evaluated in its

continuum spectrum, rather than using a simplified dichotomous approach of presence/absence.”.

Of general concern is why bruxism became so tightly woven into the current protocol. What is the importance in evaluating bruxism in a group of individuals diagnosed with PD, if it is not a disorder? REM sleep behavior disorder is often prodromal to PD, observed in more than three quarters of those with PD 21 Although PD patients with REM sleep behavior disorder are more likely to self-report sleep bruxism than PD patients without REM sleep behavior disorder, they are more likely to report each of 8 different sleep symptoms or parasomnias. 22 Moreover, self-report of sleep bruxism bears no significant relationship to sleep bruxism assessed through state-of-the-art polysomnographic recordings of sleep bruxism. 23 Thus, although the protocol authors may believe that sleep bruxism could be a prodromal marker of PD, no evidence indicates this possibility, and the proposed protocol would not evaluate the hypothesis.

Secondary aims and critical design issues. The focus on tooth wear and saliva quality seems motivated by the proposed primary investigation of bruxism. As some of the authors of the protocol have noted in a prior review paper 24, it is irreversible. Once again, without a control group, interpretation of findings from any tooth wear rating system becomes difficult. Hyposalivation has been documented as more common in people living with PD than controls 25, and hyposalivation is associated with severe erosive tooth wear in a general population sample. 26 Why is it necessary to show this relationship in a PD sample, if it is already established in a general population sample?

Failure to include a control group in the protocol design creates many interpretive problems. The protocol authors plan to conduct within-PD group analyses to achieve many of their aims. Some still show lack of understanding of procedures related to PD. For example, DAT-SPECT, when used clinically, does not provide a quantitative score indicating the degree of presynaptic dopaminergic loss. Thus, this part of secondary aim 3 is likely to be unachievable. Also, it seems likely that any PD patients who have received Deep Brain Stimulation (DBS) for control of advanced PD symptoms would have to turn their stimulators off for five nights when instrumental bruxism recording via the Grindcare device is to be done. Turning off the stimulator would likely lead to symptom exacerbation. It would not seem justifiable, especially given the unclear importance of bruxism-related aims. Patients with DBS implants should be excluded from participation.

Ultimately, without using a control group, the prevalence of bruxism cannot be interpreted using continuous or arbitrarily categorized measures. To even say that this behavior is high or low in people with PD will be impossible. Without a contrast group matched on age and sex, it will be extremely difficult to know whether TMD pain is elevated in people with PD.

Deep knowledge about PD appears to be missing from the protocol design. The protocol clearly states, “Neither patients nor the community were involved in the design or performance of this study.” This statement is not a commendable one. In PD, as in many other life-altering conditions, an increasingly loud call has been made for involvement of “patient researchers” 12 27 28. I am a professor and clinical research scientist with expertise in TMDs

	and related musculoskeletal pain conditions. I have been living with PD for well over a decade. My hope is that my comments highlight the need to include “insiders” like me who know the everyday, lived experience of the condition under investigation, in design of clinical research protocols.
--	--

## VERSION 1 – AUTHOR RESPONSE

### Concerns of the reviewer:

1. The authors have only been partially responsive to the concerns about incorrectly defining bruxism as a disorder rather than a behaviour, stating in the introduction “The most common oral movement disorder in the dental office is bruxism.”

**Our response:** we would like to thank the reviewer for her excellent point. We adjusted lines 65-66: “Dentists regularly see patients with bruxism in the dental office, which is an oral health-related issue that is not necessarily associated with systemic diseases”.

### Concerns of the reviewer:

2. Most recently, protocol coauthor Dr. Lobbezoo reaffirmed the strongly held position of this reviewer that bruxism must be evaluated in a continuous manner from polysomnographic or ambulatory instrumental methods, given problems with non- instrumental methods. These non-instrumental methods can produce only a categorical rating (possible, probable) based on clinical or patient self-reports. Nevertheless, the protocol states that one of the main parameters is the categorical “presence of bruxism (sleep and/or awake).” If protocol participants were able to use the GrindCare or BruxApp for multiple days to obtain a continuously scored measure of sleep or awake bruxism, why would the continuous data be arbitrarily dichotomized to be combined with judgments of bruxism from inferior sources?

**Our response:** we would like to thank the reviewer for her remark. Indeed, we agree that it would be unsatisfactory to only use the continuous data and transform these to a dichotomous outcome measure for bruxism. However, no guidelines have been developed yet that can be used to interpret the continuous data obtained by the GrindCare or BruxApp. Nevertheless, the continuous data will be used to evaluate our aim “To identify which factors are associated with bruxism and TMD pain in PD patients”. To that end, we can use the frequencies (i.e., the number of bruxism events per hour) as measured by the devices. Besides, and importantly in relation to this study, not every participant will be able to use the GrindCare and/or the BruxApp. Therefore, we also use clinical outcomes to answer the first aim “To investigate the presence of bruxism and TMD pain in PD patients, through objective clinical and instrumental measurements”. But to be clear, we are fully aware that this is of lesser value of this approach than, for example, polysomnography. However, this study is already an important step towards more reliable data than those collected in the pilot study of Verhoeff et al. (2018)(see lines 216-219).

### Concerns of the reviewer:

3. DAT-SPECT, when used clinically, does not provide a quantitative score indicating the degree of presynaptic dopaminergic loss. Thus, this part of secondary aim 3 is likely to be unachievable. Also, it seems likely that any PD patients who have received Deep Brain Stimulation (DBS) for control of advanced PD symptoms would have to turn their stimulators off for five nights when instrumental bruxism recording via the Grindcare device is to be done. Turning off the stimulator would likely lead to symptom exacerbation. It would not seem justifiable, especially given the unclear importance of bruxism-related aims. Patients with DBS implants should be excluded from participation.

**Our response:** We would like to thank the reviewer for her remark. However, DAT imaging is considered to be a reliable tool for investigating the presynaptic loss of dopaminergic neurons through measuring the striatal binding and has been used as an imaging measure of disease severity previously (for example Berendse et al., motor and non-motor correlates of olfactory dysfunction in Parkinson’s disease, J Neurol Sci 2011; 310: 21-4; Van Dijk et al., eur j neurol 2014; 21:388-94). As mentioned in the protocol (line 196), a semi-quantitative analysis will be used after first calculating ratios of specific to non-specific binding of the tracer for the left and right putamen and caudate nucleus, using the occipital cortex as a reference area. These regions of interest are elucidated in the protocol (line 197-198). Therefore, aim three can be answered. Besides, we would like to thank the

reviewer for her suggestion regarding patients who underwent DBS. Indeed, it is not feasible to ask patients to turn off their DBS. Only patients who received a DAT-SPECT recently are included for this part, because the moment of our inclusions cannot be far apart from the imaging. In practice, this means that there will be no patients eligible for this part of the study who are already wearing a DBS implant. However, seeing the confusion, we adjusted this in the exclusion criteria, now indicating that we will exclude patients who are wearing a DBS implant (line 277-282, table 1).

**Concerns of the reviewer:**

4. The protocol clearly states, “Neither patients nor the community were involved in the design or performance of this study.” This statement is not a commendable one. In PD, as in many other life-altering conditions, an increasingly loud call has been made for involvement of “patient researchers”

**Our response:** We would like to thank the reviewer for her righteous and sincere remark. We are in agreement with the reviewer that, indeed, the involvement of patients is of great value. However, although our statement suggests that we did not evaluate the patients' point of view, we implemented the participants' feedback that we received during the earlier published pilot study (see lines 202-203)(Verhoeff et al., 2018). In the future, we will certainly involve patients in the planning and maybe also in the performance of our studies.

**VERSION 2 – REVIEW**

<b>REVIEWER</b>	Karen Raphael New York University, Oral & Maxillofacial, Radiology and Medicine
<b>REVIEW RETURNED</b>	05-Jan-2022

<b>GENERAL COMMENTS</b>	<p>This revised protocol remains extremely ambitious and creative. The fact that four separate bodies approved the protocol does not detract from the substance of my critique, since it seems unlikely that members of prior reviewing bodies had specific expertise in the protocol's content area. Of course, compromises need to be made in any study. Yet this protocol proposes a very large and costly study. I am not confident that, as designed, aims are either clear or attainable. Nevertheless, it has generated enough interest and raised so many issues that I think BMJ-Open's readership may enjoy and learn from the exchange of ideas.</p> <p>Of note, only minor changes have been made to the text of the protocol since the prior revision, and most of these alterations reflect wording changes suggested by the second reviewer. In contrast, the most recent cover letter to the editor, in which the authors responded to prior reviewers' comments organized by a numerical listing of issues raised, shows that considerable effort was made to defend prior logic and methodology. Rather than respond to the authors' comments on a point-by-point basis, I have tried to focus on the major themes that still raise concerns for me, referencing their numerical listings from the prior response letter to the editor, where appropriate. I use the abbreviations “PD” to refer to Parkinson's Disease and “PwP” to refer to people or a person with PD.</p> <p>First, the authors thank me (#10) for reminding them of publications which contradict the proposed approach to bruxism assessment, and which update the literature on pain and oromotor movement in PD, including literature suggesting that TMD-like pain is relatively rare compared to other pain in PD. Unfortunately, the protocol has not been revised to update their literature review.</p>
-------------------------	--

	<p>The main aim of the protocol is to investigate the presence [sic] of bruxism and TMD pain in PD patients. I will cover remaining concerns related to bruxism first and TMD second. Where efficient, discussion of remaining concerns related to secondary aims will follow.</p> <p>The research literature on bruxism has gone through an important paradigm shift over the past few years, best documented via an International Consensus Meeting on Bruxism and a subsequent, well-cited publication.<sup>1</sup> Nevertheless, even some of those actively promoting the paradigm shift may return inadvertently to views on bruxism which had been held for well more than 50 years. The first part of this paradigm shift is to recognize that the evidence for bruxism as a “disorder” needing treatment is not supported to date by the best methods for assessing it. Weak self-report measures and unreliable clinical assessments with poor specificity<sup>2</sup> support bruxism’s association with oral health problems. These associations become weak or altogether disappear when using instrumental assessment methods for bruxism.</p> <p>The gold standard for assessing sleep bruxism is polysomnography with audiovisual feedback to distinguish between sleep bruxism and other sleep movement artifacts. Ambulatory methods using electromyographic (EMG) sensors such as GrindCare® are a more practical and cost-efficient instrumental substitute but cannot distinguish sleep movement artifacts from sleep bruxism. These artifacts are likely to be frequent among PwP who are at increased risk of experiencing REM sleep behavior disorder and restless leg syndrome.<sup>3</sup> Thus, the correspondence between polysomnographic assessment and EMG-based ambulatory assessment among PwP may be weaker than expected. This creates problems in interpreting the ambulatory sleep bruxism scores. Might scores generated through Grindcare® ambulatory assessment of sleep bruxism be artifactually inflated?</p> <p>For awake bruxism, no “gold standard” to assess actual wake-time behavior corresponding to polysomnography exists. The best instrumental method to date is indirect event sampling, in which respondents use a smartwatch or smartphone app to self-report their oral behavior (“teeth contact”?) occurring when prompted, and at multiple times each day. The BruxApp software has an appropriate research version.</p> <p>Self-report and clinician assessment methods are acknowledged by the authors as inferior, bearing little or no relation to instrumental methods. Moreover, the first two methods have only been used to assess “presence” or “absence” of bruxism. Until more research is conducted using the best possible assessment methods, bruxism is best considered a behaviour with the potential to be shown as a risk factor or even a possible protective factor for oral health or other health outcomes. As such, it is no more a ‘condition’ (#17) than a ‘disorder.’ It should be assessed continuously,<sup>1 4 5</sup> unless or until future research identifies a natural cut-point or step function at which point the behaviour becomes a risk factor for negative health. “Presence” of bruxism, as stated explicitly in the first aim is uninterpretable using suboptimal methods, especially without an age-matched control group. Even data from a continuously scored Grindcare® assessment of sleep bruxism requires a control group for interpretation.</p> <p>In their response to points #13,15,16, the authors recognize the limitations of self-report of bruxism, as was used in their own pilot study. Nevertheless, they seem content to rely upon clinical</p>
--	---

assessment when ambulatory measures are not feasible. We are not provided with an a priori estimate about the frequency with which PwP will agree to and/or can accurately apply the ambulatory methods. For the first secondary aim, this creates an awkward need to conduct separate analyses to predict continuous versus dichotomized assessments of sleep and awake bruxism, as assessed via ambulatory versus clinical assessments respectively. It can be argued that any discrepancy in predictors of bruxism would be reconciled in favor of predictors of continuously scored bruxism. Alternately, PwP who can use the ambulatory assessment devices are likely to differ in PD progression from those who cannot, rendering comparisons between models extremely tenuous. Thus, the use of clinical assessment or bruxism seems to be a poor use of time and resources. Although it is not discussed, clinicians assessing bruxism would need to be blinded to results of the TMD examination, created further complexity in service of an inferior bruxism assessment method. The authors may have been rightfully cautious in failing to estimate the proportion of enrolled PwP who will successfully use ambulatory bruxism assessment methods. PwP may not agree to use ambulatory bruxism assessments. For sleep bruxism, using portable EMG (Grindcare®) may be a concern for PwP, given manual dexterity problems<sup>6</sup> causing electrode placement problems as well as potential concerns about the device exacerbating well-documented poor sleep quality in PD.<sup>3 7</sup> Similarly, dexterity problems may create participants' cautions in agreeing to ambulatory assessment of awake bruxism with the Smartphone-based BruxApp. Even early in the trajectory of PD, problems with multi-tasking related to executive dysfunction makes dual task performance challenging,<sup>8</sup> and BruxApp requires interruptions that may affect daily cognitive and motor task performance particularly for PwP.

Compared to assessment of bruxism, assessment of TMDs needed to achieve the other part of the primary aim is relatively straightforward, using a well-developed standardized diagnostic method (i.e., DC/TMD<sup>9</sup>). Nevertheless, two concerns remain. First, significantly lower pressure pain thresholds have been documented in PwP,<sup>10</sup> particularly when dopaminergic medication is suboptimal and PwP are in an "off" state. Cutaneous allodynia<sup>11</sup> and hyperalgesia<sup>12</sup> occur in many PwP. These factors may lead to false-positive diagnoses of questionable clinical significance. Second, the absence of an age-matched and sex-matched control group makes interpretation of any detected rate problematic.

The importance of secondary aim 1 remains unclear to me. If rates of bruxism and/or TMD pain are not elevated in PD patients, something impossible to know without a control group, why do the authors want to examine "knowledge of factors that can influence bruxism and/or TMD pain in patients with PD" in the first secondary aim? We know from multiple other studies (e.g., <sup>13</sup>) that well-assessed bruxism (via instrumental methods) and TMD pain are highly unlikely to be related in clinical TMD samples. (The one caveat is that nonfunctional tooth contact assessed via experience sampling appears to occur more often among patients with myogenous masticatory pain.<sup>14 15</sup>) In PwP, multiple PD-related factors are proposed to be modeled, but predictors of bruxism or TMD pain cannot necessarily be remediated. For example, if stage of PD relates to either bruxism severity or TMD pain, how can a dentist or a neurologist help? If a PwP has not sought treatment for a TMD, it seems rather unlikely that the knowledge hypothetically gained by undertaking this massive, proposed study



would “prevent long and burdensome visits at the dental office.” (#11 response) Why is it clinically informative to know what predicts bruxism in PwP, if bruxism is a behaviour without clear health consequences? Given the myriad of life-altering motor and nonmotor symptoms experienced by PwP, I cannot help but question the importance of focusing on predictors of bruxism and TMD pain in PwP.

The authors cite their own case-control self-report survey as sole evidence for elevated rates of TMD pain in PwP.<sup>16</sup> The title of that paper notes that it is a pilot study. The authors may wish to consider that another study,<sup>17</sup> using ICD codes for PD diagnosis and sampling data from the National Health Insurance Research Database in Taiwan (NHIRD), found that individuals with a PD ICD code followed prospectively for up to 13 years had a two-fold increase in risk of a new temporomandibular joint-related ICD code compared to a propensity score matched control group. Muscle-based facial pain was not considered. Whether the increased relative risk reflects generally altered central pain processing in PwP<sup>11 18</sup> and propensity to multiple pain conditions is unclear. Of note, the Taiwanese study predicting temporomandibular joint-related disorders<sup>17</sup> found 32 new incidents among 6,185 (0.5%) of those initially identified with PD. In contrast, using a shorter follow-up period and stricter PD identification criteria, another NHIRD-related study<sup>19</sup> found incident musculoskeletal pain not involving the orofacial region to occur in 199 of 490 individuals with PD, representing 40.6% of their PD sample. Despite difficulty in making direct comparisons across the studies, these reports confirm my earlier concerns about the relatively low burden of TMD-related pain versus other painful conditions in PwP.

The second secondary aim is to examine salivary predictors of tooth wear in PwP. In the abstract, the justification for this aim in the context of the overall protocol is that “this can be a major consequence of bruxism.” Unfortunately, although attrition-type tooth wear may influence clinical judgments of bruxism, polysomnographic studies do not find a relationship between tooth wear and frequency of sleep bruxism events evaluated via ambulatory polysomnography.<sup>20 21</sup> Two factors are probably involved in the failure to find the expected relationship. First, the etiology of tooth wear is multifactorial, involving a combination of mechanical and chemical wear. <sup>22</sup> Second, observed tooth wear reflects a lifetime of factors causing wear; it does not solely or necessarily reflect current factors.

The final secondary aim is to examine the relation between DAT-SPECT-derived measurement of extent of dopaminergic dopamine loss and bruxism. Put aside the reasoning for this aim, which remains unclear to me. In my prior review, I queried the authors about how often DAT-SPECT is conducted in PD patients in the setting from which participants will be recruited. (#18) The revised protocol now says “The estimated percentage of additional brain imaging in newly referred patients is 40%.” This includes both MRI and DAT-SPECT and, according to the authors, is done “mainly in cases of clinical doubt.” So, we can assume that a disproportionate number of new patients seen in this clinical service who receive brain imaging are ultimately determined to not meet study inclusion criteria, in that DAT-SPECT is negative for PD or MRI identifies a different condition causing symptoms. Moreover, the authors do not break down that 40% estimate by DAT-SPECT versus MRI, but the proportion receiving DAT-SPECT must be lower than 40%. Finally, only some DAT-SPECT imaging services are likely to provide a semi-quantitative or

	<p>quantitative score. For example, in a recent U.S.-based clinical study<sup>23</sup>, a nuclear medicine specialist provided a simple judgment of whether the DAT-SPECT image's overall shape and intensity of the striatal signal was consistent or inconsistent with a diagnosis of PD. If this is also true for the planned recruitment site, achievement of this aim will not be possible. Even if quantitative or semi-quantitative scoring is available via software to rescore images, the sample size to achieve a poorly justified aim is unknown.</p> <p>Let me now move beyond specific aims to mention a few remaining concerns.</p> <p>The concern about sample size extends to still not knowing the total number of potentially eligible PwP at the single recruitment site. The plan for 382 PwP to be enrolled may be justified by power analyses but whether it approaches feasibility is unknown. The protocol lists only the number of new patients seen annually. (#20)</p> <p>I continue to assert that multivariate models predicting presence (or, better, relative frequency) of bruxism or diagnosis of TMD is underspecified (#21). To clarify: this is not a statistical issue, but a conceptual one. Carefully thinking through the relation among predictors is critical when using a forward selection procedure. Without careful a priori model specification, it is impossible to differentiate between confounding and mediation when a measure drops out of the predictive model. Think of some problems with interpreting predictive measures. For example, impulse control disorders are more likely particularly for PwP who take high doses of dopamine agonists, <sup>24</sup> <sup>25</sup> although it appears that only total levodopa-equivalent dose will be calculated in this study. Cognitive function variability is inherently limited by excluding PD patients with a score on the Montreal Cognitive Assessment scale (MoCA) necessarily less than 21 according to inclusion criteria, truncating variability. Disease stage and disease severity are likely colinear. Will both terms be included in the same model? If higher order interactions are to be considered, some logic for which interaction terms will be considered is needed. These are just some of the problems that I see with the proposed modeling for secondary aim 1.</p> <p>In sum, I certainly think that it would be worthwhile to use state-of-the-art diagnostic procedures to understand whether and to what extent PwP are at increased risk for clinically significant TMDs, especially when compared to risk for other pain conditions. The protocol jumps far beyond aiming for this basic understanding and, therefore, seems to miss a fundamental foundation.</p>
--	---

## VERSION 2 – AUTHOR RESPONSE

>>> Our response: Thank you for pointing this out to us. We have revised the title according to your suggestion (see line 3).

Thank you updating the abstract with your dissemination plan. Please also update the 'Ethics and dissemination' section of your main article to include details of your dissemination plan. See published articles for examples.

>>> Our response: Thank you for your comment. We have modified the manuscript accordingly (see lines 286-287).

Please include the planned start and end dates for the study in the methods section. We note that in your response you have stated that it isn't possible to provide a study end date to due COVID-19 related inclusion delays. Please provide the study start date and then either an estimated end date, or explain in the manuscript that the end date is currently unclear due to delays.

>>> Our response: Thank you for your comment. We have added the requested dates to the manuscript. Please note that not only the end date is an estimation, but also the start date: the regular care at the Amsterdam University Medical Centres is still scaled down due to the covid19 pandemic (see lines 140-141).

Please ensure that the information provided in your protocol article is consistent with that included in the trial registry. For example, in the manuscript, line 186, the sample size required is given as 382, but the trial registry states 246. The inclusion/exclusion criteria are also not consistent. For example, Table 1 in the manuscript states that patients with > 21 on the Montreal Cognitive Assessment will be included, but the trial registry states that patients with Montreal Cognitive Assessment score < 21 will be excluded. What happens to patients with MoCA = 21? Please update the manuscript and/or trial registry accordingly.

>>> Our response: Thank you for pointing this out. The manuscript is now aligned with the registered protocol.

Please remove the SPIRIT checklist from this submission since this is the protocol of an observational study, not a clinical trial.

>>> Our response: Thank you for this comment. The checklist has been removed.

Please ensure that reviewer comments are reflected by adequate modification to the text, not just explained in the point by point response.

>>> Our response: Notwithstanding our message to you, above, we inserted the following references as per Prof. Raphael's current and earlier reports:

- Chen YY, Fan HC, Tung MC, et al. The association between Parkinson's disease and temporomandibular disorder. *PLoS One* 2019;14(6):e0217763. doi: 10.1371/journal.pone.0217763 [published Online First: 2019/06/15] (see lines 109-112)

- Manfredini, D., Ahlberg, J., Wetselaar, P., Svensson, P., & Lobbezoo, F. (2019). The bruxism construct: From cut-off points to a continuum spectrum. *J Oral Rehabil*, 46(11), 991-997. (See lines 242-245)

Politis, M., Wu, K., Molloy, S., P, G. B., Chaudhuri, K. R., & Piccini, P. (2010). Parkinson's disease symptoms: the patient's perspective. *Mov Disord*, 25(11), 1646-1651. (See lines 70-71)

- Raphael, K. G., Santiago, V., & Lobbezoo, F. (2016). Is bruxism a disorder or a behaviour? Rethinking the international consensus on defining and grading of bruxism. *J Oral Rehabil*, 43(10), 791-798. (See lines 242-245)

- Silverdale, M. A., Kobylecki, C., Kass-Iliyya, L., Martinez-Martin, P., Lawton, M., Cotterill, S., et al. (2018). A detailed clinical study of pain in 1957 participants with early/moderate Parkinson's disease. *Parkinsonism Relat Disord*, 56, 27-32. (See lines 70-71)

### VERSION 3 – REVIEW

<b>REVIEWER</b>	Karen Raphael New York University, Oral & Maxillofacial, Radiology and Medicine
<b>REVIEW RETURNED</b>	24-Feb-2022

**GENERAL COMMENTS**

I have given the last response to my review careful consideration. By offering specific requests for revision rather than offering a general critique, I hope that satisfactory revisions to this manuscript can be made. For the latest version of the manuscript (Feb 2022) to be acceptable for publication, it is strongly advised that the discussion section be expanded to address significant limitations or cautions in interpretation of results from the study:

1. Without a control group matched on age and gender, interpretation of whether the prevalence of bruxism or TMDs is low or high in people with PD will not be possible. (I suggest that the plans for an interim analysis mentioned on pg. 9, line 63 be removed, because establishment/interpretation of prevalence will not be possible.) Even examination of tooth wear becomes problematic, because most people with PD are older adults in which tooth wear reflects a lifetime of factors; some PD medications and non-PD medications taken for other conditions associated with aging may cause xerostomia, exacerbating wear problems.

2. The authors should acknowledge that, if models predicting bruxism as a function of dichotomized bruxism (s/r or clinician assessment) differ from models based on ambulatory assessment, the latter set of models are likely to be more sensitive and accurate. Nevertheless, participants able to complete assessments may differ from those who do not complete ambulatory assessments due to differences in severity of PD symptoms. Fine motor problems which occur in PD create barriers for cell phone use and electrode placement as required for ambulatory assessments of awake and sleep bruxism. Thus, It is recommended that the authors propose to test for PD symptom differences between subgroups defined by comparing participants completing or not completing ambulatory assessments. If differences are found, this will indicate limitations to the external validity or generalizability of conclusions involving bruxism modeling.

3. The 2018 International Assessment paper on assessment of bruxism regularly added the phrase that bruxism is a masticatory muscle activity in “otherwise healthy individuals.” People with PD are certainly not “otherwise healthy.” Much as REM sleep behaviour disorder was used in the 2018 paper as an example of a situation in which masticatory muscle activity would not be considered bruxism but instead a sign of an underlying disorder, PD (for which REM sleep behaviour disorder is a well known and robust risk factor) often involves ongoing REM sleep behaviour disorder and, in later stages of levodopa-treated PD, dyskinesias, including oral dyskinesias. For people with PD, is masticatory muscle activity “bruxism” at all? This point should be added as a limitation, leading to caution in interpretation of any analyses from the study related to bruxism. At the very least, it will add to limits on external validity of any findings beyond masticatory muscle activity specifically in a PD sample.

4. Arguably the most common nonmotor symptom in PD is pain (see references from my earlier reviews). Although the complex role of dopamine in descending pain inhibition has not been fully elaborated, pain exacerbation as a characteristic of “wearing off” of dopaminergic medication in people living with PD

	<p>is well known. Moreover, research (see examples from my prior reviews) establishes that individuals living with PD have lower pain thresholds than similar individuals without PD. Thus, TMD examiners will/should pay careful attention to DC-TMD instructions to query participants who experience pain on masticatory muscle or joint palpation to ensure that any such elicited pain is “familiar pain” rather than a reflection of a general lower threshold for pain on palpation.</p> <p>Other changes not specific to the discussion section:</p> <p>5. “Condition” should be replaced with “behaviour” whenever referring to bruxism. I believe that some changes have already been made, but some terminology referring to bruxism as a “condition” remain. The term “condition” implies an unnecessary dichotomy and merely softens but does not remove the unacceptable inference that bruxism is a “disorder.”</p> <p>6. Minor revision: Page 8 line 22 states: “Neither patients nor the community were involved in the design or performance of this study.” I believe this should be corrected to read: “Neither patients nor the community were involved in the design of the study. Patients with PD will be involved in the performance of this study.”</p> <p>7. Finally, if it is possible logistically, I strongly recommended that a minor protocol revision be submitted so that clinicians making participant assessments for bruxism or TMDs are blinded to results of other assessments (either findings from instrumental ambulatory assessments or other clinician assessments, i.e., clinicians doing DC/TMD should not be the same as clinicians assessing bruxism). I would not make acceptance of the manuscript contingent upon this protocol modification, since I wish to be sensitive to practical constraints and the journal’s requirement.</p>
--	---

### VERSION 3 – AUTHOR RESPONSE

1. Without a control group matched on age and gender, interpretation of whether the prevalence of bruxism or TMDs is low or high in people with PD will not be possible. (I suggest that the plans for an interim analysis mentioned on pg. 9, line 63 be removed, because establishment/interpretation of prevalence will not be possible.) Even examination of tooth wear becomes problematic, because most people with PD are older adults in which tooth wear reflects a lifetime of factors; some PD medications and non-PD medications taken for other conditions associated with aging may cause xerostomia, exacerbating wear problems.

>>> We thank the reviewer for this comment. As to indicate the mentioned limitations, we have added the following to the Discussion: “This study does not include a control group. This limits the interpretation of whether the prevalence of bruxism or TMDs is low or high in people with PD, which will only be possible by comparing the findings with prevalences as reported in the literature. In addition, since tooth wear in older people reflects a lifetime of factors, it will be also difficult to interpret the tooth wear findings in people with PD without the possibility for a direct comparison with similar individuals without PD. Also in this case, comparisons should be sought with literature data. These issues should be considered limitations of this study.” In addition, we have added the following limitation to the list of Strengths and limitations of this study: “Because of the absence of a control group, no direct comparisons between individuals with PD and similar individuals without PD can be made.”

2. The authors should acknowledge that, if models predicting bruxism as a function of dichotomized bruxism (s/r or clinician assessment) differ from models based on ambulatory assessment, the latter set of models are likely to be more sensitive and accurate. Nevertheless, participants able to complete assessments may differ from those who do not complete ambulatory assessments due to differences in severity of PD symptoms. Fine motor problems which occur in PD create barriers for cell phone use and electrode placement as required for ambulatory assessments of awake and sleep bruxism. Thus, It is recommended that the authors propose to test for PD symptom differences between subgroups defined by comparing participants completing or not completing ambulatory assessments. If differences are found, this will indicate limitations to the external validity or generalizability of conclusions involving bruxism modeling.

>>> We thank the reviewer for this important comment, which we have elaborated in the Discussion as follows: "Importantly, participants able to complete all assessments may differ from those who cannot complete instrumental assessments due to differences in severity of their PD symptoms. Fine motor problems which occur in PD create barriers for electrode placement and cell phone use as required for instrumental assessments of sleep and awake bruxism. Therefore, we will test for PD symptom differences between subgroups defined by comparing participants completing or not completing instrumental assessments. If differences are found, this will indicate limitations to the external validity or generalizability of conclusions involving bruxism modeling." In addition, we have added the following to the Methods, paragraph on Main study parameters: "Differences in PD symptoms between those who can, and those who cannot complete the instrumental assessments will be tested as to gain insight into the external validity or generalizability of the conclusions involving bruxism modeling."

3. The 2018 International Assessment paper on assessment of bruxism regularly added the phrase that bruxism is a masticatory muscle activity in "otherwise healthy individuals." People with PD are certainly not "otherwise healthy." Much as REM sleep behaviour disorder was used in the 2018 paper as an example of a situation in which masticatory muscle activity would not be considered bruxism but instead a sign of an underlying disorder, PD (for which REM sleep behaviour disorder is a well known and robust risk factor) often involves ongoing REM sleep behaviour disorder and, in later stages of levodopa-treated PD, dyskinesias, including oral dyskinesias. For people with PD, is masticatory muscle activity "bruxism" at all? This point should be added as a limitation, leading to caution in interpretation of any analyses from the study related to bruxism. At the very least, it will add to limits on external validity of any findings beyond masticatory muscle activity specifically in a PD sample.

>>> We thank the reviewer for this comment. As to further elaborate on this important issue, we have added the following to the Discussion: "In fact, in their updated international consensus paper on bruxism, Lobbezoo et al. (2018 ) added the phrase that bruxism is a masticatory muscle activity in "otherwise healthy individuals"<sup>2</sup>. People living with PD are certainly not "otherwise healthy". In the later stages of levodopa-treated PD, dyskinesias, including oral dyskinesias, commonly occur<sup>68</sup>. Hence, the question could be raised if the masticatory muscle activity observed in people with PD is "bruxism" at all. This calls for caution in the interpretation of the bruxism-related findings of this study."

4. Arguably the most common nonmotor symptom in PD is pain (see references from my earlier reviews). Although the complex role of dopamine in descending pain inhibition has not been fully elaborated, pain exacerbation as a characteristic of "wearing off" of dopaminergic medication in people living with PD is well known. Moreover, research (see examples from my prior reviews) establishes that individuals living with PD have lower pain thresholds than similar individuals without PD. Thus, TMD examiners will/should pay careful attention to DC-TMD instructions to query participants who experience pain on masticatory muscle or joint palpation to ensure that

any such elicited pain is “familiar pain” rather than a reflection of a general lower threshold for pain on palpation.

>>> We thank the reviewer for this comment. As to clarify this important issue, we have added the following to the Discussion: “An important aspect of a TMD-pain diagnosis according to the DC/TMD is that it considers the aspect of “familiar pain” as part of the diagnostic algorithm. As such, PD-related pain characteristics like pain exacerbation due to “wearing off” of dopaminergic medication and lower pain thresholds in individuals living with PD as compared to similar individuals without PD<sup>1</sup> will be taken into account.”

Other changes not specific to the discussion section:

5. “Condition” should be replaced with “behaviour” whenever referring to bruxism. I believe that some changes have already been made, but some terminology referring to bruxism as a “condition” remain. The term “condition” implies an unnecessary dichotomy and merely softens but does not remove the unacceptable inference that bruxism is a “disorder.”

>>> We thank the reviewer for this comment. We have made the requested changes in the manuscript.

6. Minor revision: Page 8 line 22 states: “Neither patients nor the community were involved in the design or performance of this study.” I believe this should be corrected to read: “Neither patients nor the community were involved in the design of the study. Patients with PD will be involved in the performance of this study.”

>>> We thank the reviewer for this comment. We have made the requested change in the manuscript.

7. Finally, if it is possible logistically, I strongly recommended that a minor protocol revision be submitted so that clinicians making participant assessments for bruxism or TMDs are **blinded** to results of other assessments (either findings from instrumental ambulatory assessments or other clinician assessments, i.e., clinicians doing DC/TMD should not be the same as clinicians assessing bruxism). I would not make acceptance of the manuscript contingent upon this protocol modification, since I wish to be sensitive to practical constraints and the journal’s requirement.

>>> We thank the reviewer for this suggestion. We have added the aspect of blinding of the dentists clinically assessing bruxism and TMDs regarding the results of the instrumental assessments (GrindCare and BruxApp) to the Methods, at the end of the paragraph on Main study parameters: “Dentists making clinical assessments for bruxism or TMDs will be blinded to the results of the instrumental assessments (i.e., GrindCare® GC4 and BruxApp for sleep bruxism and awake bruxism, respectively).” Since the clinicians assessing bruxism will be the same ones as those assessing TMD, blinding of the results of both parts of the clinical assessment will not be possible.