Right on track: Towards improving DBS patient selection and care

Geraedts V.J.

CHAPTER 10

General discussion

Status of DBS care and considerations for improvement

The timeliness of referral for DBS is likely to become a greater issue in the nearby future. Referring neurologists may anticipate on increasingly long waiting lists and also refer patients at an earlier stage as a likely consequence of the results of the EARLYSTIM trial.[1] Better understanding of DBS referral practices could potentially improve DBS care from a referral-perspective.

Our findings show that 26% of the DBS referrals are rejected on the basis of factors that can be established prior to the referral, suggesting there is room for improvement of the referral process which in turn may reduce waiting lists of outpatient DBS centers and disappointment following rejection. Many of the current screening tools advocate a high sensitivity and low specificity to ensure that patients are not withheld a 'potentially better therapy' than oral therapy in the form of DBS.[2] However given the potential risks of DBS surgery, one should keep in mind the 'first do no harm' principle and the decision on eligibility should be made on an individual basis regardless of disappointment following justifiable rejection. There is no easy way to develop a dichotomous classification algorithm (i.e. rejected vs. accepted patients) that could aid clinicians with appropriate referrals. In fact, such a classification algorithm was attempted by our group (data not shown), but was discarded as the obtained accuracies turned out to be particularly low. The reported areas-under-the-curve are therefore best interpreted as a demonstration of the additional benefit of adding patients' expectations to a screening model, rather than absolute accuracies. A major limitation lies in the dispersity of the 'rejected class', which can be crudely subdivided into two classes: (1) patients rejected due to 'too advanced' disease fulfilling one or more exclusion criteria, or (2) patients referred 'too early' and with room for adjustment of medical (oral) therapy, i.e. not fulfilling the inclusion criteria for DBS surgery. The first subgroup ('too advanced') may have several contraindications for surgery, such as cognitive impairment, balance impairment, or medication-resistant freezing.

In contrast, the second subgroup ('too early') appears to consist of patients with a relatively good cognition and balance, but is characterized by the absence of (debilitating) motor complications whilst under optimal oral therapy (patients had thus either mild or no motor complications or suboptimal oral therapy). The patients who are ultimately accepted for DBS are in the spectrum between those more extreme subgroups which opposes a binary classification algorithm. An unfortunate but occasionally-encountered scenario during data collection was when patients were initially referred 'too early' for DBS, and ultimately ended up re-appearing for DBS screening when the disease had progressed beyond eligibility.

Not all contraindications are considered by clinicians to be equally important in all patients, as exemplified by 'unrealistic expectations' contributing to the reason to reject in 38% of rejected patients. It may be argued that 'unrealistic expectations' may be modifiable and therefore not a strict contraindication for surgery per se, whereas for example severe cognitive impairment would constitute a clear and definitive contraindication for surgery. Moreover, contraindications for surgery often clustered within patients, as exemplified by 'unrealistic expectations' being associated with the presence of other exclusion criteria. Both clinicians and patients should weigh the risks per domain (i.e. cognition, balance, etc.) and determine whether the benefits generated by DBS outweigh the individual risks.

An equally striking but different issue is that there is a 23% chance that a referred patient has unrealistic expectations of surgery. Several mechanisms may underlie this observation: (1) the patient was not or inadequately educated by the referring neurologist, (2) the patient was adequately educated by the referring neurologist but the patient retained unrealistic expectations nonetheless. Improving patient education on the potential benefits of DBS through the national Parkinson patient association may reduce this problem, whereas clinical meetings to keep referring neurologists up-to-date with the most recent developments on DBS effects would circumvent the first mechanism. Concerning preoperative expectations of DBS, two crucial questions remain: (1) to what degree should expectations be leading (or even be an exclusion criterion per se)?, and (2) what is the effect of preoperative expectations on postoperative outcomes? It is up to the physicians' discretion to answer the first question on an individual basis, although a shared-decision-making approach seems particularly appropriate with regard to elective brain surgery. Our dataset included an insufficient number of patients who had unrealistic expectations prior to surgery and nevertheless received DBS surgery to study the effect of unrealistic expectations on postoperative outcomes.

Chapter 3 attempts to answer which preoperative factors influence postoperative subjective outcomes in the form of QoL. Apparently, QoL after STN DBS is particularly heterogeneous and individually influenced, as well as dependent on both used scale and follow-up duration. Although the provided overview summarizes all available studies, a quantitative synthesis was not provided due to differences in outcome measures, study design, follow-up and likely heterogeneity between studies.[3] Moreover, several aspects influencing QoL outcomes were not addressed in any study, such as the high level of individual variation,[4] cultural influences,[5] social adjustments, and interpretability of the different metrics. The choice to classify studies based on statistical significance is clearly subject to debate,[6] however we stand by our conclusions as the reported effect sizes for all factors considered 'non-significant' were relatively small and unlikely to yield a meaningful clinical contribution after pooling of the studies and increasing the sample size.[7] Interestingly, several contra-indications for surgery

such as impaired cognition or psychiatric dysfunction have limited effects on postoperative QoL, albeit within the limits of current clinical practice. No inferences on a wider spectrum of symptoms than that is currently studied can be made. Although selection criteria for DBS eligibility are based on likelihood of 'success' such as motor improvement or absence of cognitive decline, DBS effects on QoL should be considered on an individual basis as well.

Chapter 4 has provided some suggestions for improving the efficiency of finding the optimal chronic DBS settings by demonstrating that the search space for the optimal contact point may be reduced, and by demonstrating that postoperatively higher stimulus intensity is required to induce any clinical effect (either therapeutic- or side-effect) with respect to intraoperative testing. Whether this translates into actual improvements in terms of increased clinical efficiency needs to be validated and the magnitude of effect in terms of time-gain is yet unknown. Nevertheless, optimization of DBS settings will likely become increasingly time-intensive given developments in DBS setting-modalities such as increasing number of contacts per DBS lead or directional steering (i.e. more test-options available). Faster optimization would in theory mean less visits to the hospital, shorter visits to the neurologists, fewer costs, shortened 'adjustment phase' to alterations in everyday life, potentially higher patient satisfaction and improved QoL. The proposed mechanisms behind the observed differences between intraoperative test stimulation and postoperative stimulation settings need to be studied in further detail, e.g. by validating our findings in different targets such as pallidal or thalamic DBS.

Chapter 5 indicates a clear area to improve postsurgical DBS care, by partly answering the question which factors influence postoperative satisfaction. STN DBS exerts its primary effect on motor function, and generally has no effect on symptoms unresponsive to dopaminergic treatment. Apparently, motor performance scores were not associated with postoperative satisfaction whereas non-dopaminergic dysfunction correlated to lower valuations of surgery. 'That what does not improve' therefore appears to have a bigger impact than 'that that does improve' in terms of satisfaction, although the relief of motor complications, i.e. severity of dyskinesias or 'OFF' time, was not examined in detail as all patients had similar postoperative profiles of mild or negligible motor complications. As patients' satisfaction is one of the ultimate goals of any intervention, insight into factors influencing post-intervention satisfaction is of paramount importance. Since we demonstrated that the severity of non-dopaminergic symptoms was relatively unchanged after DBS, it may be that patients retained the unrealistic expectations described in chapter 2, of wanting relief from those symptoms despite patient-education prior to surgery (although preoperative expectations were not incorporated in this study). Another hypothesis is that due to the relief of motor complications, these are no longer the most prominently debilitating symptoms

and other symptoms take on a more prominent role in patients' lives resulting in more severe valuations on patient-reported outcomes. Careful monitoring and (multidisciplinary) treatment of non-dopaminergic symptoms may be a potential target for studies targeting improving clinical care.

Patients' insight into DBS effects and subsequent improved appreciation of this intervention can be mediated by a SCT, particularly in suboptimally satisfied patients. Patients with maximum scores in terms of satisfaction do not necessarily have to be subjected to a SCT to enhance subjective valuations further, as indicated by some patients who report a decline in postoperative satisfaction following SCT (as patients with maximal scores can only retain their scores or decline on the Likert scales). However, there are other reasons to perform SCT apart from improving postoperative satisfaction, such as accurate assessment of DBS motor benefit and comparing results to the preoperative Levodopa Challenge Test to identify whether DBS settings have to be adjusted accordingly for maximal benefit. It is questionable whether the observed improvements in postoperative satisfaction are sustained over time, or whether repeated SCTs would lead to sustained patients' perception and postoperative satisfaction. These considerations should be determined on an individual basis by both treating physicians and patients. Nevertheless, we recommend to incorporate SCTs into routine postoperative care after DBS especially in case of suboptimal satisfaction.

Future perspectives with regard to DBS care

Part A of this thesis proposes means to improve DBS referral practices prior to surgery, increase the optimization of DBS settings during the early postoperative phase, and increase patients' postoperative satisfaction one year after surgery.

Future studies should identify whether more extensive education of the pros and cons of DBS surgery would lead to an improvement of referral practice and fewer unrealistic expectations of DBS, as well as investigate the cause of these unrealistic expectations in the first place. Several potential factors which are insufficiently studied with regard to predicting QoL after DBS, such as social functioning or genetic factors, need to studied in greater detail. The proposed mechanism to increase the efficiency of finding the optimal DBS settings needs to be evaluated in terms of the magnitude of actual time-gain. Furthermore, an extrapolation of the findings from chapter 4 towards different targets needs to be performed to determine the generalizability of our findings. Lastly, it needs to be studied whether increases in postoperative satisfaction after SCT are sustained over time and whether repeated SCTs may be useful if increases in satisfaction are not sustained.

The utility of quantitative EEG during the DBS screening

An overview of EEG features that correlated to PD symptoms is provided in chapter 6. Similar to chapter 3, a quantitative synthesis is not provided. Ideally, a biomarker should not have to depend on a pooling of results to show clinical utility. Spectral EEG markers have been abundantly studied, likely due both to an easier computation as well as more straightforward interpretability. The choice of spectral metric does not appear to matter much, since spectral measures are highly interrelated after all. However, there are legion opportunities to define 'connectivity' [8] and only a small subset of those has been studied in relation to PD symptoms. A comparison between multiple types of connectivity metrics may provide greater insight into the pathophysiological mechanism behind the correlation between qEEG metrics and PD symptoms. Moreover, connectivity metrics have not been properly compared amongst one another and it is currently unknown which is the 'best metric' in terms of discriminating power for any PD symptom. Even beyond single connectivity metrics, combining several metrics into coherent networks to define an EEG-profile may determine the neurophysiological signature of a patient with e.g. PD-MCI or PDD.[9]

Most studies focus on the correlation of qEEG with cognition: no longitudinal studies focussed on another domain than cognition. Spectral analyses show promise in predicting progression of cognitive (dys)function. There is limited evidence for biomarkers transcending spectral analyses to predict progression in any domain. It needs to be elucidated whether the mechanism of cognitive decline after STN DBS is similar to the cognitive deterioration in the general PD population which is not attributable to an intervention. Second, although spectral analyses have the best chance of finding their way to clinical practice given the relative ease of computation and interpretation, new and potentially more complex biomarkers should be evaluated against the current 'gold standard' of spectral analyses in order to identify whether discriminating accuracy in term of cognitive function can be further improved.

The utility of qEEG as a biomarker of cortically mediated symptoms was further demonstrated in chapter 7, which shows a correlation of qEEG with the cortical symptoms cognition and psychotic symptoms, but not with autonomic function, balance impairment of motor symptoms. It may be hypothesized that both correlations are mediated through a common mechanism, i.e. disturbances within an α-network. This would explain the correlation with α-connectivity, whereas the correlation with relative α -power would resonate throughout the other relative spectral powers as long as the correlation with α-power is sufficiently strong. Whereas the concept of an α-network in relation to cortically mediated PD symptoms is interesting, a drawback of the studied phase-based connectivity measures is that it only demonstrates a temporal relationship and provides no insight into causation. Metrics

focussing on directed entropy could possibly localize the (causal) 'driver' behind cortical dysfunction if combined with source localization rather than studying effects in sensor space. The concept of several distinct subnetworks corresponding to the distinct PD symptoms may fuel multimodal neurophysiological analyses, with differences in terms of spectral density and penetration (i.e. MEG vs. EEG), as well as differences in recording conditions depending on the symptom of interest. The use of resting-state EEG is the standard technique for studying cognition, but may be less applicable for studying motor networks for which perturbation tests may be more applicable.[10, 11] Perturbation tests for cognition are less feasible due to learning-effects and consequential attention-wandering.

Apart from the contribution to knowledge on the pathophysiological mechanism of PD symptoms, chapter 7 provides some evidence for the feasibility of using qEEG during the screening for DBS as a complementary biomarker. Earlier studies have shown that a composite score of non-dopaminergic symptoms, may provide a more complete and accurate evaluation of disease severity and progression in PD.[12, 13] Given the correlation of the qEEG measures global slowing and global desynchronization with a composite score of non-dopaminergic symptoms, these EEG markers likely reflect cortical involvement of α-synucleinopathy. Consequently, qEEG measures hold potential to contribute to the process of determining a patient's candidacy for DBS surgery. A diffuse slower and desynchronized EEG may be a warning sign for clinicians deciding on DBS eligibility and may tip the scales towards a negative recommendation for DBS surgery, whereas a fast and synchronous EEG may support a recommendation for DBS. The predictive properties of EEG slowing were demonstrated in chapter 6, and in combination with the positive results from chapter 7, would support the suggestion to use qEEG as a predictive biomarker of future cognitive decline after STN DBS. Notably, the role of qEEG in terms of prediction of future deterioration after STN DBS has not been studied yet and requires further research.

Both the concept of identification of novel biomarkers, as well as the concept of a cognitive subnetwork or a cognitive EEG-profile, was studied further in chapter 8. Whereas conventional analyses study EEG either in signal-space or in source-space, a machine learning algorithm has the potential to study the EEG in feature-space given the massive feature extraction provided by the application of an EEG feature-library. The feature library used here (*tsfresh*) does not consider inter-channel connectivity and could be extended several folds further. However, the studied algorithms nicely demonstrate the additional value of applying a feature-space beyond spectral analyses, as the accuracy of a coherent EEG profile clearly transcends that of the occipital peak frequency as a representative and easy-to-use spectral metric. A compound of numerous EEG markers may approach a cognitive subnetwork, even though the spectrum of cognitive (dys)function is relatively limited in DBS patients.

Although the use of such an extensive algorithm has the potential to identify new biomarkers and provides new insight into the pathophysiological mechanism, the goal of this study was to provide a proof-of-concept of prediction rather than to study causal mechanisms. Statements on such mechanisms should be avoided given the limitations of this study, such as the relatively small sample size, the absence of consensus-based diagnostic criteria for class labels, and the lack of external validation. All these limitations are unfortunate but were unavoidable. Regardless, we demonstrate good internal validation and consistency of our results, both in terms of diagnostic accuracy and in terms of calibration (i.e. predicted probabilities of cognitive class vs. actual cognitive function). To our knowledge, there are no other cohorts of consecutively included DBS patients that have been evaluated by means of qEEG, rendering external validation impossible.

The nature of the DBS cohort also limits the utility of diagnostic criteria for class labels, as severe cognitive dysfunction is a contraindication for both DBS surgery and referral and these patients would be rejected for the DBS screening as explained in chapter 2. The use of the 'cognitive extremes' in our cohort based on standardization of cognitive domains based on DSM-V criteria [14] was considered to be a straightforward and easily reproducible approach. The neuropsychological evaluation is currently the only gold standard to label cognition, despite its sensitivity to external influences. Using these 'cognitive extremes' limits bias to the largest extent and maximizes the distance between the two classes despite the relatively homogeneous global cognitive profile as compared to the entire possible spectrum of cognitive (dys)function in PD.

Again, the machine learning approach does not provide any indication of future cognitive deterioration and needs to be studied in a longitudinal setting to determine its clinical utility, as well as undergo external validation and assessment of clinical impact.[15, 16]

An interesting aspect of the machine learning algorithm is that it was never developed with EEG in mind. Originally, it was developed for utility in the automotive industry to study the effects of low-impact crashes on vehicles, in order to determine whether a check-engine-light has to start blinking. Its utility on EEG data resembles this check-engine-light, as a warning sign considering DBS eligibility. The data-structure of both vehicle-data and EEG data is relatively similar, which allows for similar analysis-algorithms being applied to data of different origins. The algorithm merely recognizes a signal originating from a time series and disregards the origin. Upon examining oscillations within a time series, the clinical neurophysiologist will instantly recognize a time series originating from an EEG or EMG signal, whereas the automotive engineer will not likely recognize a time series as an EEG but will sooner consider a different origin. Interpretations are influenced by prior knowledge

and area-of-expertise, which is once more highlighted by the cover of this thesis. Clinical researchers are prone to recognize the cover image as a brain with the coloured circles representing possible EEG electrode locations, or sources of brain activity (figure 10.1). Upon removing the background behind the coloured circles, an acyclic graph may be recognized by researchers proficient in graph theory research , or clinical epidemiologist interested in directed acyclic graphs to model causal relationships and confounding (figure 10.2).[17] The actual image however was based on the subway-network of the inner city of Munich (figure 10.3), something a conductor operating these metros would sooner have recognized than an EEG-system. The machine learning algorithm applied in chapter 8 was not constrained by previous knowledge other than class label and did not focus on any particular feature.

Chapter 8 also highlights the valorisation of such a multi-disciplinary approach, as clinicians are generally insufficiently proficient with the complex mathematical computations required for advanced machine learning analyses. Clinical research may highly profit from collaborating with other research fields and examining joint approaches for research aims for advanced analyses and novel modelling strategies.

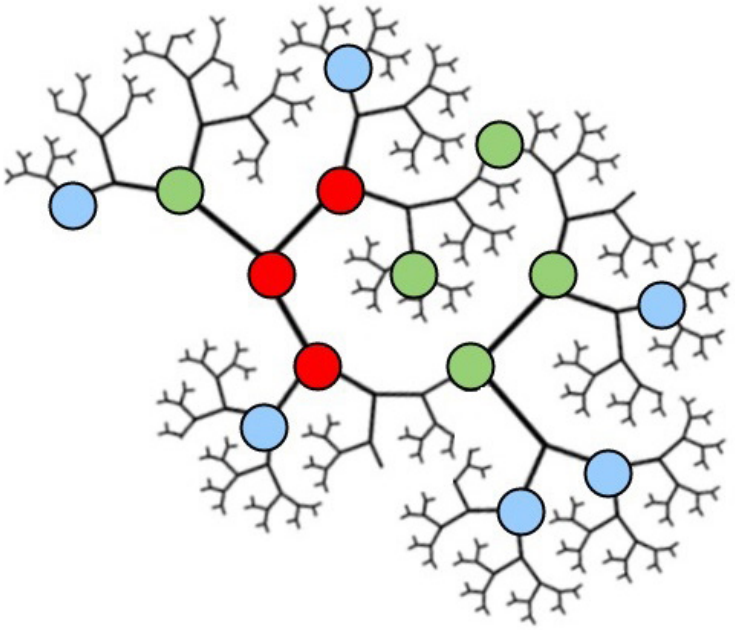

Figure 10.1 Cover image: brain with EEG electrode positions.
Red: hubs with high degree; green: nodes with lower degree; blue: leaf-nodes

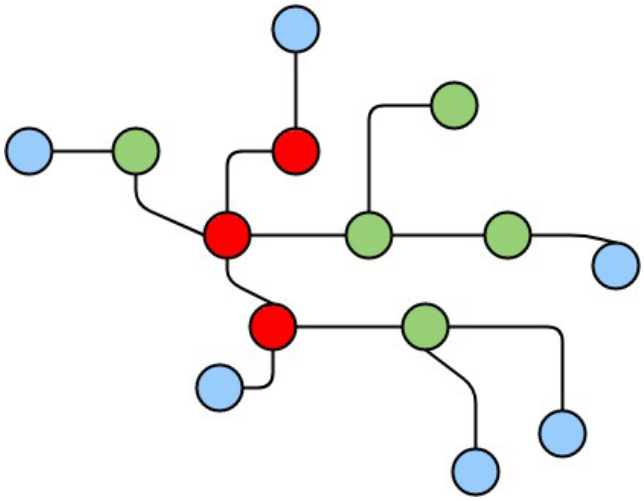

Figure 10.2 Cover image: acyclic graph

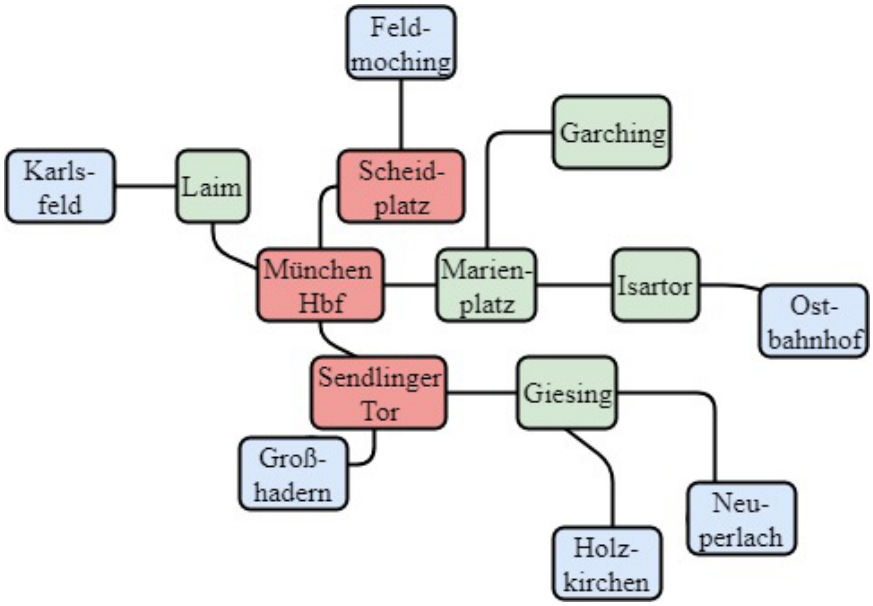

Figure 10.3 Cover image: subway-network Munich

Future perspectives with regard to using EEG during the DBS screening

Part B of this thesis demonstrates the feasibility and utility of applying qEEG during the DBS screening, particularly as a biomarker of current cognitive performance. Although some speculations on pathophysiological mechanisms can be made on the basis of chapters 6-8, the application of qEEG currently lies clearly within the domain of prediction as opposed to causality. Given the promising results, the utility of qEEG as a predictor of future deterioration after DBS needs to be determined in future studies, as well as determining the clinical impact of incorporating qEEG during the DBS screening.[15, 16] Moreover, it has to be determined which method has the greatest practical utility. Spectral biomarkers which have the most evidence-based utility based on previous literature, are easier to compute, interpret, and implement. In contrast, a compound-approach as shown in chapter 8 is more difficult to compute and implement, but appears to have a greater discriminating potential and therefore would result in greater accuracy. For the short-term future, spectral analyses may have a more immediate impact on DBS care whereas machine learning approaches need to undergo several verification and validation steps before implementation in routine clinical practices can be definitively recommended.

Concluding remarks

In conclusion, this thesis answers and raises an almost equal number of questions. Many issues may play a role in DBS screening and care. The findings presented in this thesis provide some new directions for future studies aiming to improve the screening and care of DBS patients.

An important final note to consider is a quote from Alan Alda (famous actor and science journalist) on his PD diagnosis: "it hasn't stopped my life at all". To improve patients' welfare after the PD diagnosis is the ultimate goal, and the research detailed in this thesis is only a first step along the track towards improving the screening and care of DBS patients.

References

1. Schuepbach WM, Rau J, Knudsen K, et al. Neurostimulation for Parkinson's disease with early motor complications. N Engl J Med 2013;368:610-622.
2. Coleman RR, Kotagal V, Patil PG, Chou KL. Validity and Efficacy of Screening Algorithms for Assessing Deep Brain Stimulation Candidacy in Parkinson Disease. Mov Disord Clin Pract 2014;1:342-347.
3. Ioannidis JPA, Patsopoulos NA, Rothstein HR. Reasons or excuses for avoiding meta-analysis in forest plots. BMJ (Clinical research ed) 2008;336:1413-1415.
4. Donaldson GW, Moinpour CM. Individual differences in quality-of-life treatment response. Medical care 2002;40:Iii39-53.
5. Molzahn AE, Kalfoss M, Schick Makaroff K, Skevington SM. Comparing the importance of different aspects of quality of life to older adults across diverse cultures. Age and Ageing 2010;40:192-199.
6. Altman DG, Bland JM. Statistics notes: Absence of evidence is not evidence of absence. BMJ 1995;311:485.
7. Thiese MS, Ronna B, Ott U. P value interpretations and considerations. J Thorac Dis 2016;8:E928-E931.
8. van Diessen E, Numan T, van Dellen E, et al. Opportunities and methodological challenges in EEG and MEG resting state functional brain network research. Clin Neurophysiol 2015;126:1468-1481.
9. Geraedts VJ, van Hilten JJ, Contarino MF, Tannemaat MR. Unravelling the Parkinson's disease network: Taking the connectome beyond the brain. Clin Neurophysiol 2019;130:2017-2018.
10. Yang Y, Guliyev B, Schouten AC. Dynamic Causal Modeling of the Cortical Responses to Wrist Perturbations. Front Neurosci 2017;11:518-518.
11. Friston KJ. Functional and effective connectivity: a review. Brain connectivity 2011;1:13-36.
12. van der Heeden JF, Marinus J, Martinez-Martin P, van Hilten JJ. Importance of nondopaminergic features in evaluating disease severity of Parkinson disease. Neurology 2014;82:412-418.
13. van der Heeden JF, Marinus J, Martinez-Martin P, van Hilten JJ. Evaluation of severity of predominantly non-dopaminergic symptoms in Parkinson's disease: The SENS-PD scale. Parkinsonism Relat Disord 2016;25:39-44.
14. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (5th ed. ed.). Arlington: VA: American Psychiatric Publishing, 2013.
15. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J 2014;35:1925-1931.
16. Steyerberg EW, Moons KGM, van der Windt DA, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS Med 2013;10:e1001381-e1001381.
17. Ferguson KD, McCann M, Katikireddi SV, et al. Evidence synthesis for constructing directed acyclic graphs (ESC-DAGs): a novel and systematic method for building directed acyclic graphs. International journal of epidemiology 2019.