

Article

# Dealing With Side Effects of Deep Brain Stimulation: Lessons Learned From Stimulating the STN

**Markus Christen**, University of Zurich and University of Notre Dame  
**Merlin Bittlinger**, Charité-Universitätsmedizin Berlin  
**Henrik Walter**, Charité-Universitätsmedizin Berlin  
**Peter Brugger**, University Hospital Zurich  
**Sabine Müller**, Charité-Universitätsmedizin Berlin

Deep brain stimulation (DBS) is increasingly investigated as a therapy for psychiatric disorders. In the ethical evaluation of this novel approach, incidence and impact of side effects (SE) play a key role. In our contribution, we analyze the discussion on SE of DBS of the subthalamic nucleus (STN)—a standard therapy for movement disorders like Parkinson's disease (PD)—based on 66 case reports, 69 review papers, and 347 outcome studies from 1993 to 2009. We show how the DBS community increasingly acknowledged the complexity of STN-DBS side effects. Then we discuss the issue of study quality and the methods used to assess SE. We note that some side effects are the subject of conflicting evaluations by the different stakeholders involved. This complicates the ethical controversy inherent in any novel treatments for diseases that involve psychiatric aspects. We delineate how the lessons from STN-DBS could guide future DBS applications in the field of psychiatry.

**Keywords:** deep brain stimulation, neuroethics, Parkinson's disease, psychiatry, side effects, subthalamic nucleus

Since the advent of medicine, the maxim “do not harm your patient” forms the core of the ethos for physicians. This is reflected in the principle of nonmaleficence as one of the four principles of biomedical ethics (Beauchamp and Childress 2009). Nevertheless, it is broadly acknowledged that harmful side effects (SE) of therapies have to be weighed up against their beneficial effects. The deliberation of harms and benefits becomes more difficult if side effects of novel therapeutic approaches have to be considered, of which the nature, extent, and incidence are not yet known.

Deep brain stimulation (DBS) is an example for a novel therapeutic approach. Its roots go back to the early 1950s (Hariz et al. 2010), and it emerged in its current form in the 1980s as an alternative for ablative surgery in movement disorders and an experimental therapy in chronic pain (Siefried and Blons 1997). DBS became an established therapy for Parkinson's disease (PD) and other movement disorders in the last decade (Benabid et al. 2009). Side effects have been discussed since the advent of DBS, but the sensibility for them and the appreciation of their complexity is increasing. This also reflects the maturation of the therapy.

The term “side effect” does not provide per se an ethical orientation how to deal with a specific therapy. Therefore,

terms like “adverse events” or “sequelae” (adverse events that count as long-term negative consequences) should be avoided as long as the negative evaluation of the side effect is not clear. For assessing particular side effects, they can be classified along the two dimensions predictability and evaluation (Table 1).

As Table 1 shows, only one type of side effects (predictable and clearly outweighing the benefits) bears the clear “ethical message” that the therapy should not be applied or that research on this therapy should be stopped. Thus, the determination of what counts as an adverse event involves both a measurement problem and an evaluation problem, which can be entangled in the process of developing the therapy (Müller and Christen 2011). DBS for PD patients is a paradigmatic example for outlining that problem, since first, the predictability of side effects for individual patients is difficult; second, the evaluation of some side effects differs significantly between patients, their relatives, and physicians (Müller and Christen 2011); and third, both the disease (Kulisevsky et al. 2008) and alternative therapeutic approaches (medication or surgery; see Voon et al. 2006 and Olanow 2002) may involve effects similar to those of DBS.

This research has been supported by the Swiss Academy of Medical Sciences (Käthe-Zingg-Schwichtenberg-Fonds) and by the Federal Ministry of Education and Research, Germany (project 01 GP 0804).

Address correspondence to Markus Christen, PhD, MSc, University of Zurich, Institute of Biomedical Ethics, Pestalozzistrasse 24, CH-8032 Zurich, Switzerland, or Department of Psychology, 118 Haggard Hall, University of Notre Dame, Notre Dame, IN 46556, USA. E-mail: christen@ethik.uzh.ch

**Table 1. Ethical requirements depending on the predictability and the evaluation of the side effects of a therapy**

Evaluation	Predictability of side effects (SE) of a therapy	
	Predictable	Not predictable
The SE of the therapy clearly outweigh its therapeutic effects.	Do not begin/stop the therapy.	Ensure sensibility for novel SE.
There are conflicting evaluations of the SE by different stakeholders.	Define the authority to decide about the usage of the therapy.	Ensure involvement of different stakeholder’s perspectives during the development of the therapy.
The therapeutic effects clearly outbalance the SE.	Define a procedure to take individual variability of the impact of the SE into account.	Define a procedure to decide whether a novel SE is classified as unproblematic or not.

The following investigation is based on a comprehensive analysis of the research literature on DBS in the nucleus subthalamicus (STN), currently the preferred target for DBS in PD. This analysis covers 66 case reports, 69 review papers, and 347 outcome studies from 1993 to 2009 that emerged from an extensive search in the following databases: CPCI-S, Embase, Francis, Medline, PsychINFO, and SCI-expanded (the reference lists are available as supporting online material).

As the STN is part of various thalamo-cortical circuits (Marani et al. 2008), the relatively high incidence of cognitive and affective side effects after STN DBS compared to other DBS targets is not surprising (Hariz et al. 2008). The way the DBS community dealt with this issue is thus a paradigmatic case for analyzing SE measurement and evaluation in the course of the establishment of novel therapies. Understanding this process may support the ethical analysis of the current application of DBS to a variety of psychiatric disorders (for an overview about psychiatric DBS, see Krack et al. 2010).

**THE COMPLEXITY OF ADVERSE EVENTS**

Since the early 1990s, the STN has been investigated as a potential DBS target both in animal and clinical studies. In 1993, the first case was published in a French journal (Pollak et al. 1993). Several case reports and outcome studies followed, and since the late 1990s the number of publications on STN-DBS has increased steadily (Müller and Christen 2011). The number of STN-DBS-related issues discussed in the literature has grown rapidly, whereas a bibliometric investigation demonstrated that case reports spearhead the transdisciplinary communication about DBS (Christen and Müller 2011).

To handle the complexity of issues that are discussed in our literature body of 482 STN-DBS publications, we have sorted them into 18 issue classes (Table 2). For each class, we have evaluated all tests used in the outcome literature to measure the respective phenomena and the wording used to describe the corresponding SE. Note that not every issue class is directly related to SE. This is true especially for studies on the neuronal basis of DBS effects (usually in-

vestigated by PET) or about the cost-effectiveness of DBS. Furthermore, the boundaries between some issue classes are less clear and require predefinitions. For example, we have classified studies about language fluency as “cognitive” (in accordance with the current neuropsychological understanding).

Each publication (case report, review, outcome study) was attributed to one or several issue classes with regard to the topics discussed and the methods used. Thereby we did not take into account possible causal relations between certain issues. For example, many issues have an impact on the quality of life (Q). Nevertheless, a study whose primary focus, for example, was insomnia was only classified as “I,” not as “Q.” For analyzing the time course of the publication praxis, it was necessary to build bigger groups. Therefore, we have grouped the 18 issue classes into four groups as follows:

1. Group: Understanding therapeutic effects: F, M, V.
2. Group: Medical and technical intervention issues: O, P, T.
3. Group: Main affective, behavioral and cognitive side effects of DBS: B, C, D, L, Q.
4. Group: Other issues: A, E, I, K, N, S, W.

The histogram in Figure 1 shows the time course of the different groups of issues in the DBS literature. The absolute numbers of publications per year belonging to one of the four issue groups are displayed for the years 1993 to 2009. The analysis reveals that—after the first, pioneering years with very few studies—issues on main affective, behavioral, and cognitive side effects quickly appeared in the literature and have been the dominant group since 2003. This finding is corroborated by an analysis of DBS posters presented at conferences (Christen and Müller 2011). Although one has to take into account that this analysis is not sensible for the valuation of these effects (i.e. whether they are considered to be unproblematic or not), this finding somehow contrasts with several statements in the literature, such that the DBS community would often ignore the neurobehavioral consequences of the therapy (e.g., Burn and Tröster 2004).

**Table 2. Issue classes of therapeutic effects and side effects present in the STN-DBS literature**

Group	Abbr.	Issue class	Examples of side effects
Group 1	F	Functional studies, i.e., studies that involve PET or other methods for investigating the causal effect of DBS	—
	M	Motor effects	Apraxia, axial symptoms, dyskinesia, dystonia, gait disorders, motor fluctuations
	V	Issues related to Levodopa and other medication	Dopamine dysregulation syndrome, changes in LEDD
Group 2	O	Operation/surgery related issues	Hemorrhage, hematoma, ischemia, surgical complications, infections
	P	Patient issues, i.e., issues related to patient selection, patient management, rehabilitation	—
	T	Effects related to the technology (device)	Battery problems, electrode break, local infections, lead fracture, pulse generator malfunction
Group 3	B	Behavioral effects, i.e., effects that concern abnormal behavior	Compulsive shopping, hypersexuality, hypomania, pathological gambling, suicide (attempts)
	C	Effects on cognition (reasoning, memory, etc.)	Cognitive decline, confusion, dementia, memory decline, verbal fluency change
	D	Depression, anxiety, apathy and other mood effects	Ahedonia, apathy, depression, mood changes, sadness
	L	Language, i.e., effects regarding the general speech ability and motor aspects of speech	Aphasia, dysarthria, hypophonia, speech impairment, voice freezing
	Q	Quality of life and social aspects	Disability in daily living, decreased life satisfaction, partnership problems
Group 4	A	Effects regarding the autonomous nervous system, autonomous functioning	Drooling, dysphagia, hyperhidrosis
	E	Emotion recognition changes	Difficulties of emotion discrimination, difficulties of face perception, hyperemotivity
	I	Insomnia, i.e., effects related to sleep	Drowsiness, fatigue, insomnia, sleep disorders
	K	Cost issues (German: <i>Kosten</i> ), i.e., cost-benefit studies, cost-effectiveness of DBS, etc.	—
	N	Other neurological effects	Epilepsy, postural imbalance, seizures
	S	Effects regarding sensory systems	Blurred vision, parasthesia, visual hallucinations
	W	Weight and energy intake changes	Abnormal weight gain, binge eating, obesity

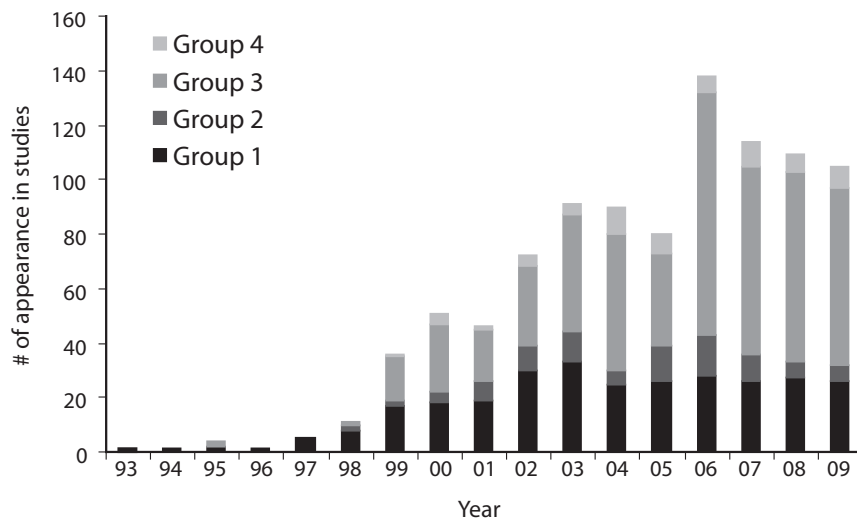
Note. Only selected examples of side effects using the wording found in the publications are displayed.

### QUALITY OF STN-DBS STUDIES

The sensibility for new side effects in the process of maturation of a novel therapy is a critical issue—and we may say that the DBS community has passed this test successfully. Another issue is the quality of the studies. Although it is well known that novel therapies start with isolated case studies that usually lack quality criteria like randomization or blinding, at some point the urge for better studies is raised. DBS did not deviate from this development path, and various authors have discussed the issue of study quality (e.g., Woods et al. 2006) and proposed standards for improving study quality (e.g., Morrison et al. 2000). We investigated the study quality for all outcome studies that involved at least one issue of group 3 (i.e., B, C, D, L, or Q). For that,

we expanded the criteria for level of evidence assignment proposed by Martinez-Martin and Deuschl (2007) using a rating system that involves several aspects being considered as relevant for study quality (e.g., regarding follow-up time) by the DBS community.<sup>1</sup>

1. Each study was assigned with points as follows: study was prospective: +1; study was case-controlled with at least 20 participants in each branch: +2 (+1 if less than 20 participants in either branch); study was randomized: +1; test evaluation was blinded: +1; study involved more than one center: +1; the presurgery and postsurgery assessment of the neuropsychiatric tests were made in the “best” (pre: med-on/ post: med-on, stim-on) condition of the patient: +2 (+1 if the assessments were made pre and post); the study involved at least 20 patients: +1; the study had a follow-up



**Figure 1.** The histogram displays the total number of issues (compare with Table 2) addressed in the publications about STN-DBS (case reports, reviews, and outcome studies) pooled in four groups (group 1: F, M, V; group 2: O, P, T; group 3: B, C, D, L, Q; group 4: A, E, I, K, N, S, W; see text).

As Figure 2a demonstrates, the quality range of the studies is broad. Somewhat surprising is the fact that the mean quality of the outcome studies has not increased significantly since 2000 (Figure 2b; earlier studies were not taken into account due to their low numbers). Although the absolute number of high-quality studies of group 3 has increased, this is shrouded by the also increasing number of outcome studies of poor quality.

This absent increase of the average study quality is not per se problematic, as long as the community is able to differ between good and poor studies. To investigate whether this is the case, we calculated a citation coefficient based on DBS review papers about the outcome studies of group 3.<sup>2</sup> Then we performed a correlation analysis of the relationship of this citation coefficient and the quality rating for each study. The result is a (weak) positive correlation of the citation coefficient with the quality rating (Pearson's correlation coefficient: 0.29). That means that high-quality papers tend to be cited more often in the reviews. This can be interpreted as a hint for a higher appreciation of high-quality studies of group 3 by the DBS community.

time of at least 12 months: +1; the study involved tests of at least three issue classes: +1.

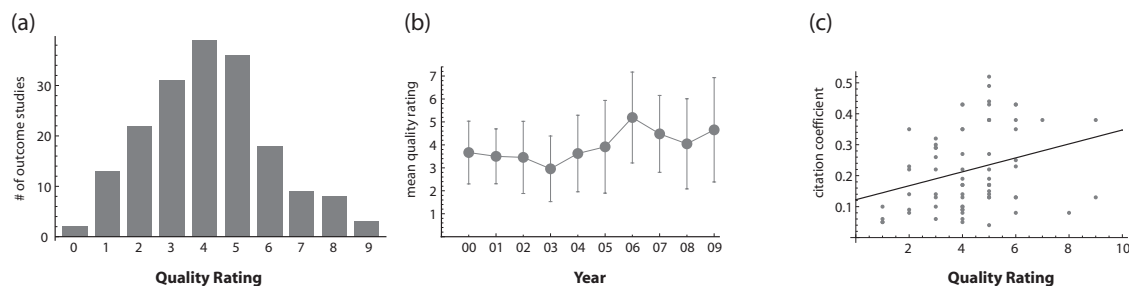
2. The citation coefficient was construed as follows: We counted the appearance of outcome studies in the analytic part of reviews that either performed a meta-analysis following established standards (e.g., Cochrane) or were at least systematically evaluated (i.e., we excluded merely narrative reviews; thus, we considered 23 reviews). The citation of outcome papers is weighted with the probability of being able to be cited due to the year of publication to take into account, such that a paper, e.g., published in 2006 cannot be cited in a review published in 2004. Thus, each outcome paper received a citation coefficient value between 0 and 1. For the correlation analysis, only outcomes with nonzero citation coefficient have been analyzed.

## MEASURING ADVERSE EVENTS

After analyzing the attention for side effects of group 3 and the quality of the studies investigating them, we investigated a third issue: To what extent do the studies capture "relevant" side effects, i.e. those reflecting serious ethical issues (see Table 1)? This point requires a closer look to the methods and tests used in the outcome studies. We listed all tests used in the 347 outcome studies and attributed them to one of the 18 issue groups. By far most of the tests were assigned to one of the five issue classes B, C, D, L, and Q (group 3), whereas the internal distribution is very uneven. Figure 3a demonstrates that more than half of all methods applied are tests regarding cognitive issues. Also, the number of uses of the tests themselves is remarkably uneven. Very few tests are used regularly. Furthermore, the probability that a test is used in a study for neuropsychological outcome assessment does not correspond completely to the four standards proposed in the literature (Defer et al. 1999; Morrison et al. 2000; Pillon 2002; Saint-Cyr et al. 2000). For example, the Hopkins Verbal Learning Test and the Odd Man Out Test, both recommended in all four standards, are comparably rarely used. This may indicate a learning effect by the community, as tests better than the ones initially recommended are available that measure similar constructs.

For the ethical evaluation it is of particular interest which perspectives are represented in the tests, as conflicting evaluations of side effects often result from different perspectives of stakeholders. For analyzing this point, we have classified all tests as follows:

- I. Test scores that result from the evaluation of the patient's performance by a trained evaluator.
- II. Test scores that result from a self-assessment of the patient.



**Figure 2.** (a) Quality rating distribution of the outcome studies of group 3, (b) time course of the mean quality rating of outcome studies (group 3), and (c) correlation between the citation coefficient of studies (which reflects the appreciation of papers by the authors of reviews) with the quality rating of the studies. The chart also includes the linear approximation of the correlation.

III. Test scores that result from an interrogation etc. from closely related persons of the patient (family members, caregivers).

In order to avoid biases due to low-quality studies, we have only investigated those studies that achieved a quality rating of at least 5 (see footnote 1, 182 studies). We have counted the number of different tests, the number of test executions (i.e., in how many studies the test was used), and the number of patients that have been tested by these methods. The cumulative numbers for the three classes I, II, and III are displayed in Figure 3b.

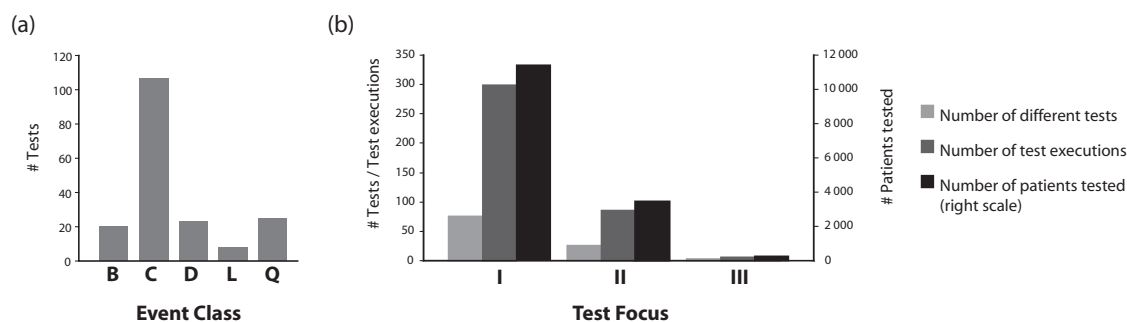
We see a clear dominance of category I tests, whereas the usage of category III tests is basically nonexistent. Thus, the perspective on side effects is very biased in the DBS literature. This finding might explain the “satisfaction gap” between the physician’s and the patient’s expectation that is discussed in the literature (Agid et al. 2006).

#### THE ETHICS OF ADVERSE EVENTS: CONCLUSIONS AND RECOMMENDATIONS FOR PSYCHIATRIC DBS

What is the impact of this in-depth analysis of the literature about side effects of STN-DBS for the ethical debate about

the application of DBS in psychiatry? Compared to many somatic diseases, harm–benefit assessments for psychiatric therapies are complicated by at least three problems: First, for most psychiatric disorders no clear correlation with a specific neurological dysfunction is proven. Second, many interventions affect various neuronal mechanisms—e.g., selective serotonin reuptake inhibitors (SSRI) have effects not only on the serotonin metabolism, but also on the neurogenesis in the hippocampus (Santarelli et al. 2003). Third, the evaluation of both the disease and the beneficial and negative therapy effects depends much more strongly on subjective evaluations compared to somatic medicine. For example, neither patients nor their relatives nor their physicians would doubt that toothache is painful, whereas hypomania is evaluated differently by different stakeholders (see, e.g., the examples in Krug et al. 2010). “Clear-cut” cases (predictable side effects that clearly outbalance therapeutic effects) are probably rather rare in psychiatric diseases.

This is important, since the introduction of DBS to psychiatry is driven also by the expectation that it will improve the understanding of the causes of these diseases and that it will be a causal therapy. Already the usage of DBS for the treatment of movement disorders was accompanied



**Figure 3.** (a) Number of tests per issue class (B: behavioral, C: cognitive, D: depression and other mood issues, L: language, Q: quality of life). (b) Number of different tests (first bar in each group), number of accumulated test executions (middle bar), and total number of patients tested (left bar, right scale) with methods of either class I (test scores generated by evaluator), II (self-assessment of patient), or III (test scores emerge from persons affiliated to the patient).

by the narrative that DBS is more precise than its alternatives, completely reversible, and individually scalable. Although this is to a large extent true, the problems of measuring and evaluating side effects do not vanish. On the contrary, our analysis revealed that the availability of a more precise tool triggered research on the mechanisms behind the effects of DBS on cognitive functions, mood, and behavior and thus tends to increase the spectrum of potential SE to look at. If DBS will play an important role in psychiatry, we cannot expect that the SE spectrum will become smaller compared to that implicated by the alternatives.

However, we have found a well-developed sensibility for side effects in the DBS community. Nevertheless, the side effects are not yet measured and evaluated sufficiently. Our analysis reveals that the majority of methods used investigate subtle cognitive changes that may be statistically significant but whose relevance for the patients is unclear. Only a minority of investigations focus on the self-assessment of the patients, and even less on the assessments of their caregivers. This methodological bias implies blindness for certain side effects. We expect that this problem will be aggravated if DBS is used to treat psychiatric disorders as depression or addiction, since interpersonal relationships play a crucial role in overcoming these disorders.

Finally, the quality of studies that promote the extended use of DBS gives cause for concern, although we note that our rating system does not take into account that the requirements for quality may differ between studies if they addressed different types of outcomes, such that not all quality items are required for a specific study. After all, progress is recognizable and the community is somewhat able to discriminate between good and bad studies, yet it has not managed to cut down the continuous generation of low-quality contributions.

In summary, the ethical evaluation of side effects of STN-DBS must not abstract from the measurement and evaluation problems that constitute the definition of what counts as a “side effect.” The role of ethicists is not only to safeguard against the “bad effects” of therapies. They should also point at blind spots in clinical studies and widen the perspective on all sorts of effects of new therapies.

## REFERENCES

- Agid, Y., M. Schüpbach, M. Gargiulo, et al. 2006. Neurosurgery in Parkinson's disease: the doctor is happy, the patient less so? *Journal of Neural Transmission Suppl.* 70: 409–414.
- Beauchamp, T. L., and J. F. Childress. 2009. *The principles of biomedical ethics*, 6th ed. Oxford: Oxford University Press.
- Benabid, A. L., S. Chabardes, J. Mitrofanis, and P. Pollak. 2009. Deep brain stimulation of the subthalamic nucleus for the treatment of Parkinson's disease. *Lancet Neurology* 8(1): 67–81.
- Burn, D. J., and A. I. Tröster. 2004. Neuropsychiatric complications of medical and surgical therapies for Parkinson's disease. *Journal of Geriatrics, Psychiatry and Neurology* 17: 172–180.
- Christen, M., and S. Müller. 2011. Single cases promote knowledge transfer in the field of DBS. *Frontiers in Integrative Neuroscience* 5: article 13.
- Defer, G. L., H. Widner, R. M. Marié, P. Rémy, and M. Levivier. 1999. Core Assessment Program for Surgical Interventional Therapies in Parkinson's Disease (CAPSIT-PD). *Movement Disorders* 14: 572–584.
- Hariz, M. I., S. Rehncrona, N. P. Quinn, J. D. Speelman, C. Wensing, and the Multicentre Advanced Parkinson's Disease Deep Brain Stimulation Group. 2008. Multicenter study on deep brain stimulation in Parkinson's disease: An independent assessment of reported adverse events at 4 years. *Movement Disorders* 23(3): 416–421.
- Hariz, M. I., P. Blomstedt, and L. Zrinzo. 2010. Deep brain stimulation between 1947 and 1987: The untold story. *Neurosurgery Focus* 29(2): E1.
- Krack, P., M. I. Hariz, C. Baunez, J. Guridi, and J. A. Obeso. 2010. Deep brain stimulation: from neurology to psychiatry? *Trends in Neuroscience* 33(10): 474–484.
- Krug, H., O. Müller, and U. Bittner. 2010. Technological interventions in the self? An ethical evaluation of deep brain stimulation relating to patient narratives. *Fortschritte der Neurologie Psychiatrie* 78: 644–651.
- Kulisevsky, J., J. Pagonabarraga, B. Pascual-Sedano, C. García-Sánchez, and A. Gironell. 2008. Prevalence and correlates of neuropsychiatric symptoms in Parkinson's disease without dementia. *Movement Disorders* 15;23 (13): 1889–1896.
- Marani, E., T. Heida, E. A. Lakke, and K. G. Usunoff. 2008. The subthalamic nucleus. Part I: development, cytology, topography and connections. *Advances in Anatomy, Embryology, and Cell Biology* 198: 1–113.
- Martinez-Martin P, and G. Deuschl. 2007. Effect of medical and surgical interventions on health-related quality of life in Parkinson's disease. *Movement Disorders* 22(6): 757–765.
- Morrison, C. E., J. C. Borod, M. F. Brin, et al. 2000. A program for neuropsychological investigation of deep brain stimulation (PNIDBS) in movement disorder patients: Development, feasibility, and preliminary data. *Neuropsychiatry, Neuropsychology and Behavioral Neurology* 13(3): 204–219.
- Müller, S., and M. Christen. 2011. Deep brain stimulation in Parkinsonian patients—Ethical evaluation of cognitive, affective, and behavioral sequelae. *AJOB Neuroscience* 2(1): 3–13.
- Olanow, C. W. 2002. Surgical therapy for Parkinson's disease. *European Journal of Neurology* 9(suppl. 3): 31–39.
- Pillon, B. 2002. Neuropsychological assessment for management of patients with deep brain stimulation. *Movement Disorders* 17(suppl. 3): S116–S122.
- Pollak, P., A. L. Benabid, C. Gross, et al. 1993. Effets de la stimulation du noyau sous-thalamique dans la maladie de Parkinson. *Revue Neurologique (Paris)* 149(3): 175–176.
- Saint-Cyr, J. A., and L. L. Trépanier. 2000. Neuropsychologic assessment of patients for movement disorder surgery. *Movement Disorders* 15(5): 771–783.

- Santarelli, L., M. Saxe, C. Gross, et al. 2003. Requirement of hippocampal neurogenesis for the behavioral effects of antidepressants. *Science* 301: 805–809.
- Siegfried, J., and S. Blons. 1997. *The neurosurgical treatment of Parkinson's disease and other movement disorders*. London: William & Wilkins Europe Ltd.
- Voon, V., C. Kubu, P. Krack, J. L. Houeto, and A. I. Tröster. 2006. Deep brain stimulation: Neuropsychological and neuropsychiatric issues. *Movement Disorders* 21(suppl. 14): S305–S326.
- Woods, S. P., J. D. Rippeth, E. Conover, C. L. Carey, T. D. Parsons, and A. I. Tröster. 2006. Statistical power of studies examining the cognitive effects of subthalamic nucleus deep brain stimulation in Parkinson's disease. *Clinical Neuropsychologist* 20(1): 27–38.