APPENDIX

Appendix 2. Included references details

Assessment Scales — Depression / Anxiety

Author, Year (Country)	Ardouin-d	BDI	BRMES	BSI-d	GDS	HAD-d	HDRS	IOWA-d	MADRS	MINI-d	NMSQuest-d	POMS-d	SCL-90-R-d	UPDRS 1,3	Zung-d	AMDP-AT	Ardouin-a	BAI	BAS	BSI-a	HAD-a	HAMA	IOWA-a	MINI-a	NMSQuest-a	SCL-90-R-a	STAI-s	STAI-t	Zung-a	STN	GPi	Follow-up	Follow-up time	Comparator
Alegret, 2004 (ES) (a)		x																												✓		✓	6m, 1y	
Altug, 2011 (TR) (a,e)				x																	x									✓		✓	3m, 6m	
Auclair-Ouellet, 2011 (CA) (a,l)		x																			x									✓		✓	6m, 1y	
Berney, 2002 (CH/CA) (a,h)							x		x(d)																					✓		✓	3-6m	
Bordini, 2007 (US) (a)					x																									✓		✓	6m	
Daniele, 2003 (IT) (a,e,F)															x														x	✓		✓	4m, 6m, 1y, 18m	
De Gaspari, 2006 (IT) (a)		x																												✓		✓	15m	
Denheyer, 2009 (CA) (a)		x																												✓		✓	16m	
Derost, 2007 (FR) (a,g,i)							x																							✓		✓	6m	
Drapier, 2005 (FR) (a)									x																					✓		✓	1y	
Dujardin, 2004 (FR/CA/BE) (a,A)									x							x														✓		✓	3m	
Fasano, 2010 (IT) (a)															x														x	✓		✓	8y	
Funkiewiez, 2006 (FR/UK) (a,i,o)		x																												✓		✓	3m	
Gervais-Bernard, 2009 (FR) (a)		x																												✓		✓	1y, 5y	
Heo, 2008 (KR) (a)		x																												✓		✓	6m, 1y	
Houeto, 2002 (FR)							x(J)		x(c)														x(J)	x(c)						✓		✓	19m	
Houeto, 2006 (FR) (a)									x													x								✓		✓	6m, 2y	
Huebl, 2011 (DE/UK) (n)		x																												✓		✓	3m	
Kaiser, 2008 (AT) (a,e,h,D)		x										x	x													x	x	x		✓		✓	3m, 6m, 1y, 3y	
Kalteis, 2006 (AT) (a,e,D)		x	x									x	x									x				x	x	x		✓		✓	3w, 9w, 3m, 6m, 1y	
Kishore, 2010 (IN) (a,e)		x			x																x									✓		✓	1y, 3y, 5y	
Krack, 2003 (FR) (a)		x																												✓		✓	1y, 3y, 5y	
Krause, 2004 (DE) (k)														x																✓		✓	30m	
Lhommée, 2012 (FR) (a,j,I)	x(c)	x							x(c)								x(c)	x						x(c)						✓		✓	1y	
Martínez-Martín, 2002 (ES) (a)					x																	x								✓		✓	6m	
Merello, 2008 (AR) (a)							x																							✓		✓	6m, 1y	
Nazzaro, 2011 (US) (c)									x																x					✓		✓	1y	
Ory-Magne, 2007 (FR) (a,e)									x																					✓		✓	1y, 2y	
Perozzo, 2001 (IT) (d)		x																									x	x		✓		✓	6m	
Perriol, 2006 (FR) (c)										x																				✓		✓	1y	
Saint-Cyr, 2000 (US/CA) (a,f)					x																									✓		✓	6m	
Schadt, 2006 (US) (a,h,m)		x																												✓		✓	23m	
Schneider, 2010 (DE/US) (a,e,f)							x																							✓		✓	5-10d, 18-24d, 3-4m	
Schoenberg, 2008 (US) (d)																											x	x		✓		✓	5m	
Simuni, 2002 (US) (d)				x																										✓		✓	6m	
Temel, 2007 (NL) (b,g)		x																												✓		✓	3m, 6m	
Tröster, 2003 (US) (a)		x																												✓		✓	3,5m	
Witjas, 2007 (FR) (a)		x																												✓		✓	1y	
Yamada, 2006 (JP) (a,j)														x																✓		✓	3m	
Zibetti, 2007 (IT) (a,e)														x																✓		✓	1y, 2y	
Zibetti, 2009 (IT) (a,e,E)		x																									x	x		✓		✓	4m, 1y, 3y	
Zibetti, 2011 (IT) (a,e,E)		x																									x	x		✓		✓	1y, 5y, 9y	
Morrison, 2004 (US) (a,j)					x																									✓		✓	3m	2
Oyama, 2011 (JP) (a,f,O)															x															✓		✓	2-4w	2
Smeding, 2006 (NL) (b)									x			x																		✓		✓	6m	2
Wang, 2009 (CN) (a,e,G,H)							x(C)								x															✓		✓	1w, 2m, 5m, 11m, 17m	2
Witt, 2008 (DE/AT) (b,B)		x							x								x													✓		✓	6m	2
York, 2008 (US) (a)		x	x																	x							x	x		✓		✓	6m	2
Capecci, 2005 (IT) (a,e)		x																												✓		✓	1y, 2y	3
Drapier, 2006 (FR) (a,e)									x							x														✓		✓	3m, 6m	3
Péron, 2010 (FR/CH) (a)									x																					✓			35m	1,3
Montel, 2008 (FR) (a)									x	x(c)																				✓			1y	2,3
Castelli, 2008 (IT) (a)		x																									x	x		✓			3y	3
McDonald, 2012 (US/UK) (l)							x														x									✓			1y	3
Fields, 1999 (US) (n,M,N)		x										x					x														✓	✓	3m	
Ghika, 1998 (CH) (d)		x																													✓	✓	3m	
Loher, 2002 (CH/DE) (q)							x																								✓	✓	3m, 1y	
Burchiel, 1999 (US)*		x																												✓	✓	✓	1y	
Weaver, 2009 (US)*		x																												✓	✓	✓	6m	2
Ardouin, 1999 (FR) (a,g,K,L)		x																												✓	✓	✓	3m, 6m	4
Follett, 2010 (US) (a,m,p)		x																												✓	✓	✓	2y	4
Rothlind, 2007 (US) (a)		x																									x	x		✓	✓	✓	15m	4
Volkmann, 2001 (DE) (a)							x																							✓	✓	✓	6m, 1y	4

Psychometric instruments used are highlighted with "x". Stimulation target is marked with "✓" in "STN" and/or "GPi" columns. Follow-up studies have "✓" in the respective column and

APPENDIX

comparison studies are codified by 1 to 4 so different comparators can be distinguished. Grey shading denotes references excluded from analysis (non- comparable data).

Notes: 1 to 4 corresponds to comparators coding; 1: healthy control group, 2: medical treatment control group, 3 eligible for surgery control group and 4: GPi comparison group; * references with STN and GPi data not discriminated: no further analysis; (a) original data is pre M(SD) and post M(SD): changeM calculated as (postM - preM); changeSD calculated as $\sqrt{[(preSD2 + posSD2 - 2 \times r \times preSD \times posSD) / n]}$; (b) original data is pre M(SD) and change M(SD); (c) original data not quantitative/comparable (percentage of patients): no further analysis; (d) original data not quantitative/comparable (qualitative description): no further analysis; (e) within each period of time, the longest follow-up was selected for the analysis; (f) "x to y months" type follow-up: y months assumed; (g) original data reported by groups: separately considered for the analysis; (h) original data reported on total sample and by groups: total sample considered; (i) SD calculated from SE as (SE x $\sqrt{n}$); (j) original data reported in on and off state: only on considered; (k) original data not quantitative/comparable (no dispersion measure): no further analysis; (l) mean (SD) assumed; (m) SD calculated from 95% CI as [(upper limit - lower limit) / 3.92] x $\sqrt{n}$; (n) original data reported individually: preM(SD) and changeM(SD) calculated; (o) graphical data; (p) intention-to-treat analysis; (q) original data not quantitative/comparable (percentage of change): no further analysis; (A) cognitive outcomes compared with control group; depression and anxiety assessed only in patients; so, follow-up STN-DBS study design assumed; (B) "positive change scores indicate clinical improvement; data are (...) mean (SD) (...) for changes between baseline (before DBS) and 6 months": – changeM assumed; (C) HDRS not consecutively assessed: "depression was evaluated (...) using the Self-Rating Depression Scale (...); every patient whose SDS score showed a mild depression, or more, was evaluated again (...) using the Hamilton Depression Scale"; (D) partial duplicates: BDI, POMS-d, STAI-s, STAI-t, SCL-90-R-d and SCL-90-R-a data from Kaiser, 2008; BRMES and HAMA data from Kalteis, 2006; (E) partial duplicates:

3 years follow-up data from Zibetti, 2009; 9 years follow-up data from Zibetti, 2011 (and the respective preoperative data for each one); (F) data from n=20 (whole sample) and from n=9 (18 months follow up sample); evaluation moments at 3, 6, 12 and 18 months; n=20 preoperative data considered for short-term follow-up analysis; n=9 preoperative data considered for mid- term follow-up analysis; (G) stimulation device was turned on 4 weeks after the surgery: postoperative moments converted to post-DBS moments by subtracting 1 month; (H) "depression severity index" = "accumulative scores of each item"/"maximum scores of the scale": mean x 80 and SD x 80 assumed; (I) "the assessments took place (...) 12 months (...) later, with the exception of the cognitive status, which was controlled 3 months after surgery"; "outcome measures" = "motor function" + "cognitive status" + "psychiatric history" + "mood and behavioral modifications: ardouin scale" + "acute non-motor fluctuations": 1 year follow up assumed to mood evaluation; (J) results separated by groups "identical", "ameliorated" and "aggravated": not comparable with other studies; (K) Partial duplicate: 4 groups: STN versus GPI and Paris versus Grenoble: GPI in Grenoble, GPI versus STN comparison in Grenoble, STN in Paris, GPI in Paris and GPI versus STN comparison in Paris included; STN in Grenoble duplicated; (L) 4 groups: STN versus GPI and Paris versus Grenoble: only 57 in a total of 62 patients performed BDI assessment and the distribution by groups was not indicated: total n assumed for each group. (M) staged DBS; evaluation times were "1 month before first surgery, 2 months following first surgery (unilateral), and 3 months following second surgery (bilateral)": 3 months follow-up assumed; (N) "Test-retest interval was about 3 months between baseline and post-unilateral electrode placement evaluation, and 4 months between post-unilateral and post-bilateral electrode placement evaluations. This occurred with the exception of one patient who on separate occasions had the lead and pulse generator repositioned following bilateral operation, resulting in a 22-month lapse between neuropsychological assessments after first and second DBS electrode placement.": global 3 months follow-up assumed; (O) pre- and postoperative evaluations performed in patients group; only 1 evaluation in control group: postoperative cross-sectional analysis assumed.

Test-retest coefficient (r) was 0,66 for BAI[83], 0,64 (short term) and 0,75 (mid- and long- term) for BDI[81], 0,79 for BSI-a[85], 0,84 for BSI-d[85], 0,94 for GDS[84], 0,98 for HAD-a[88], 0,99 for HAD-d[88], 0,87 for HDRS[79], 0,56 for MADRS[80], 0,4 for STAI-s[90], 0,86 for STAI-t[90] and 0,651 for UPDRS I,3[86]. 0,98 assumed for BRMES[87]. 0,75 assumed for POMS-d[89]. Conservative value of 0,56 was assumed for SCL-90-R-d and Zung-d. Conservative value of 0,4 was assumed for AMDP-AT, BAS, HAMA, SCL-90-R-a and Zung-a.

Abbreviations: in "follow-up" column, w, m and y refers to weeks, months and years, respectively; pre: preoperative data; post: postoperative data; M: mean; SD: standard deviation; SE: standard error; 95%CI: 95% confidence interval; r: test-retest correlation coefficient; n: sample size; "change" refers to the postop - preop temporal change; "difference" refers to the STN - comparison group difference; AMDP-AT: association for methodology and documentation In psychiatry, anxiety part; Ardouin-a and Ardouin-d: "anxiety" and "depressive mood" items of the Ardouin scale, respectively; BAI: Beck anxiety inventory; BAS: brief scale for anxiety; BDI: Beck depression inventory; BRMES: Bech-Rafaelsen Melancholia Scale; BSI-a and BSI-d: anxiety and depression scales of the brief symptom inventory, respectively; GDS: geriatric depression scale; GPi: globus pallidus, pars interna; HAD-a and HAD-d: anxiety and depression parts of the hospital anxiety and depression scale, respectively; HAMA: Hamilton anxiety scale; HDRS: Hamilton depression rating scale; IOWA-a and IOWA-d: anxiety and depression parts of the IOWA scales of personality change, respectively; MADRS: Montgomery-Asberg depression rating scale; MINI-a and MINI-d: "general anxiety"/"anxiety disorders" and "major depression episode/disorder" items of the "mini international neuropsychiatric interview, respectively; NMSQuest-a and NMSQuest-d: items "anxiety" and "feeling sad" of the non motor symptom questionnaire, respectively; POMS-d: profile of mood states, depression domain; SCL-90-R-a and SCL-90-R-d: anxiety and depression domains of the symptom checklist-90-revised; STAI-s and STAI-t: state and trait (respectively) anxiety inventory; STN: subthalamic nucleus; UPDRS I,3: unified parkinson's disease rating scale,

part I, item 3 "depression"; Zung-a and Zung-d: Zung self-rating anxiety and depression scales, respectively. Country abbreviations according to ISO 3166-1 decoding table.

APPENDIX