Diagnosis and treatment of prolactinomas: the patient's perspective anno 2025

Trigt, V.R. van

CHAPTER 7

Subtle Cognitive Impairments and Psychological Complaints in Patients with Prolactinoma Despite Biochemical Control

Published: Journal of Clinical Endocrinology and Metabolism. 2025 Jun 13:dgaf355

Victoria R. van Trigt[1], Cornelie D. Andela[1,2], Leontine E.H. Bakker[1], Steffanie C.M. Brama[1], Lotte E. Schmidt[1], Florian M. Sneekes[1], Margot W. Zeelenberg[1], Sasja D. Huisman[1], Stephanie E.E.C. Bauduin[3], Olaf M. Dekkers[1], Marco J.T. Verstegen[4], Wouter R. van Furth[4], Iris C.M. Pelsma[1*], Nienke R. Biermasz[1*]

* *Shared last authors*

Author affiliations

1. Department of Medicine, Division of Endocrinology, and Center for Endocrine Tumors Leiden, Leiden University Medical Center, Leiden, the Netherlands
2. Department of Rehabilitation Medicine, Leiden University Medical Center, the Netherlands
3. Department of Psychiatry, Leiden University Medical Center, Leiden, the Netherlands
4. Department of Neurosurgery, Leiden University Medical Center, University Neurosurgical Center Holland, Leiden, The Netherlands

ABSTRACT

Purpose

To assess cognitive functioning and psychological complaints in patients with biochemically controlled prolactinoma.

Methods

Cross-sectional study comparing otherwise healthy patients treated for prolactinoma to age-, gender-, and education-level-matched controls. The cognitive assessment included eight tests assessing memory, verbal fluency, processing speed, selective attention, and executive functioning. Additionally, patients completed seven validated questionnaires on psychological complaints. Generalized estimating equations were performed. P-values <0.050 were considered significant.

Results

Sixty patients (controlled on dopamine agonists, n=30; in surgical remission, n=30), among whom 41 (68.3%) female, aged 42.3±11.7 years old, were compared to 60 matched controls. Patients scored significantly lower on assessments for verbal memory (fewer words on Verbal Learning Test of Rey: β=-1.8, 95%CI -2.7, -1.0), selective attention (fewer correct digits on Digit Deletion Test (β=-8.8, 95%CI -16.2, -0.2), longer time on Trial Making Test A (β=5.2, 95%CI 3.2, 7.2)), and processing speed (fewer correct substitutions on Digit-Symbol Substitution Test: β-4.2, 95%CI -8.2, -0.2). Furthermore, patients reported higher degrees of apathy (Apathy Scale: β=2.4, 95%CI 0.6, 4.1), irritability (Irritability Scale: β=2.2, 95%CI 0.3, 4.1), fatigue (Fatigue Severity Scale: β=6.7, 95%CI 2.7, 10.8), and anxiety and depressive symptoms (Hospital Anxiety and Depression Scale, anxiety: β=1.1, 95%CI 0.1, 2.1; depression: β=2.7, 95%CI 0.8, 2.7). Tests assessing executive functioning and task switching were comparable in patients and matched controls.

Conclusions

Compared to matched controls, patients treated for prolactinoma showed subtle cognitive impairments (i.e., memory, attention, and processing speed) and reported more psychological complaints. Physicians should be aware of these impairments and address them adequately.

INTRODUCTION

Patients treated for prolactinoma suffer from various physical symptoms, including headaches, galactorrhea, and hypogonadism – leading to subfertility and menstrual cycle disturbances in females [1]. Besides physical symptoms and complaints, patients frequently report a cognitive and psychological burden [2].

Treatment is predominately pharmacological with dopamine agonists (DA), although the most recent consensus statement suggests consideration of transsphenoidal surgery (TSS) for non-invasive prolactinomas [1]. Both interventions are effective in achieving normoprolactinemia [3], yet their effect on cognitive and psychological functioning remains largely unknown. A recent literature review including 18 mostly cross-sectional studies with high risks of bias, indicated improvement, but not always normalization of mental wellbeing after biochemical normalization [4]. Song et al. studied cognitive functioning in patients with prolactinoma and demonstrated that patients in surgical remission (n=20) showed better response activation and inhibition compared to patients with active prolactinoma (n=20) [5]. Moreover, amelioration of cognitive functions was observed in a heterogeneous group of patients operated for pituitary adenomas, including 12 patients with a prolactinoma [6]. Furthermore, Montalvo et al. demonstrated that cabergoline use was associated with improvement of cognitive functioning (i.e., processing speed, working memory, visual learning and problem solving) in a small group of patients with prolactinoma (n=7) [7]. A main limitation of these studies is the lack of control groups (accounting for confounders) and comparison between treatment modalities.

Considering the self-reported burden of decreased cognitive and psychological functioning prior to and after treatment, insight into cognitive functioning and psychological complaints after surgery or medical treatment would be valuable. Therefore, this cross-sectional study reports on cognitive functioning and psychological complaints in patients with prolactinoma treated medically or surgically who do not have overt psychopathological comorbidity, in comparison to matched healthy controls. Based on clinical experience and previous findings, treated patients were hypothesized to demonstrate remaining impairments in cognitive functioning and psychological complaints, and that these impairments would positively correlate with prolactin levels at diagnosis, hypopituitarism, and symptoms of anxiety and depression, which improve with longer durations of biochemical control/remission. Furthermore, patients in surgical remission were hypothesized to have less impairments than DA-controlled patients, as DAs may cause cognitive and psychological side effects.

METHODS

Study design and participants

This cross-sectional study compared cognitive functioning and patient-reported psychological complaints in patients with prolactinoma (18-70 years old) without any confirmed psychiatric comorbidity to matched controls. The study was approved by the Science Committee (W2020.020), and all participants gave digital informed consent.

Two patient groups were studied: normoprolactinemic patients with prolactinoma (prolactin below the upper limit of normal (xULN)) (1) controlled on a stable DA dose for ≥6 weeks, or (2) in surgical remission ≥6 months after TSS. Prolactinoma diagnosis was based on the combination of symptomatic hyperprolactinemia, a pituitary mass on magnetic resonance imaging (MRI), and exclusion of non-tumorous causes of hyperprolactinemia. Healthy controls were matched 1:1 to the patients based on age (<10-year age difference), gender, and education level (i.e., low, medium, and high, based on the guidelines of Statistics Netherlands [8]) at the time of the cognitive assessment. Controls were recruited either through referrals by patients or via advertisements if the patients could not provide a control. All participants were compensated for travel costs and controls who were recruited via advertisements received a 20-euro gift voucher.

Exclusion criteria were current pregnancy, current or past drug or alcohol abuse, use of medication known to reduce cognitive functioning (e.g., opiates, benzodiazepines, antihistamines), previous (pituitary) radiotherapy, major comorbidity (e.g., severe kidney, liver, cardiac, systemic inflammatory disease, malignancy), neurological pathology (e.g., cerebrovascular accident, cerebral trauma, dementia, epilepsy), (history of) any psychiatric condition (e.g. obsessive compulsive disorder, anxiety disorder, depression, attention deficit disorder, attention deficit hyperactivity disorder, burn-out). Strict exclusion of *all* psychopathologies was performed to avoid bias. Healthy controls had no physical or psychiatric conditions, no medication use (contraceptives were accepted), and no current or past alcohol or drug abuse.

Prolactinoma treatment

Patients were treated at the outpatient clinic of the Leiden University Medical Center (LUMC), a tertiary referral center for pituitary care, according to international guidelines, following a previously described Value-Based Health Care (VBHC) care pathway [1, 9, 10]. Most DA-treated patients were on cabergoline which was up-titrated if needed, aiming at the minimal dose to maintain normoprolactinemia [1]. Surgically treated patients underwent endoscopic TSS in a dedicated care protocol described previously [11]. The surgical indication was typically DA intolerance. Hypopituitarism was assessed dynamically upon clinical indication and substituted following international guidelines [12].

Study procedures

All participants completed an online set of validated questionnaires, i.e., patient-reported outcomes measures (PROMs), in the week prior to the cognitive assessment. All participants underwent an extensive cognitive assessment performed by one out of five trained researchers (SCMB, LES, FMS, MWZ, VRvT) following a standardized protocol, as shown in Supplementary File 1 [13]. All assessments were conducted under quiet conditions in a separate hospital room. Patients and their matched controls were evaluated at the same time of day to account for circadian fluctuations in cognitive functions. Two researchers independently scored cognitive test performance. Disagreements were resolved by discussion. In case of persisting disagreement, the majority vote was selected through consultation of a third researcher. The second and third correctors were blinded for the type of participant.

Cognitive assessment

The cognitive assessment consisted of 8 tests (outlined below) and lasted 60-75 minutes in total. Detailed information on the tests and their scoring is provided in Supplementary Table 1 [13].

The Verbal Learning Test of Rey

Fifteen unrelated words were shown and read to the participant, to be reproduced by the participant [14]. The test consists of five rounds of reproduction: immediate reproduction (first four rounds), and a delayed reproduction (after 20 minutes). This study reports reproduction round 1, 2, 4 and the delayed reproduction - providing the most relevant information. More reproduced words indicate better verbal memory (scale 0-15).

WAIS Digit Span Task

A series of digits are read aloud by the investigator to be reproduced by the participant in three rounds: digit span forward (identical order), digit span backward (reversed order), and sequencing (ascending order). Higher scores per round indicate better verbal memory and working memory (scale 0-16 per round) [15].

Rey Complex Figure Test

Participants copy a complex figure and reproduce the figure from memory after 3 minutes (immediate recall) and 30 minutes (delayed recall) [16]. Higher scores indicate better visuospatial memory (scale 0-36).

WAIS Digit-Symbol Substitution Test

Participants substitute as many numbers with indicated symbols as possible within two minutes [17]. More correctly substituted numbers indicate better selective attention and processing speed (scale 0-135).

Digit Deletion Test

Participants are presented with a form containing 800 digits (numbers 1-9), in which they cross out numbers 3 and 7 diagonally and underline number 4. The number of correctly edited, incorrectly edited and missed numbers in three minutes are counted [18]. More correctly edited numbers indicate better selective attention and processing speed (scale 0-240).

Trail Making Test (TMT)

In TMT-A, participants connect circles with numbers (numbers 1-25) in ascending order. In TMT-B, participants connect numbers (number 1-13) and letters (letter A-L) in alternating order [19]. Shorter duration to perform the task and fewer mistakes indicate better selective attention (TMT-A), and cognitive flexibility (TMT-B).

D-KEFS Tower Test (TT)

Participants replicate nine towers by stacking five differently sized disks onto three wooden pegs, never stacking larger disks on top of smaller disks, and only moving one disk at a time [20]. A total performance score, average time to first step, time-per-step-ratio, step-accuracy-ratio, and rule-violations-per-item-ratio are noted. Higher scores indicate better executive functioning (scaled scores range 0-10).

FAS

Participants produce as many words as possible in one minute beginning with an F, A, and S, respectively [21]. More correct-, and fewer incorrect words and repetitions indicate better executive functioning and verbal fluency (no maximum score).

PROMs

Participants completed an online set of either six or seven validated PROMs (duration approximately 60 minutes). The PROMs are summarized below. Detailed explanations and scoring are provided in Supplementary Table 2 [13].

The first questionnaire, *Leiden Bothers and Needs Pituitary (LBNQ-Pituitary)* was only completed by patients to assess clinical characteristics, including relevant questions concerning psychological and cognitive functioning. This questionnaire measures pituitary-disease burden (Bothers) and the Needs for attention for these symptoms by the treating physician. It consists of five separate subscales and a total score for Bothers and Needs, respectively (33 items, scale 0-100 with higher scores indicating higher disease burden) [22].

The following six questionnaires were completed by all participants. The *Apathy Scale (AS)* and *Irritability Scale* (IPS) measure various aspects of apathy and irritability, respectively (14 items, scale 0-42 with higher scores indicating greater apathy or irritability, respectively) [23, 24]. Scores ≥14 on either instrument indicate participants

being apathic or irritable, respectively. The *Fatigue Severity Scale (FSS)* measures the effect of fatigue on health-related quality of life (HR-QoL) (9 items, scale: 9-63, with higher scores indicating more fatigue) [25]. The *Hospital Anxiety and Depression Scale (HADS)* describes the severity of depressive symptoms and anxiety (14 items, total score scale 0-21)) [26, 27], with a separate subscale for depressive symptoms and anxiety (sub scores ≥8: clinically relevant depressive symptoms or anxiety). The *Irrational Beliefs Inventory-50* (IBI-50) measures the extent of irrational beliefs including five subscales (50 items, scale 0-250, with higher scores indicating stronger irrational beliefs) [28]. The *Dutch Clinical Personality Questionnaire* (DCPQ) measures six personality traits: Negativism, Shyness, Extraversion, Narcissism, and Severe Psychopathology (120 items, scale 0-40 per trait, with higher scores indicating a higher chance of having the trait) [29].

Study parameters

Patient demographics, biochemical analyses, and radiologic examinations were derived from our prospective database [30]. Demographic information from controls and additional information from all participants were acquired during a short interview as part of the cognitive assessment protocol (Supplementary File 1 [13]). Adenoma remnants were subdivided by current size: micro <10mm, macro 10-40mm, and giant >40mm. Disease duration was defined as the time from diagnosis to cognitive assessment (months). Duration of biochemical control/remission was defined as duration from sustained normoprolactinemia to cognitive assessment (months). For this study, subtle cognitive impairment was defined as significantly worse scores compared to healthy controls, without deviations >2.0SD from the control group's mean.

Data analysis

Data were analyzed using IBM SPSS statistics 29 (IBM Corp. Armonk, NY, USA) and reported as mean ± standard deviation (SD) or median [interquartile range (IQR)] depending on the normality of the data distribution for continuous variables, or frequency (percentage) for dichotomous variables. For comparison of baseline characteristics between patients and controls, an independent T-test was used for continuous variables, and a χ^2 test for categorical data.

The primary analysis compared the outcomes of the cognitive assessment and PROMs of treated patients with (1:1) matched controls using generalized estimating equations (GEEs) accounting for matching without correction for additional factors as baseline characteristics were not statistically nor clinically relevantly different. Z-scores and 95% confidence intervals (95%CI) for cognitive test results were calculated based on the patients' matched controls to account for differences in age, gender, education, and time of testing.

The secondary analyses compared 1) the outcomes of the cognitive assessment and PROMs of patients controlled on DA with patients in surgical remission using generalized

linear models (GLMs) of Z-scores. 95%CI and β were reported. 2) Subsequently, factors influencing the cognitive assessment were estimated by multilinear regression analyses using Z-scores for the outcomes of the cognitive assessment as dependent factors, and serum prolactin at diagnosis, hypopituitarism (yes/no), duration of biochemical control/remission, and total HADS score as predictive factors. Appropriateness assumptions were evaluated using scatter plots, probability-probability plots, residual statistics, and Cooks tests. Baseline prolactin and biochemical control/remission duration were log-transformed because they were not normally distributed. Regression analyses were only performed for outcomes of the cognitive assessment that differed between patients and controls in the primary analysis to limit the number of analyses. Again, 95%CI and standardized β were reported.

A p-value of <0.050 was considered significant, due to the exploratory nature of the study and the underlying association of the endpoints, as the tests cover partially overlapping cognitive domains. To avoid overcorrection, it was more appropriate to present outcomes with interval estimations, evaluating general patterns, than to merely perform hypothesis testing with correction for multiplicity [31].

RESULTS

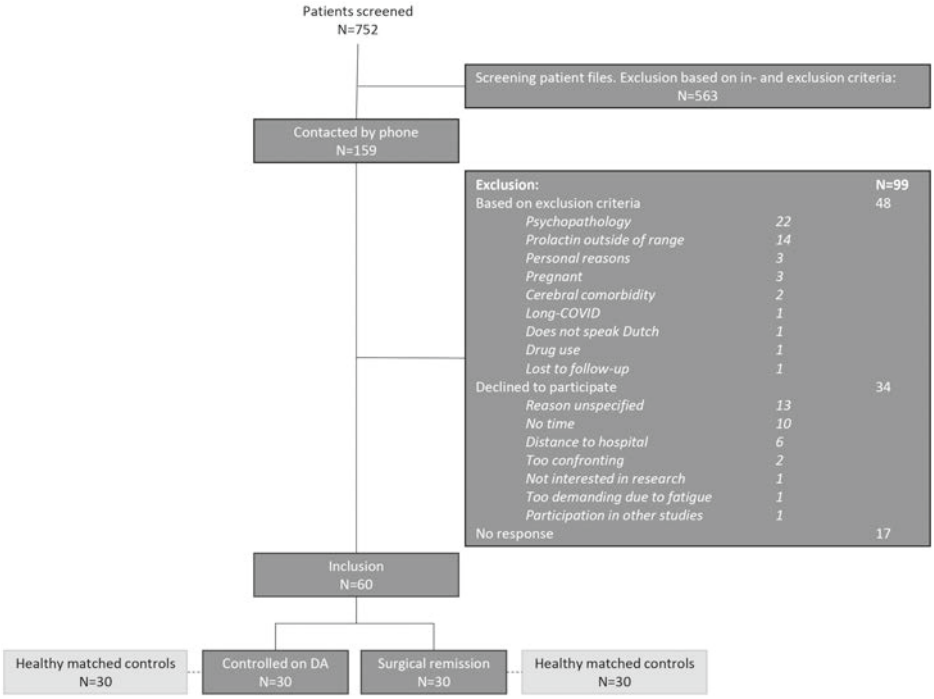
1. Full cohort

Inclusion of patients and clinical characteristics

The flowchart of patient inclusion is depicted in Figure 1. Eligible patients were invited to participate by phone or email (n=159). In total, 142 patients were successfully contacted, of whom 48 were excluded based on in- and exclusion criteria (among which 22 due to psychopathology), and 34 patients declined to participate, primarily citing lack of time or travel distance as reasons (unspecified reason, n=13).

Thus, 60 patients were included (41 females (68.3%)), mean age 42.3±11.7 years, whose characteristics are summarized in Table 1. The patients' education levels were classified as either high (n=37, 61.7%), medium (n=19, 31.7%), or low (n=4, 6.7%). Of these, 30 patients were controlled on DA (17 females (56.7%)), and 30 patients were in surgical remission (24 females (80.0%)). Concerning the DA-group, 27 patients were on cabergoline (90.0%), and 20 patients used DA for >2 years (66.7%). The mean current DA dose was 0.62±0.50 mg/week. The surgical group underwent surgery 26 [12-39] months prior to cognitive assessment. Twenty-eight (93.3%) surgical patients had received prior DA-treatment but were off medication since surgery. At the time of assessment, prolactin levels were 0.4±0.3xULN. The median duration of biochemical control/remission was 31 [13-77] months.

Figure 1 Flowchart of patient inclusion



DA, dopamine agonist.

Table 1 Characteristics at the time of cognitive assessment

		All patients N=60	All matched healthy controls N=60	Patients controlled on DA N=30	Healthy controls (DA) N=30	Patients in surgical remission N=30	Healthy controls (surgery) N=30
General characteristics							
Gender (female)		41 (68.3)	41 (68.3)	17 (56.7)	17 (56.7)	24 (80.0)	24 (80.0)
Age[a,b] (years)		42.3±11.7	40.6±12.7	45.5±12.6	44.2±13.4	39.2±9.9	37.0±11.1
Education	Low	4 (6.7)	4 (6.7)	0 (0.0)	0 (0.0)	4 (13.3)	4 (13.3)
	Medium	19 (31.7)	19 (31.7)	12 (40.0)	12 (40.0)	7 (23.3)	7 (23.3)
	High	37 (61.7)	37 (61.7)	18 (60.0)	18 (60.0)	19 (63.3)	19 (63.3)
Previous night sleep[a]	h : min	7:00±1:33	7:23±1:02	6:53±1:22	7:16±1:00	7:07±1:43	7:31±1:03
Hormonal contraceptives[a]	Systemic	6 (10.0)	15 (25.0)	2 (6.7)	5 (16.7)	4 (13.3)	10 (33.3)
	Local	6 (10.0)	6 (10.0)	3 (10.0)	2 (6.7)	3 (10.0)	4 (13.3)
Caffeine (units/week)[a]		20.6±17.1	20.7±18.0	24.0±19.8	20.0±13.2	16.2±12.8	21.4±15.5
Current smoking[a]		8 (13.3)	10 (16.7)	3 (10.0)	5 (16.7)	5 (16.7)	5 (16.7)
Alcohol (units/week)[a]	≤2	41 (68.3)	29 (48.3)	20 (66.7)	18 (60.0)	21 (70.0)	11 (36.7)
	3-7	11 (18.3)	23 (38.3)	5 (16.7)	8 (26.7)	6 (20.0)	15 (50.0)
	8-14	6 (10.0)	5 (8.3)	3 (10.0)	2 (6.7)	3 (10.0)	3 (10.0)
	>14	2 (3.3)	3 (5.0)	2 (6.7)	2 (6.7)	0 (0.0)	1 (3.3)
Disease characteristics							
Control/ remission duration[c] (m)		31 [13-77]	-	41 [14-111]	-	29 [13-65]	-
Disease duration[d] (m)		60 [28-118]	-	54 [24-126]	-	60 [60-112]	-
Prolactin at diagnosis (xULN)[e]		4.6 [2.8-23.8]	-	15.0 [2.9-29.2]	-	4.0 [2.6-7.2]	-
Current prolactin (xULN)		0.4±0.3	-	0.4±0.3	-	0.5±0.2	-
Tumor size at diagnosis	Not visible[f]	2 (3.3)	-	2 (6.7)	-	0 (0.0)	-
	Micro	30 (50.0)	-	11 (36.7)	-	19 (63.3)	-
	Macro	24 (40.0)	-	14 (46.7)	-	10 (33.3)	-
	Giant	4 (6.7)	-	3 (10.0)	-	1 (3.3)	-
Current tumor size[g]	Not visible	30 (51.7)	-	2 (6.7)	-	28 (100.0)	-
	Micro remnant	15 (25.9)	-	15 (50.0)	-	0	-
	Macro remnant	13 (22.4)	-	13 (43.3)	-	0	-
	Giant remnant	0	-	0	-	0	-
Current hypopituitarism[h]	Any Axis	12 (20.0)	-	8 (26.7)	-	4 (13.3)	-
	LH/FSH	8 (13.3)	-	7 (23.3)	-	1 (3.3)	-

Table 1 Characteristics at the time of cognitive assessment *(continued)*

		All patients N=60	All matched healthy controls N=60	Patients controlled on DA N=30	Healthy controls (DA) N=30	Patients in surgical remission N=30	Healthy controls (surgery) N=30
	TSH	4 (6.7)	-	2 (6.7)	-	2 (6.7)	-
	ACTH	2 (3.3)	-	1 (3.3)	-	1 (3.3)	-
	GH	1 (1.7)	-	0	-	1 (3.3)	-
	AVP	1 (1.7)	-	0	-	1 (3.3)	-
Treatment							
Current (DA)	Yes	30 (50.0)	-	30 (100.0)	-	0	-
	Cabergoline	27 (90.0)	-	27 (90.0)	-	0	-
	Quinagolide	1 (3.3)	-	1 (3.3)	-	0	-
	Bromocriptine	2 (6.7)	-	2 (6.7)	-	0	-
Previous	DA	58 (96.7)	-	30 (100.0)	-	28 (93.3)	-
	TSS[i]	32 (53.3)	-	2 (6.7)	-	30 (100.0)	-
	Radiotherapy	0	-	0	-	0	-
Duration DA treatment	< 6 m	6 (10.3)	-	1 (3.3)	-	5 (17.9)	-
	6 m-2 y	15 (25.9)	-	8 (26.7)	-	7 (25.0)	-
	> 2 y	36 (62.1)	-	20 (66.7)	-	16 (57.1)	-
	Unknown	1 (1.7)	-	1 (3.3)	-	0	-
Current DA dose[j] (mg/week)		0.62±0.50	-	0.62±0.50	-	0	-
Time since surgery (months)		25 [12-39]	-	31 [22-31]	-	26 [12-39]	-

Data reported as value (%), mean ± standard deviation or median [interquartile range]. *ACTH*, adrenocorticotropic hormone; *AVP*, arginine vasopressin; *DA*, dopamine agonist; *FSH*, follicle-stimulating hormone; *GH*, growth hormone; *h*, hours; *m*, month; *min*, minutes; *NA*, not applicable; *LH*, luteinizing hormone; *TSH*, thyroid-stimulating hormone; *TSS*, transsphenoidal surgery; *xULN*, times upper limit of normal (females 23.3ug/L, males 15.2 ug/L); *y*, years.

[a] Characteristics were statistically compared between (1) all patients and all controls, (2) patients controlled on DA and their matched healthy controls, and (3) patients in surgical remission and their matched healthy controls. No significant differences were found (p>0.050).

[b] Statistical testing was performed despite matching, as 10 years age difference was accepted.

[c] Defined as time from achievement of sustained normoprolactinemia to cognitive assessment, available for n=55 (patients controlled on DA: n=28, patients in surgical remission: n=27).

[d] Defined as time from diagnosis to cognitive assessment.

[e] Data available for 54 patients (patients controlled on DA: n=26, patients in surgical remission: n=28).

[f] Both patients developed a visible adenoma throughout the disease course.

[g] Data available for 58 patients; postoperative MRI was not yet performed in 2 surgically treated patients.

[h] All patients were substituted adequately according to international guidelines, except for one asymptomatic male with mild hypogonadism in the DA group who refused testosterone substitution.

[i] Two patients in surgical remission had undergone two surgeries.

[j] In equivalents of cabergoline: bromocriptine 2.5mg/day or quinagolide 75ug/day were considered to be the equivalent of cabergoline 0.5mg/week.

Hypopituitarism of any axis was present in 12 (20.0%) patients, concerning the gonadotropic axis in 8 patients (13.3%). Hypopituitarism was substituted adequately in all but one asymptomatic, mildly hypogonadotropic male refusing testosterone treatment.

None of the 30 surgically treated patients had an adenoma remnant on most recent MRI (missing data, n=2, 6.7%). Concerning DA-treated patients, a microadenoma remnant was present in 15 patients (50.0%), a macroadenoma remnant in 13 patients (43.3%), and no remnant was visible in 2 patients (6.7%).

The self-reported disease burden as measured by LBNQ-Pituitary, indicated the highest burden in the *Physical and Cognitive Complaints*, followed by the *Mood Symptoms* domain for both DA-treated and surgically treated patients. An overview of LBNQ-Pituitary scores is presented in Table 2. Concerning individual questions, twenty-four (40.0%) patients indicated mood disturbances due to their pituitary disease. Memory and concentration problems were indicated by 37 (61.7%) and 28 (46.7%) patients, respectively.

Table 2 self-reported disease burden for patients as measured by LBNQ-Pituitary

LBNQ-Pituitary		**All patients N=60**	**Patients controlled on DA N=30**	**Patients in surgical remission N=30**
Bothers	Mood problems	6.9 [0.0-29.9]	12.5 [0.0-29.9]	5.6 [0.0-31.3]
	Negative illness perception	2.1 [0.0-12.5]	4.2 [0.0-13.5]	0.0 [0.0-9.4]
	Issues in sexual functioning	0.0 [0.0-12.5]	0.0 [0.0-18.8]	0.0 [0.0-14.1]
	Physical and cognitive complaints	16.7 [2.8-38.9]	18.1 [2.8-38.9]	15.3 [0.0-22.4]
	Issues in social functioning	0.0 [0.0-10.0]	0.0 [0.0-12.5]	0.0 [0.0-10.0]
	Total	8.7 [1.1-22.7]	8.3 [2.5-23.5]	9.1 [0.8-22.2]
Needs	Mood problems	8.3 [0.0-25.0]	8.3 [0.0-22.9]	6.9 [0.0-31.9]
	Negative illness perception	0.0 [0.0-15.6]	2.1 [0.0-16.7]	0.0 [0.0-13.5]
	Issues in sexual functioning	0.0 [0.0-12.5]	0.0 [0.0-14.1]	0.0 [0.0-12.5]
	Physical and cognitive complaints	11.1 [0.7-36.1]	11.1 [2.1-38.2]	12.5 [0.0-37.5]
	Issues in social functioning	0.0 [0.0-7.5]	0.0 [0.0-11.3]	0.0 [0.0-2.5]
	Total	8.0 [1.0-22.5]	8.0 [1.3-24.2]	7.6 [0.8-23.7]

Data reported as median [interquartile range]. *LBNQ-Pituitary,* Leiden Bothers and Needs Pituitary.

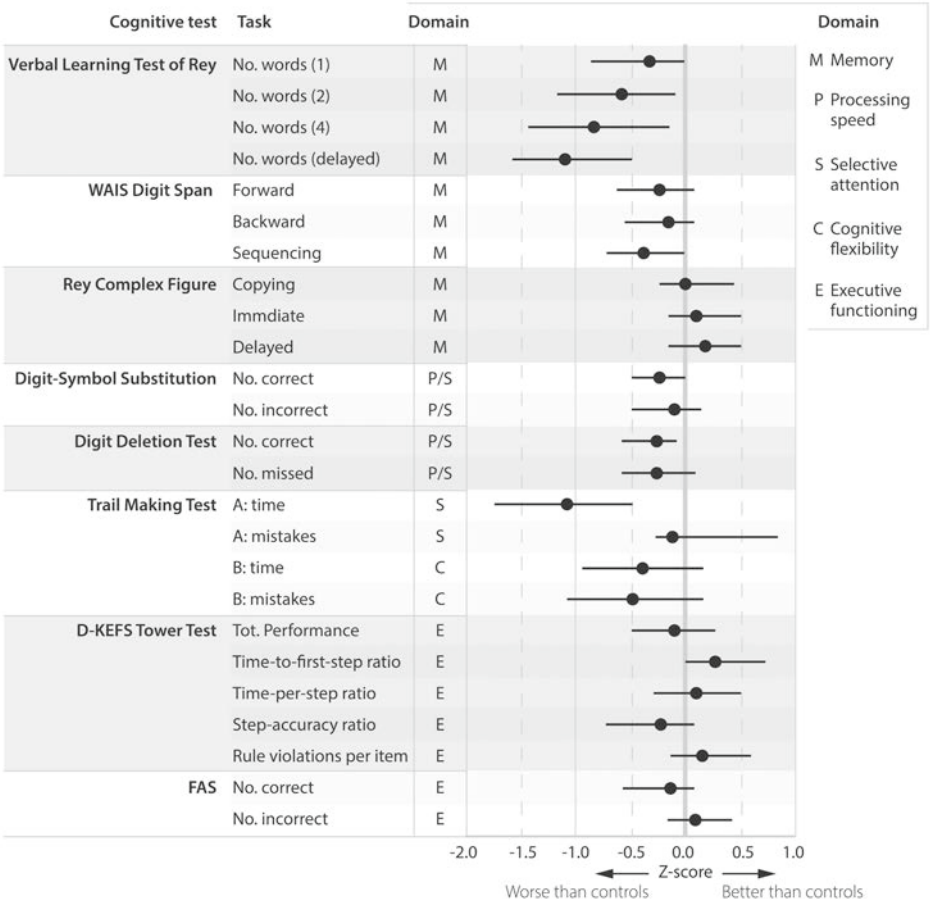
1. Full cohort

Cognitive functioning

The outcomes of the cognitive assessment are summarized in Figure 2 and Supplementary Table 3 [13]. Using Z-scores for the overall comparison of patients to controls, patients scored 0.2 to 1.1SD lower than controls on tasks measuring verbal memory (Rey Complex Figure Test, WAIS Digit Span Task). Patients performed significantly worse on all attempts of the Verbal Learning Test of Rey (fewer reproduced words, most pronounced on the delayed attempt: β=-1.8 (95%CI -2.7, -1.0), p<0.001). Patients performed -0.2 to -1.1SD worse on tasks assessing attention, with fewer correct substitutions on WAIS Digit-Symbol Substitution Test (β=-4.2 (95%CI -8.2, -0.2), p=0.040), and more time

(β=5.2 (95%CI 3.2, 7.2), p<0.001) and a higher number of mistakes on TMT-A (β=0.1 (95%CI 0.0, 0.2), p=0.028)). Patients performed -0.2 to -0.4SD worse on tasks assessing processing speed (Digit deletion test: fewer correctly deleted digits (β=-8.8 (95%CI -16.2, -1.4), p=0.019), and WAIS Digit-Symbol Substitution Test). Patients performed similar to matched controls on tasks assessing visuospatial memory (Rey Complex Figure Test), cognitive flexibility (TMT-B) and executive functioning (TT, FAS).

Figure 2 Results of the cognitive assessment for biochemically controlled patients after medical treatment or surgery (n=60)



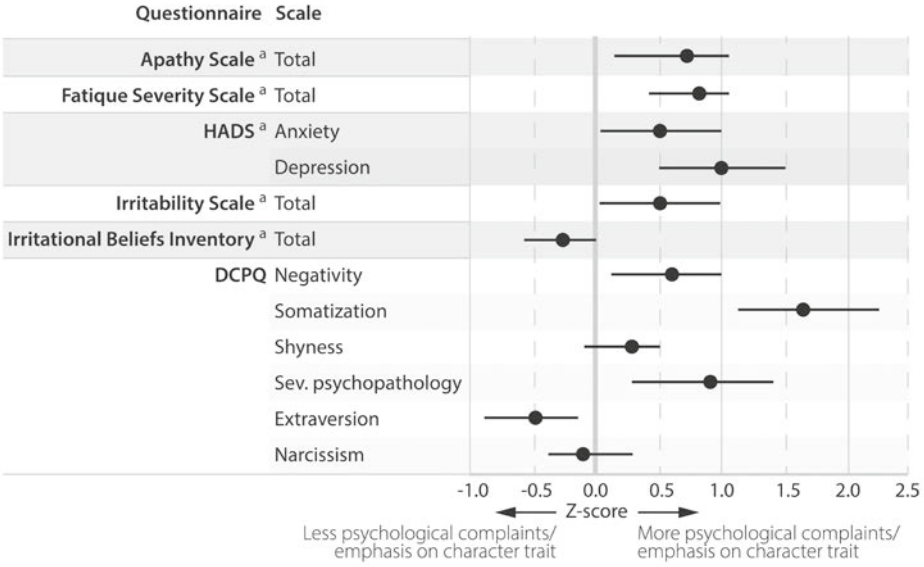
The cognitive domain tested is indicated per test. Data are displayed as Z-scores for patients based on matched healthy controls, and reported as mean with 95% confidence intervals. Lower Z-scores indicate patients scoring worse than controls, and vice versa. For the Verbal Learning Test of Rey the first, second, fourth and delayed attempt are shown, as these provide the most relevant information. *No.*, number; *Tot.*, total.

Psychological complaints and personality traits

As shown in Figure 3 and Supplementary Table 4 [13], patients generally reported more psychological symptoms than controls. Patients reported significantly more apathy (AS: β=2.4 (95%CI 0.6, 4.1), P=0.009), more fatigue (FSS: β=6.7 (95%CI 2.7, 10.8), p<0.001), more irritability (IPS: β=2.2 (95%CI 0.3, 4.1), p=0.024), more anxiety (HADS anxiety score: β=1.1 (95%CI 0.1, 2.1), p=0.034), and more depressive symptoms (HADS depression score: β=1.7 (95%CI 0.8, 2.7), p<0.001). Using the HADS, clinically relevant symptoms of depression were observed in 5 (8.3%) patients, and anxiety in 15 (25.4%) patients, respectively. There were no significant differences regarding irrational beliefs (IBI-50) between patients and controls.

Concerning maladaptive personality traits (DCPQ), patients exhibited higher levels of negativism (β=2.8 (95%CI 0.8, 4.9), p=0.007), somatization (β=7.3 (95%CI 5.0, 9.5), p<0.001), severe psychopathology (β=1.9 (95%CI 0.8, 3.1), p=0.001), and lower levels of extraversion (β=-3.7 (95%CI -6.7, -0.7), p=0.015) compared to controls. Degrees of shyness and narcissism were comparable between patients and controls.

Figure 3 Results of patient-reported outcome measures concerning psychological complaints and maladaptive personality traits for biochemically controlled patients after medical treatment or surgery (n=60)



Data are displayed as Z-scores for patients based on matched healthy controls, and reported as mean with 95% confidence intervals. Lower Z-scores indicate patients reporting less psychological complaints or emphasis on a character trait than controls, and vice versa. *DCPQ*, Dutch Clinical Personality Questionnaire; *HADS*, Hospital Anxiety and Depression Score; *sev.*, Severe.

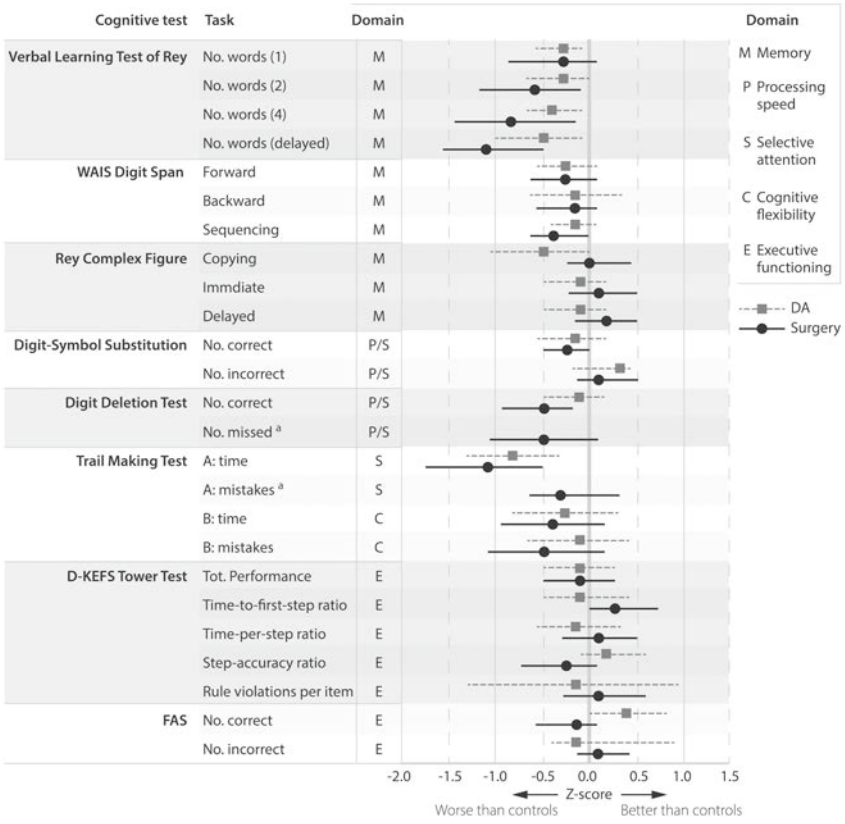
[a] Data missing for one surgically treated patient who did not complete all patient-reported outcome measures (n=59).

2. Comparison of medically and surgically treated patients

Cognitive functioning

Baseline characteristics and results of the LBNQ-Pituitary for the treatment groups are shown in Table 1 and 2, respectively. None of the patient groups had mean Z-scores >2.0SD below the control group's mean on any outcome of the cognitive assessment. Surgically- and DA-treated patients generally showed comparable results on the cognitive tests, as depicted in Figure 4. Differences were marginal and distributed randomly across cognitive domains, thus most likely resulting from multiple testing. An overview of absolute test results for the patient groups and controls, and Z-scores of the outcomes of the cognitive assessment are provided in Supplementary Table 5 and 6, respectively (13).

Figure 4 Results of the cognitive assessment for patients controlled on DA (n=30) and patients in surgical remission (n=30).



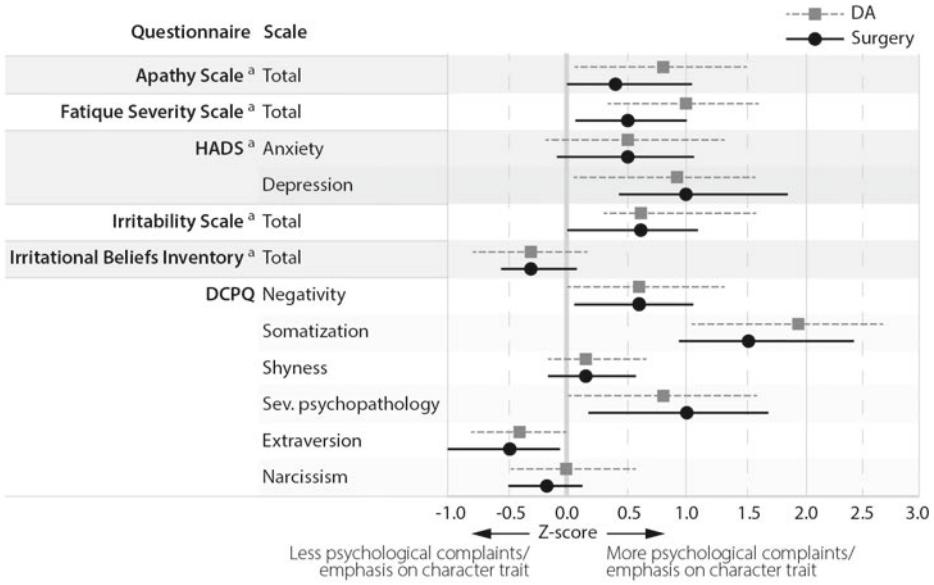
Data are displayed as Z-scores for patients based on matched healthy controls, and reported as mean with 95% confidence intervals. Lower Z-scores indicate patients scoring worse than controls, and vice versa. Differences in disease characteristics may exist between patients controlled on DA and patients in surgical remission as they were not matched. For the Verbal Learning Test of Rey the first, second, fourth and delayed attempt are shown, as these provide the most relevant information. *DA*, dopamine agonist; *No.*, number; *Tot.*, total.

[a] Z-score could not be calculated for patients controlled on DA as the mean and standard deviation were zero for the matched controls.

Psychological complaints and maladaptive personality traits

None of the patient groups had mean Z-scores >2.0SD below the control group's mean on any PROM. Surgically and DA-treated patients scored comparably on all PROMs. Concerning maladaptive personality traits, both patient groups scored highest on somatization, as depicted in Figure 5. An overview of absolute results for the patient groups and controls, and Z-scores for PROMs are provided in Supplementary Table 7 and 8, respectively [13].

Figure 5 Results of patient-reported outcome measures concerning psychological complaints and maladaptive personality traits for patients controlled on DA (n=30) and patients in surgical remission (n=30)



Data are displayed as Z-scores for patients based on matched healthy controls, and reported as mean with 95% confidence intervals. Lower Z-scores indicate patients reporting less psychological complaints or emphasis on a character trait than controls, and vice versa. Differences in disease characteristics may exist between patients controlled on DA and patients in surgical remission as they were not matched. *DCPQ*, Dutch Clinical Personality Questionnaire; *HADS*, Hospital Anxiety and Depression Score; *sev.*, Severe.

[a] Data missing for one surgically treated patient who did not complete all patient-reported outcome measures.

3. Factors of influence for cognitive functioning

Z-scores of cognitive tests were not correlated with prolactin at diagnosis, duration of biochemical control/remission, presence of pituitary deficiencies, or total HADS scores. Outcomes of multilinear regression analyses are shown in Supplementary Table 9 [13].

DISCUSSION

Clinical experience suggested that cognitive and psychological complaints are prevalent among patients with prolactinoma and can persist after disease control. This study was the first to compare a large group of patients with biochemically controlled prolactinoma to matched controls, enabling additional comparison of treatment modalities, using Z-scores. Patients with diagnosed psychological complaints were excluded to avoid bias, potentially leading to underestimation of cognitive impairments and psychological complaints. Even in this selected group, our clinical observations were confirmed - with 40% of the patients self-reporting mood disturbances, and half of the patients self-reporting memory and concentration problems. This study demonstrated subtle cognitive impairments, and more psychological complaints and maladaptive personality traits in patients compared to controls.

Previous research in untreated patients with prolactinoma demonstrated cognitive impairments, including impairments in verbal memory, working memory, attention and executive functioning [32-35]. An overview of available studies on cognitive functioning in prolactinoma cohorts is provided in Table 3. Two studies prospectively examining cognitive functioning showed improvement of response activation (using EEG measurements), and processing speed, working memory, visual learning and reasoning, and problem solving after treatment [6, 7]. One cross-sectional study found no differences in reaction times and accuracy between surgically treated patients and controls [5]. The results of our cross-sectional study in patients with biochemically controlled prolactinoma showed subtle cognitive impairments (i.e., memory, attention, processing speed), suggesting that these impairments may persist after biochemical normalization. We hypothesized cognitive functioning would be impaired to a greater extent in DA-controlled patients compared to patients in surgical remission, as DAs may induce (cognitive) side effects. However, after correction for age, gender, and education level, DA-controlled patients and patients in surgical remission scored similarly on the cognitive tasks. This finding might be explained by good tolerance to DA in the medically treated patients, since DA side effects were an indication for neurosurgical intervention.

Although comparison of cognitive impairments between pituitary diseases is difficult due to differences in cognitive assessment protocols, similar impairments in verbal learning, attention, and processing speed were found in patients with long-term (on average 13 years) remission of Cushing's Disease (CD) [36]. By contrast, no cognitive impairments were found in patients after long-term remission of acromegaly or non-functioning pituitary adenoma (NFPA) [37]. As the present study was cross-sectional, and relatively short-term after intervention, long-term cognitive outcomes in patients with treated prolactinoma remain unknown. The observed impairments may still be reversible (with appropriate rehabilitation). Therefore, monitoring cognitive complaints and referring to cognitive rehabilitation facilities is important for this group, in whom these problems are less acknowledged than in CD. Monitoring cognitive complaints can be performed during routine clinical visits by using (open-ended) screening questions (e.g., "how is memory going?", "how is concentrating, such as in following the plot of a movie or book?", Do you need more time to process information and/or solve problems?") potentially supported by using validated questionnaires (e.g., Cognitive failure questionnaire, LBNQ-Pituitary). Furthermore, the opinion of a patient's partner/spouse on these issues should be included when available.

Concerning psychological complaints, higher degrees of apathy, fatigue, irritability, anxiety, and depression were reported by patients compared to controls, which was in line with a recent systematic review on HR-QoL, indicating the most pronounced impairments in the mental health domain [4]: more fatigue, poorer sleep quality, and shorter sleep duration [2, 38, 39]. Albeit equivocal, symptoms generally improved – without normalization – after biochemical control [3, 4, 9, 40, 41]. Our group previously studied psychological complaints and personality traits using the same PROMs in patients with acromegaly (n=68), CD (n=51), and NFPA (n=60) in long-term biochemical control [36, 37, 42], enabling comparison of Z-scores. Psychological complaints in patients with biochemically controlled prolactinoma were similar to complaints in CD, acromegaly and NFPA. Patients with treated prolactinoma, acromegaly and NFPA scored 0.5 to 1.0SD worse than matched controls, whereas patients with CD scored 1.0 to 2.0SD worse than matched controls (37, 42). Furthermore, these previous studies demonstrated more maladaptive personality traits (i.e., negative affect, lack of positive affect, somatic arousal) in treated patients with CD and acromegaly compared to controls [36, 42]. Athanasoulia et al. observed more neuroticism, increased sensitivity to negative emotional stress, increased fear of uncertainty, and higher degrees of socially desirable behavior in patients with prolactinoma (controlled on DA and active) compared to healthy controls [37]. Thus, previous and present findings suggest that psychological complaints are prevalent after biochemical normalization in patients with pituitary adenoma – irrespective of the (type of) hormonal hypersecretion. Therefore, these psychological complaints, and personality traits, should be addressed in both patients with active disease, and after biochemical control. Monitoring these psychological complaints can be supported by validated PROMs (e.g., LBNQ-Pituitary, HADS), and

patients should be referred to a psychologist/social worker and/or self-management programs upon indication [43].

The underlying mechanisms of subtle cognitive impairments and psychological complaints may be multifactorial. Firstly, hyperprolactinemia might induce (ir) reversible alterations in the brain, since alterations in brain activity and brain structures have been found in patients with active prolactinoma (Table 3) [5, 33-35]. Furthermore, these structural alterations (i.e., decreased grey matter volume of the left hippocampus, left orbitofrontal cortex, right middle frontal cortex and right interior frontal cortex) were negatively correlated with verbal memory and executive functioning [35]. Additionally, prolactin levels and cognitive impairments were positively associated in some [32-34], yet not all studies [5]. In our cohort, prolactin levels at diagnosis did not correlate with the cognitive assessment after treatment, potentially due to (partial) reversibility, variable sensitivity to the effects of prolactin, or a lack of power. Secondly, dopamine – an important neuroregulator of working memory and cognitive control, amongst others – may impact cognitive function [44]. The dopaminergic tone in prolactinoma may be altered bidirectionally, with hyperprolactinemia suppressing dopamine due to dopaminergic neurons becoming refractory, and medical treatment increasing dopaminergic tone [45]. The current use of DAs did not clearly affect cognitive functioning in our cohort, as similar outcomes were observed in DA- and surgically treated patients. However, ongoing effects of prior DA-treatment in the surgically treated patients cannot be excluded. Furthermore, surgically and medically treated patients differed in some disease characteristics, which could not completely be accounted for statistically: surgically treated patients were generally DA-intolerant, and the DA-treated patients may have had irresectable/larger tumors. The role of hypopituitarism remains unclear. One study in patients with CD in remission reported that hypopituitarism was associated with more impairments in cognitive functioning [42]. By contrast, the present and previous studies observed no correlation between treated hypopituitarism and cognitive impairments, with previous studies indicating mild-to-no objectifiable deficits in treated primary and secondary adrenal insufficiency (46), and unconvincing effects of sex-hormone replacement on cognitive functioning in (mostly elderly) individuals [47, 48]. Lastly, it is well known that mood disorders can influence cognitive functioning [49], although this association was not observed in the present study, potentially due to exclusion of patients with diagnosed psychopathology. Taken together, underlying mechanisms of cognitive impairments require further analysis.

Multiplicity is inevitable when assessing cognition and psychopathology, as it involves assessment of multiple domains using separate testing instruments. Multiple testing corrections were not appropriate due to underlying association of the endpoints. Despite this being the largest prolactinoma population assessed for cognition and psychological complaints to date, power to detect subtle associations might be lacking. Nevertheless, clear trends were observed in patients with biochemically controlled prolactinoma.

The current study is a starting point for further, in-depth exploration of cognitive functioning. Future studies should include patients who underwent primary surgical treatment without DA pretreatment to examine DA effects, and longitudinal pre- and post-treatment testing to assess the degree of reversibility of cognitive impairments and psychological complaints in both surgically and medically treated patients. Additionally, functional, and structural brain MRI studies can provide insight into the course of (potentially persisting) cerebral alterations. Moreover, patients with prolactinoma should be compared to patients with other chronic (hormonal) conditions to elucidate prolactin- or dopamine-specific effects on the brain. Lastly, the added value of screening tools and cognitive rehabilitation programs should be formally evaluated.

In conclusion, patients with biochemically controlled prolactinoma demonstrated subtle cognitive impairments, pertaining to verbal memory, attention, and processing speed, compared to controls. Furthermore, patients reported more psychological complaints and maladaptive personality traits compared to controls. Physicians should be aware of these impairments and complaints and address these issues appropriately. Further research should encompass longitudinal assessments and the evaluation of the potential added value of cognitive rehabilitation and psychological support programs.

Table 3 Overview of studies reporting on cognitive functioning in patients with prolactinoma

Author (year)	Subjects	Controls	Design	Outcomes
Psaras (6) (2011)	Patients: N=106 patients with a pituitary adenoma (female: 69 (65%)), among which 12 patients with a prolactinoma (female: NR) Exclusion: Craniopharyngioma, metastasis, psychiatric or neurological disorders, major comorbidity known to affect neurocognitive functions (kidney or liver disease), insufficient hearing ability, uncorrected visual abnormalities	None	Longitudinal study: Measurement before and 3- and 12-months post-surgery Assessment: D2 Letter Cancellation test (selective attention), WAIS Digit Span Task (working memory), Trail-Making-test A (attentional speed), Intelligence Structure Test-2000 (episodic memory)	All: Improvement of concentration, working memory, and attentional speed within 3 months. Improvement of episodic memory after 12 months Prolactinoma: Removal of suprasellar extension was the most important factor for improvement of neurocognitive functions
Montalvo [7] (2018)	Patients: Microprolactinoma N=7 (female: 6 (86%)) Exclusion: Mental illness, prolactin-elevating drugs, substance use, mental retardation, dementia, other hyperprolactinemia causing diseases, pregnancy	None	Longitudinal study: Measurement before and 6-12 months after starting cabergoline Assessment: Brief Assessment of Cognition in Schizophrenia-Symbol Coding (speed of processing), Category Fluency-Animal Naming test (speed of processing), Trail Making Test Part A (selective attention), Continuous Performance Test-Identical Pairs (attention and vigilance), WMS-III Spatial Span (working memory), University of Maryland Letter-Number Span (working memory), Hopkins Verbal Learning Test-Revised (verbal learning), Brief Visuospatial Memory Test-Revised (visual learning) and Neuropsychological Assessment Battery-Mazes (reasoning and problem solving)	Improvement of processing speed, working memory, visual learning, reasoning and problem-solving after starting cabergoline
Yao [34] (2018)	Patients: Currently untreated prolactinoma, with hyperprolactinemia, ≥9 years of Education N=32 (female: 32 (100%)) Exclusion: Left-handed, in pubertal stage, significant VFD, history of neurological or psychiatric disorders, acquired brain injury, substance abuse, smoking >10 cigarettes/day, medication use (incl. DA and oral contraceptives)	Healthy volunteers matched for age, sex, education, and handedness N=26	Cross-sectional study Assessment: Structural MRI Wisconsin Card Sorting Test (executive functioning), Picture Recall Test, Visual Recognition Test and Story Recall tests (all assessing nonverbal or verbal memory)	Patients scored worse on verbal memory and executive functioning. Furthermore, patients showed a decreased grey matter volume of the left hippocampus, left orbitofrontal cortex, right middle frontal cortex and right interior frontal cortex. Impairments of verbal memory and executive functioning were associated with gray matter volume of the left hippocampus and right middle frontal cortex.

Table 3 Overview of studies reporting on cognitive functioning in patients with prolactinoma *(continued)*

Author (year)	Subjects	Controls	Design	Outcomes
Bala [31] (2022)	Patients: Treatment naïve prolactinoma N=27 (female: 15 (56%)) Exclusion: History of neurologic or psychiatric disorders, use of medication (incl. DA, oral contraceptives), pregnancy, substance abuse, significant VFD	Healthy, not pregnant, no medication. Matched based on sex, age, education, handedness, mood level N=27	Cross-sectional study Assessment: D2 Test of Attention and Color Trails Test (selective attention and concentration), Color Trails Test (spatial screening and working memory), Elevator Counting (sustained attention), Visual Elevator (attentional switching), Telephone Search and Telephone Search while Counting (visual screening and divided attention), Digit Span (auditory-verbal working memory), Symbol Span (spatial working memory)	Patients scored worse on selective attention, spatial screening and spatial working memory, auditory-verbal working memory, attentional switching, visual screening and divided attention. Higher prolactin levels were associated with lower scores in some working memory and attention assessments.
EEG studies				
Cao [33] (2020)	Patients: Currently untreated prolactinoma, 20-50 years old, >9 - ≤15 years of education. N=21 (female: 14 (67%) Exclusion: Radiotherapy, craniotomy, history of neurologic or psychiatric disorders, comorbidities that could impair cognitive function (severe kidney, liver or heart disease), coma, infections, epilepsy, hydrocephalus, CSF leakage, substance abuse, medication use (incl. oral contraceptives) insufficient hearing ability, uncorrected visual abnormalities	Healthy volunteers matched for age and education level N=21	Cross-sectional study Assessment: EEG, Visual Go/Nogo task	Impaired response activation and inhibition with lower frontal theta activity and occipital alpha activity in both Go and Nogo conditions in PRL compared to healthy controls. Mediator model suggested that the relationship between frontal theta power and inhibitory ability was mediated by increased prolactin levels. Negative relationship between prolactin levels and gray matter volume of left hippocampus and right inferior frontal cortex.

Table 3 Overview of studies reporting on cognitive functioning in patients with prolactinoma *(continued)*

Author (year)	Subjects	Controls	Design	Outcomes
Song [5] (2020)	Patients: Prolactinoma resistant or intolerant to long-term drug treatment Pre-surgery: N=20 (female: 10 (50%)) Six months post-surgery: N=20 (female: 10 (50%)) Exclusion: History of craniotomy or radiotherapy, neurologic or psychiatric disorders, comorbidity that can affect cognitive function (kidney, liver, or heart disease), severe complications (e.g., hydrocephalus, CSF leak), substance abuse, medication use (incl. oral contraceptives)	Healthy volunteers matched for age, gender, and education N=20	Cross-sectional study Assessment: EEG, Visual Go/Nogo task	Longer reaction times and lower accuracy in active PRL compared to post-surgical PRL and healthy controls. Generally, no difference between reaction times and accuracy in post-surgical PRL compared to healthy controls No correlation between prolactin levels and amplitude of P300 in both Go and Nogo conditions
Cao [32] (2023)	Patients: Currently untreated prolactinoma resistant to long-term medical therapy N=26 (female: 18 (69%) Exclusion: Radiotherapy, craniotomy, history of neurologic or psychiatric disorders, comorbidities that could impair cognitive function (severe kidney, liver, or heart disease), coma, infections, epilepsy, hydrocephalus, CSF leakage, substance abuse, medication use (incl. DA and oral contraceptives) insufficient hearing ability, uncorrected visual abnormalities	Healthy volunteers matched for age and education level N=26	Cross-sectional study Assessment: EEG, Color-shape Switching Task (task switching)	Patients showed longer reaction time in switch trials and larger switch costs. Weaker frontal theta activity and disrupted frontoparietal connectivity in patients compared to healthy controls. Higher prolactin levels were associated with larger decrease in cognitive performance

Data are presented as value (%) or mean ± standard deviation, unless stated otherwise. *CSF*, cerebrospinal fluid; *DA*, dopamine agonist; *EEG*, electroencephalogram; *NR*, not reported; *TSS*, transsphenoidal surgery; *VFD*, visual field defects.

REFERENCES

1. Petersenn S, Fleseriu M, Casanueva FF, Giustina A, Biermasz N, Biller BMK, et al. Diagnosis and management of prolactin-secreting pituitary adenomas: a Pituitary Society international Consensus Statement. Nat Rev Endocrinol. 2023;19(12):722-40.
2. Andela CD, Niemeijer ND, Scharloo M, Tiemensma J, Kanagasabapathy S, Pereira AM, et al. Towards a better quality of life (QoL) for patients with pituitary diseases: results from a focus group study exploring QoL. Pituitary. 2015;18(1):86-100.
3. Zamanipoor Najafabadi AH, Zandbergen IM, de Vries F, Broersen LHA, van den Akker-van Marle ME, Pereira AM, et al. Surgery as a Viable Alternative First-Line Treatment for Prolactinoma Patients. A Systematic Review and Meta-Analysis. J Clin Endocrinol Metab. 2020;105(3):e32-41.
4. Castle-Kirszbaum M, Biermasz N, Kam J, Goldschlager T. Quality of life in Prolactinoma: A systematic review. Pituitary. 2024;27(3):239-47.
5. Song J, Cao C, Wang Y, Yao S, Catalino MP, Yan D, et al. Response Activation and Inhibition in Patients With Prolactinomas: An Electrophysiological Study. Frontiers in Human Neuroscience. 2020;14.
6. Psaras T, Milian M, Hattermann V, Gerlach C, Honegger J. Executive functions recover earlier than episodic memory after microsurgical transsphenoidal resection of pituitary tumors in adult patients--a longitudinal study. J Clin Neurosci. 2011;18(10):1340-5.
7. Montalvo I, Llorens M, Caparrós L, Pamias M, Torralbas J, Giménez-Palop O, et al. Improvement in cognitive abilities following cabergoline treatment in patients with a prolactin-secreting pituitary adenoma. Int Clin Psychopharmacol. 2018;33(2):98-102.
8. Standaard Onderwijsindeling 2016, edition: 2019/'20: Centraal Bureau voor de Statistiek, Den Haag/Heerlen; [Available from: https://www.cbs.nl/nl-nl/onze-diensten/methoden/classificaties/onderwijs-en-beroepen/standaard-onderwijsindeling--soi--/standaard-onderwijsindeling-2016.
9. van Trigt VR, Bakker LEH, Pelsma ICM, Zandbergen IM, Jentus MM, Kruit MC, et al. The Changing Treatment Paradigm for Prolactinoma—A Prospective Series of 100 Consecutive Neurosurgical Cases. The Journal of Clinical Endocrinology & Metabolism. 2024.
10. van Trigt VR, Zandbergen IM, Pelsma ICM, Bakker LEH, Verstegen MJT, van Furth WR, et al. Care trajectories of surgically treated patients with a prolactinoma: why did they opt for surgery? Pituitary. 2023;26(5):611-21.
11. van Furth WR, de Vries F, Lobatto DJ, Kleijwegt MC, Schutte PJ, Pereira AM, et al. Endoscopic Surgery for Pituitary Tumors. Endocrinol Metab Clin North Am. 2020;49(3):487-503.
12. Fleseriu M, Hashim IA, Karavitaki N, Melmed S, Murad MH, Salvatori R, et al. Hormonal Replacement in Hypopituitarism in Adults: An Endocrine Society Clinical Practice Guideline. The Journal of Clinical Endocrinology & Metabolism. 2016;101(11):3888-921.
13. van Trigt VR, Andela CD, Bakker EHL, et al. Supplementary Tables and Figures belonging to Subtle cognitive impairments and psychological complaints inpatients with prolactinoma despite biochemical control (2025). figshare. Journal Contribution. 2025. Doi: 10.6084/m9.figshare.29448281.v1
14. Rey A. L'examen clinique en psychologie. [The clinical examination in psychology.]. Oxford, England: Presses Universitaries De France; 1958. 222- p.
15. Wechsler D. Wechsler Adult Intelligence Scale - Fourth Edition: Technical and Interpretive Manual. San Antonio: Pearson. 2008.
16. Rey A. L'examen psychologique dans les cas d'encéphalopathie traumatique. (Les problems.). [The psychological examination in cases of traumatic encepholopathy. Problems.]. Archives de Psychologie. 1941;28:215-85.
17. van der Elst W, van Boxtel MP, van Breukelen GJ, Jolles J. The Letter Digit Substitution Test: normative data for 1,858 healthy participants aged 24-81 from the Maastricht Aging Study (MAAS): influence of age, education, and sex. J Clin Exp Neuropsychol. 2006;28(6):998-1009.
18. Dekker RMJ, Dekker PH. De ontwikkeling van vijf nieuwe Nederlandstalige tests. PITS: Leiden. 2007.
19. Reitan RM. Trail Making Test: Manual for administration and scoring: Reitan Neuropsychology Laboratory; 1986.
20. Delis DKE, J. K. Delis-Kaplan Executive Function System (D-KEFS). San Antonio: The Psychological Corporation. 2001.
21. Benton AL, Hamsher KD. Multilingual Aphasia Examination. Iowa City: University of Iowa. 1976.
22. Andela CD, Scharloo M, Ramondt S, Tiemensma J, Husson O, Llahana S, et al. The development and validation of the Leiden Bother and Needs Questionnaire for patients with pituitary disease: the LBNQ-Pituitary. Pituitary. 2016;19(3):293-302.
23. Sergio E. Starkstein, M.D., Ph.D. ,, Gustavo Petracca, M.D. ,, Erán Chemerinski, M.D. , and, Janus Kremer, M.D. Syndromic Validity of Apathy in Alzheimer's Disease. American Journal of Psychiatry. 2001;158(6):872-7.

24. Anjan Chatterjee, M.D., MPH ,, Karen E. Anderson, M.D., M.S. ,, Carol B. Moskowitz, M.S. ,, Willard A. Hauser, M.D. , and, Karen S. Marder, M.D., MPH. A Comparison of Self-Report and Caregiver Assessment of Depression, Apathy, and Irritability in Huntington's Disease. The Journal of Neuropsychiatry and Clinical Neurosciences. 2005;17(3):378-83.
25. Fisk JD, Ritvo PG, Ross L, Haase DA, Marrie TJ, Schlech WF. Measuring the functional impact of fatigue: initial validation of the fatigue impact scale. Clin Infect Dis. 1994;18 Suppl 1:S79-83.
26. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand. 1983;67(6):361-70.
27. Michopoulos I, Douzenis A, Kalkavoura C, Christodoulou C, Michalopoulou P, Kalemi G, et al. Hospital Anxiety and Depression Scale (HADS): validation in a Greek general hospital sample. Ann Gen Psychiatry. 2008;7:4.
28. Koopmans PCSR, Timmerman I, Emmelkamp PM. The irrational Beliefs Inventory (IBI): development and psychometric evaluation. European Journal of Psychological Assessment. 1994;10:15-27.
29. Barelds DPH, Luteijn F. NKPV NEderlands Klinische Persoonlijkheidsvragenlijst. Amsterdam: Boom Test Uitgevers. 2015.
30. Zandbergen IM, Zamanipoor Najafabadi AH, Pelsma ICM, van den Akker-van Marle ME, Bisschop P, Boogaarts HDJ, et al. The PRolaCT studies - a study protocol for a combined randomised clinical trial and observational cohort study design in prolactinoma. Trials. 2021;22(1):653.
31. Schulz KF, Grimes DA. Multiplicity in randomised trials I: endpoints and treatments. Lancet. 2005;365(9470):1591-5.
32. Bala A, Dziedzic T, Olejnik A, Marchel A. Attention and working memory in patients with prolactinomas: a case-control study. Sci Rep. 2022;12(1):22565.
33. Cao C, Wen W, Chen A, Wang S, Xu G, Niu C, et al. Neuropsychological Alterations of Prolactinomas' Cognitive Flexibility in Task Switching. Brain Sci. 2023;13(1).
34. Cao C, Wen W, Liu B, Ma P, Li S, Xu G, et al. Theta oscillations in prolactinomas: Neurocognitive deficits in executive controls. Neuroimage Clin. 2020;28:102455.
35. Yao S, Song J, Gao J, Lin P, Yang M, Zahid KR, et al. Cognitive Function and Serum Hormone Levels Are Associated with Gray Matter Volume Decline in Female Patients with Prolactinomas. Front Neurol. 2017;8:742.
36. Tiemensma J, Kokshoorn NE, Biermasz NR, Keijser BJ, Wassenaar MJ, Middelkoop HA, et al. Subtle cognitive impairments in patients with long-term cure of Cushing's disease. J Clin Endocrinol Metab. 2010;95(6):2699-714.
37. Tiemensma J, Biermasz NR, van der Mast RC, Wassenaar MJ, Middelkoop HA, Pereira AM, et al. Increased psychopathology and maladaptive personality traits, but normal cognitive functioning, in patients after long-term cure of acromegaly. J Clin Endocrinol Metab. 2010;95(12):E392-402.
38. Athanasoulia AP, Ising M, Pfister H, Mantzoros CS, Stalla GK, Sievers C. Distinct dopaminergic personality patterns in patients with prolactinomas: a comparison with nonfunctioning pituitary adenoma patients and age- and gender-matched controls. Neuroendocrinology. 2012;96(3):204-11.
39. Ernersson Å, Bachrack-Lindström M, Landberg E, Ekman B, Wahlberg J. Reduced Health Related Quality of Life, Increased Fatigue, and Daytime Sleepiness in Women with Hyperprolactinemia. Horm Metab Res. 2023;55(4):266-72.
40. Buckman MT, Kellner R. Reduction of distress in hyperprolactinemia with bromocriptine. Am J Psychiatry. 1985;142(2):242-4.
41. Lobatto DJ, Zamanipoor Najafabadi AH, de Vries F, Andela CD, van den Hout WB, Pereira AM, et al. Toward Value Based Health Care in pituitary surgery: application of a comprehensive outcome set in perioperative care. Eur J Endocrinol. 2019;181(4):375-87.
42. Tiemensma J, Biermasz NR, Middelkoop HA, van der Mast RC, Romijn JA, Pereira AM. Increased prevalence of psychopathology and maladaptive personality traits after long-term cure of Cushing's disease. J Clin Endocrinol Metab. 2010;95(10):E129-41.
43. Andela CD, Repping-Wuts H, Stikkelbroeck N, Pronk MC, Tiemensma J, Hermus AR, et al. Enhanced self-efficacy after a self-management programme in pituitary disease: a randomized controlled trial. Eur J Endocrinol. 2017;177(1):59-72.
44. Cools R, D'Esposito M. Inverted-U-shaped dopamine actions on human working memory and cognitive control. Biol Psychiatry. 2011;69(12):e113-25.
45. Ben-Jonathan N, Hnasko R. Dopamine as a prolactin (PRL) inhibitor. Endocr Rev. 2001;22(6):724-63.
46. Harbeck B, Kropp P, Kreitschmann-Andermahr I. Glucocorticoid replacement therapy for primary and secondary adrenal insufficiency and their impact on cognition. Front Endocrinol (Lausanne). 2023;14:1153595.
47. Resnick SM, Matsumoto AM, Stephens-Shields AJ, Ellenberg SS, Gill TM, Shumaker SA, et al. Testosterone Treatment and Cognitive Function in Older Men With Low Testosterone and Age-Associated Memory Impairment. Jama. 2017;317(7):717-27.
48. Andy C, Nerattini M, Jett S, Carlton C, Zarate C, Boneu C, et al. Systematic review and meta-analysis of the effects of menopause hormone therapy on cognition. Front Endocrinol (Lausanne). 2024;15:1350318.
49. Semkovska M, Quinlivan L, O'Grady T, Johnson R, Collins A, O'Connor J, et al. Cognitive function following a major depressive episode: a systematic review and meta-analysis. The Lancet Psychiatry. 2019;6(10):851-61.

SUPPLEMENTARY MATERIALS

Supplementary Table 1 Description of cognitive assessments

Cognitive assessment	Cognitive domains	Procedure
The Verbal Learning Test of Rey [1]	Verbal memory	Fifteen words depicted on cards are shown one by one and simultaneously read aloud by the investigator. The participant is asked to remember and reproduce the words. The task is repeated four times (immediate reproduction) and after 15-20 min (delayed reproduction).
Weschler Adult Intelligence Score (WAIS) Digit Span Test [2]	Verbal memory, working memory	Sequences of numbers read aloud by researcher: 1. Forward: participant recalls in the same order 2. Backward: participant recalls in reverse order 3. Sequencing: participant recalls from low to high
Rey Complex Figure Test [3]	Visuospatial memory, visuospatial construction	1. The participant is asked to copy a complex figure, using a pencil and paper 2. Immediate reproduction from memory (after 3 min) 3. Delayed reproduction from memory (after 30 min)
WAIS Digit-Symbol Substitution Test [4]	Selective attention, processing speed	The participant is asked to substitute as many numbers with indicated symbols as possible within 2 minutes.
Digit Deletion Test [5]	Selective attention, processing speed	The participant is presented with a form containing 800 numbers (1-9), in which they are asked to cross out numbers 3 and 7 diagonally and underline number 4. Time limit: 3 minutes.
Trail Making Test [6]	Task A: selective attention. Task B: cognitive flexibility	The participant is asked to connect as fast and accurately as possible: Task A: circles with digits (1-25) in ascending order Task B: circles with digits (1-13) and letters (a-l) alternatingly in the right sequence. In case of mistakes the participant has to correct the mistake before proceeding.
D-KEFS Tower Test [7]	Executive functioning, (including inhibition of impulsive responses)	The participants is asked to replicate nine towers by stacking five differently sized disks onto three wooden pegs, never stacking larger on top of smaller disks and only moving one disk at a time, only using one hand.
FAS test [8]	Executive functioning (including verbal fluency)	The participant is asked to orally produce as many words as possible starting with an F, A and S, respectively, in 60 seconds. Words have to be official Dutch words, and proper nouns were not accepted.

Results	Interpretation and scoring details
For immediate (1-4) and delayed reproductions: 1. Number of correctly reproduced words (0-15) 2. Number of repeated words 3. Number of incorrect words (intrusions) This study only reports on immediate reproduction rounds 1, 2, 4 and the delayed reproduction, as these provide the most relevant information	More correctly reproduced words and fewer repeated and incorrect words indicate better verbal memory. If the participant repeats a word by "thinking aloud", this was not scored as a repeat. If the researcher is not sure if the participant wants to state the word (again), the participant is asked if the word should be noted or not.
1. Number of correct sequences in forward test (0-16) 2. Number of correct sequences in backward test (0-16) 3. Number of correct sequences in sequencing test (0-16)	Higher scores indicate better verbal memory and working memory. Self-corrections are allowed and not scored as mistakes.
18 items are scored per round. For each item: 2 points: accurate and correct location 1 point: accurate or correct location ½ point: recognizable, yet inaccurate, and incorrect location 0 points: neither accurate nor correct location 1. Total points copying figure (0-36) 2. Total points immediate reproduction (0-36) 3. Total points delayed reproduction (0-36)	Higher scores indicate better visuospatial memory and visuospatial construction In case of doubt about scoring, an agreement is made based on discussion and consensus with the second (and in case of persistent disagreement) a third corrector.
1. Total number of correct symbols (0-135) 2. Total number of incorrect symbols (0-135)	More correctly substituted symbols and fewer mistakes indicate better attention and processing speed. If unclear whether a symbol is drawn correctly, an agreement is made amongst the first, second (and third) correctors on how to score that particular deviation to ensure all participants are scored consistently.
1. Total number of correctly edited numbers (0-240) 2. Total number of incorrectly edited numbers (0-800) 3. Total number of missed numbers (0-240)	More correctly edited numbers and fewer incorrectly edited and missed numbers indicate better selective attention and processing speed. If unclear whether a digit is crossed out or underlined, agreements are made amongst the first, second (and third) correctors on how to score that particular deviation to ensure all participants were scored consistently.
For Task A and Task B, respectively: 1. Completion time (no time cap) 2. Number of mistakes	Less time and fewer mistakes indicate better selective attention (Task A) and cognitive flexibility and devided attention (Task B). Self-corrections are allowed and are not scored as mistakes if corrected before correction by the investigator.
1. Total performance score: score for number of steps needed to correctly complete a tower 2. Average time for first step 3. Time-per-step-ratio 4. Step-accuracy-ratio 5. Total number rule violations 6. Rule-violation-per-item-ratio	Each subscore is standardized for age using reference values provided by the manufacturer, with the population mean being 10 points. Higher scores indicate better executive functioning. A step is only scored when the participant lifts a disc from a peg, and released it completely on a different peg. A rule violation is not counted as a step, as errors have to be undone.
For each letter: 1. Total number of correct words 2. Total number of repeated words 3. Total number of incorrect words	More correct words and fewer repeated or incorrect words indicate better executive functioning. Only officially existing Dutch words are accepted. In case of doubt, a dictionary was consulted. Proper nouns are scored as incorrect, as explained to the participant prior to starting the task.

Supplementary Table 2 Description of used patient-reported outcome measures

PROM	Description	Scales
Leiden Bother and Needs Pituitary Questionnaire [9]	33 items covering 5 subscales: - Mood problems (9 items) - Sexual functioning (4 items) - Negative illness perceptions (6 items) - Physical and cognitive complaints (9 items) - Social functioning issues (5 items)	Bothers scoring: 5-point scale: Not at all = 0 Extremely = 4 Needs scoring: 5-point Likert scale: Not important = 0 Extremely important = 4
Apathy Scale (AS) [10]	14 items measuring the degree of apathy during previous two weeks	4-point scale: Items 1-8: Strongly present = 0 Not at all present = 3 Items 9-14: Not at all present = 0 Strongly present = 3
Fatigue Severity Scale (FSS) [11]	9 items measuring the influence of fatigue on cognitive, psychosocial and physical functioning	7-point scale: Totally disagree = 1 Totally agree = 7
Hospital Anxiety and Depression Scale (HADS) [12]	14 items measuring feelings of anxiety or depression in the past 4 weeks	4-point scale: Scales vary per question
Irritability Scale (IPS) [13]	14 items measuring irritability	4-point scale: Not at all = 0 Often = 3
Irrational Beliefs Inventory-50 (IBI-50) [14]	50 items measuring 5 subscales: - Avoidance (10 items) - Rigidity (14 items) - Worrying (12 items) - Need for approval (7 items) - External control (7 items)	5-point scale: Strongly disagree = 1 Strongly agree = 5
Dutch Clinical Personality Questionnaire [15]	120 items measuring 6 subscales, each containing 20 items: - Negativism - Somatization - Shyness - Severe psychopathology - Extraversion - Narcissism	3 options: ? (uncertain) = 1 Correct = 0 or 2 (depending on the question) Incorrect = 0 or 2 (depending on the question)

PROM patient-reported outcome measure.

Results	Scoring	*Interpretation*
Bothers and Needs, respectively, concerning: - Mood problems (0-100) - Sexual functioning (0-100) - Negative illness perceptions (0-100) - Physical and cognitive complaints (0-100) - Social functioning issues (0-100) - Total score (0-100)	Total Bothers: (sum of all bothers/ 132) x 100 Total Needs: (sum of all needs/ 132) x 100	Higher scores indicate more bothers by the complaint or higher need for attention by the healthcare provider, respectively.
Total AS score (0-42)	Total AS score: sum of all 14 items	<14: low apathy score ≥ 14: high apathy score
Total FSS score (9-63)	Total FSS score: sum of all 9 items	Higher scores indicate a larger impact of fatigue on daily functioning.
Total anxiety score (0-21) Total depression score (0-21) Total HADS score (0-42)	Total anxiety score: sum of all oddly numbered items Total depression score: sum of all evenly numbered items Total HADS score: sum of all 14 items	For the subscales: 0-7: no anxiety or depression, respectively 8-10: indication for clinically relevant anxiety or depression, respectively.
Total IPS score (0-42)	Total IPS score: sum of all items	<14: low irritability score ≥ 14: high irritability score
- Avoidance (10-50) - Rigidity (14-70) - Worrying (12-60) - Need for approval (7-35) - External control (7-35) - Total IBI-50 score (50-250)	Score per subscale: sum of all items belonging to that subscale Total IBI-50 score: sum of all 50 items	Higher scores indicate a higher degree of irrational thinking.
- Negativism (0-40) - Somatization (0-40) - Shyness (0-40) - Severe psychopathology (0-40) - Extraversion (0-40) - Narcissism (0-40)	Score per subscale: sum of all items belonging to the subscale	Higher scores on a subscale indicate more emphasis on the personality trait.

Supplementary Table 3 Results of the cognitive assessment for biochemically controlled patients and matched controls

Cognitive test		Patients N=60	Matched healthy controls N=60	Coefficient β	95% confidence interval	P-value
Verbal Learning Test of Rey[a]	1: correct	6.8±2.0	7.5±2.2	-0.7	-1.4, 0.0	0.038*
	1: intrusions	0.3±0.5	0.4±0.7	-0.2	-0.3, 0.0	0.069
	1: repetitions	0.4±0.6	0.6±1.5	0.0	-0.2,0.2	0.745
	2: correct	9.0±2.9	10.2±2.5	-1.2	-2.1, -0.3	0.010*
	2: intrusions	0.2±0.6	0.3±0.8	-0.2	-0.4, 0.1	0.163
	2: repetitions	1.2±1.8	1.2±1.8	-0.1	-0.7, 0.6	0.883
	4: correct	11.4±2.4	12.6±2.1	-1.2	-1.9, -0.5	<0.001*
	4: intrusions	0.1±0.3	0.1±0.3	0.0	-0.1, 0.1	0.739
	4: repetitions	1.3±1.3	1.7±1.7	-0.4	-0.9, 0.2	0.201
	Delayed: correct	9.8±3.1	11.6±2.5	-1.8	-2.7, -1.0	<0.001*
	Delayed: intrusions	0.2±0.7	0.3±0.6	0.0	-0.2, 0.2	0.752
	Delayed: repetitions	0.8±1.0	1.3±1.9	-0.5	-1.1, 0.0	0.054
WAIS Digit Span Task	Forward	8.6±1.7	9.1±1.7	-0.5	-1.0, 0.0	0.074
	Backward	8.6±1.5	8.8±1.4	-0.3	-0.8, 0.2	0.281
	Sequencing	8.3±1.8	8.9±2.3	-0.6	-1.3, 0.2	0.122
Rey Complex Figure Test	Copying	34.8±1.4	35.0±1.4	-0.2	-0.7, 0.3	0.408
	Immediate recall	22.6±5.3	22.6±5.4	-0.1	-1.9, 1.8	0.941
	Delayed recall	22.3±5.2	22.2±5.4	0.0	-1.9, 1.9	0.983
WAIS Digit-Symbol Substitution Test	Correct	74.2±13.1	78.4±16.1	-4.2	-8.2, -0.2	0.040*
	Incorrect	0.1±0.4	0.3±0.7	-0.2	-0.4, 0.0	0.846
Digit Deletion Test	Correct	119.6±22.8	128.4±26.5	-8.8	-16.2, -1.4	0.019*
	Incorrect	0.1±0.3	0.0±0.1	0.1	0.0, 0.1	0.173
	Missed	5.2±4.6	4.2±3.8	1.0	-0.5, 2.4	0.185
Trail Making Test	A: Time (s)	27.2±8.4	22.0±5.7	5.2	3.2, 7.2	<0.001*
	A: Mistakes	0.2±0.4	0.0±0.2	0.1	0.0, 0.2	0.028*
	B: Time (s)	58.7±21.9	51.5±15.3	4.7	-1.5, 10.9	0.135
	B: Mistakes	0.4±0.8	0.3±0.5	0.2	-0.1, 0.4	0.202
D-KEFS Tower Test	Performance score[b]	11.2±2.4	11.5±2.3	-0.3	-1.2, 0.6	0.497
	Average time to first step[b]	10.8±2.7	10.4±2.6	0.4	-0.6, 1.3	0.441
	Time-per-step-ratio[b]	10.1±2.4	10.1±2.2	0.0	-0.8, 0.7	0.932
	Step-accuracy-ratio[b]	9.0±2.5	9.0±2.6	0.0	-0.9, 0.9	1.000
	Rule-violations-per-item-ratio[b]	10.4±1.4	10.5±0.6	0.0	-0.4, 0.3	0.924
FAS	Correct, total	35.6±12.2	34.8±11.9	0.8	-3.2, 4.8	0.687
	Incorrect, total	1.1±1.5	1.0±1.5	0.0	-0.5, 0.6	0.906
	Repetitions, total	0.6±1.0	0.4±0.8	0.2	-0.1, 0.5	0.270

Data reported as value (%) or mean ± standard deviation. * Indicate significant differences (p<0.050). *S* seconds.

[a] The first, second and fourth round of immediate reproduction, and delayed round of reproduction were shown as these provide the most relevant information.

[b] Standardized scores are reported, with higher scores indicating better performance and score 10 being the age-stratified population mean.

Supplementary Table 4 Patient-reported psychological complaints and maladaptive personality traits for biochemically controlled patients and matched controls

Patient-reported outcome measure		Patients N=60	Matched healthy controls N=60	Coefficient β	95% confidence interval	P-value
Apathy Scale[a]	Total	12.8±6.0	10.4±3.7	2.4	0.6, 4.1	0.009*
	Score ≥ 14, n (%)	27 (45.8)	13 (21.7)			
Fatigue Severity Scale[a]	Total	29.5±13.2	22.8±9.7	6.7	2.7, 10.8	<0.001*
HADS[a]	Anxiety score	5.4±3.8	4.3±2.2	1.1	0.1, 2.1	0.034*
	Anxiety ≥8, n (%)	15 (25.4)	5 (8.3)			
	Depression score	3.6±3.5	1.9±1.8	1.7	0.8, 2.7	<0.001*
	Depression ≥8, n (%)	5 (8.3)	0			
	Total HADS score	9.0±6.4	6.1±3.4			
Irritability Scale[a]	Total	9.3±7.0	7.1±3.7	2.2	0.3, 4.1	0.024*
	Score ≥14, n (%)	15 (25.4)	2 (3.3)			
Irrational Beliefs Inventory[a]	Total	129.8±17.3	133.9±14.8	-4.2	-9.5, 1.1	0.123
Dutch Clinical Personality Questionnaire	Negativism	9.5±7.3	6.7±4.7	2.8	0.8, 4.9	0.007*
	Somatization	12.4±8.2	5.1±3.9	7.3	5.0, 9.5	<0.001*
	Shyness	11.4±8.8	9.5±7.9	1.9	-0.5, 4.4	0.118
	Severe psychopathology	3.8±4.4	1.9±2.2	1.9	0.8, 3.1	0.001*
	Extraversion	17.6±8.7	21.3±7.9	-3.7	-6.7, -0.7	0.015*
	Narcissism	15.3±7.9	15.9±7.0	-0.6	-3.2, 2.1	0.682

Data reported as value (%) or mean ± standard deviation. * Indicate significant differences (p<0.050). *HADS* Hospital Anxiety and Depression Scale.

[a] Data missing for one surgically treated patient, who did not complete all patient-reported outcome measures (n=59).

Supplementary Table 5 Results of the cognitive assessment for patients controlled on medication and patients in surgical remission and matched healthy controls

Cognitive test		Patients controlled on DA N=30	Matched healthy controls (DA) N=30	Patients in surgical remission N=30	Matched healthy controls (surgery) N=30
Verbal Learning Test of Rey[a]	Attempt 1: correct	6.9±1.8	7.7±2.4	6.7±2.3	7.3±2.0
	Attempt 1: intrusions	0.3±0.5	0.4±0.6	0.3±0.5	0.5±0.7
	Attempt 1: repetitions	0.5±0.6	0.3±0.5	0.4±0.6	0.4±0.6
	Attempt 2: correct	9.2±2.7	10.1±2.7	8.9±3.1	10.3±2.2
	Attempt 2: intrusions	0.3±0.7	0.3±0.7	0.1±0.4	0.4±0.8
	Attempt 2: repetitions	1.3±1.7	1.0±1.5	1.0±1.8	1.4±2.1
	Attempt 4: correct	11.2±1.9	12.1±2.3	11.6±2.8	13.0±1.8
	Attempt 4: intrusions	0.1±0.3	0.1±0.3	0.1±0.3	0.1±0.3
	Attempt 4: repetitions	1.4±1.3	1.4±1.9	1.1±1.4	1.9±1.6
	Delayed: correct	9.4±3.1	10.9±2.7	10.1±3.0	12.3±2.1
	Delayed: intrusions	0.2±0.6	0.3±0.7	0.2±0.8	0.2±0.6
	Delayed: repetitions	1.0±1.1	1.1±1.1	0.6±0.9	1.6±2.4
WAIS Digit Span Task	Forward	8.6±1.6	9.1±1.7	8.6±1.7	9.1±1.6
	Backward	8.4±1.5	8.7±1.2	8.7±1.5	9.0±1.6
	Sequencing	8.2±1.6	8.5±2.3	8.4±2.0	9.2±2.3
Rey Complex Figure Test	Copying	34.9±1.2	35.3±0.8	34.7±1.6	34.7±1.8
	Immediate recall	21.2±5.1	22.0±5.8	24.0±5.1	23.3±5.0
	Delayed recall	20.8±5.3	21.6±6.0	23.6±4.8	22.9±4.8
WAIS Digit-Symbol Substitution Test	Correct	74.0±13.9	76.9±13.4	74.4±12.5	79.8±18.5
	Incorrect	0.1±0.3	0.4±0.9	0.2±0.5	0.2±0.5
Digit Deletion Test	Correct	119.6±22.6	122.8±25.9	119.5±23.3	133.9±26.4
	Incorrect	0.1±0.3	0.0±0.0	0.0±0.0	0.0±0.2
	Missed	4.9±3.4	4.6±4.2	5.6±5.6	3.8±3.4
Trail Making Test	A: Time	27.8±8.3	23.0±6.2	26.5±8.6	20.9±5.1
	A: Mistakes	0.2±0.4	0.0±0.0	0.1±0.3	0.1±0.3
	B: Time	60.8±21.7	56.9±15.3	56.7±22.4	51.1±15.1
	B: Mistakes	0.4±0.7	0.3±0.5	0.4±0.9	0.2±0.5

Supplementary Table 5 Results of the cognitive assessment for patients controlled on medication and patients in surgical remission and matched healthy controls *(continued)*

Cognitive test		Patients controlled on DA N=30	Matched healthy controls (DA) N=30	Patients in surgical remission N=30	Matched healthy controls (surgery) N=30
D-KEFS Tower Test	Performance score[b]	11.5±2.2	11.8±2.1	10.9±2.5	11.2±2.6
	Average time to first step[b]	10.8±2.8	10.9±2.3	10.8±2.6	9.9±2.8
	Time-per-step-ratio[b]	10.4±2.0	10.7±1.7	9.7±2.6	9.5±2.5
	Step-accuracy-ratio[b]	9.3±2.7	8.6±2.9	8.7±2.4	9.3±2.2
	Rule-violations-per-item-ratio[b]	10.4±1.8	10.5±0.6	10.5±0.6	10.4±0.6
FAS	Correct, total	36.7±11.5	32.1±10.9	34.5±12.9	37.5±12.4
	Incorrect, total	1.1±1.7	0.8±1.0	1.1±1.3	1.2±1.9
	Repetitions, total	0.6±0.7	0.5±0.8	0.6±1.2	0.33±0.8

Data are reported as mean ± standard deviation.

[a] The first, second and fourth round of immediate reproduction, and delayed round of reproduction were shown as these provide the most relevant information.

[b] Standardized scores are reported, with higher scores indicating better performance and score 10 being the age-stratified population mean.

Supplementary Table 6 Results of the cognitive assessment for patients controlled on medication and patients in surgical remission

Cognitive test		Patients controlled on DA N=30	Patients in surgical remission N=30	Coefficient β	95% confidence interval	P-value
Verbal Learning Test of Rey[a]	Attempt 1: correct	-0.3	-0.3	0.0	-0.5, 0.5	0.907
	Attempt 1: intrusions	-0.2	-0.3	0.1	-0.3, 0.5	0.747
	Attempt 1: repetitions	0.2	-0.1	0.4	-0.1, 0.8	0.154
	Attempt 2: correct	-0.3	-0.6	0.3	-0.3, 0.9	0.304
	Attempt 2: intrusions	-0.1	-0.3	0.3	-0.1, 0.7	0.148
	Attempt 2: repetitions	0.2	-0.2	0.4	-0.1, 0.9	0.139
	Attempt 4: correct	-0.4	-0.8	0.4	-0.2, 1.0	0.196
	Attempt 4: intrusions	-0.1	0.0	-0.1	-0.6, 0.4	0.717
	Attempt 4: repetitions	-0.1	-0.5	0.5	0.1, 0.9	0.014*
	Delayed: correct	-0.5	-1.1	0.5	-0.1, 1.2	0.111
	Delayed: intrusions	-0.1	0.0	-0.1	-0.7, 0.5	0.735
	Delayed: repetitions	-0.1	-0.4	0.3	-0.1, 0.7	0.104
WAIS Digit Span Task	Forward	-0.3	-0.3	0.0	-0.5, 0.5	0.971
	Backward	-0.2	-0.2	0.0	-0.6, 0.6	0.974
	Sequencing	-0.2	-0.4	0.2	-0.2, 0.6	0.350
Rey Complex Figure Test	Copying	-0.5	0.0	-0.5	-1.1, 0.1	0.089
	Immediate recall	-0.1	0.1	-0.3	-0.8, 0.2	0.240
	Delayed recall	-0.1	0.2	-0.3	-0.8, 0.2	0.220
WAIS Digit-Symbol Substitution Test	Correct	-0.2	-0.3	0.1	-0.4, 0.5	0.742
	Incorrect	-0.3	-0.1	-0.2	-0.5, 0.2	0.344
Digit Deletion Test	Correct	-0.1	-0.5	0.4	0.0, 0.9	0.059
	Incorrect[b]	-	0.5	-	-	-
	Missed	0.1	0.5	-0.5	-1.1, 0.2	0.162
Trail Making Test	A: time	0.8	1.1	-0.3	-1.1, 0.4	0.414
	A: Mistakes[b]	-	0.3	-	-	-
	B: Time	0.3	0.4	-0.1	-0.8, 0.6	0.754
	B: Mistakes	0.1	0.5	-0.3	-1.1, 0.5	0.408
D-KEFS Tower Test	Performance score[c]	-0.1	-0.1	0.0	-0.5, 0.5	0.920
	Average time to first step[c]	-0.1	0.3	-0.4	-0.9, 0.1	0.147
	Time-per-step-ratio[c]	-0.2	0.1	-0.2	-0.8, 0.3	0.402
	Step-accuracy-ratio[c]	0.2	-0.3	0.5	0.0, 1.0	0.046*
	Rule-violations-per-item-ratio[c]	-0.2	0.1	-0.4	-1.5, 0.7	0.486
FAS	Correct, total	0.4	-0.2	0.7	0.2, 1.2	0.012*
	Incorrect, total	0.2	-0.1	0.3	-0.3, 1.0	0.333
	Repetitions, total	0.1	0.3	-0.2	-0.7, 0.4	0.590

Data are reported as mean Z-scores. *CI* 95% confidence interval. * Indicate significant differences (p<0.050).

[a] The first, second and fourth round of immediate reproduction, and delayed round of reproduction were shown as these provide the most relevant information.

[b] Z-scores could not be calculated as the mean and standard deviation were zero for matched controls.

[c] Standardized scores are reported, with higher scores indicating better performance and score 10 being the age-stratified population mean.

Supplementary Table 7 Patient-reported psychological complaints and maladaptive personality traits for patients controlled on medication and patients in surgical remission and matched controls

Patient-reported outcome measure		Patients controlled on DA N=30	Matched healthy controls (DA) N=30	Patients in surgical remission N=30	Matched controls (surgery) N=30
Apathy Scale [a]	**Total**	**12.6±6.0**	10.1±3.3	12.9±6.1	10.8±4.0
	Score ≥ 14, n (%)	12 (40.0)	6 (20.0)	15 (51.7)	7 (23.3)
Fatigue Severity Scale [a]	Total	29.7±13.7	22.2±7.7	29.4±12.9	23.4±11.4
HADS [a]	Anxiety	5.2±4.1	4.1±2.1	5.5±3.6	4.4±2.3
	Anxiety ≥8, n (%)	9 (30.0)	3 (10.0)	6 (20.7)	3 (10.0)
	Depression	3.9±3.9	2.2±2.0	3.3±3.1	1.5±2.3
	Depression ≥8, n (%)	3 (10.0)	0	2 (6.9)	0
Irritability Scale [a]	Total	8.9±8.0	6.8±3.2	9.7±6.0	7.4±4.2
	Score ≥14, n (%)	6 (20.0)	1 (3.3)	9 (31.0)	1 (3.3)
Irrational Beliefs Inventory [a]	Total	126.5±18.6	130.7±128.5	113.1±15.5	137.2±15.4
Dutch Clinical Personality Questionnaire	Negativism	9.1±7.8	6.2±4.6	9.9±6.8	7.1±4.9
	Somatization	11.4±7.6	4.7±6.5	13.4±8.8	5.5±4.2
	Shyness	12.1±9.5	10.0±8.3	10.7±8.3	8.9±7.7
	Severe psychopathology	3.7±4.7	2.0±2.2	4.0±4.1	1.8±2.2
	Extraversion	16.6±8.7	20.2±8.5	18.6±8.7	22.4±7.3
	Narcissism	15.3±8.7	15.1±6.1	15.3±7.2	16.6±7.8

Data reported as value (%), or mean ± standard deviation. *HADS* Hospital Anxiety and Depression Scale.
[a] Data missing for one surgically treated patient, who did not complete all questionnaires (n=29).

Supplementary Table 8 Patient-reported psychological complaints and maladaptive personality traits for patients controlled on medication and patients in surgical remission

Patient-reported outcome measure		Patients on controlled on DA N=30	Patients in surgical remission N=30	Coefficient β	95% confidence interval	P-value
Apathy Scale[a]	Total	0.8	0.4	0.3	-0.6, 1.1	0.560
Fatigue Severity Scale[a]	Total	1.0	0.5	0.5	-0.3, 1.2	0.224
HADS[a]	Anxiety	0.5	0.5	0.0	-0.9, 0.9	0.958
	Depression	0.9	1.1	-0.2	-1.2, 0.7	0.630
Irritability Scale[a]	Total	0.6	0.4	0.1	-0.9, 1.1	0.862
Irrational Beliefs Inventory[a]	Total	-0.3	-0.3	0.0	-0.6, 0.6	0.905
Dutch Clinical Personality Questionnaire	Negativism	0.6	0.5	0.0	-0.7, 0.8	0.927
	Somatization	1.9	1.5	0.3	-0.8, 1.3	0.645
	Shyness	0.2	0.2	0.0	-0.5, 0.6	0.955
	Severe psychopathology	0.8	1.0	-0.3	-1.3, 0.8	0.628
	Extraversion	-0.4	-0.5	0.1	-0.5, 0.6	0.760
	Narcissism	0.0	-0.2	0.2	-0.4, 0.8	0.493

Data are reported as mean Z-scores. P-values<0.050 are considered statistically significant. *HADS* Hospital Anxiety and Depression Scale.

[a] Data missing for one surgically treated patient who did not complete all questionnaires (n=29).

Supplementary Table 9 Multilinear regression for the Verbal Learning Test of Rey, Trail Making Task A, Digit Deletion Test and Digit-Symbol Substitution Test for patients in biochemical remission

	Standardized coefficient β	95% confidence interval	P-value
Verbal Learning Test of Rey			
First attempt: number correct			
Duration of biochemical control/remission (months)	0.0	-0.7, 0.6	0.925
Prolactin at diagnosis (xULN)	-0.2	-0.8, 0.3	0.323
HADS total score	-0.1	-0.1, 0.0	0.602
Any pituitary insufficiency (yes/no)	0.0	-0.9, 0.8	0.904
Second attempt: number correct			
Duration of biochemical control/remission (months)	0.0	0.0, 0.0	0.857
Prolactin at diagnosis (xULN)	0.1	0.4, 0.8	0.555
HADS total score	0.0	-0.1, 0.1	0.845
Any pituitary insufficiency (yes/no)	0.0	-1.0, 1.0	0.989
Fourth attempt: number correct			
Duration of biochemical control/remission (months)	-0.2	-1.3, 0.4	0.283
Prolactin at diagnosis (xULN)	0.0	-0.7, 0.6	0.855
HADS total score	0.2	-0.1, 0.0	0.238
Any pituitary insufficiency (yes/no)	0.1	-0.6, 1.5	0.360
Delayed attempt: number correct			
Duration of biochemical control/remission (months)	0.0	-1.0, 0.7	0.780
Prolactin at diagnosis (xULN)	0.0	-0.7, 0.7	0.973
HADS total score	-0.2	-0.1, 0.0	0.360
Any pituitary insufficiency (yes/no)	0.1	-0.8, 1.4	0.562
Trail Making Test			
Task A: time			
Duration of biochemical control/remission (months)	0.1	-0.6, 1.4	0.452
Prolactin at diagnosis (xULN)	0.0	-0.7, 0.9	0.892
HADS total score	0.1	0.0, 0.1	0.386
Any pituitary insufficiency (yes/no)	0.2	-0.4, 2.2	0.153
Digit Deletion Test			
Number correct			
Duration of biochemical control/remission (months)	0.1	-0.5, 0.7	0.759
Prolactin at diagnosis (xULN)	0.0	-0.5, 0.4	0.859
HADS total score	0.0	0.0, 0.0	0.925
Any pituitary insufficiency (yes/no)	-0.1	-0.8, 0.7	0.853
Digit-Symbol Substitution Test			
Number correct			
Duration of biochemical control/remission (months)	-0.2	-0.8, 0.2	0.296
Prolactin at diagnosis (xULN)	-0.1	-0.5, 0.2	0.445
HADS total score	-0.2	-0.1, 0.0	0.365
Any pituitary insufficiency (yes/no)	0.1	-0.4, 0.9	0.391

Multilinear regression analyses were performed for the tasks that were found to differ between patients and matched healthy controls. The duration of biochemical control/remission and prolactin at diagnosis were log-converted. *HADS* Hospital anxiety and depression, *xULN* times upper limit of normal.

REFERENCES

1. Rey, A., *L'examen clinique en psychologie*. The Clinical Examination in Psychology, 1958: p. 222-222.
2. Wechsler, D., *Wechsler Adult Intelligence Scale - Fourth Edition: Technical and Interpretive Manual*. San Antonio: Pearson, 2008.
3. Rey, A., *L'examen psychologique dans les cas d'encéphalopathie traumatigue (les problems)*. Archives de Psychologie, 1941. 28: p. 998-1009.
4. van der Elst, W., et al., *The Letter Digit Substitution Test: normative data for 1,858 healthy participants aged 24-81 from the Maastricht Aging Study (MAAS): influence of age, education, and sex*. J Clin Exp Neuropsychol, 2006. 28(6): p. 998-1009.
5. Dekker, R.M.J. and P.H. Dekker, *De ontwikkeling van vijf nieuwe Nederlandstalige tests*. PITS: Leiden, 2007.
6. Reitan, R.M., *Trail Making Test: Manual for administration and scoring*. 1986: Reitan Neuropsychology Laboratory.
7. Delis, D.K.E. and K. J., *Delis-Kaplan Executive Function System (D-KEFS)*. San Antonio: The Psychological Corporation, 2001.
8. Benton, A.L. and K.D. Hamsher, *Multilingual Aphasia Examination*. Iowa City: University of Iowa, 1976.
9. Andela, C.D., et al., *The development and validation of the Leiden Bother and Needs Questionnaire for patients with pituitary disease: the LBNQ-Pituitary*. Pituitary, 2016. 19(3): p. 293-302.
10. Sergio E. Starkstein, M.D., Ph.D. ,, et al., *Syndromic Validity of Apathy in Alzheimer's Disease*. American Journal of Psychiatry, 2001. 158(6): p. 872-877.
11. Fisk, J.D., et al., *Measuring the functional impact of fatigue: initial validation of the fatigue impact scale*. Clin Infect Dis, 1994. 18 Suppl 1: p. S79-83.
12. Michopoulos, I., et al., *Hospital Anxiety and Depression Scale (HADS): validation in a Greek general hospital sample*. Ann Gen Psychiatry, 2008. 7: p. 4.
13. Anjan Chatterjee, M.D., MPH ,, et al., *A Comparison of Self-Report and Caregiver Assessment of Depression, Apathy, and Irritability in Huntington's Disease*. The Journal of Neuropsychiatry and Clinical Neurosciences, 2005. 17(3): p. 378-383.
14. Koopmans, P.C.S.R., I. Timmerman, and P.M. Emmelkamp, *The irrational Beliefs Inventory (IBI): development and psychometric evaluation*. European Journal of Psychological Assessment, 1994. 10: p. 15-27.
15. Barelds, D.P.H. and F. Luteijn, *NKPV NEderlands Klinische Persoonlijkheidsvragenlijst*. Amsterdam: Boom Test Uitgevers, 2015.