Cognitive performance after postoperative pituitary radiotherapy: a dosimetric study of the hippocampus and the prefrontal cortex

Pauline Brummelman1*, Margriet G.A. Sattler2*, Linda C. Meiners3, Martin F. Elderson1,4, Robin P.F. Dullaart1, Gerrit van den Berg1, Janneke Koerts5, Oliver Tucha6, Bruce H.R. Wollenbuttel1,4, Alfonsus C.M. van den Bergh2 and André P. van Beek1

1Department of Endocrinology, 2Department of Radiation Oncology, 3Department of Radiology, 4LifeLines Cohort Study & Biobank, University Medical Center Groningen, University of Groningen 5Department of Clinical and Developmental Neuropsychology, University of Groningen, Groningen, The Netherlands

*both authors contributed equally.

European Journal of Endocrinology 2012; 166 (2): 171-179
Chapter 5

ABSTRACT

Objective  The hippocampus and prefrontal cortex (PFC) are important for memory and executive functioning and are known to be sensitive to radiotherapy (RT). Radiation dosimetry relates radiation exposure to specific brain areas. The effects of various pituitary RT techniques were studied by relating detailed dosimetry of the hippocampus and PFC to cognitive performance.

Methods  In this cross-sectional design, 75 non-functioning pituitary macroadenoma (NFA) patients (61±10 years) participated and were divided into irradiated (RT+, n=30) and non-irradiated (RT-, n=45) groups. The RT+ group (who all received 25 fractions of 1.8 Gy; total dose: 45 Gy) consisted of three RT technique groups: three-field technique, n=10; four-field technique, n=15; and five-field technique, n=5. Memory and executive functioning were assessed by standardized neuropsychological tests. A reconstruction of the dose distributions for the three RT techniques was made. The RT doses on 30, 50, and 70% of the volume of the left and right hippocampus and PFC were calculated.

Results  Cognitive test performance was not different between the four groups, despite differences in radiation doses applied to the hippocampi en PFC. Age at RT, time since RT, and the use of thyroid hormone varied significantly between the groups; however, they were not related to cognitive performance.

Conclusion  This study showed that there were no significant differences on cognitive performance between the three-, four- and five field RT groups and the non-irradiated patient group. A dose-response relationship could not be established, even with a radiation dose that was higher on most of the volume of the hippocampus and PFC in case of a four-field RT technique compared to the three- and five-field RT techniques.
INTRODUCTION

Patients with a non-functioning pituitary macroadenoma (NFA) may receive postoperative radiotherapy (RT) for local control [1,2]. However, the safety of RT to the brain has been questioned because of concerns related to second tumor induction, cerebrovascular disease and increased mortality [3]. In addition, accelerated cognitive decline may develop over many years following RT. The incidence and severity of this complication is dependent on radiation fraction dose, total dose and volume, but can also be influenced by other disease- and treatment-related factors [4]. Susceptibility of different brain regions is also reported to vary [5-7]. In addition, time since RT is considered to be an important factor because deterioration in cognitive functions might appear only after a few years [8]. Furthermore, both younger and older patients carry a greater risk of cognitive impairment from RT [9].

Patients who received RT for primary brain tumors performed worse on tests for executive functioning [10]. Other groups reported poor memory performance in irradiated low-grade glioma patients [11]. However, controversy remains because studies on the effects of RT on cognition are usually difficult to interpret, giving differences in tumor localization, which are likely to affect various cognitive domains. In addition, Armstrong et al. described that radiation effects appear to be severe only in a minority of patients. Further, risk of cognitive impairment was found to be related to direct and indirect effects of cancer type, concurrent clinical factors, and premorbid risk factors [9].

In pituitary adenoma patients, Noad et al. [12] found that patients who received postoperative RT performed worse on executive functioning when compared with patients who underwent surgery alone. However, we [13] and others [14-17] did not find an effect of RT on the cognitive performance of patients treated for pituitary disease.

Although older literature supports a role for RT-induced cognitive decline, it remains to be established whether modern pituitary RT techniques result in poorer cognitive performance. However, detailed dosimetric RT studies in relation to objective measures of cognition are lacking in humans. These studies offer the opportunity to relate radiation exposure of prespecified brain areas to cognitive performance. Precise RT dose–volume reconstructions in the brain allow the comparison of the radiation exposure of radiation-sensitive brain areas of different RT techniques. In this context, the temporal lobe/hippocampus and the prefrontal cortex (PFC) appear to be especially relevant. The temporal lobe is crucial for the acquisition of new information as well as its storage and retrieval. The hippocampus within the temporal lobe is important for declarative memory, i.e. the conscious recollection of facts and events [18]. Specifically, the left and right hippocampus appear to have different memory functions [19,20]. The hippocampal granule cell layer, which undergoes neural regenesis, seems to be more sensitive to RT than glial or neural cells in the brain, as shown in animal models [21]. The PFC is of great importance for executive functioning (i.e. planning, cognitive flexibility and inhibition [22]) [23], which seem to decrease after RT [10,12].
Previously, we found no major influence of pituitary RT on cognition in patients with NFA [24]. However, smaller effects could not be excluded. Therefore, we decided to refine the strategy to relate the radiation dose to radiosensitive brain areas (i.e. the hippocampus and PFC) to cognitive test performance typically associated with these brain areas. In addition, we studied which RT technique was superior at limiting dose to the hippocampus and PFC.

**METHODS AND MATERIALS**

**Patients**

In this cross-sectional study, patients were recruited for participation at the Endocrine Outpatient Clinic of the University Medical Center Groningen (UMCG), a tertiary referral center for pituitary surgery in The Netherlands. Inclusion criteria were age ≥ 18 years, treatment for NFA, and regular follow-up in our endocrine outpatient clinic (i.e. at least once a year). The diagnosis of NFA was based on two criteria: the presence of a pituitary macroadenoma (>1 cm) on magnetic resonance imaging (MRI) and the absence of overproduction of any of the pituitary hormones. Pituitary deficiencies were defined according to generally accepted guidelines. Biochemical control of adequacy of the hormonal substitution treatment was judged by the physicians responsible for the care of participating patients, with the use of free thyroxine, insulin-like growth factor 1, and testosterone measurements where necessary. All patients included in present analysis underwent transsphenoidal surgery (TSS) as a primary treatment, in some cases followed by a second surgical procedure if a large remnant, accessible only by surgery, persisted. We only report data of patients who underwent TSS because this is a standard initial treatment in most cases. Patients who underwent a craniotomy more often had larger tumors, necessitating more aggressive treatment, which affect postoperative comorbidity and potentially also cognitive performance. To assure that acute posttreatment effects had resolved, patients were included only at least 6 months after surgery or RT. Furthermore, their hormone replacement schedules had to be stable during these preceding months. Patients were not eligible for participation if they had a (prior) neurological or psychiatric condition, if they had impairments of vision or hearing, or a restriction in hand function, which was expected to interfere with test performance. In addition, patients were not eligible when they were recently diagnosed with a chronic disease or a depressive disorder, indicated by letters of the attending physicians, and in case of pregnancy or an addictive disorder. Prior to their regular visit at our endocrine clinic, patients were approached by telephone to participate. Of a total of 173 NFA patients visiting our endocrine clinics, 75 consecutive patients who received TSS as primary treatment with or without a three-, four- or five-field RT techniques were tested between September 2008 and December 2009. The baseline characteristics of the entire cohort (n=173) did not differ from the presented study population, confirming the representativeness of our study population (data not shown). Approval was given by the medical ethics review committee of the UMCG.
Radiotherapy

Fractionated external beam RT was given with linear accelerators (4-18 MV) between 1987 and 2008 (n=30). In this time period, pituitary RT was performed using three-, four-, or five-field techniques (Table 1). The three-field technique replaced the two lateral opposed field technique because the older two-field technique irradiated a large volume of normal brain with an equivalent or even higher dose of what was applied to the tumor, with reports of brain necrosis as a result of that. The three-field technique consisted of two lateral fields and one vertex field. Since the availability of three-dimensional (3D) radiation treatment planning systems in the 1990s of the previous century, non-coplanar radiation techniques became possible. With non-coplanar techniques, RT fields do not overlap in the same plane and therefore overdose to normal tissue can be minimized. The three- and five-field irradiation techniques used in this study were coplanar, but the four-field technique was non-coplanar. This four-field technique was planned to spare the temporal lobes.

For all the three multiple field RT techniques, similar dose prescriptions were used: 1.8 Gy with a total dose of 45 Gy given in 25 fractions. The median overall radiation treatment time was 35 days (range 31-37 days). From 1987 to 1990, the RT dose of the tumor/pituitary adenoma was prescribed at the tumor encompassing isodose (n=2). From 1991 onwards, the RT dose was prescribed at a central point in the tumor/pituitary adenoma according to the recommendations of the Internal Commission on Radiation Units and Measurements (n=28).

The dose distribution effects of the three different RT techniques were reconstructed using a planning computed tomography (CT) and MRI scan of the head of a 64-year-old patient that served as a model for the whole patient group. The gross tumor volume (i.e. pituitary adenoma remnant) and brain areas at risk (i.e. the left and right hippocampus and PFC) were delineated on MRI with a slice thickness of 1 mm. A planning target volume was generated by adding a 3D margin of 10 mm around the gross tumor volume. The normal tissues at risk were delineated with the help of an experienced neuroradiologist (L C Meiners) according to the established neuroanatomical boundaries [25-27]. To allow a direct comparison, the 3D conformal RT treatment plans were generated using the Pinnacle treatment planning system (version 8.0h). The 3D dose distribution of the left and right hippocampus and PFC were based on a 3D reconstruction of the CT scan. For a description of the RT exposure of the left and right hippocampus and the PFC, the RT dose received at 30, 50 and 70% of these brain area volumes are given. These percentages were chosen because at 30, 50 and 70%, the largest differences between the three different RT techniques in radiation dose were seen. These percentage points give a better idea of the dose distributions compared with minimum, maximum, mean and SDS.
Table 1. Clinical characteristics of patients treated for non-functioning pituitary macroadenoma (NFA) with different radiotherapy (RT) techniques (three, four, and five fields) and without RT (RT-). Data are median and interquartile range, absolute numbers or percentage.

<table>
<thead>
<tr>
<th></th>
<th>Three fields</th>
<th>Four fields</th>
<th>Five fields</th>
<th>RT-</th>
<th>P value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>10</td>
<td>15</td>
<td>5</td>
<td>45</td>
<td></td>
</tr>
<tr>
<td>Basic characteristics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>61 (58-71)</td>
<td>59 (53-70)</td>
<td>55 (46-61)</td>
<td>63 (57-70)</td>
<td>0.236</td>
</tr>
<tr>
<td>Sex (males/females)</td>
<td>9/1</td>
<td>10/5</td>
<td>3/2</td>
<td>30/15</td>
<td>0.491</td>
</tr>
<tr>
<td>Educational level (1/2/3/4/5/6/7)</td>
<td>1/1/0/1/5/1/1</td>
<td>0/2/0/5/6/1/1</td>
<td>0/0/0/2/1/0</td>
<td>0/5/0/7/19/13/1</td>
<td>0.407</td>
</tr>
<tr>
<td>Surgery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at surgery (years)</td>
<td>48 (45-56)</td>
<td>53 (47-63)</td>
<td>42 (27-48)</td>
<td>57 (51-65)</td>
<td>0.008</td>
</tr>
<tr>
<td>Average time since surgery (years)</td>
<td>13 (10-15)</td>
<td>5 (3-8)</td>
<td>14 (12-19)</td>
<td>6 (2-8)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>&lt; 1 – 5 years (number %)</td>
<td>0</td>
<td>8 (53)</td>
<td>0</td>
<td>21 (47)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>5 – 10 years</td>
<td>1 (10)</td>
<td>6 (40)</td>
<td>0</td>
<td>17 (38)</td>
<td></td>
</tr>
<tr>
<td>&gt; 10 years</td>
<td>9 (90)</td>
<td>1 (7)</td>
<td>5 (100)</td>
<td>7 (16)</td>
<td></td>
</tr>
<tr>
<td>Patients with second surgery (%)</td>
<td>2 (20)</td>
<td>1 (7)</td>
<td>1 (20)</td>
<td>3 (7)</td>
<td>0.478</td>
</tr>
<tr>
<td>RT</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age at RT (years)</td>
<td>48 (45-58)</td>
<td>55 (50-64)</td>
<td>43 (29-50)</td>
<td>NA</td>
<td>0.040</td>
</tr>
<tr>
<td>Average time since RT (years)</td>
<td>13 (10-14)</td>
<td>3 (2-4)</td>
<td>12 (11-17)</td>
<td>NA</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Hormonal substitution</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of hormone replacements</td>
<td>0/1/2/3/4/5</td>
<td>0/1/0/6/3/0</td>
<td>1/3/2/5/4/0</td>
<td>0/2/1/1/1/0</td>
<td>10/9/7/10/0</td>
</tr>
<tr>
<td>Glucocorticoid (%)</td>
<td>80</td>
<td>80</td>
<td>40</td>
<td>53</td>
<td>0.119</td>
</tr>
<tr>
<td>Thyroid hormone (%)</td>
<td>100</td>
<td>67</td>
<td>80</td>
<td>51</td>
<td>0.026</td>
</tr>
<tr>
<td>Growth hormone (%)</td>
<td>50</td>
<td>33</td>
<td>40</td>
<td>24</td>
<td>0.418</td>
</tr>
<tr>
<td>Sex hormone (%)</td>
<td>70</td>
<td>60</td>
<td>40</td>
<td>56</td>
<td>0.712</td>
</tr>
<tr>
<td>Desmopressin (%)</td>
<td>10</td>
<td>13</td>
<td>20</td>
<td>11</td>
<td>0.940</td>
</tr>
</tbody>
</table>

* P value between the four groups by Kruskal-Wallis or Chi-square tests. NA: not applicable.

a: Duncan’s post hoc test: significant differences between the five field group on the hand and the four field and RT- groups on the other hand.

b: Duncan’s post hoc test: significant differences between the three field and five field groups on the one hand and the four field and RT- groups on the other hand.

c: Duncan’s post hoc test: significant differences between the four field and five field groups.

d: Duncan’s post hoc test: significant differences between the three field and five field groups on the one hand and the four field group on the other hand.
Cognitive tests
Aspects of verbal memory were assessed with the 15 Words Test (15 WT) which is a Dutch equivalent of the Rey Auditory Verbal Learning Test [28]. In this test, 15 words were presented five times. After each trial, patients were asked to name immediately the words they remembered. This allowed the calculation of three different scores describing immediate memory:

i) The short-term memory score is based on the number of words patients were able to name after the first presentation of the word list.

ii) The total memory score represents the total number of words patients remembered over the five trials.

iii) The learning score describes the difference between the number of words remembered in the third trial in comparison with the first trial. Besides immediate memory, delayed memory was measured.

iv) The delayed memory score is based on the number of words patients could recall after a period of about 30 minutes.

Executive functioning was assessed using the Ruff Figural Fluency Test (RFFT) [29]. In this test, patients were presented with sheets of paper on which 35 squares were printed, each with a fixed pattern of five dots. The test consisted of five parts, which differed with regard to the designs. While the configurations of dots are the same in the first three parts of the test, two types of distractions are added in two of these parts. In the last two parts, the configurations of the dots are different and without distractions. The participant was asked to produce as many different designs as possible by connecting two or more dots in each square with straight lines. The time for each part was restricted to 1 minute so that the total test time was 5 minutes. Responses were scored with regard to the total number of unique designs generated over the five parts. The perseverative errors score represents the total number of repetitions of the same design drawn. The interrater variability (two independent raters) was determined by Pearson’s $r$ and was 0.99 for both total unique designs and perseverative errors. The error ratio is calculated by the total number of perseverative errors divided by the total number of unique designs.

Questionnaires and protocol
A common questionnaire on demographic and health-related data was used with special attention for educational level, social status, full-time/part-time employment, social security benefit, comorbidity, use of medicine, cardiovascular risk factors, traumatic brain injury, and dementia. Education level was determined by using a Dutch education system, comparable to the International Standard Classification of Education (ISCED) [30]. This scale ranges from 1 (elementary school not finished) to 7 (university level). The Hospital Anxiety and Depression
Chapter 5

Scale (HADS) consists of 14 items and measures anxiety and depression [31]. Each item is scored as a number, with a maximum score for each subscale (anxiety or depression) of 21. Higher scores indicate more severe anxiety or depression.

In fixed order, the test protocol was as follows: i) the 15 WT: direct recall; ii) the RFFT; iii) a common questionnaire to assess baseline information; iv) physical examination: length, weight, blood pressure, waist circumference, hip circumference and compliance to the test situation; v) the HADS; and finally vi) the 15 WT: delayed recall. The assessment took ~ 40 min and was performed directly after or just before patients’ visit to the outpatient clinic. All testing and scoring of tests were performed by trained personnel.

**Reference data: healthy control subjects**

The performances of patients were compared with Dutch controls. Normative data for the HADS were derived from Spinhoven et al. [31]. In their study, psychometric properties of the HADS were assessed in six different groups of Dutch subjects (n=6165): i) a random sample of younger adults (18-65 years; n=199); ii) a random sample of elderly subjects of 57-65 years of age (n=1901); iii) a random sample of elderly subjects of 66 years and older (n=3293); iv) a sample of consecutive general practice patients (n=112); v) a sample of consecutive general medical outpatients with unexplained somatic symptoms (n=169); and vi) a sample of consecutive psychiatric outpatients (n=491). In all six groups, an authorized Dutch translation of the HADS was used. General population mean and SDS were used from 18 to 65 years and >65 years to calculate Z-scores.

Reference data for the 15 WT were derived from control subjects of the Maastricht Aging Study. In this cohort, 1780 healthy participants between 24 and 81 years were evaluated on a Dutch adaptation of the Rey Verbal Learning Test, the 15 WT [28]. Regressions models given by the authors were used to determine accurate Z-scores. The final test scores were controlled for age, sex, and education. Reference data for the RFFT were derived from a sample (n=10.289) of the LifeLines Cohort Study [32]. Reference groups were stratified by a matrix of eight education levels and 13 age levels (half decades from 20 to 85 years). Each cluster consisted on average of 120 subjects. RFFT forms were analyzed by a computerized pattern recognition program. There was a good internal consistency (n=373) between computerized rating and human rating for unique designs (Cronbach’s α = 0.99) and perseverative errors (Cronbach’s α = 0.97). Using a Bland-Altman analysis, a near perfect level of agreement was found between these two rating methods with intraclass correlation for unique designs (0.99) and perseverative errors (0.96). Using the mean and SDS for each reference group, we standardized our patient scores by converting it into Z-scores.
Statistical analyses
The analyses were all carried out using the PASW (SPSS, Inc., Armonk, NY, USA) statistics package. Demographic data are presented as median and interquartile range, frequencies, or percentages. We compared data of four groups of patients: i) patients who received TSS and a three-field RT technique; ii) patients who received TSS and a four-field RT technique; iii) patients who received TSS and a five-field RT technique; and iv) patients who only received TSS. Categorical variables were analyzed by using χ² tests. The non-parametric Kruskal-Wallis test was used for all continuous variables who failed to meet the normality assumption. The two-tailed α level of <0.05 was considered statistically significant. In case of statistical differences between the groups on demographic or cognitive data, Duncan’s method was used as post-hoc test [33].

RESULTS

Study population
Seventy-five NFA patients (52 men and 23 women, age 61 ± 10 years) participated in this study. Thirty patients received TSS and postoperative pituitary RT (RT+ group), whereas 45 patients did not receive RT (RT- group). The RT+ group consisted of three RT technique groups: three-field technique, n=10; four-field technique, n=15; and five-field technique, n=5. Patients’ characteristics are given in Table 1. No differences in age at time of study, sex, or educational level were found between the four groups. In our cohort, the older three- and five-field RT techniques were applied earlier in time, and from 2001, these techniques were replaced by a four-field technique. Consequently, significant differences between groups were found for average time since RT. Furthermore, patients in the four-field group were on average older at the time of RT compared with those in the five-field group (Table 1). Hormonal substitution for pituitary deficiencies were similar between groups with the exception of thyroid hormone that was given less frequently in patients without RT. Feelings of anxiety and depression were comparable between the four groups and not indicative of clinical anxiety and depression (data not shown). Social status, full-time/part-time employment, social security benefit, and comorbidity were all comparable between the four groups at time of assessment (data not shown).

Radiotherapy
The RT dose distributions on transversal, coronal and sagittal CT scan images are shown for the different RT techniques (Fig. 1). Estimated RT dosimetric data (derived from Dose-Volume Histograms, plots not shown) revealed increased radiation dose exposure in the four-field RT technique; doses received at 50 and 70% of the hippocampus and PFC were upto seven fold higher compared with the three- and five-field RT techniques (Table 2). Furthermore, the
patient model used in this study had a tumor that had a deviation to the right. Therefore, slightly higher doses were found on the right hippocampus compared with the left hippocampus.

Figure 1 The RT dose distributions on transversal, coronal, and sagittal CT scan images are shown for different RT techniques. The color areas shown on the CT scan images represent different RT isodose areas: red = 49.5–51.8 Gy; orange = 48.2 Gy; green = 42.8 – 45 Gy; light blue = 40 Gy; dark blue = 20 – 30 Gy; and white = 5-10 Gy. The color lines shown on the CT scan images represent different delineated structures: red line, planning target volume; yellow line, gross tumor volume/ pituitary adenoma remnant; light blue line, prefrontal cortex; and pink line, frontal cortex.
Table 2: Estimated dosimetric data memory performance and executive functioning of patients treated for non-functioning pituitary macroadenoma (NFA) with different radiotherapy (RT) techniques (three, four and five fields) and without RT.

<table>
<thead>
<tr>
<th>N</th>
<th>Three fields</th>
<th>With different RT techniques</th>
<th>Without RT</th>
<th>P value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Three fields</td>
<td>Four fields</td>
<td>Five fields</td>
</tr>
<tr>
<td></td>
<td></td>
<td>10</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td>Estimated dosimetric data (RT dose received at 30/50/70% of the volume of:)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left hippocampus (Gy)</td>
<td>23.0/3.2/2.0</td>
<td>21.8/14.9/13.5</td>
<td>19.6/4.1/2.9</td>
<td></td>
</tr>
<tr>
<td>Right hippocampus (Gy)</td>
<td>28.5/4.4/2.0</td>
<td>29.8/15.2/14.2</td>
<td>19.0/4.6/2.8</td>
<td></td>
</tr>
<tr>
<td>Prefrontal cortex (Gy)</td>
<td>23.0/19.0/3.6</td>
<td>26.5/18.2/17.3</td>
<td>25.6/16.2/4.9</td>
<td></td>
</tr>
<tr>
<td>Memory performance (15 Words Test; mean (SD))</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Short-term memory</td>
<td>-0.13 (1.42)</td>
<td>-0.65 (0.90)</td>
<td>-0.39 (1.23)</td>
<td>-0.20 (1.09)</td>
</tr>
<tr>
<td>Total memory</td>
<td>-0.46 (1.66)</td>
<td>-1.33 (1.07)</td>
<td>-0.92 (0.72)</td>
<td>-0.62 (1.15)</td>
</tr>
<tr>
<td>Learning score</td>
<td>-0.23 (1.17)</td>
<td>-0.70 (0.93)</td>
<td>-0.16 (0.90)</td>
<td>-0.22 (1.04)</td>
</tr>
<tr>
<td>Delayed memory</td>
<td>0.00 (1.39)</td>
<td>-0.96 (1.21)</td>
<td>-1.26 (0.67)</td>
<td>-0.86 (1.19)</td>
</tr>
<tr>
<td>Executive functioning (Ruff Figural Fluency test; mean (S.D.))</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unique designs</td>
<td>-0.52 (1.08)</td>
<td>-1.19 (1.08)</td>
<td>-0.43 (1.39)</td>
<td>-0.56 (1.13)</td>
</tr>
<tr>
<td>Perseverative errors</td>
<td>-0.87 (0.73)</td>
<td>-0.60 (0.89)</td>
<td>-0.61 (0.67)</td>
<td>-0.33 (1.46)</td>
</tr>
<tr>
<td>Error ratio</td>
<td>-0.80 (0.48)</td>
<td>-0.31 (1.40)</td>
<td>-0.59 (0.71)</td>
<td>-0.09 (1.78)</td>
</tr>
</tbody>
</table>

*P values by Kruskal-Wallis ANOVA. Cognitive performance data are given as Z-scores.
Cognitive tests

Cognitive functioning Z-scores are given in Table 2 for memory performance and executive functioning. No significant differences were found on the 15 WT and the RFFT between the four groups; no RT dose-volume effect on cognition was found. Longer time since RT (three- and five-field techniques) or older age at RT (four-field technique) did not result in poorer cognitive test results.

DISCUSSION

This study showed that there were no significant differences between the three-, four- and five-field RT groups and the non-irradiated patient group. Therefore, a dose-response relationship could not be established, even with a radiation dose that was higher on most of the volume of the hippocampus and PFC in case of a four-field RT technique compared with the three- and five-field techniques.

To our knowledge, this is the first study that related detailed dosimetric data of pituitary RT to cognitive performance. Our results confirm our previous data [13,24] and that of others [14-17] that modern dose regimens of pituitary RT does not appear to have a major influence on cognition, but also extend these by showing an absence of a relationship between radiation dose and cognitive performance. For our study, we used a homogeneous patient group of patients with NFA, thereby excluding confounding results by excess hormone or other treatment-related factors [34]. Further, our study lacks the inherent weakness of studies on patients with primary brain tumors or gliomas. Because of the fixed position of the pituitary tumor that is not affecting cortical or subcortical areas of the brain, cognitive test results were not influenced by localization of both brain tumor and subsequent targeted focal RT.

This study showed a larger cumulative radiation dose on most of the volume of the left and right hippocampus and PFC with a four-field technique. Although this radiation technique was developed to circumvent excess radiation exposure to the brain - especially both temporal lobes - by avoiding radiation fields overlap, our reconstruction showed that this technique is not preferred with regard to RT exposure to the hippocampus and PFC. Despite this excess radiation no poorer cognitive performance was observed in this patient group. Arguably, patients with four-field RT had a shorter follow up time (3 years) when compared with the three- and five-field techniques (13 and 12 years respectively). For this reason, cognitive dysfunction may not have developed. However, at slightly higher radiation doses, some have reported a decrease in executive functioning already after six months in patients with primary brain tumors [10]. It is evident that long-term follow-up is necessary for this patient group [4].

In contrast to our study, Noad et al. found differences between irradiated and non-irradiated pituitary patients with regard to cognitive functioning. They tested 71 patients with pituitary tumors treated with surgery with or without RT (25 fractions of 1.8 Gy) on quality of life and
several cognitive functions (memory, attention and executive functioning) [12]. In addition to an impairment in cognitive function regardless of treatment type, they reported a significantly worse performance on executive function (measured with the Stroop test) in the RT+ group compared with patients receiving only surgery. As noted by the authors, their finding may be explained by chance. Jalali et al. [35] also found that RT applied to the left temporal lobe in patients with tumors of low malignant potential (craniopharyngioma, cerebellar astrocytoma, optic pathway glioma, and cerebral low-grade glioma) were predictive of cognitive decline. In addition, Douw et al. [36] report a decline in attentional functioning in patients with low grade gliomas who received RT. Among the many differences between these studies and ours, the most explanatory are the higher radiation doses used (> 54 Gy in 30 fractions), the different tumor pathology, and the young age of the patient group (median 13 years). Comparison with the above mentioned studies [10,35,36] suggests that there may be a kind of threshold for RT to injure the brain and cause cognitive impairment. It is likely that only above this threshold a dose response relationship can be found. The age difference between the patients of Jalali’s study and ours is probably also essential, in that a developing brain is more likely to be affected by RT. The absence of a RT dose–cognitive response relationship in our study suggest that all applied RT techniques seem to operate within safe RT dose boundaries.

In accordance with our finding that pituitary RT was not associated with reduced cognition, others also found no differences between patients treated for pituitary tumors with or without RT [14-17]. In a recent review by Loeffler & Shih [37], it was stated that the overall rate of treatment-related adverse effects (secondary tumors, visual complications, neurological symptoms, strokes, general, and mental health) is low and that only hypopituitarism is to be expected following RT in pituitary adenoma patients. However, there are indications that this risk does not exceed the risk on hypopituitarism in patients who only received surgery [2]. Unfortunately, cognition received little attention in this recently published paper by Loeffler. However, Lawrence et al. [4] recently reviewed the published data regarding RT-induced brain injury and found very limited evidence that brain RT in 2 Gy fractions causes irreversible cognitive decline in adults with primary and metastatic brain tumors.

Some study limitations are to be made. In this study, only a small number of patients were included in some subgroups yielding low statistical power and potentially excluding detection of differences that were not large. This study showed that there were no significant differences on cognitive test performance between the three-, four- and five field RT groups and the non-irradiated patient group. Therefore, it seems plausible that RT might at most have some very subtle impact on cognition. However, some reservations are justified. Ideally, to detect such a small size effect, a prestudy analytical design with regard to group sizes should be made. Literature provides no data on estimated RT-induced cognitive decline in humans with fractionated RT limited to a total dose of 45 Gy given in 25 fractions. In addition, the disease prevalence is low (especially considering that not all received RT), creating little
room for extensive prestudy analytical design. Instead, we approached all patients that were eligible within September 2008 and December 2009, which is relatively small, and therefore has statistical limitations.

Further, it should also be taken into account that it is plausible that not everyone shares the same vulnerability to damage of RT. For instance, Armstrong et al. [9] reviewed that the apolipoprotein E (ApoE) genotype could prove to be a premorbid risk factor for greater RT-induced damage, possibly related to its risk for development of neurofibrillary tangles and plaque neuritis [38]. In our group, we found no differences in ApoE genotype between the four groups with normal allele frequencies of apoE4. Finally, to precisely define the volume of the hippocampus and PFC, we used imaging of a 64-year-old patient. The exact volume of the hippocampus and PFC differs inter-individually. As it is a common practice in dosimetric studies, we estimated dosimetric data of this model and extrapolated this to our patients.

Although we found that the multiple RT field techniques do not appear to have a major influence on cognitive performance, radiotherapeutic treatment techniques continue to develop to reduce radiation to healthy tissue. Stereotactic RT (SRT) is one of the new RT techniques and is currently used for the treatment of pituitary tumors in our center. SRT is expected to be a promising alternative treatment method to deliver the radiation dose precisely. This technique is expected to spare normal tissue more than the multiple field external beam RT. The first SRT study results are promising according to tumor control and clinical status (including changes in neurological status, such as visual function and endocrinological function) at a median follow up of 25.5 months [39]. Long-term effects on cognitive functioning need to be awaited. In the (nearby) future, alternative forms can be expected like hippocampal sparing RT or even limbic circuit sparing, neural stem cell sparing and neural progenitor cell sparing RT [5-7]. This may be especially important for patients with benign tumors, with a long life expectancy like NFA patients.

In conclusion, no dose-response relationship could be established, confirming that current multiple field RT techniques and fractionated radiation dose regimens do not result in major differences in cognitive performance involving the hippocampus and PFC in NFA patients.

Declaration of interest
The authors declare that there is no conflict of interest that could be perceived as prejudicing the impartiality of the research reported.

Funding
This research did not receive any specific grant from any funding agency in the public, commercial or not-for-profit sector.
Acknowledgements
We like to thank Dr W J Sluiter for his help in the statistical analyses. The help of A Klop is greatly appreciated. We thank Rob Bieringa, Joost Keers, René Oostergo, Rosalie Visser, and Judith Vonk for their work related to data collection and validation of the RFFT norm scores derived from the LifeLines Cohort Study. The authors are grateful to the study participants, the staff from the LifeLines Cohort Study and Medical Biobank Northern Netherlands, and the participating general practitioners and pharmacists. The LifeLines Cohort Study, and generation and management of GWAS genotype data for the LifeLines Cohort Study, is supported by the Netherlands Organization of Scientific Research NWO (grant 175.010.2007.006), the Economic Structure Enhancing Fund (FES) of the Dutch government, the Ministry of Economic Affairs, the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the Northern Netherlands Collaboration of Provinces (SNN), the Province of Groningen, the University Medical Center Groningen and the University of Groningen, the Dutch Kidney Foundation, and the Dutch Diabetes Research Foundation.
REFERENCES


