Recovery of pituitary function in the late-postoperative phase after pituitary surgery: results of dynamic testing in patients with pituitary disease by insulin tolerance test 3 and 12 months after surgery

C Berg, T Meinel, H Lahner, K Mann and S Petersenn
Department of Endocrinology and Division of Laboratory Research, University of Duisburg-Essen, 45122 Essen, Germany
(Correspondence should be addressed to C Berg who is now at Division of Endocrinology, Medical Center, University of Duisburg-Essen, Hufelandstraße 55, 45122 Essen, Germany; Email: christian.berg@uni-due.de)

Abstract
The insulin tolerance test (ITT) is considered the gold standard for assessment of GH and ACTH reserve in patients with pituitary disease following pituitary surgery and is usually performed after 6–12 weeks. However, abnormal axes may not be completely recovered by then. The aim of this study was to evaluate dynamic testing 3 and 12 months after transsphenoidal pituitary surgery.

Design and patients: Serial dynamic testing was performed in 36 patients (13 women, age 18–78) at 3 and 12 months after transsphenoidal surgery.

Results: Compared with 3-month results, median GH peak levels during ITT after 12 months increased by 38% (P<0.05). In patients initially classified as GH deficiency (GHD), median GH peak increased after 12 months by 23% (P<0.05). At 3 and 12 months, 36% (13/36) and 47% (17/36) were GH sufficient respectively. Median cortisol peak levels after 12 months increased by 17% (P<0.01) compared with 3-month ITT. In ACTH-insufficient (AI) patients, peak cortisol levels increased significantly by 12% (P<0.05) at 12 months, and in ACTH-sufficient patients, peak cortisol levels increased significantly by 13% (P<0.05). At 12 months, there was recovery from AI in 11% of the patients, and recovery from GHD in 11% of patients.

Conclusions: Serial dynamic testing results in a change in classification by ITT results in a relevant proportion of patients. Dynamic testing should be repeated during follow-up.

Introduction
Investigation of pituitary insufficiency is an important aspect in the follow-up of patients with pituitary surgery. Whereas patients with adrenal insufficiency may develop severe symptoms, GH deficiency (GHD) in adults typically presents with subtle signs. GHD may be clinically characterized by changes in body composition, impaired psychological well-being, reduction in bone mineral density, and metabolic alterations in lipids and insulin resistance, which GH replacement reverses (1–3). Under these circumstances, the diagnosis in a suspected patient is purely biochemical, based on an insufficient GH peak response in one or two stimulation tests, depending on the number of other pituitary hormone deficiencies (4).

The insulin tolerance test (ITT) has become the ‘gold standard’ to determine the need for cortisol and GH replacement in patients with hypothalamic–pituitary disease (5, 6). It assesses the integrated central and peripheral response to a stressful event caused by hypoglycemia. GH and cortisol responses <3 µg/l and <500 nmol/l respectively have been defined as evidence of deficiency (5–7). After transphenoidal surgery (TS), evaluation for pituitary insufficiency is usually performed after an interval of 2–3 months to allow recovery of altered pituitary function (8). However, subsequent testing is less defined, as the potential for subsequent improvement due to re-expansion of the compressed normal pituitary is unknown. The aim of this study was to evaluate the diagnostic utility of an additional provocative testing by ITT 12 months after pituitary surgery in patients with hypothalamic–pituitary disease.

Subjects and methods
Patients
In this retrospective analysis, we analyzed data from 36 patients (13 women, age 18–78, median body mass index (BMI) 26.5 (19.4–43.7) kg/m²) routinely
investigated by ITT at the University Hospital Essen 3 and 12 months after TS of a sellar mass (26 nonfunctioning adenomas, 5 prolactinomas, 3 cranio-opharyngiomas, and 2 meningioma). All tumors were larger than 1 cm in size. Exclusion criteria were recombinant GH replacement therapy prior to or during the evaluation period. GHD patients included in the study either refused or had contraindications for GH replacement. All patients who had pituitary hormonal deficiencies other than GH were on appropriate and stable replacement therapy. Furthermore, all patients had routine magnetic resonance imaging evaluation of the sellar region at 3 and 12 months. Patients with relevant tumor remnants or recurrence were excluded from the analysis. All patients gave informed written consent to have their data analyzed in this study.

Methods

For ITT, patients received 0.15 IU/kg of regular insulin (Actrapid Novo Nordisk, Mainz, Germany) i.v. to achieve blood glucose levels below 2.2 mmol/l and until symptoms of hypoglycemia developed. Patients on chronic corticosteroid replacement therapy (generally 10–15 mg hydrocortisone per day) received their last dosage at 1400 h the day before testing, resulting in a drug restriction period of at least 18 h. Blood samples for GH, cortisol, and glucose were taken at −10, 0, 15, 30, 45, 60, 90, and 120 min. None of the patients required dextrose replacement due to massive hypoglycemia, and hypoglycemia was evident in all patients within the first 30 min. Further assessment of anterior pituitary function was performed by baseline hormonal testing as well as by provocative tests as required. TSH deficiency was defined by low serum-free thyroxine level without appropriate elevation in serum TSH. In males, secondary hypogonadism was defined by low serum testosterone with inappropriately low gonadotropin level; and in postmenopausal females, the presence of low serum estradiol level without a rise in gonadotropin level; in premenopausal females, by amenorrhea in testosterone with inappropriately low gonadotropin level; and in patients with polyuria and polydipsia, confirmed by inappropriately low serum gonadotropin concentrations. Vasopressin deficiency was suspected in patients with polyuria and polydipsia, and confirmed by low urinary osmolality during thirst test.

Serum GH levels were determined by a chemiluminescence immunometric assay (Immulite 2000 assay, Siemens AG, Erlangen, Germany). All samples from each individual patient were analyzed together. The assay was calibrated against the WHO 1st international standard (80/505) for human GH. Intra- and interassay coefficients of variation (CV) values for a low point of the standard curve were 5.4 and 7.9% respectively. For ITT, a peak GH response below 3 μg/l established the diagnosis of severe GHD. Serum cortisol levels were determined by competitive immunoassay, using commercial kits (Advia Centaur, Bayer). The analytical sensitivity of the assay was 5.5 nmol/l.

Intra-assay variations as CV for various cortisol values were 3.7% (107.1 nmol/l), 3.1% (155.3 nmol/l), 2.9% (391.0 nmol/l), 3.8% (759.6 nmol/l), and 3.0% (1025.0 nmol/l) respectively. Interassay variations for the cortisol concentrations mentioned above were 5.5, 3.8, 3.1, 1.9, and 4.0% respectively. A peak cortisol <500 nmol/l was used to define adrenal insufficiency. All other parameters were determined by routine methods.

Statistical analysis

Results (median (range)) are expressed as absolute values for GH and cortisol. GraphPad Prism 4.0 software for Windows (GraphPad Software, San Diego, CA, USA) was used for statistical analysis. Spearman’s rank correlation analysis was used to determine relationship between variables. For further statistical analysis, Wilcoxon’s matched pairs test and the Mann–Whitney test were performed where appropriate.

Results

Investigation of the somatotropic and the corticotropic axes by ITT 3 months after pituitary surgery

GH Three months after pituitary surgery, median GH peak level during ITT was 1.9 (0.1–32.3) μg/l. In 36% (13/36) of the patients, GH peak response was >3 μg/l, and patients were classified as GH sufficient (GHS). In these patients, median GH peak was 13.0 (3.5–32.3) μg/l. In 64% (23/36) of the patients, GH peak response was <3 μg/l, and patients were classified as GH deficient (GHD). Their median GH peak was 1.2 (0.1–2.7) μg/l. Total insulin-like growth factor 1 was 92 (14–352) μg/l and was significantly lower in GHD patients (88 (14–182) μg/l) than that in the GHS patients (159 (20–352) μg/l, P <0.05). Proven dysfunction of 1, 2, 3, or 4 pituitary axes other than GH (including diabetes insipidus) was present in 11 (29%), 10 (27%), 10 (27%), and 5 (14%) respectively of all patients. BMI was not significantly different in GHS and GHD patients (25.7 (20.1–41.2) versus 27.4 (19.4–43.7) kg/m² (P=0.58)).

Cortisol Median cortisol peak level during ITT was 445 (10–811) nmol/l. In 44% (16/36) of the patients, peak cortisol was >500 nmol/l, and these patients were classified as ACTH sufficient (AS). Median cortisol peak in AS patients was 575 (503–811) nmol/l. In 56% (20/36) of the patients, ACTH insufficiency (AI) was present with a median cortisol peak of 311 (10–476) nmol/l. Basal cortisol was 342 (11–690) nmol/l, significantly lower in AI patients than in AS patients (255 (11–484) vs 363 (262–690) nmol/l, P<0.01). Proven dysfunction of 1, 2, 3, or 4 pituitary axes other
than ACTH (including diabetes insipidus) respectively was present in 13 (36%), 8 (22%), 10 (28%), and 5 (14%) respectively of all patients.

**Results of dynamic re-testing at 12 months**

**Glucose** There was no difference in median nadir of blood glucose at 3- and 12-month ITT (2.01 (0.89–2.2) vs 2.04 (0.83–2.17) mmol/l, *P* = NS). Intra-individual variation in nadir glucose levels was 9.0 ± 5.3%.

**GH** The results were grouped according to the classification by ITT at 3 months. There was a high correlation for GH peak values in 3- and 12-month ITT results (*r* = 0.89, *P* < 0.0001). Twelve months after pituitary surgery, median GH peak level in ITT increased by 38% (*P* < 0.05; Fig. 1a). In patients who were classified as GHD at 3 months, median GH peak at 12 months increased by 23% (*P* < 0.05) compared with 3 months (Fig. 1b). In GHS patients, median GH levels increased by 10% (*P* = NS; Fig. 1b). While at 3 months, only 36% of patients were GHS, 47% of patients were GH sufficient after 12 months, as four patients (11%) restored GH reserve (Fig. 2), and none developed new GHD during follow-up. The lower 95% percentile peak GH levels at 3 months (*x̄* ± 2S.D.) of those patients who recovered during follow-up were 0.3 µg/l.

**Cortisol** Three- and 12-month results were highly correlated (*r* = 0.85, *P* < 0.0001). At 12 months, median cortisol peak levels in ITT increased significantly by 17% (*P* < 0.01) compared with 3-month results (Fig. 3a). Basal cortisol at 12 months was not significantly different from basal cortisol at 3 months. When re-tested at 12 months, median cortisol peak levels in AI patients significantly increased by 12% (*P* < 0.05), and median cortisol peak levels in AS patients significantly increased by 13% (*P* < 0.05; Fig. 3b). Fifty-six percent of patients were AS at 12 months compared with 44% at 3 months. Four patients restored ACTH reserve between 3 and 12 months (Fig. 4). None of the patients developed new AI at 12 months. The lower 95% percentile peak cortisol levels at 3 months (*x̄* ± 2S.D.) of those patients who recovered during follow-up were 230 nmol/l.

**Other pituitary deficits** At 3 months, proven dysfunction of 0, 1, 2, or 3 pituitary axes other than GH and cortisol (including diabetes insipidus) was present in 15 (42%), 8 (22%), 7 (19%), and 0 (0%) respectively (Fig. 5). After 12 months, proven dysfunction of 0, 1, 2, or 3 pituitary axes other than GH and cortisol was present in 20 (56%), 10 (28%), 5 (14%), and 1 (3%) respectively. Median number of deficits other than GH and cortisol was 0.5 (0–2) at 3 months compared with 0 (0–3) at 12 months (*P* = 0.38). In 17% (6/36) of the patients, one pituitary axis other than GH
or cortisol had recovered after 12 months, and in 3% (1/36) of the patients, two axes had recovered. In one patient, a second deficient axis was prevalent at 12 months, and another patient revealed three deficient axes at 12 months compared with two deficient axes at 3 months.

**Discussion**

Accurate postoperative assessment of the patient after transsphenoidal pituitary surgery includes monitoring of the anterior or posterior pituitary function, which may be adversely affected or improved by this operation. In our center, the post-surgical period includes an early postoperative period (immediately after surgery through the following first 4 weeks) and subsequently a late-postoperative period. This study focused on the late-postoperative assessment, and monitoring of GH and ACTH reserve in patients 3 and 12 months after surgery. So far, it is assumed that a single assessment of GH and ACTH reserve in patients with a sellar mass at the beginning of the late-postoperative phase is predictive of long-term outcome in a patient who is otherwise stable and has no tumor relapse (8, 9). Hence, in most centers, dynamic testing for GH and ACTH reserve is usually performed after an interval of 2–3 months to allow recovery of altered pituitary function, determining the need for life-long replacement therapy. There are few data yet suggesting that serial stimulation testing beyond this time is necessary as the potential for subsequent improvement due to re-expansion of the compressed normal pituitary is not very well characterized. The outcome of initial hypopituitarism is unclear, because once a clear deficiency has been demonstrated, patients are often left on replacement therapy without further assessing possible recovery from late postoperative hypopituitarism. Of note, abnormal axes may not be recovered completely within the first months after surgery.

This is the first study showing that in the late-postoperative phase after pituitary surgery, there is...
ongoing recovery of pituitary function demonstrated by a significant improvement of GH and cortisol peak levels in ITT during follow-up. We have found that dynamic re-testing at 12 months after pituitary surgery in comparison to a single ITT at 3 months detects a recovery from ACTH and GH insufficiency, in each 11%.

Evidence suggests that the main mechanism of hypopituitarism in patients with pituitary adenomas is compression and destruction of the normal pituitary gland by the expanding mass, focal necrosis by compression of the portal vessels in the pituitary stalk, either secondary to direct effects of the expanding tumor mass or to raised intrasellar pressure (10, 11). TS is the treatment of choice in the majority of patients with functioning and nonfunctioning pituitary adenomas. Apart from removing the adenoma, this technique makes it possible to preserve normal pituitary tissue in many cases. Even though it may induce new pituitary insufficiencies, it has also been shown to improve initial pituitary hypofunction in nonfunctioning adenomas (12–15) and in hormone-secreting tumors (16). After surgery, the incidence and degree of hypopituitarism depend on a number of factors, including the size of the original tumor, the degree of infiltration, and the experience of the surgeon. Given the regeneration potential of pituitary tissue, recovery after surgical decompression may be anticipated in large or long-standing tumors (17). To focus on individual recovery of pituitary function, our study was restricted to patients who underwent TS for macroadenomas, performed in one specialized center for neurosurgery only.

The rate of recovery within the first year after surgery has been investigated in several studies. In a study by Webb et al. (17), 48% of 93 patients with at least one pituitary hormone deficiency preoperatively regained some pituitary function within 1–6 months. In another study by Colao et al. (18), pituitary function improved in 13% of patients after 1 year with only 3.6% of patients recovering from GHD as the most common deficiency. In a study by Arafah (12), patients with non-functioning pituitary macroadenomas recovered from AI in 38%, from hypogonadism in 32%, from hypothyroidism in 57%, and from GHD in 15% within 2–3 months. However, none of these studies investigated recovery of GH and ACTH reserve by serial dynamic testing during later postoperative follow-up.

It is known that ACTH is the hormone that is most frequently recovering (17). Both ACTH and GH reserve can be tested by different methods like glucagon test (19), CRH test (20), and ACTH test (21) for cortisol, and arginine test (22, 23), GHRH + arginine test (24–27), and GHRP-6 test (28, 29) for GH, but for both hormone axes the ITT is the gold standard test (4, 7). The diagnosis of cortisol deficiency in the chronic phase after TS is challenging because there is often a drug interference with hormonal testing. Therefore, patients on glucocorticoids were re-assessed at 3- and 12-month follow-up visits by withholding hydrocortisone therapy for at least 18 h to exclude an influence of corticosteroids on ITT results. When re-tested at 12 months by ITT, we found a significant increase in median cortisol peak levels and subsequently an increase in numbers of AS patients. Cortisol secretion improved not only in AI but also in AS patients, pointing to an overall ongoing recovery of the ACTH-producing cells of the pituitary probably due to re-expansion of the compressed normal pituitary. Interestingly, basal cortisol levels did not improve and were unable to reveal possible recovery.

The pituitary hormone least likely to recover is usually GH (8). The optimal postoperative time frame to assess for and begin GHD therapy is not yet established (8). A number of studies have found persistent GHD to be common after pituitary surgery (30). In one study, the overall incidence of GHD was 80.2% on testing as early as 3 months following pituitary surgery (31). However, there is evidence that recovery of pituitary function begins immediately after surgery (10). Since the time frame for recovery of GH function varies postoperatively, the optimal time for initiating therapy is uncertain. When re-testing by provocative testing at 12 months, we found a significant increase in median peak GH levels and subsequently an increase in the numbers of GHS patients. Although not statistically significant, probably due to the number of patients tested, GH secretion improved not only in GHD patients but also in GHS patients, pointing to an overall ongoing recovery of the GH-producing cells of the pituitary. Further studies are needed to evaluate the time frame and extend of GH recovery.

This study comprises a population with only large pituitary adenomas. While it offers a closely monitored and homogeneous study population, these patients may not be fully representative for all patients undergoing transsphenoidal pituitary surgery. However, the study population well reflects the group of patients in whom monitoring and improvement of long-term outcome is the most prominent clinical issue.

Within-subject variability of ITT results may be a limitation of our study. Reproducibility of the ITT has rarely been examined, particularly in patients with hypopituitarism, and available data are conflicting. In a study by Vestergaard et al. performed in healthy adults, the CV for peak cortisol levels during ITT was 10% (31). In a study by Pfeifer et al., the within-subject variation for cortisol in repeated ITT testing reached 21.5% in healthy subjects. For hypopituitary patients, it even reached 41.6% (range 3.5–92.7%) (32). GH is known to show even greater variability in repeated ITTs, especially in healthy subjects (32). However, in both studies, a single ITT did not misclassify hypopituitary patients and was therefore adequate for clinical decisions regarding GH and/or cortisol replacement. Nevertheless, in light of this rather large variability, special attention should be paid to the group of patients with cortisol peak values between 400 and 500 nmol/l. Seventeen percent of patients in our study had peak
values in this gray zone. If these patients had been tested again at the same time point, there is a possibility that they would have reached different, possibly higher cortisol values due to the within-subject variation of the test results. However, with random fluctuation of test results, one would expect changes in both directions with both lower and higher levels at repeated testing. Whereas we observed improvement of the GH and ACTH function in a relevant number of patients, none of the patients demonstrated deterioration to insufficiency, pointing to true recovery of pituitary function rather than spontaneous variability of stimulated peak values. We found that only patients with an initial peak cortisol > 230 nmol/l and peak GH > 0.3 μg/l respectively were able to recover, so we suggest 230 nmol/l for cortisol and 0.3 μg/l for GH as the cut-offs for re-testing patients later postoperatively.

In conclusion, a relevant number of patients improved their pituitary function during the late-postoperative follow-up. However, more studies are needed to confirm the extent of normalization during long-term follow-up and potential predictive factors. We would anticipate a greater number of recoveries to be identified if the clinician is aware of this possibility and is re-testing the patient with borderline results during later follow-up; this would prevent unnecessary lifelong substitution in a number of patients, which in itself may be harmful.

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 gratefully acknowledge the excellent support of the personnel at the Metabolic Laboratory, Division of Endocrinology, University Hospital of Essen, Germany.

References

4 Consensus Guidelines for the Diagnosis and Treatment of Adults with Growth Hormone Deficiency. Summary statement of the growth hormone research society workshop on adult growth hormone deficiency. Journal of Clinical Endocrinology and Metabolism 1998 83 379–381.


28 Petersenn S, Jung R & Bell FU. Diagnosis of growth hormone deficiency in adults by testing with GHRP-6 alone or in combination with GHRH: comparison with the insulin tolerance test. *European Journal of Endocrinology* 2002 **146** 667–672.


