Accepted Manuscript published as EJE-19-0732.R1. Accepted for publication: 09-Jan-2020

The effect of treatment on quality of life in patients with acromegaly: a prospective study.

TLC Wolters*[1], SHPP Roerink*[1,5], RBTM Sterenborg[1], MAEM Wagenmakers[1,2], O Husson[3,4], JWA Smit[1], ARMM Hermus[1], RT Netea-Maier[1].

*TLCW and SHPPR equally contributed to this manuscript

[1]Department of Internal Medicine, Division of Endocrinology, Radboud University Medical Center, Nijmegen, The Netherlands

[2]Department of Internal Medicine, Center for Lysosomal and Metabolic Diseases, Erasmus MC University Medical Center Rotterdam, Rotterdam, The Netherlands

[3]Department of Psychosocial Research and Epidemiology, Netherlands Cancer Institute – Antoni van Leeuwenhoek Hospital, Amsterdam, The Netherlands.

[4]Department of Clinical Studies, Institute of Cancer Research and Royal Marsden NHS Foundation Trust, London, UK.

[5]Department of Internal Medicine, Rijnstate Hospital, Arnhem, The Netherlands

Short title: Course of Quality of Life in Acromegaly.

Keywords: Health-related Quality of Life, Acromegaly, AcroQoL, RAND-36, ASE

Word count: 4901

Correspondence: Romana T. Netea-Maier (Romana.Netea-Maier@radboudumc.nl)

Abstract

Objective. Acromegaly has a negative influence on health-related quality of life (HRQoL). Previous studies provide limited information on the course of HRQoL during treatment. This study aims to assess the effect of treatment on the course of HRQoL at six predefined time points.

Design. This prospective study examines HRQoL in treatment-naive patients before and during the first 2.5 years of acromegaly treatment.

Methods. Therapy-naive acromegaly patients completed three validated questionnaires (RAND-36, AcroQoL and the Appearance Self Esteem (ASE)) at six predetermined time points before, during and after treatment. Outcomes were correlated to IGF1 levels and disease control status.

Results. Twenty-seven acromegaly patients completed the questionnaires at all time points. After treatment, all patients had controlled acromegaly. Scores of RAND-36 domains *General health*, *Vitality* and *Health change*, and all AcroQoL dimensions (except for *Relations*) improved during treatment ($P \leq 0.003$); the largest changes were detected during the first year. Gender influenced HRQoL scores, since AcroQoL scores significantly improved in males, but not in females. Over time, IGF1 levels were negatively correlated with HRQoL. After 2.5 years of follow-up, HRQoL of controlled patients was still lower than in the general population.

Conclusion. HRQoL of acromegaly patients was considerably reduced at diagnosis. Disease control was associated with an improvement of HRQoL scores. Males showed a more pronounced improvement than females. The largest changes were detected in the first year of treatment. However, HRQoL during and after treatment remained impaired in acromegaly patients, emphasizing the need of additional support.

Accepted Manuscript published as EJE-19-0732.R1. Accepted for publication: 09-Jan-2020

Introduction

Acromegaly is a rare endocrine disorder, caused by excessive production of Growth Hormone (GH) and Insulin-like-Growth Factor 1 (IGF1), which causes significant morbidity and, if left untreated, increased mortality (1). The prevalence lies between 28 and 137 cases per 1.000.000 people and the incidence ranges between 2 and 11 cases per 1.000.000 people per year (2). Due to the gradual progression of the disease and the frequent delay in diagnosis, patients are exposed to GH and IGF1 excess for a long period of time (3). GH and IGF1 stimulate growth of various tissues in the human body, such as connective tissue, bone and skin, which results in characteristic features as facial and hand disproportions (4-6). Remarkably, patients in long-term remission remain self-conscious about their facial appearance when compared to age- and gender-matched controls (7). Moreover, patients are at risk for a multitude of hormonal, cardiovascular, metabolic and respiratory complications, which are known to affect patients' well-being and functioning, and frequently result in an impaired health-related quality of life (HRQoL) (8-11).

Cross-sectional studies reported an impaired HRQoL in untreated acromegaly patients compared to treated acromegaly patients, but also in treated patients compared to normative values obtained in the general population or to non-acromegalic controls (matched for age, gender, social economic status and/or demographic characteristics) (12-16). In addition, prospective studies stated that HRQoL improved after treatment, irrespective of the used treatment modality, but remained lower compared to the general population (9, 16-20). However, the course of HRQoL during treatment has not been studied in detail, since previous prospective studies performed only two or at most three measurements. To date, a study evaluating HRQoL at more than three predetermined time points with a follow-up of more than two years in unselected untreated acromegaly patients which applied strict criteria of disease control (GH <0.4 ug/l during an oral glucose tolerance test (oGTT)) (21) has not been performed. Therefore, the aim of this study was to prospectively assess HRQoL at six predefined time points before, during, and after treatment in order to examine the impact of acromegaly treatment on HRQoL in more detail. Understanding the course of specific HRQoL domains over time may help

health care providers to anticipate to and proactively address specific problems in order to improve HRQoL of patients with acromegaly.

Materials and methods

Patients

All untreated adult patients with acromegaly who visited the outpatient clinic of the Radboud University Medical Center (Nijmegen, The Netherlands) between July 2012 and June 2016 were eligible to be included in this study. Active acromegaly was diagnosed according to the current international consensus: an increased IGF1 level (>2 SD above the age- and sex-adjusted mean) and an insufficient suppression of serum GH levels (≥0.4 μg/l) during an oGTT (1). Magnetic resonance imaging (MRI) of the pituitary gland was performed in each patient to identify a pituitary adenoma. The study design is depicted in Figure 1. At baseline (T_0), body length and weight were measured, non-fasted venous blood was drawn to determine IGF1 and GH levels, and information was obtained regarding estimated disease duration and presence of diabetes mellitus, hypertension, dyslipidemia, and pituitary hormone disturbances. For 2.5 years, patients visited our center every 6 months (T_1-T_5), and non-fasted IGF1 levels, disease status (controlled or uncontrolled) and weight were determined. At each visit, patients filled out three questionnaires (for details see below). HRQoL scores at T_2 and T_5 were of specific interest to examine the influence of surgery and the course of HRQoL over time.

After diagnosis, standard care was pre-treatment with a long-acting somatostatin receptor analogue (SSA; Lanreotide Autogel® in all patients) for 6 months, followed by endoscopic endonasal transsphenoidal adenomectomy (EETA). This treatment protocol is based on beneficial effects of short-term biochemical control on acromegaly-related comorbidities (e.g. diabetes mellitus, sleep apnea syndrome) and hence renders a lower peri-operative risk (22-24). In addition, SSA pre-treatment is suggested to have favourable effects on tumour size and structure, and improves short-term (and possible long-term) biochemical control after surgery (22, 25).

Patients who were not suitable for surgery were treated with primary medical therapy. In general, a disproportionately high perioperative risk (based on comorbidities), the expectation of no or very little

Accepted Manuscript published as EJE-19-0732.R1. Accepted for publication: 09-Jan-2020

benefit of pituitary surgery with regard to biochemical or local tumor control, or refusal by the patient are contra-indications for pituitary surgery in our center.

If biochemical control was not obtained by SSA monotherapy, the GH-receptor antagonist Pegvisomant (PEGV) or a dopamine agonist (DA) was added. In case of residual or recurrent disease after surgery, medical therapy was restarted postoperatively. When possible, patients underwent a second surgical intervention. In case of persistent IGF1 levels above the reference range despite maximal tolerable medical therapy, patients underwent radiotherapy.

Surgical control was defined as postoperative IGF1 levels within the sex- and age-adjusted reference range, combined with a random GH level <1 µg/l or a sufficient suppression of serum GH levels (GH <0.4 µg/l) during an oGTT, performed approximately four months after surgery, without use of GH- or IGF1 lowering drugs (1, 21, 26). *Biochemical control* was defined as IGF1 levels within the sex- and age-adjusted reference range with use of GH- or IGF1 lowering drugs (21). Both surgically controlled and biochemically controlled patients are considered *controlled patients. Active acromegaly despite treatment* was defined as elevated IGF1 levels despite treatment (surgery, radiotherapy, medication). *Adrenal insufficiency* (AI) was defined as a serum morning cortisol <100 nmol/l, after withdrawal of glucocorticoids for 24 h, or a maximal cortisol response <550 nmol/l during an insulin tolerance test (ITT) (27). Subclinical AI was defined as normal morning cortisol levels with an insufficient response (cortisol <550 nmol/l) during an ITT. Women were defined as *postmenopausal* when gonadotrophin levels were in the postmenopausal range and/or when they were older than 55 years. In premenopausal women and men, *hypogonadism* was defined as estrogen or total testosterone levels below the reference range. *Hypothyroidism* was defined as free thyroxin (fT4) serum levels <8 pmol/l (institutional reference range 8–22 pmol/l). *Hypopituitarism* was defined as the presence of one or more of the aforementioned pituitary hormonal deficiencies.

This study was conducted in accordance with the Declaration of Helsinki and approved by our local ethical committee (CMO region Arnhem-Nijmegen; 2012-131). All subjects signed informed consent prior to participation.

Hormone assays

Serum IGF1 and GH levels were determined using a chemiluminescent immunometric assay (Liaison, DiaSorin, Saluggia, Italy).

Questionnaires

Patients were asked to complete three health-related quality of life questionnaires (generic as well as a disease-specific) at six points in time: at diagnosis and subsequently every 6 months during a 2.5 year period [Figure 1].

RAND-36

The RAND-36 item Health Survey is a multidimensional general health assessment questionnaire that comprises 36 questions to evaluate nine aspects of health: physical functioning, social functioning, physical role limitation, emotional role limitation, mental health, bodily pain, vitality, general health perception, and health change. The RAND-36 uses subdimension scores, ranging from 0 to 100. A high score reflects a high HRQoL regarding that specific dimension. In this study, the Dutch translation of the RAND-36 was used (28). The RAND-36 is identical to the SF-36, except for the addition of the ninth subdimension *Health Change* and a slightly different scoring algorithm for the subscale *Bodily Pain* and *General Health Perceptions* (29, 30).

The RAND-36 scores obtained in our cohort were compared to scores from a cohort of 180 citizens (age 45-54 years) from the general Dutch population; these data were published in 1996 and are regarded as normative values (30). In addition, we compared scores from our patients to scores from two Dutch cohorts that were obtained more recently (2009-2015). One cohort consists of 162 healthy controls from the Nijmegen area with a mean age of 53.9 years who completed the RAND-36 questionnaire in 2009-2010 and 2014, respectively, while participating in two earlier studies by our group (7, 31). The last cohort comprises 1223 adult patients (mean age 53.9 years) who completed a rehabilitation program because of underlying diagnoses as chronic pain, brain injury, chronic fatigue syndrome, arthritis or various neurological problems (32). This cohort is used to compare scores from patients with treatment-naive and controlled acromegaly, which is regarded a chronic disease (11, 12, 33, 34), to scores that are representative for patients suffering from a chronic disease or disability.

 Accepted Manuscript published as EJE-19-0732.R1. Accepted for publication: 09-Jan-2020

AcroQoL

The AcroQoL is a disease-specific HRQoL questionnaire that consists of 22 questions; the answers are formulated as a Likert scale from 1 to 5. The lower the score, the larger the negative disease-related impact on quality of life. The AcroQoL is divided into two main categories: *Physical* and *Psychological* functioning. *Psychological* functioning is subdivided into two subdimensions: *Appearance* and *Personal Relationships*. Minimum and maximum scores for the *Physical* dimension range from 8 to 40 points. Both subdimensions of the *Appearance* scale can range from 7 to 35 points. Final scores of the (sub)dimensions and *Total* scores are converted to a scale from 0 to 100 (12, 35). This study used the validated Dutch translation of the AcroQoL (36). It has been suggested to use the three-subscales of the AcroQoL questionnaire *(Physical, Relations* and *Appearance*) instead of the *Total* score alone to provide a more specific representation of HRQoL (37). Scores of all subscales plus the *Total* score were used in this study.

ASE

To assess patients' satisfaction with their appearance, the Appearance Self-Esteem (ASE) scale, part of the Self-report State Self Esteem scale, was used. The Appearance subscale includes 5 questions, with answers based on a 5-point Likert scale ranging from 'not at all' (1) to 'extremely' (5); a score of 30 points corresponds to complete satisfaction with one's appearance (38).

Statistical analyses

Data were analysed with SPSS 25.0. Data are represented as numbers with percentages for categorical variables and as means with SD or as medians with minimum and maximum values for continuous variables, depending on the normality of the distribution, which was tested by the Shapiro-Wilk test. IGF1 levels were log-transformed prior to statistical analysis. The subgroup analysis based on gender and disease status was predefined based on previous literature (9, 39-42).

Baseline characteristics for men and women were tested for between-group differences using the independent samples T-test for normally distributed variables, and the Mann-Whitney U-test with additional Hodges-Lehmann tests to observe differences in confidence intervals for non-normally

distributed variables. Differences between categorical variables were tested with the Fisher's exact test.

Spearman rank correlation was used to determine correlations. For correlations between a continuous and a dichotomous variable, point biserial correlations (R_{pb}) were determined.

Prospective data were analyzed with multilevel models or the Friedman's two-way analysis, depending on the normality of the distribution. Non-normally distributed data were log-transformed and the derived residuals were tested for normality. If log-transformation did not result in normally distributed residuals, non-parametric tests were used. We used *time point* as factor for analyses on the total group of patients, and *time point* and *gender* as factors in the subgroup analyses. For comparisons based on disease status, we used *disease status* as factor. The Hodges-Lehman test was used to determine median differences between measurements in nonparametric tests. For categorical values generalized linear models with likelihood ratios were used.

All tests were two-tailed. For T-tests, Mann-Whitney U-tests or Fisher's exact tests, P-values of <0.05 were considered statistically significant. Correction for multiple testing was performed using the Holm-Bonferroni correction in all multilevel tests for repeated measurements over time and for comparisons between HRQoL scores between subgroups based on disease status (treatment-naive, controlled, and active despite treatment). After applying Holm-Bonferroni correction, p-values were considered significant when $p<0.05/(15\text{-rank}+1)$ for repeated measurements over time and $p<0.05/(3\text{-rank}+1)$ for the disease status. To calculate correlation coefficients on repeated observations within subjects, the method of Bland and Altman was used (43).

Results

Subject characteristics (Table 1)

Thirty-two patients were eligible for participation; four patients refused because of time constraints (N=2) or lack of willingness to participate in medical research (N=2), and one patient did not comprehend the Dutch language. The remaining 27 newly diagnosed treatment-naïve acromegaly

Accepted Manuscript published as EJE-19-0732.R1. Accepted for publication: 09-Jan-2020

patients (of which 12 males) were included and had a mean age of 51.0 ± 2.4 years. Females were older than males (55.8 ± 3.3 vs. 45 ± 3.3 years; P=0.03). Twenty patients had a macro-adenoma (>1 cm; 74.1%), six patients had a micro-adenoma and one patient was diagnosed with a GH-releasing hormone-producing bronchial intermediate-grade neuroendocrine tumor (NET).

Hypertension was present in 12 patients (44.4%), diabetes mellitus type 2 in five patients (18.5%) and dyslipidemia in four patients (15.4%) at T_0. The prevalence of DM and dyslipidemia did not change during the study. However, three (25%) patients were cured from hypertension postoperatively (after T_2).

One patient completed a female-to-male gender transition prior to participation in this study (44) and is regarded as a man in the analysis. During the whole study, four out of 162 questionnaires (2.7%) were missing due to cancelled visits and two patients skipped T_1 because they did not underwent pretreatment.

Disease control and acromegaly management (Supplementary Table 3)

Twenty-three patients (85.2%) were pre-treated with a SSA, for a mean duration of six months (range 5-11), followed by EETA. PEGV was added to the pre-treatment in one patient, and a DA in another, because of insufficiently controlled IGF1 levels with SSA monotherapy (Supplementary Table 3). One patient refused SSA pre-treatment and underwent EETA 6 weeks after baseline. The patient with the bronchial NET underwent a partial lobectomy without pre-treatment eight weeks after diagnosis. These two patients consequently skipped the second measurement time point (T_1). Two women did not undergo EETA. In one woman, surgery was not performed because of her old age and the rather mild activity of her acromegaly, which responded well to SSA treatment. The other woman had a non-resectable giant adenoma and a disproportionally increased peri-operative risk. They were primarily treated with a SSA and a SSA combined with PEGV, respectively. At T_2, 16 of the 25 surgically treated patients (64%) were in surgical control and nine patients (36%) had residual or recurrent disease and were treated by medication. Four patients repeatedly had normal IGF1 levels combined with a mildly disturbed oGTT (GH nadir 0.7-0.8 µg/l). Since their IGF1 values fell in the reference

range, they were considered surgically controlled patients in the analysis, according to the current Endocrine Society clinical practice guideline (1).

A second surgical procedure was performed in one patient between T_3 and T_4, after which she was surgically controlled. Next to treatment with a SSA and PEGV, one patient underwent gamma knife radiosurgery two weeks after T_2 and one patient underwent postoperative stereotactic radiotherapy between T_3 and T_4.

At T_5, the disease was adequately controlled in all patients: 17 patients (63%) were surgically controlled, the remaining 10 patients (37%) were biochemically controlled. Six used SSA monotherapy, one a SSA combined to a DA, and three a SSA combined with PEGV. The group mean IGF1 level at T_0 was 99.8 (95% confidence interval (CI) 85.7-114) nmol/l with a gender- and age-corrected SD score (SDS) of 9.6 (7.5-11.7) and decreased to 24.3 (21.8-26.8) nmol/l at T_5 with a SDS of 0.9 ((0.5-1.3); P<0.001).

Hormonal deficiencies

At T_0, eight men had hypogonadism (29.6%). One male had a primary hypogonadism as a result of a bilateral orchidopexy in childhood, and the other seven males had unsubstituted secondary hypogonadism. Eleven women were postmenopausal (73.3%). At T_2, three men had recovered from secondary hypogonadism. Four hypogonadal men were substituted with a stable dose of testosterone for at least three months and one male with mild and asymptomatic hypogonadism refused substitution therapy. One premenopausal woman developed secondary amenorrhea combined with estrogen values below the reference range after post-operative radiation therapy between T_4 and T_5.

At T_0, two patients had a history of hypothyroidism, one primary and one secondary, and were adequately substituted with levothyroxine for at least 3 months prior to the inclusion. One patient developed an autoimmune thyroiditis-related hypothyroidism (anti-TPO levels >1000 U/ml) between T_0 and T_1. At T_1, all patients were adequately substituted with levothyroxine for >3 months.

At T_0, one patient had secondary adrenal insufficiency treated with hydrocortisone replacement therapy and another patient had subclinical adrenal insufficiency with normal basal cortisol levels but an insufficient response (cortisol <550 nmol/l) to insulin-induced hypoglycemia during an ITT (27)

 Accepted Manuscript published as EJE-19-0732.R1. Accepted for publication: 09-Jan-2020

and required hydrocortisone substitution during stress situations. One female developed subclinical adrenal insufficiency after a second surgical approach between T_4 and T_5.

Questionnaire scores

RAND-36

Although scores of all RAND-36 dimensions improved between T_0 and T_5, only *General health, Physical role limitation, Vitality*, and *Health change* showed a significant increase after correction for multiple testing [Figure 2; Supplementary Table 1]. In general, scores of females seemed to show more variation between different time points and were lower at T_0 compared to males, although this was not statistically significant. At T_5 however, men scored higher at the subscale *Mental health* (79.3 95% CI 68.8-89.9 vs. 68.8 (60.3-77.3); P=0.03) and *Bodily pain* (89.1 (78.7-99.5) vs. 70.6 (58.1-83.1); P=0.03) [Figure 3; Supplementary Table 2]. The subdimensions *General health, Vitality* and *Health change* showed the most variable scores over time [Figure 2 & 3]. In females, the strongest improvements in HRQoL scores were observed from T_0 towards T_1 or T_2; after that time point HRQoL scores stabilized or slightly decreased. This pattern was paralleled with a steep increase of the *Health change* score towards T_2, and a mild decline thereafter. In men, a more gradual improvement was observed, although the *Health change* score also peaked at T_2 and declined again thereafter [Figure 3].

Disease status influenced the outcomes of *Physical role limitation, Vitality, Bodily pain, General health* and *Health change.* Patients with controlled disease reported higher scores compared to patients with untreated acromegaly (all $P \leq 0.006$). Compared to patients with active disease despite treatment, patients with controlled disease had higher scores for the dimensions *Physical role limitation, General health* and *Health change* (all $P \leq 0.009$).

When compared to scores obtained from previously published cohorts, patients scored worse at all time points compared to normative data from healthy controls with the same age (7, 29, 31). When compared to former rehabilitation patients (32), acromegaly patients scored comparable at baseline, but better at T_5 [Figure 4A].

AcroQoL

AcroQoL scores improved between T_0 and T_5, (P≤0.003), except for *Relations* [Figure 3, Figure 4B & Supplementary Table 2]. The largest changes took place between T_0 and T_2 for the dimensions *Total* (P<0.001)*, Physical* (P<0.001)*, Appearance* (P=0.001) and *Psychological* (P=0.004). The dimension *Relations* did not significantly change during treatment [Supplementary Table 1].

Assessing gender-specific AcroQoL scores, all scores (except *Relations*) improved in males (P≤0.002), but did not significantly change in females after correction for multiple testing [Supplementary Table 2]. At T_0, there were no differences in the AcroQoL scores between male and female patients; at T_5 men scored higher on the dimensions *Physical* (78.4 (67.5-89.3) vs. 61 (48.1-73.9); P=0.04) and *Relations* (89 (83.3-94.7) vs. 76.5 (67.1-84.3); P=0.02) compared to women.

During the whole study, patients with controlled disease had higher scores in all AcroQoL dimensions compared to patients with untreated disease (P<0.001), except for *Relations* (P=0.051). The same accounts when comparing controlled patients to patients with active disease despite treatment for the dimensions *Physical, Psychological* and *Appearance* (P≤0.003).

ASE

ASE scores showed an improvement during 2.5 years of treatment in the whole group, but this difference was not statistically significant after correction for multiple testing. Similar to the RAND-36 and AcroQoL results, females tended to score lower than males at most time points. At T_0, the ASE score was not statistically different in females and males, but at T_5, males scored higher (22. 3 (20.6-24) vs. 18.4 (16.5-20.3); P=0.004). Across the whole study duration, patients with controlled disease had higher ASE scores compared to patients with untreated disease (P=0.006).

Correlations

The changes in IGF1 levels over time showed a negative correlation with the changes in RAND-36 *General Health* (R-0.26; P=0.003), *Physical role limitation* (R-0.19; P 0.027), *Vitality* (R-0.21; P=0.019), and *Health change* (R-0.36; P<0.001). Changes in IGF1 levels were also correlated with

 Accepted Manuscript published as EJE-19-0732.R1. Accepted for publication: 09-Jan-2020

AcroQoL *Total* (R-0.21; P=0.02), *Physical* (R-0.25; P=0.005), and *Appearance* (R-0.18; P=0.04) scores and with ASE scores (R-0.18; P=0.04).

Summarizing, all HRQoL scores improved during follow-up, except for AcroQoL *Relations*. The largest changes were detected between T_0 and T_2. Apart from lower AcroQoL *Physical* scores in patients with hypopituitarism, HRQoL scores did not differ between patients with or without hypopituitarism.

Discussion

This study is the first to prospectively assess HRQoL in unselected, consecutive treatment-naïve patients with acromegaly before, during and after treatment at six predetermined time points. Our main finding is that HRQoL scores of both generic (RAND-36, ASE) and disease-specific (AcroQoL) questionnaires improved during follow-up, particularly during the first 6 to 12 months of treatment, after which HRQoL remained stable. However, generic HRQoL remained below the levels that were found in a healthy reference population of the same age.

Importantly, compared to patients with active acromegaly, better RAND-36 and AcroQoL scores were observed in controlled patients, with comparable HRQoL scores in surgically and biochemically controlled patients after 2.5 years of follow-up. This suggests that, in the first 2.5 years after diagnosis, normalization of circulating IGF1 levels, rather than the treatment modality used for disease control, affects HRQoL in patients with acromegaly.

Multiple (mostly cross-sectional) studies have reported a reduced HRQoL in acromegaly patients (regardless of disease status) compared to healthy controls (9, 13-17, 45, 46). Our study stands out because of its six predetermined time points, the inclusion of treatment-naïve patients only and its relatively long follow-up (2.5 years). Only one prospective study had a longer follow up (five years), but conducted only two measurements and did not include treatment-naive patients (17). Studies on treatment-naïve patients had a short follow-up duration (one year) during which disease control was not reached in all patients, and conduced only two or three measurements (19, 20).

We compared RAND-36 scores of our cohort to scores obtained from citizens from the general Dutch population with a similar age (7, 29, 31) and from Dutch rehabilitation patients (32). Despite the clear improvement during the follow-up, our patients scored lower on all RAND-36 subdimensions (except for *Health change*) compared to healthy controls. Compared to former rehabilitation patients, acromegaly patients scored comparable at baseline, but better at T_5.

Although normative values for the AcroQoL and ASE are not available for the general Dutch population, we expect that our patients also score lower on these questionnaires, given their lower RAND-36 scores and the lower HRQoL that has consistently been reported in (treated) acromegaly (12-16).

Multiple factors have been postulated to contribute to the persistently decreased HRQoL in treated acromegaly patients, such as persistent acromegaly-related comorbidities, changes in physical appearance (7) and psychosocial consequences of the disease. These finding support the concept that acromegaly, even in the context of surgical or biochemical hormonal control has the character of a chronic disease and a comparable negative impact on HRQoL (11, 12, 34).

It is not clear whether or to what extent disease status influences HRQoL (40). Some authors have reported worse HRQoL in patients with active disease compared to controlled disease (39, 41), whereas others found no differences (15, 47). In addition, there is little agreement on the relation between IGF1 levels and HRQoL since both weak negative (17, 18, 39, 48) or no (10, 15, 40, 45, 46) associations have been reported. However, we observed multiple negative correlations between changes in IGF1 levels and changes in HRQoL scores.

Last, the treatment modality may influence HRQoL. In general, achievement of remission is reported to increase HRQoL (18, 45, 49), although different treatment modalities are suggested to have distinct effects. surgery is reported to have a positive (19, 20) or neutral (47) effect on HRQoL. HRQoL was reported to improve, but not normalize, by treatment of patients with active acromegaly (treatment-naive or after surgery) with a SSA (40, 48, 50, 51), regardless of disease control (14, 17, 52, 53).

 Accepted Manuscript published as EJE-19-0732.R1. Accepted for publication: 09-Jan-2020

In our study, only two patients were primarily treated with a SSA respectively a SSA combined with PEGV; it was therefore not possible to assess the distinct effects of surgery or medication as primary approach. The same accounts for the effects of radiotherapy, which has been reported to be a negative predictor of HRQoL over time (9, 11, 40, 54, 55).

The inconsistent results regarding the influence of IGF1 levels, disease status and treatment modality on HRQoL might result from interference of comorbidities and other acromegaly-related persistent changes, which are known to significantly impact patients' well-being (11, 40, 49).

Patient characteristics are also known to impact on HRQoL. Hormonal deficiencies are associated in general with a lower HRQoL (46), but reports on the influence of hypopituitarism on HRQoL in acromegaly are discrepant (15, 17, 40, 42); most studies did not observe a relation between presence of hypopituitarism and HRQoL (40). In our study HRQoL scores did not differ in patients with and without pituitary hormone deficiencies, except for lower AcroQoL *Physical* scores in patients with hypopituitarism.

Gender also influences HRQoL; a significant improvement of AcroQoL scores was only observed in males in our study. Female gender was associated with lower HRQoL scores compared to male gender at all time points, which is in line with earlier studies (15, 16, 40, 42). Lower scores in women have been reported for various HRQoL questionnaires. The presence of a chronic health condition also has a stronger negative impact on HRQoL in females than in males. This has been attributed to differences in coping strategy, social economic status and the increased prevalence of depressive symptoms in women (56).

The largest changes in HRQoL scores were observed during the first year after diagnosis, thereafter scores stabilized at T_2 (six months after EETA) towards the end of the follow-up. This period is characterized by the induction of a period of SSA pretreatment, followed by surgery in most patients. The subsequent treatment-related changes that are involved likely impact on the physical, psychosocial and functional level. Consequently, the decrease in serum IGF1 levels was most pronounced in this timeframe. Earlier prospective studies found the largest changes after three months (19) and six months of treatment (45), respectively, which supports our findings.

Interestingly, the *Health change* score increased earlier in females compared to males, which may be caused by the (not statistically significant) larger proportion of females with disease control at T_1 and T_2.

The stabilization, or for some scales modest decline, in HRQoL scores that was observed after T_2 may be explained by the impact of irreversible acromegaly-related consequences as joint complaints, changed appearance and neurocognitive & psychological problems (7, 12-14, 16, 57).

In accordance with previous studies, *Appearance* was the most affected AcroQoL scale with the lowest scores both before and after treatment (7, 15, 17), and the subscale *Relations* was the less affected scale in these patients (58).

This study has some limitations. The major limitation is the relatively small cohort of patients, which makes it difficult to perform subgroup analyses to elucidate factors (e.g. comorbidities, treatment modality) that influence HRQoL in acromegaly and limits the reliability of our multilevel model analysis. However, given the low incidence of acromegaly and the scarcity of treatment-naive patients, it is difficult to obtain larger groups of patients. However, our results are consistent with previous reports and the residuals of our mixed models analyses were normally distributed, which indicates a normal distribution of our data. Regardless, influence of heterogeneity with regard to comorbidities, pituitary hormonal status or treatment modality cannot be ruled out. A second limitation is that since HRQoL might change on the longer-term, a longer follow-up period than in our study may represent long-term HRQoL in patients with controlled acromegaly more reliably.

In addition, our patients completed the same HRQoL questionnaires six times, which introduces the risk of 'response shift', which means that answers to a QoL assessment can change during follow-up in the absence of change in objective circumstances (59). This may be explained by changes in the patients' internal standards of interpretation (for example because of adaptation to certain limitations or circumstances) or changes in the way they prioritize or value different areas of life. Questionnaire response behavior can also be biased by current mood or context effects, for example, the answered questionnaires in an earlier part of the study may influence or changes their answer to subsequent questionnaires ('priming') (59). Importantly, previous studies reported good reliability and sensitivity

 Accepted Manuscript published as EJE-19-0732.R1. Accepted for publication: 09-Jan-2020

for change for the AcroQoL (10). The subscales of the SF-36 (which are nearly identical to the RAND-36 subscales) were found to be sufficiently reliable for use in repeated measurement designs since the within-subject subscale reliabilities ranged from acceptable to good (60). However, these potential sources of bias should be kept in mind when interpreting our results.

In conclusion, acromegaly patients reported an impaired HRQoL at diagnosis, which improved during the first 2.5 years of treatment, but did not normalize. The most pronounced changes were observed during the first year of treatment. Perception of HRQoL appeared to be gender-specific and disease control is associated with a better HRQoL. Active individualized management is recommended with an emphasis on improvement and maintenance of HRQoL in order to limit negative effects of disease-related complications, improve physical and psychosocial functioning, and coping strategies in acromegaly 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 investigator-initiated study was supported by an unrestricted research grant from IPSEN Pharmaceuticals.

Acknowledgements

We thank NM Rokx, IF Mustafajev, I Mommers and AFJ de Haan for their excellent help in the conductance of this study.

References

1. Katznelson L, Laws ER, Jr., Melmed S, Molitch ME, Murad MH, Utz A, Wass JA, Endocrine S. Acromegaly: an endocrine society clinical practice guideline. The Journal of clinical endocrinology and metabolism. 2014;99(11):3933-51.
2. Lavrentaki A, Paluzzi A, Wass JA, Karavitaki N. Epidemiology of acromegaly: review of population studies. Pituitary. 2017;20(1):4-9.
3. Colao A, Grasso LFS, Giustina A, Melmed S, Chanson P, Pereira AM, Pivonello R. Acromegaly. Nat Rev Dis Primers. 2019;5(1):20.
4. Colao A, Ferone D, Marzullo P, Lombardi G. Systemic complications of acromegaly: epidemiology, pathogenesis, and management. Endocrine reviews. 2004;25(1):102-52.
5. Ben-Shlomo A, Melmed S. Acromegaly. Endocrinology and metabolism clinics of North America. 2008;37(1):101-22, viii.
6. Hoevenaren IA, Wagenmakers MA, Roerink SH, Netea-Maier RT, Ulrich DJ, Maal TJ. Three-dimensional soft tissue analysis of the hand: a novel method to investigate effects of acromegaly. Eur J Plast Surg. 2016;39(6):429-34.
7. Roerink SH, Wagenmakers MA, Wessels JF, Sterenborg RB, Smit JW, Hermus AR, Netea-Maier RT. Persistent self-consciousness about facial appearance, measured with the Derriford appearance scale 59, in patients after long-term biochemical remission of acromegaly. Pituitary. 2015;18(3):366-75.
8. Lavrentaki A, Paluzzi A, Wass JA, Karavitaki N. Epidemiology of acromegaly: review of population studies. Pituitary. 2016.
9. Vandeva S, Yaneva M, Natchev E, Elenkova A, Kalinov K, Zacharieva S. Disease control and treatment modalities have impact on quality of life in acromegaly evaluated by Acromegaly Quality of Life (AcroQoL) Questionnaire. Endocrine. 2015;49(3):774-82.
10. Webb SM, Badia X, Surinach NL, Spanish AcroQol Study G. Validity and clinical applicability of the acromegaly quality of life questionnaire, AcroQoL: a 6-month prospective study. European journal of endocrinology / European Federation of Endocrine Societies. 2006;155(2):269-77.
11. Gadelha MR, Kasuki L, Lim DST, Fleseriu M. Systemic Complications of Acromegaly and the Impact of the Current Treatment Landscape: An Update. Endocrine reviews. 2019;40(1):268-332.
12. Biermasz NR, Pereira AM, Smit JW, Romijn JA, Roelfsema F. Morbidity after long-term remission for acromegaly: persisting joint-related complaints cause reduced quality of life. The Journal of clinical endocrinology and metabolism. 2005;90(5):2731-9.
13. Biermasz NR, van Thiel SW, Pereira AM, Hoftijzer HC, van Hemert AM, Smit JW, Romijn JA, Roelfsema F. Decreased quality of life in patients with acromegaly despite long-term cure of growth hormone excess. The Journal of clinical endocrinology and metabolism. 2004;89(11):5369-76.
14. Postma MR, Netea-Maier RT, van den Berg G, Homan J, Sluiter WJ, Wagenmakers MA, van den Bergh AC, Wolffenbuttel BH, Hermus AR, van Beek AP. Quality of life is impaired in association with the need for prolonged postoperative therapy by somatostatin analogs in patients with acromegaly. European journal of endocrinology / European Federation of Endocrine Societies. 2012;166(4):585-92.
15. T'Sjoen G, Bex M, Maiter D, Velkeniers B, Abs R. Health-related quality of life in acromegalic subjects: data from AcroBel, the Belgian registry on acromegaly. European journal of endocrinology / European Federation of Endocrine Societies. 2007;157(4):411-7.
16. Andela CD, Scharloo M, Pereira AM, Kaptein AA, Biermasz NR. Quality of life (QoL) impairments in patients with a pituitary adenoma: a systematic review of QoL studies. Pituitary. 2015;18(5):752-76.
17. Kyriakakis N, Lynch J, Gilbey SG, Webb SM, Murray RD. Impaired quality of life in patients with treated acromegaly despite long-term biochemically stable disease: Results from a 5-years prospective study. Clinical endocrinology. 2017;86(6):806-15.
18. Paisley AN, Rowles SV, Roberts ME, Webb SM, Badia X, Prieto L, Shalet SM, Trainer PJ. Treatment of acromegaly improves quality of life, measured by AcroQol. Clinical endocrinology. 2007;67(3):358-62.
19. Milian M, Honegger J, Gerlach C, Psaras T. Health-related quality of life and psychiatric symptoms improve effectively within a short time in patients surgically treated for pituitary

Accepted Manuscript published as EJE-19-0732.R1. Accepted for publication: 09-Jan-2020

tumors--a longitudinal study of 106 patients. Acta Neurochir (Wien). 2013;155(9):1637-45; discussion 45.
20. Fujio S, Arimura H, Hirano H, Habu M, Bohara M, Moinuddin FM, Kinoshita Y, Arita K. Changes in quality of life in patients with acromegaly after surgical remission - A prospective study using SF-36 questionnaire. Endocrine journal. 2017;64(1):27-38.
21. Giustina A, Chanson P, Bronstein MD, Klibanski A, Lamberts S, Casanueva FF, Trainer P, Ghigo E, Ho K, Melmed S, et al. A consensus on criteria for cure of acromegaly. The Journal of clinical endocrinology and metabolism. 2010;95(7):3141-8.
22. Colao A, Auriemma RS, Pivonello R. The effects of somatostatin analogue therapy on pituitary tumor volume in patients with acromegaly. Pituitary. 2016;19(2):210-21.
23. Yang C, Li G, Jiang S, Bao X, Wang R. Preoperative Somatostatin Analogues in Patients with Newly-diagnosed Acromegaly: A Systematic Review and Meta-analysis of Comparative Studies. Scientific reports. 2019;9(1):14070.
24. Fougner SL, Bollerslev J, Svartberg J, Oksnes M, Cooper J, Carlsen SM. Preoperative octreotide treatment of acromegaly: long-term results of a randomised controlled trial. European journal of endocrinology / European Federation of Endocrine Societies. 2014;171(2):229-35.
25. Dagistanli FK, Ozkaya HM, Kucukyoruk B, Biceroglu H, Metin D, Gazioglu N, Oz B, Kadioglu P, Ozturk M. Preoperative Somatostatin Analogue Treatment Might Trigger Apoptosis and Autophagy in Tumor Tissues of Patients with Acromegaly: A Pilot Study. Experimental and clinical endocrinology & diabetes : official journal, German Society of Endocrinology [and] German Diabetes Association. 2018;126(3):168-75.
26. Melmed S, Bronstein MD, Chanson P, Klibanski A, Casanueva FF, Wass JAH, Strasburger CJ, Luger A, Clemmons DR, Giustina A. A Consensus Statement on acromegaly therapeutic outcomes. Nature reviews Endocrinology. 2018;14(9):552-61.
27. Arlt W, Allolio B. Adrenal insufficiency. Lancet. 2003;361(9372):1881-93.
28. Van der Zee K SR. The measurement of the general health situation with the RAND-36: A manual.1993:[28 p.].
29. van der Zee KI, Sanderman R. The Measurement of the General Health with the RAND-36, A Handbook. Groningen, The Netherlands: Research Institute SHARE, UMCG; 2012.
30. VanderZee KI, Sanderman R, Heyink JW, de Haes H. Psychometric qualities of the RAND 36-Item Health Survey 1.0: a multidimensional measure of general health status. International journal of behavioral medicine. 1996;3(2):104-22.
31. Wagenmakers MA, Netea-Maier RT, Prins JB, Dekkers T, den Heijer M, Hermus AR. Impaired quality of life in patients in long-term remission of Cushing's syndrome of both adrenal and pituitary origin: a remaining effect of long-standing hypercortisolism? European journal of endocrinology / European Federation of Endocrine Societies. 2012;167(5):687-95.
32. Krops LA, Jaarsma EA, Dijkstra PU, Geertzen JH, Dekker R. Health Related Quality of Life in a Dutch Rehabilitation Population: Reference Values and the Effect of Physical Activity. PloS one. 2017;12(1):e0169169.
33. Romijn JA. The chronic syndromes after previous treatment of pituitary tumours. Nature reviews Endocrinology. 2016;12(9):547-56.
34. Adelman DT, Liebert KJ, Nachtigall LB, Lamerson M, Bakker B. Acromegaly: the disease, its impact on patients, and managing the burden of long-term treatment. Int J Gen Med. 2013;6:31-8.
35. Webb SM, Prieto L, Badia X, Albareda M, Catala M, Gaztambide S, Lucas T, Paramo C, Pico A, Lucas A, et al. Acromegaly Quality of Life Questionnaire (ACROQOL) a new health-related quality of life questionnaire for patients with acromegaly: development and psychometric properties. Clinical endocrinology. 2002;57(2):251-8.
36. Badia X, Webb SM, Prieto L, Lara N. Acromegaly Quality of Life Questionnaire (AcroQoL). Health and quality of life outcomes. 2004;2:13.
37. Felt JM, Jr., Depaoli SA, Pereira AM, Biermasz NR, Tiemensma J. Total score or subscales in scoring the acromegaly quality of life questionnaire: using novel confirmatory methods to compare scoring options. European journal of endocrinology / European Federation of Endocrine Societies. 2015;173(1):37-42.
38. Heatherton TF PJ. Development and validation of a scale for measuring state self-esteem. . J pers Soc Psychol. 1991;60:895-910.

39. Matta MP, Couture E, Cazals L, Vezzosi D, Bennet A, Caron P. Impaired quality of life of patients with acromegaly: control of GH/IGF-I excess improves psychological subscale appearance. European journal of endocrinology / European Federation of Endocrine Societies. 2008;158(3):305-10.
40. Geraedts VJ, Andela CD, Stalla GK, Pereira AM, van Furth WR, Sievers C, Biermasz NR. Predictors of Quality of Life in Acromegaly: No Consensus on Biochemical Parameters. Frontiers in endocrinology. 2017;8:40.
41. Trepp R, Everts R, Stettler C, Fischli S, Allemann S, Webb SM, Christ ER. Assessment of quality of life in patients with uncontrolled vs. controlled acromegaly using the Acromegaly Quality of Life Questionnaire (AcroQoL). Clinical endocrinology. 2005;63(1):103-10.
42. van der Klaauw AA, Kars M, Biermasz NR, Roelfsema F, Dekkers OM, Corssmit EP, van Aken MO, Havekes B, Pereira AM, Pijl H, et al. Disease-specific impairments in quality of life during long-term follow-up of patients with different pituitary adenomas. Clinical endocrinology. 2008;69(5):775-84.
43. Bland JM AD. Calculating correlation coefficients with repeated observations: Part 1-Correlation within subjects. Bmj. 1995;310:446.
44. Roerink S, Marsman D, van Bon A, Netea-Maier R. A missed diagnosis of acromegaly during a female-to-male gender transition. Archives of sexual behavior. 2014;43(6):1199-201.
45. Sardella C, Lombardi M, Rossi G, Cosci C, Brogioni S, Scattina I, Webb SM, Gasperi M, Martino E, Bogazzi F. Short- and long-term changes of quality of life in patients with acromegaly: results from a prospective study. Journal of endocrinological investigation. 2010;33(1):20-5.
46. Webb SM, Badia X. Quality of Life in Acromegaly. Neuroendocrinology. 2016;103(1):106-11.
47. Karaca Z, Tanriverdi F, Elbuken G, Cakir I, Donmez H, Selcuklu A, Durak AC, Dokmetas HS, Colak R, Unluhizarci K, et al. Comparison of primary octreotide-lar and surgical treatment in newly diagnosed patients with acromegaly. Clinical endocrinology. 2011;75(5):678-84.
48. Chin SO, Chung CH, Chung YS, Kim BJ, Kim HY, Kim IJ, Kim JG, Kim MS, Kim SY, Lee EJ, et al. Change in quality of life in patients with acromegaly after treatment with octreotide LAR: first application of AcroQoL in Korea. BMJ open. 2015;5(6):e006898.
49. Tseng FY, Chen ST, Chen JF, Huang TS, Lin JD, Wang PW, Huey-Herng Sheu W, Chang TC, Acromegaly Registry Study G. Correlations of clinical parameters with quality of life in patients with acromegaly: Taiwan Acromegaly Registry. Journal of the Formosan Medical Association = Taiwan yi zhi. 2019.
50. Caron PJ, Bevan JS, Petersenn S, Houchard A, Sert C, Webb SM, Group PI. Effects of lanreotide Autogel primary therapy on symptoms and quality-of-life in acromegaly: data from the PRIMARYS study. Pituitary. 2016;19(2):149-57.
51. Sonino N, Scarpa E, Paoletta A, Fallo F, Boscaro M. Slow-release lanreotide treatment in acromegaly: effects on quality of life. Psychother Psychosom. 1999;68(3):165-7.
52. Strasburger CJ, Karavitaki N, Stormann S, Trainer PJ, Kreitschmann-Andermahr I, Droste M, Korbonits M, Feldmann B, Zopf K, Sanderson VF, et al. Patient-reported outcomes of parenteral somatostatin analogue injections in 195 patients with acromegaly. European journal of endocrinology / European Federation of Endocrine Societies. 2016;174(3):355-62.
53. Hua SC, Yan YH, Chang TC. Associations of remission status and lanreotide treatment with quality of life in patients with treated acromegaly. European journal of endocrinology / European Federation of Endocrine Societies. 2006;155(6):831-7.
54. van der Klaauw AA, Biermasz NR, Hoftijzer HC, Pereira AM, Romijn JA. Previous radiotherapy negatively influences quality of life during 4 years of follow-up in patients cured from acromegaly. Clinical endocrinology. 2008;69(1):123-8.
55. Kreitschmann-Andermahr I, Buchfelder M, Kleist B, Kohlmann J, Menzel C, Buslei R, Koltowska-Haggstram M, Strasburger C, Siegel S. Predictors of Quality of Life in 165 Patients with Acromegaly: Results from a single-center study. Endocrine practice : official journal of the American College of Endocrinology and the American Association of Clinical Endocrinologists. 2017;23(1):79-88.

Accepted Manuscript published as EJE-19-0732.R1. Accepted for publication: 09-Jan-2020

56. da Rocha NS, Schuch FB, Fleck MP. Gender differences in perception of quality of life in adults with and without chronic health conditions: the role of depressive symptoms. J Health Psychol. 2014;19(6):721-9.
57. Martin-Rodriguez JF, Madrazo-Atutxa A, Venegas-Moreno E, Benito-Lopez P, Galvez MA, Cano DA, Tinahones FJ, Torres-Vela E, Soto-Moreno A, Leal-Cerro A. Neurocognitive function in acromegaly after surgical resection of GH-secreting adenoma versus naive acromegaly. PloS one. 2013;8(4):e60041.
58. Solomon E, Branisteanu D, Dumbrava A, Solomon RG, Kiss L, Glod M, Preda C. Executive functioning and quality of life in acromegaly. Psychol Res Behav Manag. 2019;12:39-44.
59. Blome C, Augustin M. Measuring change in quality of life: bias in prospective and retrospective evaluation. Value Health. 2015;18(1):110-5.
60. Kelly A, Rush J, Shafonsky E, Hayashi A, Votova K, Hall C, Piccinin AM, Weber J, Rast P, Hofer SM. Detecting short-term change and variation in health-related quality of life: within- and between-person factor structure of the SF-36 health survey. Health and quality of life outcomes. 2015;13:199.

Legends Tables & Figures

Table 1. Patient characteristics at baseline (T_0). Values are displayed as mean ± SD or median and range, depending on the normality of the distribution. P-values (males vs. females) were calculated using the independent samples T-test or the Mann-Whitney U-test depending on the normality of the distribution. Categorical parameters are displayed as number with percentage; differences were calculated with Fisher's exact test. GH: Growth Hormone; IGF1: insulin-like growth factor 1; SDS: standard deviation score; GHRH: Growth Hormone Releasing Hormone; NET: neuroendocrine tumour.

Figure 1. Overview of study and measurements. RAND-36: Research and Development-36 item Health Survey; AcroQol: Acromegaly Quality of Life questionnaire; ASE: Appearance Self-esteem questionnaire.; IGF1: Insulin-like Growth Factor 1; PreT: pre-operative treatment with a Somatostatin analogue, Pegvisomant and/or a dopamine agonist.

Figure 2. HRQoL score in total group of patients during 2.5 years of treatment. A-B: RAND-36; C: AcroQoL; D: ASE. At each time point, the mean is displayed. RAND-36: Research and Development-36 item Health Survey; AcroQol: Acromegaly Quality of Life questionnaire; ASE: Appearance Self-esteem questionnaire.

Figure 3. HRQoL scores in male and female patients during 2.5 years of treatment. A-B: RAND-36; C: AcroQoL; D: ASE. At each time point, the mean is displayed. RAND-36: Research and Development-36 item Health Survey; AcroQol: Acromegaly Quality of Life questionnaire; ASE: Appearance Self-esteem questionnaire.

Figure 4. Radar plots with RAND-36 (A) and AcroQoL (B) scores from patients at T_0 and T_5 and from age-matched healthy controls (7, 28, 31) and former rehabilitation patients (32). At each time point, the mean is displayed. RAND-36: Research and Development-36 item Health Survey; AcroQol: Acromegaly Quality of Life questionnaire; ASE: Appearance Self-esteem questionnaire.

 Accepted Manuscript published as EJE-19-0732.R1. Accepted for publication: 09-Jan-2020

Supplementary data

Supplementary Table 1. Prospective HRQoL questionnaire scores total group of patients. RAND-36: Research and Development -36 item Health Survey; AcroQoL: Acromegaly quality Of Life questionnaire; IGF1: Insulin-like Growth Factor 1; SDS: standard deviation score; ASE: Appearance Self-esteem questionnaire. Values are displayed as mean with 95% confidence interval. Disease status is depicted as number with percentage. All P-values were calculated using a multilevel model. *: P-value is significant after Holm-Bonferroni correction for multiple testing.

Supplementary Table 2. HRQoL questionnaires scores in males and females at T_0, T_2, and T_5. RAND-36: Research and Development-36 item Health Survey; AcroQoL: Acromegaly quality Of Life questionnaire; IGF1: Insulin-like Growth Factor 1; SDS: standard deviation score; ASE: Appearance Self-esteem questionnaire. Values are displayed as mean with 95% confidence interval. Differences between males and females at T_0 and at T_5 were calculated with a Mann-Whitney U test. Differences between T_0 and T_2 and between T_0 and T_5 in the subgroups of patients were calculated with a multilevel linear model. *: P-value is significant after Holm-Bonferroni correction for multiple testing.

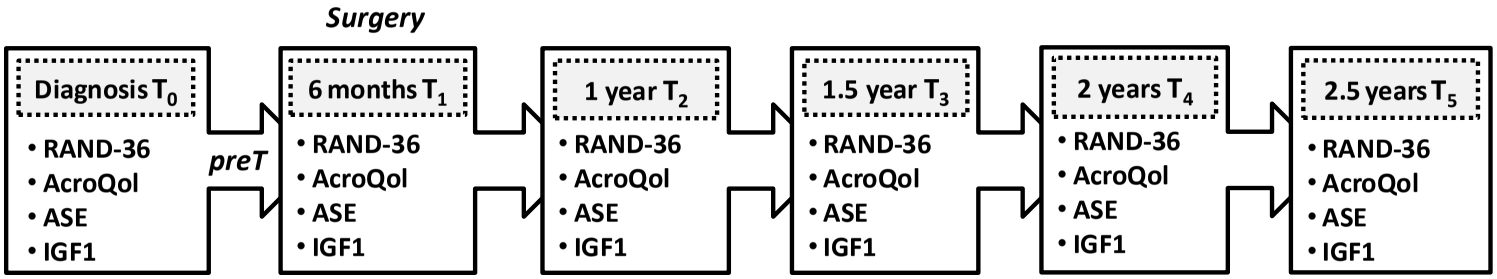
Supplementary Table 3. Course of treatment and IGF1 levels in individual subjects. IGF1: insulin-like growth factor 1; O: data not available; M: medical treatment; MS: medical treatment with a somatostatin analogue (SSA); MP: medical treatment with pegvisomant (PEGV); MD: medical treatment with a dopamine agonist; MS + MP: medical treatment SSA combined with PEGV; MS + MD: medical treatment SSA combined with a dopamine agonist; S: surgery; SRT: stereotactic radiotherapy; GRS: gamma knife radiosurgery; BC: biochemical control with use of medication; reS: second surgical procedure; RD: residual or recurrent disease.

In patients with medical pretreatment, surgery and cessation of medical treatment took place simultaneously. The moment or timespan of treatment in months is displayed between parentheses. They grey marked areas indicate the time period in which the patient had controlled acromegaly (i.e. IGF1 values in the sex- and age-adjusted reference range). O*: IGF1 levels at this study visit were missing. Patients were regarded as ‘controlled’ or ‘uncontrolled’ based on whether the IGF1 levels

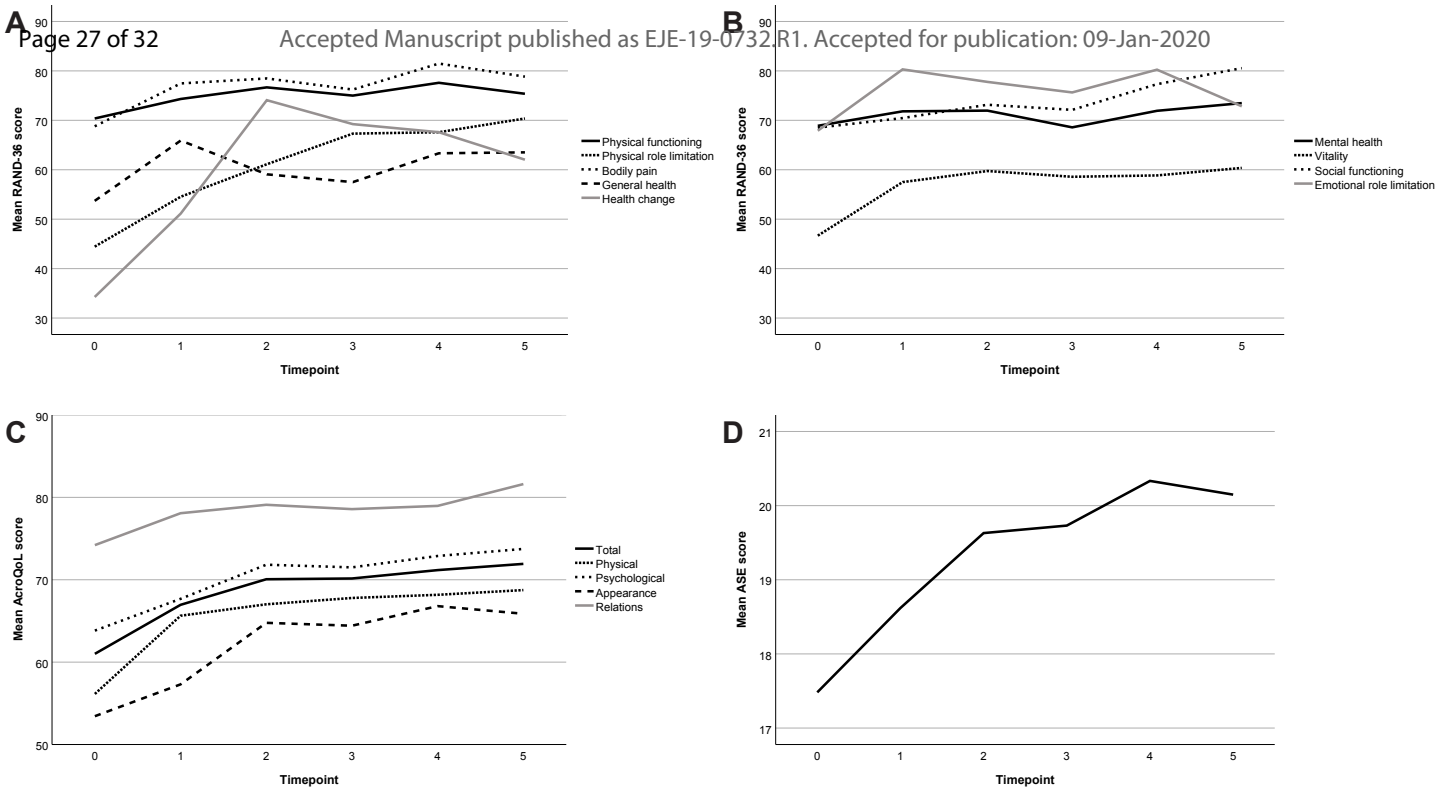
that were measured before and after the omitted study visits were in the normal range for age and sex; O**: these patients did not undergo medical pretreatment prior to surgery and therefore did not undergo study visit T_1.

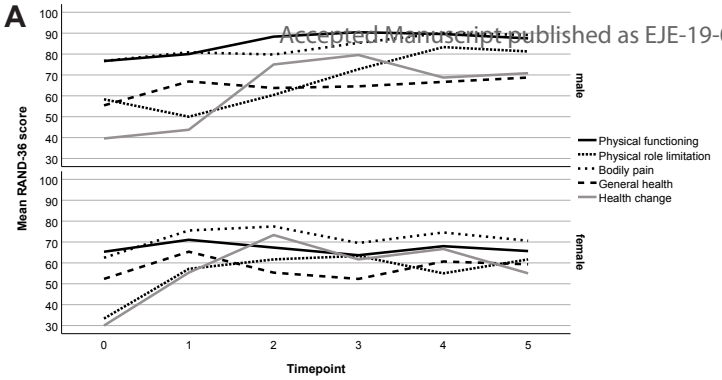
Accepted Manuscript published as EJE-19-0732.R1. Accepted for publication: 09-Jan-2020

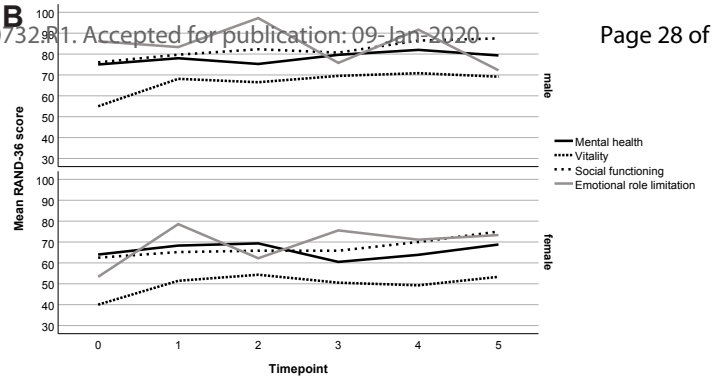
Table 1. Patient characteristics	**T_0 (N=27)**	**Males (N=12)**	**Females (N=15)**	**P**
Gender (N male, %)	12 (44.4)	-	-	
Age (years)	51.0 ± 13.1	45 ± 11.6	55.8 ± 12.6	0.03
Body Mass Index (kg/m^2)	29 ± 5.1	29.6 ± 3.8	28.6 ± 6.0	0.62
GH (µg/l)	8.1 (1-127.7)	10.6 (2.4-29.8)	8.1 (1-127.7)	0.76
IGF1 (nmol/l)	97.3 (40.6-208)	112.6 (61.9-208)	84.6 (40.6-146)	0.06
IGF1 SDS	7.6 (3.5-23.2)	11.2 (6.8-23.2)	6.2 (3.5-16.7)	0.001
Duration of symptoms until diagnosis (years)	7.0 (2.0 – 28)	5 (2-20)	10 (2-28)	0.15
Tumor type (N, %)				
Microadenoma	6 (22.2)	3 (25)	3 (20)	
Macroadenoma	20 (74.1)	8 (66.7)	12 (80)	0.65
GHRH-producing NET	1 (3.7)	1 (8.3)	0 (0)	
Comorbidities (N, %)				
Hypertension	12 (44.4)	4 (33.3)	8 (53.3)	0.44
Dyslipidemia	4 (15.4)	1 (9.1)	3 (20)	0.61
Diabetes mellitus	5 (18.5)	3 (25)	2 (13.3)	0.63
Hypogonadism	8 (29.6)	8 (66.7)	0 (0)	<0.001
Hypocortisolism	2 (7.4)	2 (16.6)	0 (0)	0.19
Hypothyroidism	2 (7.4)	1 (8.3)	1 (6.7)	1

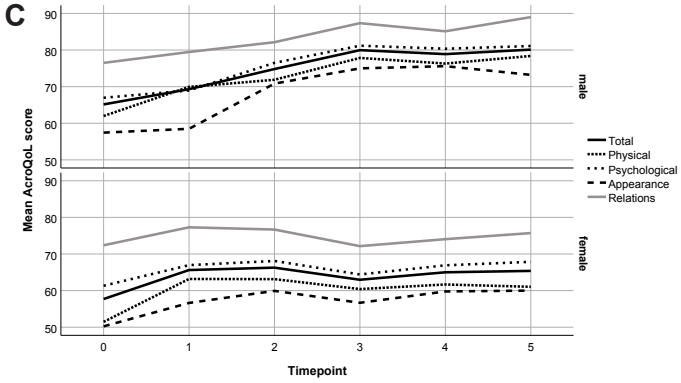


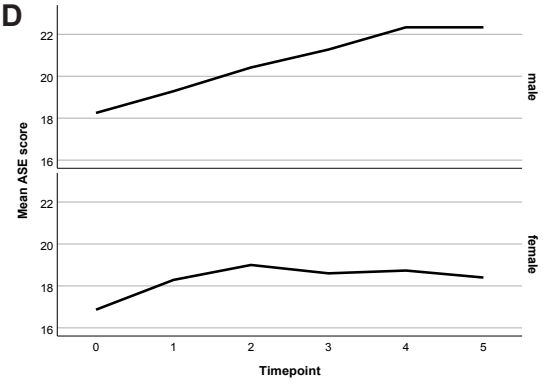
Accepted Manuscript published as EJE-19-0732.R1. Accepted for publication: 09-Jan-2020



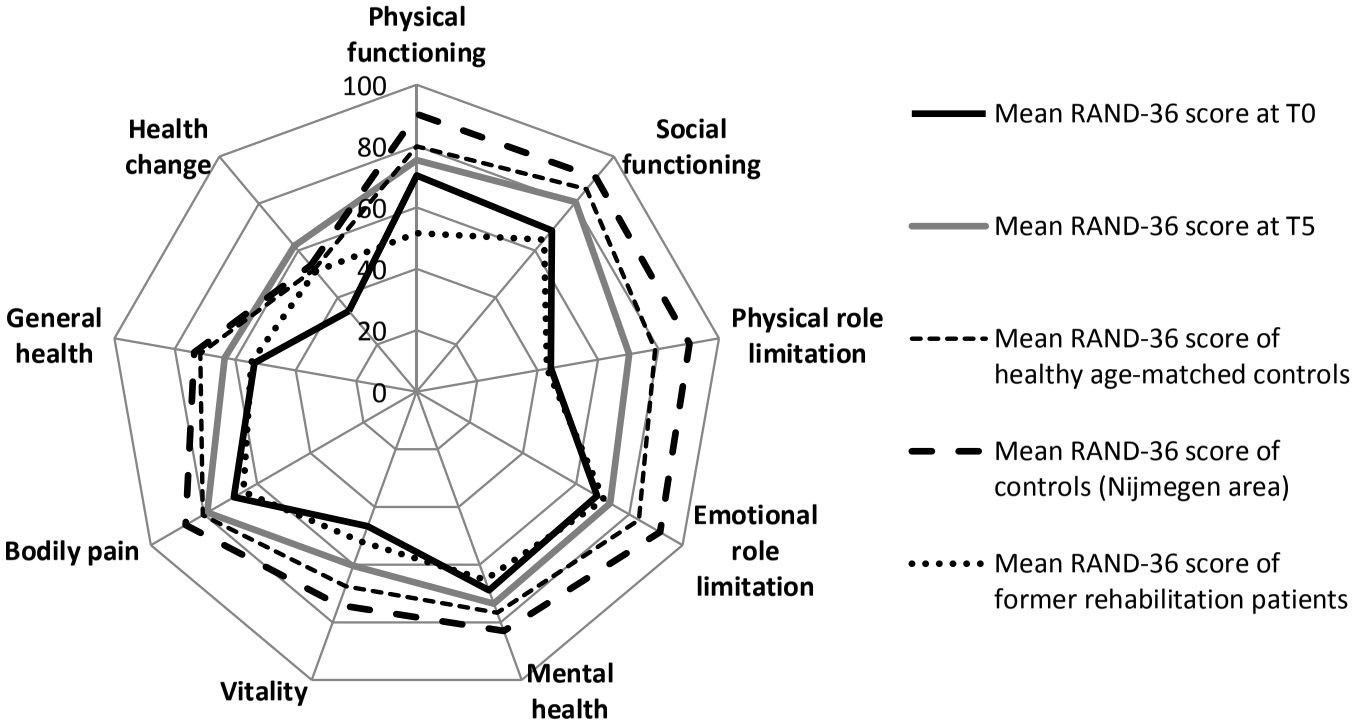




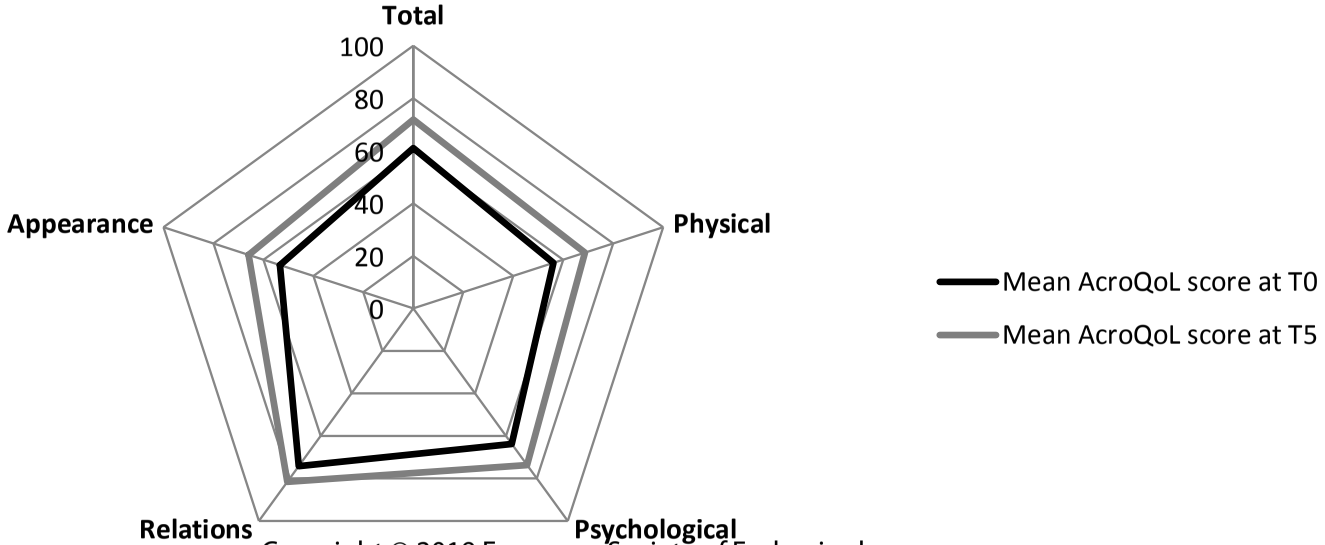




A Change in RAND-36 scores over time



B Change in AcroQoL scores over time



Supplementary Table 1 **Prospective HRQoL scores**	**T_0** **Baseline (N=27)**	**T_1** **(N=22)**	**T_2** **(N=27)**	**T_3** **(N =26)**	**T_4** **(N=27)**	**T_5** **(N=27)**	**P-value** **$T_0 - T_2$**	**P-value** **$T_0 - T_5$**
RAND-36 score								
Physical functioning	70.4 (59.3-81.5)	74.3 (63.2-85.5)	76.7 (66.9-86.4)	75 (63.8-86.2)	77.6 (66.6-88.6)	75.4 (63-87.8)	0.03	0.16
Social functioning	68.5 (57.5-79.5)	70.5 (60.7-80.2)	73.1 (63.7-82.6)	72.1 (62.1-82.1)	77.3 (67.8-86.8)	80.6 (72.1-89)	0.26	0.04
Physical role limitation	44.4 (26.8-62.1)	54.6 (36.2-7.9)	61.1 (42.6-79.6)	67.3 (51.7-82.9)	67.6 (50.7-84.5)	70.4 (53-87.7)	0.03	0.004
Emotional role limitation	67.9 (51.8-84)	80.3 (64.7-95.9)	77.8 (61.8-93.7)	75.6 (58.8-92.5)	80.3 (67.4-93.1)	72.8 (56.9-88.8)	0.22	0.68
Mental health	68.9 (60.7-77)	71.8 (64.6-79)	72 (65.4-78.6)	68.6 (60.2-77)	71.9 (64.6-79.3)	73.5 (67-80)	0.32	0.14
Vitality	46.7 (37.4-55.9)	57.5 (48.1-66.9)	59.7 (52.6-66.9)	58.6 (48.8-68.4)	58.9 (48.5-69.2)	60.4 (51.6-69.1)	<0.001*	0.002*
Bodily pain	68.8 (58-79.5)	77.5 (68.9-86)	78.5 (69.9-87)	76.2 (68-84.4)	81.5 (74.5-88.5)	78.8 (70.2-87.5)	0.02	0.06
General health perceptions	54.4 (47.4-61.5)	65.9 (58.4-73.5)	59 (50.8-67.3)	57.5 (49.6-65.4)	63.9 (55.5-72.3)	64.3 (55.4-73.2)	0.12	0.002*
Health change	34.3 (24.7-43.8)	51.1 (39.1-63.2)	74.1 (66.1-82.1)	69.2 (57.7-80.8)	67.6 (57-78.2)	62 (52.4-71.7)	<0.001*	<0.001*
AcroQoL score								
Total score	61 (53.8-68.2)	67 (59.5-74.4)	70.1 (63.4-76.7)	70.4 (63.4-77.4)	71.5 (64.2-78.7)	71.9 (65.4-78.4)	<0.001*	<0.001*
Physical	56.1 (47.6-64.7)	65.6 (56.6-74.7)	67 (58.9-75.2)	67.8 (59-76.6)	68.2 (59.6-76.7)	68.7 (60-77.5)	<0.001*	<0.001*
Psychological	63.8 (56.4-71.3)	67.7 (59.6-75.8)	71.8 (65.2-78.4)	71.8 (64.9-78.8)	73.3 (66-80.7)	73.7 (67.5-80)	0.004	0.003*
Appearance	53.4 (45.4-61.5)	57.3 (48.1-66.5)	64.8 (56.6-73)	64.4 (55.4-73.4)	67.3 (59.1-75.6)	65.9 (58-73.7)	0.001*	0.001*
Relations	74.2 (65.5-83)	78.1 (70-86.2)	79.1 (73-85.2)	79.3 (73-85.6)	79.4 (72-86.7)	81.6 (75.9-87.3)	0.10	0.21
ASE score	18.1 (16.1-20.2)	18.6 (16.6-20.6)	19.4 (17.9-21.3)	19.7 (18-21.5)	20.3 (18.1-22.6)	20.2 (18.7-21.6)	0.047	0.047
IGF1 (nmol/l)	99.8 (85.7-114)	41.9 (28.7-55.2)	31.4 (24.9-37.9)	25.5 (21.5-29.5)	24.3 (21.7-26.9)	24.3 (21.8-26.8)	<0.001*	<0.001*
IGF1 SDS	9.6 (7.5-11.7)	3.4 (1.6-5.1)	1.8 (1-2.7)	1 (0.4-1.6)	0.9 (0.5-1.2)	0.9 (0.5-1.3)	<0.001*	<0.001*
BMI	29 (27-31)	28.7 (26.6-28.6)	29.8 (27.3-32.3)	29.5 (27-32)	29.4 (27-31.7)	28.7 (26.4-31)	0.09	0.13
Disease status								
Untreated (%)	27 (100)	0	0	0	0	0		
Surgical control (%)	0 (0)	0	16 (59.3)	16 (59.3)	17 (63)	17 (63)	<0.001*	<0.001*
Biochemical control (%)	0 (0)	11 (44)	3 (11.1)	7 (25.9)	9 (33.3)	10 (37)		
Active despite treatment (%)	0 (0)	14 (56)	8 (29.6)	4 (14.8)	1 (3.7)	0		

Accepted Manuscript published as EJE-19-0732.R1. Accepted for publication: 09-Jan-2020

Supplementary Table 2. Male-female comparison	**T_0 Males (N =12)**	**T_2 Males (N =12)**	**T_5 Males (N = 12)**	**P Males T_0-T_2**	**P Males T_0-T_5**	**T_0 Females (n = 15)**	**T_2 Females (N =12)**	**T_5 Females (n = 15)**	**P Females T_0–T_2**	**P Females T_0-T_5**	***P Female - male T_0***	***P Female - male T_2***	***P Female - male T_5***
RAND-36 score													
Physical functioning	76.7 (61.1-92.3)	88.3 (80.3-96.4)	87.5 (76.1-99)	*0.007*	*0.003**	65.3 (48.5-82.2)	67.3 (51.8-82.9)	65.7 (45.4-85.9)	*0.6*	*0.86*	*0.33*	*0.03*	*0.12*
Social functioning	76 (58.6-93.5)	82.3 (67.7-96.8)	87.5 (75.8-99.2)	*0.31*	*0.42*	62.5 (47.2-77.8)	65.8 (53.2-78.5)	75 (62.5-87.5)	*0.56*	*0.25*	*0.2*	*0.05*	*0.11*
Physical role limitation	58.3 (30.1-86.5)	60.4 (31.3-89.5)	81.3 (56.7-105.8)	*0.85*	*0.024*	33.3 (9.5-57.1)	61.7 (34.5-88.8)	61.7 (35.6-87.8)	*0.005*	*0.03*	*0.14*	*0.9*	*0.3*
Emotional role limitation	86.1 (67.1-105.2)	97.2 (91.1-103.3)	72.2 (43.9-100.5)	*0.35*	*0.26*	53.3 (29.4-77.3)	62.2 (35.3-89.1)	73.3 (52.2-94.5)	*0.4*	*0.18*	*0.05*	*0.03*	*0.94*
Mental health	75 (62.7-87.3)	75.3 (62.7-87.8)	79.3 (68.8-89.9)	*0.96*	*0.56*	64 (52.4-75.6)	69.3 (61.6-77)	68.8 (60.3-77.3)	*0.2*	*0.22*	*0.2*	*0.19*	*0.03*
Vitality	55 (41-69)	66.5 (56.6-76.4)	69.2 (58.4-79.9)	*0.004*	*0.04*	40 (27.3-52.7)	54.3(44-64.7)	53.3 (40.2-66.5)	*0.004*	*0.06*	*0.13*	*0.09*	*0.1*
Bodily pain	76.7 (61.1-92.3)	79.8 (61.5-98.1)	89.1 (78.7-99.5)	*0.62*	*0.18*	62.5 (46.8-78.1)	77.4 (69.4-85.8)	70.6 (58.1-83.1)	*0.008*	*0.08*	*0.18*	*0.27*	*0.03*
General health perceptions	57.1 (45.7-68.5)	63.8 (49.9-77.6)	68.8 (54.5-83)	*0.14*	*0.12*	52.3 (42.4-62.3)	55.3 (44.2-66.4)	60.7 (48.2-73.2)	*0.45*	*0.015*	*0.71*	*0.36*	*0.39*
Health change	39.6 (23.8)	75 (63.3-86.7)	70.8 (55.9-85.7)	*<0.001**	*<0.001**	30 (17-43)	73.3 (61.1-85.6)	55 (42-68)	<0.001*	*<0.001**	*0.41*	*0.98*	*0.08*
AcroQoL score													
Total score	65.1 (55.5-74.8)	74.8 (67.3-82.3)	80.1 (74.7-85.5)	*0.008*	*<0.001**	57.7 (46.5-68.9)	66.3 (55.4-77.1)	65.4 (55.1-75.6)	*0.008*	*0.1*	*0.34*	*0.28*	*0.05*
Physical	62 (50.4-73.6)	71.9 (59.6-84.2)	78.4 (67.5-89.3)	*0.01*	*<0.001**	51.5(38.4-64.6)	63.1 (51.4-74.9)	61 (48.1-73.9)	*0.001**	*0.02*	*0.24*	*0.27*	*0.04*
Psychological	67 (57.5-76.4)	76.5 (69.5-83.5)	81.1 (74.8-87.4)	*0.02*	*0.002**	61.3 (49.3-73.3)	68.1 (57.2-79)	67.9 (58.2-77.6)	*0.07*	*0.45*	*0.73*	*0.41*	*0.05*
Appearance	57.4 (47-67.9)	70.8 (62-79.7)	73.2 (64.6-81.8)	*0.01*	*0.002**	50.2 (37.5-63)	59.9 (46.6-73.3)	60 (47.6-72.4)	*0.04*	*0.19*	*0.46*	*0.27*	*0.26*
Relations	76.5 (63.9-89.1)	82.1 (74.8-89.4)	89 (83.3-94.7)	*0.2*	*0.05*	72.4 (58.9-85.9)	76.7 (66.7-86.6)	76.5 (67.1-84.3)	*0.27*	*0.83*	*0.59*	*0.54*	*0.02*
ASE score	19.7 (16.2-23.1)	20.4 (17.6-23.2)	22.3 (20.6-24)	*0.51*	*0.11*	16.9 (14.2-19.5)	19 (15.7-21.3)	18.4 (16.5-20.3)	*0.04*	*0.36*	*0.13*	*0.23*	*0.004*
IGF1 (nmol/l)	114.8 (89.9-139.8)	31.2 (23.3-39.2)	27.4 (23.4-31.4)	*<0.001**	*0.002**	87.8 (72-103.6)	31.5 (21.2-41.8)	21.9 (18.9-24.8)	*<0.001**	*<0.001**	*0.06*	*0.69*	*0.014*
IGF1 SDS	12.9 (9.5-16.3)	1.6 (0.4-2.8)	1.2 (0.5-1.8)	*<0.001**	*<0.001**	7 (5-8.9)	2 (0.6-3.3)	0.7 (0.3-1.2)	*<0.001**	*0.002**	*0.001*	*0.45*	*0.2*

Supplementary Table 3. Course of treatment & IGF1 levels			IGF1 (nmol/l)					
No	**Sex**	**Treatment**	**T_0**	**T_1**	**T_2**	**T_3**	**T_4**	**T_5**
1	M	MS (0-6) + S: cured	208	O*	O*	4.6	15.1	13
2	M	MS (0-6) + S: cured	80.7	48.9	34.8	28.4	30.1	28
3	M	MS (0-6) + S: cured	111	21.6	O	23.8	22.4	29.8
4	M	MS (0-7) + S: cured	61.9	41	25.1	34.4	23.8	29.4
5	M	MS (0-6) + S: cured	66.7	28.6	8.5	26.2	23.6	34.1
6	F	MS (0-6) + S: cured	108	26.3	23.6	16	13.3	18.2
7	F	MS (0-6) + S: cured	77.2	41.4	20.4	23.3	25.3	22.4
8	F	MS (0-6) + S: cured	97.3	19.3	23.5	27.5	26.9	25.5
9	F	1. MS (0-6) + S: RD 2. MS (14-30): BC (16-30)	79.8	40.3	36.5	16.2	25.8	28.3
10	F	MS (0-6) + S: cured	118.7	9	24.7	19.5	16.1	19.4
11	M	S (2): cured	122.1	O**	34.2	O*	32.3	30.4
12	F	Primary MS (0-30): BC (6-30)	40.6	22	24.7	27.1	26.3	28
13	F	1. MS (0-6) + S: RD 2. MS (13-30): BC (15-30)	83.7	19.3	48.3	13.6	15	17.8
14	F	MS (0-6) + S: cured	63.8	15.1	11	10.6	11.4	11.7
15	F	1. MS (0-6) + S: RD 2. MS (12-24) + P (15-22) + reS (24): cured	118.7	126	54.4	48	27.5	25.8
16	M	1. MS (0-6) + S: RD 2. MS (11-30): BC (12-30)	109.5	32.4	27.4	26.6	25.9	20.3
17	F	1. MS (0-6) S: RD 2. MS (8-30) + MP (11-30) + SRT (20-21): BC (29-30)	146	114.1	83.2	37.4	36.7	25.9
18	M	1. S (1.5): RD 2. MS (2.5-30): BC (14-30)	132.5	O**	40.7	31.4	29.7	28
19	F	MS (0-11) + S: cured	53.5	15.2	17.2	15	17.2	15.6
20	F	1. MS (0-6) + S: RD 2. MS (8-30): BC (10-30)	85.2	O*	20.2	22.7	23	21.1
21	M	1. MS + MD (0-6) + S: RD 2. MS + MD (14-30): BC (24-30)	95.7	73.9	34.9	35.4	29.7	34
22	M	1. MS (0-6) + MP (4-6) + S: RD 2. MS (9-30) + MP (10-30) + GRS (12): BC (15-30)	114.2	59.9	45.9	24.2	23	27
23	M	MS (0-6) + S: cured	130.4	45	40.3	32	28.8	21.9
24	F	MS (0-6) + S: cured	105.8	32.1	31.6	27.6	33.3	28.7
25	M	MS (0-5) + S: cured	145.4	38.5	20.6	30.3	29.6	33
26	F	MS started at T_0, thereafter lost to follow-up	84.5		Lost to follow-up			
27	F	Primary MS (0-30) + MP (9-30): BC (22-30)	84.6	52.3	36.1	42.5	26.2	19.3
28	F	MS (0-6) + S: cured	54.4	22.2	16.8	18.7	17.6	17.5