International Multicenter Validation Study of the SAGIT® Instrument in Acromegaly

Andrea Giustina,[1] Marcello D. Bronstein,[2] Philippe Chanson,[3] Stephan Petersenn,[4] Felipe F. Casanueva,[5] Caroline Sert,[6] Aude Houchard,[6] Shlomo Melmed[7]

1. Institute of Endocrine and Metabolic Sciences, Vita-Salute San Raffaele University and IRCCS San Raffaele Hospital, Milan, Italy

2. Hospital das Clínicas da Universidade de São Paulo, São Paulo, Brazil

3. Université Paris-Saclay, Inserm, Physiologie et Physiopathologie Endocriniennes, Assistance Publique-Hôpitaux de Paris, Hôpital Bicêtre, Service d'Endocrinologie et des Maladies de la Reproduction, Le Kremlin-Bicêtre, France

4. ENDOC Centre for Endocrine Tumours, Hamburg, Germany

5. Instituto Salud Carlos III, Santiago de Compostela University – CIBER de Fisiopatología de la Obesidad y Nutrición (CIBERobn), Santiago de Compostela, Spain

6. Ipsen, Boulogne-Billancourt Cedex, France

7. Pituitary Center, Cedars-Sinai Medical Center, Los Angeles, CA, USA

Study Registration: (ClinicalTrials.gov NCT02539927)

Correspondence and reprint requests: Professor Andrea Giustina, MD, Institute of Endocrine and Metabolic Sciences, IRCCS San Raffaele Hospital, Via Olgettina 60, 20132 Milano, Italy. E mail: giustina.andrea@hsr.it

Funding: This study was sponsored by Ipsen

Disclosures

AG has been a consultant for Genevant, Ipsen, Pfizer and Recordati and has received research grants from Ipsen, Pfizer and Novartis. **MB** has received research grants for clinical trials as principal investigator from Crinetics, Ipsen and Novartis; is a member of steering committees and advisory boards for Chiasma, Ipsen and Novartis; and is a speaker for Ipsen and Novartis. **PC** has received unrestricted research and educational grants from Ipsen, Novartis and Pfizer (on behalf of the Service of Endocrinology and Reproductive Diseases, Hôpitaux Universitaires Paris-Sud and Association Recherche Endocrinologie Bicêtre); has served as an investigator for clinical trials funded by Novartis, Pfizer, Ipsen, Italpharmaco, Antisense and Chiasma and has given lectures for Ipsen, Novartis and Pfizer (all fees and honoraria paid to PC's Institution or Research Association). **SP** has been a speaker and advisory board member for Ipsen, Novartis and Pfizer. **FC** has given lectures and acted as an advisor for Pfizer and Ipsen. **CS** and **AH** are employees of Ipsen. **SM** has received research grants from and acted as an investigator for Pfizer, Ipsen, and has acted as a consultant for Chiasma, Crinetics and Ionis.

Abstract

Context

The SAGIT® instrument (SAGIT) has been developed to enable accurate characterization of acromegaly disease activity.

Objective

Evaluate the ability of SAGIT to discriminate between acromegaly disease control status.

Design

Multicenter, non-interventional, prospective and retrospective, longitudinal study.

Settings and Patients

Academic and private clinical practice sites; patients aged ≥18 years with diagnosis of controlled (n=109) or non-controlled (n=105) acromegaly, assessed by clinical global evaluation of disease control (CGE-DC) questionnaire, investigator therapeutic decision and international guidelines. Control status was not determined at baseline for 13 patients. As a limited number of patients were enrolled retrospectively (N=9), all presented analyses are based on the prospective population (N=227).

Methods

Patients were assessed over a two-year follow-up period. Classification and regression tree (CART) analyses were performed to investigate how the SAGIT components at baseline (signs/symptoms [S], associated comorbidities [A], GH levels [G], IGF-1 levels [I], tumor features [T]) discriminate between controlled and non-controlled acromegaly.

Results

Baseline mean subscores S, G, I and T, were significantly lower in patients with CGE-DC controlled acromegaly compared with CGE-DC non-controlled acromegaly. SAGIT components I and G for CGE-DC and S, A, G, I and T for the clinician's therapeutic decision were retained by CART analyses. For international guidelines, only SAGIT component I was retained. The risk for undergoing at least one treatment change during the study for patients with CGE-DC non-controlled acromegaly relative to CGE-DC controlled acromegaly was 3.44 times greater.

Conclusion

The SAGIT instrument is a valid and sensitive tool to comprehensively and accurately assess acromegaly severity.

Key words: Acromegaly control, acromegaly management, clinician-reported outcomes, SAGIT® instrument

INTRODUCTION

Acromegaly is a rare disorder generally caused by a growth hormone (GH)-secreting pituitary adenoma, which results in GH excess and elevated insulin-like growth factor-1 (IGF-1) (1,2). The prevalence ranges between 2.8 and 13.7 per 100,000 people (3,4). The disease is characterized by skeletal and soft tissue changes to the extremities and face over time, including frontal skull bossing, joint changes, vertebral fractures, jaw malocclusion and overbite, and skin thickening (1,5-7). Other symptoms include headaches, sweating and swelling (1). Major comorbidities associated with acromegaly include altered carbohydrate metabolism, hypertension, sleep apnea, heart disease, hypopituitarism and neoplasms, mainly due to GH and/or IGF-1 hypersecretion (1,5,6,8-11). Comorbidities are associated with reduced quality of life (QoL) and increased mortality (12). Treatment by surgery or medical therapy, aims to normalize GH and/or IGF-1 levels, achieving hormonal and symptom control, and decreasing the risk of developing associated comorbidities (13,14). Life expectancy among patients with acromegaly can be comparable to that of the general population following successful restoration of biochemical (GH/IGF-1) control (5,9,15). Therefore, it is important to avoid therapeutic inertia and minimize the time taken to achieve biochemical control, while also accurately monitoring clinical aspects of the disease (14).

The SAGIT® instrument (SAGIT) has been developed to help practicing endocrinologists to better characterize disease activity, providing a precise classification of acromegaly severity (SAGIT® instrument available online as a **supplementary resource** (16)). SAGIT reflects key components associated with management of acromegaly, namely signs and symptoms (S), associated comorbidities (A), GH levels (G), IGF-1 levels (I), and tumor features (T). Each of the five components is scored by a clinician, as follows: SAGIT-S (0–4), SAGIT-A (0–6), SAGIT-G (0–4), SAGIT-I (0–3), SAGIT-T (0–5). Therefore, global SAGIT scores may range between 0–22, with higher scores representing increased disease severity. In 2016, a pilot study involving a targeted population of endocrinologists and patients with acromegaly confirmed the acceptability, utility, and ease of use of SAGIT (17). Subsequently, a clinical validation study was initiated, and baseline data reported (18). Results from the baseline analysis highlighted discrepant investigator-evaluated disease control status, disease activity, hormonal control and therapeutic decisions (**supplementary infographic** (19)).

Here, the final results from the SAGIT validation study are reported. The primary objective of the study was to define and validate the scoring of SAGIT by: (i) evaluating the ability of the instrument to discriminate between acromegaly disease control status and (ii) defining acromegaly staging based on clinical, biochemical and tumor parameters derived from the instrument.

METHODS

Study design

The SAGIT validation study (ClinicalTrials.gov NCT02539927) was a multicenter, non-interventional, prospective and retrospective, longitudinal study. The follow-up period was two years, including a minimum of three visits and a maximum of six visits per patient, as follows: a baseline visit, a maximum of four follow-up visits (carried out during the period up to 12 months after baseline) and an end-of-study/discontinuation visit (carried out between 12 months and up to 24 months after baseline) (**Figure 1**). The SAGIT validation study was conducted in accordance with the Declaration of Helsinki and Good Pharmacoepidemiology Practice. Study documentation was approved by institutional review boards.

Patients

Treatment-naïve and non-treatment-naïve male and female adults aged ≥18 years were eligible for participation in the SAGIT validation study if they had a diagnosis of acromegaly confirmed by: (i) IGF-1 levels above the age-adjusted upper limit of normal (ULN); (ii) lack of suppression of GH levels to <0.4 µg/L after a 75 g oral glucose load (patients without diabetes) or random GH levels >1.0 µg/L (patients with diabetes); (iii) a pituitary adenoma visualized with magnetic resonance imaging (MRI). Patients were additionally required to have data needed to complete assessments with SAGIT in their medical records. Patients were excluded if they had acute or severe disease (including acromegaly) that was not controlled and required intensive treatment. The validation study was designed to enrol balanced numbers of patients with and without disease control. The numbers of patients with controlled and non-controlled acromegaly (according to a clinical global evaluation of disease control [CGE-DC] questionnaire) were monitored at the time of the inclusion visit. Patients were fully informed, agreed to participate in the validation study, and signed an informed consent form to authorize their doctor to collect medical data about acromegaly from their medical records.

Data collection and evaluation

Electronic case report forms (eCRFs) were completed by investigators for each patient during each visit unless otherwise stated. The data collected varied slightly according to the type of visit, but they all included: a SAGIT score; current medication/surgery/radiotherapy treatment(s) for acromegaly (including start date, frequency and dosage); dates and values of blood tests (GH and IGF-1), MRI examinations used to complete SAGIT at the given visit and the Acromegaly Quality of Life (AcroQoL) questionnaire. AcroQoL is a validated disease-specific self-assessment questionnaire that allows the routine monitoring of QoL in patients with acromegaly (20). In addition, a CGE-DC questionnaire was included in the eCRFs to assess acromegaly control status. The CGE-DC questionnaire was specifically adapted for this study from the Clinical Global Impressions Scale, a commonly used clinician-rated measure of symptom severity (21). The disease control status was classified by CGE-DC as either

'controlled', 'non-controlled' or 'yet to be clarified' and was a subjective evaluation of patient health status based on the overall perception and medical knowledge the investigator had about the patient's acromegaly at a given visit. The same investigator also completed an investigator therapeutic decision form. The therapeutic decision could be: (i) continue the current treatment(s)/no treatment initiation; (ii) intensify the current treatment(s)/initiate a treatment(s); or (iii) decrease/downgrade the current treatment(s). CGE-DC and the investigator's therapeutic decision had to be performed by the same investigator throughout the whole duration of the validation study for a given patient to avoid inter-rater variability. Patients whose disease control status was yet to be clarified were excluded from the primary objective analyses.

Acromegaly disease control was classified based on the Acromegaly Consensus Group criteria to define acromegaly disease control (22,23). Analyses using evaluation based on international guidelines reduced variability across countries and between investigators. Acromegaly was considered as active if all three of the following criteria were satisfied: (i) random GH ≥1 µg/L or nadir GH after oral glucose tolerance test (OGTT) ≥0.4 µg/L; (ii) elevated IGF-1 levels (>100% ULN); (iii) considered clinically active based on the investigator's assessment (assessed by the same local investigator during the entire validation study). Acromegaly was considered controlled if both of the following criteria were satisfied: (i) random GH <1 µg/L or nadir GH after OGTT <0.4 µg/L; (ii) age normalized IGF-1. Patients with acromegaly that could not be classified as active or controlled were defined as 'not classified'. These patients were excluded from the primary objective analyses.

Statistical analyses

It was estimated that a total enrollment of at least 200 patients, with a minimum of 100 patients with controlled acromegaly and 100 patients without disease control, would allow the primary objective of the study to be achieved. The sample size of the study was calculated to estimate the discriminant accuracy of the classification algorithm, as follows: 82 patients would be required in each group to estimate an area under the receiver operating characteristic (ROC) curve of 0.8 with a 95% confidence interval (CI) of 0.15 width (24). Aiming for a total of 100 patients per group (controlled and non-controlled, as per CGE-DC) was estimated to ensure the recruitment of at least 82 evaluable patients in both groups.

Classification rules of SAGIT were established to investigate how the five components associated with management of acromegaly (S, A, G, I and T) could discriminate between controlled and non-controlled acromegaly using the three classifications (CGE-DC disease control status, the investigator's therapeutic decision and international guidelines). These rules were determined using a classification and regression tree (CART) analysis. A CART analysis consists of building a sequence of binary decisions to classify patients into different categories.

To validate the decision tree and evaluate its quality, a leave-one-out cross-validation was performed. The whole sample was split into n number of parts (n ≤6 to ensure a reasonable sample size in each part), and n-1 parts were used for learning ('learning sample') while the part left-out was used for validation ('validation sample'). CART analyses were performed using baseline CGE-DC disease control status, the investigator's therapeutic decision and an evaluation of acromegaly control based on international guidelines as outcomes.

The ability of the baseline SAGIT disease control status (according to each of the CART analysis) to predict treatment changes was evaluated using a multivariate stepwise logistic regression analysis (entry of variable at a 20% significance level and retention of variable at a 5% level). For each regression, the following predictors were included: baseline CGE-DC disease control status according to CART, age, sex, body mass index, time since acromegaly diagnosis and treatment-naïve status at baseline.

A specific methodology was developed to obtain ROC curves for decision trees resulting from the CART analysis, as described (25). The Youden index (maximum of sensitivity + specificity -1) was computed for all possible cut-off points of the ROC curve. The cut-off point corresponding to the Youden index was selected as the optimal cut-off to separate into the different disease control status subgroups.

A meaningful improvement in QoL was defined as a change in AcroQoL global score above or equal to the responder threshold (increase of ≥50% of the baseline standard deviation [SD] of the score, which corresponded to 9.9) (26).

For continuous variables, summary statistics included number of available observations, number of missing values, arithmetic mean, 95% CI of the mean, SD, median, first and third quartiles (Q1 and Q3) and the range (minimum, maximum). For categorical or discrete variables, percentage numbers were presented, including the 95% CI when applicable. Missing data were displayed but not accounted for in the denominator for the percentage calculation. Significance of the between-group difference was established if the 95% CIs of the two subgroups were not overlapping.

Post hoc analyses to determine SAGIT global scores were constructed from the sum of the five components that comprise the instrument. SAGIT global score thresholds for classification based on CGE-DC disease control status, the investigator's therapeutic decision and an evaluation of acromegaly based on international guidelines were also carried out.

Analyses were originally planned to be conducted on the retrospective population (association between patients with controlled versus non-controlled acromegaly and the occurrence of a significant clinical event and/or treatment change[s]) in addition to the prospective population.

However, due to the limited number of patients enrolled retrospectively (N=9), unless stated otherwise all presented analyses are based on the prospective population.

RESULTS

Patient disposition

A total of 24 centers in 9 countries (Belgium, Brazil, France, Germany, Italy, Spain, the Netherlands, the UK, and the USA) participated; 252 patients were screened, and 227 were enrolled. Of these patients, 109 had CGE-DC controlled acromegaly, while 105 patients had CGE-DC non-controlled acromegaly and 13 patients' control status was not determined at baseline. In total, 33 patients were treatment-naïve (CGE-DC non-controlled, n=31; disease status not determined, n=2) and 194 non-treatment naïve (CGE-DC controlled, n=109; CGE-DC non-controlled, n=74; disease status not determined, n=11). A patient was considered treatment naïve if they had not received any previous pituitary surgery, radiotherapy, or medications for acromegaly. The most frequent prior medical or surgical events were pituitary tumor removal (59.5%), hypertension (28.2%), goiter (21.1%) and diabetes mellitus (12.8%).

Patient demographics

Within the enrolled population, mean age was 51.6 years (SD: 12.7) and 52.9% were female. Time since diagnosis was <1 year for 28.2% of patients and ≥1 year for 71.8%. Median time since diagnosis was 3.9 years (95% CI: [5.6, 7.5]). Full details of the patient population and baseline data were reported in the SAGIT® validation study (18). Baseline patient disposition and demographics are reported in **Supplementary Table 1** (27).

Primary analyses

SAGIT scores from baseline to end of study

An overview of the mean SAGIT subscores at baseline by CGE-DC disease control status is provided in **Figure 2A**. At baseline, mean subscores S, G, I and T, were significantly lower in patients with CGE-DC controlled acromegaly compared with CGE-DC non-controlled acromegaly, while mean subscore A was similar between both disease status subgroups. Secondary analyses revealed that the mean change from baseline to the end-of-study of SAGIT subscores S, A and T were similar between CGE-DC controlled acromegaly and CGE-DC non-controlled acromegaly (**Figure 2B**). Subscores G and I maintained statistical significance and exhibited the most pronounced changes in CGE-DC non-controlled acromegaly. A complete breakdown of SAGIT subscores by CGE-DC disease control status at baseline and end-of-study are presented in **Supplementary Table 1 and 2**, respectively (27).

Discriminant accuracy of SAGIT to classify acromegaly according to disease control status as defined by CGE-DC

The CART analysis selected SAGIT components that best discriminated between controlled and non-controlled acromegaly. Using acromegaly control status as defined by CGE-DC as the binary outcome, SAGIT components G (GH concentration) and I (IGF-1 concentration) were retained by the CART analysis to classify acromegaly as controlled or non-controlled (**Figure 3**).

According to the tree depicted in **Figure 3**, the rules were as follows: if SAGIT-I=0, 90% of patients had acromegaly classified as controlled. If SAGIT-I=1 and SAGIT-G=0, 75% of patients had acromegaly classified as controlled. If SAGIT-I=1 and SAGIT-G=1, 2, 3 or 4, 64% of patients had acromegaly classified as non-controlled. Lastly, if SAGIT-I=2 or 3, 98% of patients had acromegaly classified as non-controlled.

Classification derived from the CART analysis revealed that most patients were correctly classified: 93.0% of patients with CGE-DC controlled acromegaly and 87.0% of patients with CGE-DC non-controlled acromegaly were correctly classified by CART (**Table 1A**). The sensitivity and specificity for each node of the classification tree with CGE-DC as the binary outcome were used to build a ROC curve. The area under the curve (AUC) was 0.92 (sensitivity: 0.92; specificity: 0.87).

As defined by the investigator's therapeutic decision

A similar CART analysis was performed using the investigator's therapeutic decision as an outcome. All SAGIT components were retained by the CART analysis to classify patients based on whether they had intensified/initiated treatment for acromegaly or whether their treatment had remained unchanged (**Figure 4**).

According to the tree depicted in **Figure 4**, the simplified rules were as follows: if SAGIT-G=0, 1 or 2 and SAGIT-I=0, 91% of patients' treatment remained unchanged. If SAGIT-G=0, 1 or 2 and SAGIT-I=1, 2 or 3, 71% of patients' treatment also remained unchanged. If SAGIT-G=3 or 4 and SAGIT-A=0, 1 or 2, 70% of patients' treatment had to be intensified or initiated. Lastly, if SAGIT-G=3 or 4 and SAGIT-A=3, 4 or 5, for 100% of patients, treatment had to be intensified or initiated.

Classification derived from the CART analysis revealed that for the majority of patients, the therapeutic decision was correctly classified: 88.1% of patients whose treatment remained unchanged and 75.0% of patients whose treatment had to be intensified or initiated were correctly classified by CART (**Table 1B**). The sensitivity and specificity for each node of the classification tree with the investigator's therapeutic decision as the binary outcome were used to build a ROC curve.

The AUC was 0.84 (sensitivity: 0.89; specificity: 0.74), a lower performance than the one achieved by using CGE-DC.

As defined by published consensus guidelines

The evaluation of acromegaly based on international guidelines was dichotomized into controlled acromegaly or active acromegaly. The resulting classification tree is depicted in **Figure 5**. SAGIT component I was retained by the CART analysis to classify acromegaly as active or controlled. The algorithm established that if SAGIT-I=0 then acromegaly should be considered controlled, and if I ≥1, acromegaly should be considered active. Classification from the CART analysis revealed that disease status was correctly classified in all instances (**Table 1C**). The sensitivity and specificity for each node of the classification tree with international guidelines as the binary outcome were used to build a ROC curve. The AUC was 1.00 (sensitivity: 1.00; specificity: 1.00).

Secondary analyses

Predictive accuracy of SAGIT disease control status (CART analysis) on treatment changes

Overall, 44.1% (95% CI: [37.5, 50.8]) of patients experienced at least one treatment change during the SAGIT validation study, with a statistically significantly lower incidence in the CGE-DC controlled group (14.7%; 95% CI: [8.6, 22.7]) compared with the CGE-DC non-controlled group (74.3%; 95% CI: [64.8, 82.3]). The risk to have at least one treatment change during the study for patients with CGE-DC non-controlled acromegaly relative to patients with CGE-DC controlled acromegaly (reference group) was 3.44 times greater (95% CI: [2.4, 5.0]).

SAGIT disease control status according to CART analysis was a statistically significant predictor of the occurrence of treatment changes. The odds of treatment change during the study were 9 times higher, according to the SAGIT disease control group from CART (using CGE-DC) in non-controlled acromegaly compared with controlled acromegaly (odds ratio 9.3; 95% CI: [5.0, 17.2]; p<0.0001). When the SAGIT disease control group from CART (using investigator's therapeutic decision) was used as the outcome, the odds of treatment change were 20 times higher in patients with treatment classified as intensified/initiated, compared with treatment classified as unchanged (odds ratio 20.5; 95% CI: [8.8, 47.7]; p<0.0001). Lastly, the odds of treatment change during the study were 6 times higher according to the SAGIT disease control group from CART in active acromegaly (using evaluation based on international guidelines; adjusted for age and sex) compared with acromegaly classified as controlled (odds ratio 6.5; 95% CI: [3.5, 11.9]; p<0.0001).

Quality of life

Patient QoL at baseline, change from baseline to end-of-study visit, and the percentage of patients with a meaningful improvement between the two visits, are presented by CGE-DC disease control status in **Table 2**. At baseline, standardized mean global scores obtained with the AcroQoL questionnaire were 66.3 (95% CI: [62.7, 69.8]) in patients with CGE-DC controlled acromegaly versus 55.4 (95% CI: [51.7, 59.2]) in patients with CGE-DC non-controlled acromegaly (**Table 2**). The 10-point difference between patients with CGE-DC controlled and CGE-DC non-controlled acromegaly can be considered clinically relevant and was observed for all dimensions of QoL (28). At the end-of-study, standardized mean global scores obtained with the AcroQoL questionnaire were 67.6 (95% CI: [63.4, 71.9]) in patients with CGE-DC controlled acromegaly versus 58.6 (95% CI: [53.6, 63.6]) in patients with CGE-DC non-controlled acromegaly. The percentage of patients with a meaningful improvement in their QoL was statistically significantly lower when acromegaly was classified as CGE-DC controlled at baseline (13.6%; 95% CI: [7.0, 23.0]) compared with CGE-DC non-controlled at baseline (32.6%; 95% CI: [23.0, 43.3]).

Emerging acromegaly features during this study

Significant clinical events were rare and occurred in three patients (1.3%). Two events occurred in patients receiving lanreotide: the event of another tumor diagnosis, also reported as two significant adverse events (large intestine polyp and rectal polyp, both considered unrelated to lanreotide treatment), and one event of diabetes onset reported in a patient who already had impaired fasting glucose at the initiation of lanreotide treatment.

In total, 12 adverse events were recorded and included four non-fatal serious adverse events in two patients receiving lanreotide. In one patient a large intestine polyp and rectal polyp led to hospitalization (also considered a significant clinical event) and the other patient was also hospitalized due to a large intestine polyp and colon adenoma. None of the serious adverse events were considered related to treatment and both patients recovered. Finally, none of the eight non-serious adverse events were reported in more than two patients: asthenia (n=2), alopecia (n=2), bradycardia (n=1), condition aggravated (n=1), nausea (n=1) and diarrhea (n=1). All non-serious adverse events were classified as related to lanreotide treatment.

Post-hoc analyses: SAGIT global score change from baseline

The mean SAGIT global score at baseline was significantly lower in CGE-DC controlled acromegaly (mean 3.6; 95% CI: [3.2, 4.1]) versus CGE-DC non-controlled acromegaly (mean 10.4; 95% CI: [9.7, 11.2]) (**Supplementary Figure 1A** (27)); by the end of the study, mean change from baseline in SAGIT global score was significantly greater in CGE-DC non-controlled acromegaly (mean change -4.2; 95% CI: [-5.5, -2.8]) than CGE-DC controlled acromegaly (mean change -0.3; 95% CI: [-1.0, 0.4]) (**Supplementary Figure 1B** (27)).

At baseline, acromegaly treatment that was initiated or intensified had a significantly higher mean SAGIT global score (mean 10.3; 95% CI: [9.2, 11.4]) versus acromegaly treatment that remained unchanged (mean 5.7; 95% CI: [5.0, 6.3]) (**Supplementary Figure 2A** (27)). By the end-of-study assessment, mean SAGIT global scores had decreased, irrespective of the investigator's therapeutic decision. However, the decrease in SAGIT global score from baseline was greater when acromegaly treatment was initiated/intensified (mean change -5.0; 95% CI: [-6.6, -3.5]) versus treatment that remained unchanged (mean change -0.6; 95% CI: [-1.3, 0.1]) (**Supplementary Figure 2B** (27)).

At baseline, acromegaly defined as controlled according to international guidelines had a significantly lower mean SAGIT global score (mean 2.9; 95% CI: [2.4, 3.3]), compared with acromegaly defined as active (mean 11.0; 95% CI: [10.1, 11.9]) (**Supplementary Figure 3A** (27)). A significantly greater decrease in mean SAGIT global score from baseline was measured in active acromegaly (mean change -5.6; 95% CI: [-7.0, -4.2]) versus controlled acromegaly (mean change -0.3; 95% CI: [-1.3, 0.7]) (**Supplementary Figure 3B** (27)).

ROC curve analyses were performed to determine SAGIT global score threshold. According to the Youden index (0.72; sensitivity: 0.91; specificity: 0.82), a SAGIT global score of 6 was the optimal cut-off for analyses performed using CGE-DC (AUC 0.94) (**Figure 6A**). Similarly, a SAGIT global score of 7 (Youden index: 0.49; sensitivity: 0.75; specificity: 0.73) was the optimal cut-off using the investigator's therapeutic decision (AUC 0.79) (**Figure 6B**). Finally, ROC curve analyses also revealed that a SAGIT global score of 7 (Youden index: 0.87; sensitivity: 0.89; specificity: 0.98) was the optimal cut-off (AUC 0.97) for an evaluation of acromegaly based on international guidelines (**Figure 6C**).

DISCUSSION

Constantly updated international guidelines enable correct evidence-based management of acromegaly in a standardized way (13,14). However, dissemination of this knowledge may vary depending on geographical region as well as the level of expertise of different clinical centers (29). Thus, appropriate follow-up and clinical management is not totally assured for all patients with acromegaly.

The SAGIT® instrument (SAGIT) has been designed to assist disease staging for clinicians in different settings by assessing treatment response and adapting appropriate disease management in a standardized manner (17,18). This international study was undertaken to validate SAGIT and enable its integration into daily clinical practice.

Interestingly, four of the five SAGIT components (signs and symptoms [S], GH levels [G], IGF-1 levels [I], and tumor size [T]) showed significant subscore differences, depending on the CGE-DC disease control status at baseline. Patients with CGE-DC controlled acromegaly showed significantly lower IGF-1 levels, lower GH levels and smaller tumor size than patients with CGE-DC non-controlled acromegaly. This finding is of clinical interest since, while confirming the key role of biochemical parameters in defining disease control, also indicates that tumor size can have an important role in this definition (5). In fact, it has been shown that medical treatment with somatostatin receptor ligands (SRLs) can attain tumor shrinkage but not necessarily in conjunction with biochemical control (30,31). Moreover, patients with CGE-DC controlled acromegaly had on average, one less sign/symptom than patients with CGE-DC non-controlled acromegaly. Also, this finding is relevant since it strengthens the role of physical examination and accurate clinical history collection in the management of acromegaly. The non-significant differences in associated comorbidities (SAGIT component A) between patients with controlled and non-controlled acromegaly may have alternative explanations: not all comorbidities may be reversible with biochemical control, such as sleep apnea (32) and may even worsen despite biochemical control, as can be the case with vertebral fractures (33), diabetes (34) and osteoarthritis (35). Alternatively, disease duration, particularly in the undiagnosed phase may vary considerably between patients due to persistent diagnostic delay (36) and may affect more than the biochemical control per se (i.e. development of comorbidities). These data raise the need for earlier detection (which may be aided by SAGIT) of acromegaly to avoid permanent changes by comorbidities and may indicate many patients do not receive sufficient treatment of their comorbidities. Therefore, patients with low scores in SAGIT components S, G, I and T and a high score in A, may require specific follow-up and intervention in addition to receiving treatment to establish biochemical control.

A component of the primary objective was to evaluate the discriminant accuracy of SAGIT to predict disease control status. IGF-1 and GH concentrations for CGE-DC and all components for the clinician's therapeutic decision were retained by CART analyses. This latter finding supports the role of tumor

size as a determinant of the clinical decision-making process as previously reported in an international survey (37). With international guidelines, only IGF-1 concentration was retained. Classification of disease control status by SAGIT was consistent with each of the three classifications; 93.0% of patients with CGE-DC controlled acromegaly and 87.0% of patients with CGE-DC non-controlled acromegaly were correctly classified by CART, with IGF-1 and GH concentrations the strongest performing components. These findings confirm the key role of biochemical control in defining acromegaly disease control (22).

When classifying therapeutic decisions by CART, all five components of SAGIT contributed to the classification algorithm, although GH concentration had the greatest influence. These results imply that the instrument could have a major role in assisting clinicians worldwide in therapeutic decision making by integrating all components that need to be considered. For example, choosing the optimal drug approach, including tumor size and considering all comorbidities (5). Most therapeutic decisions were correctly classified by CART, although performance was better for patients whose treatment remained unchanged (88.1% correctly classified) versus patients whose treatment had to be intensified or initiated (75.0% correctly classified). This finding is of interest and likely identifies the subgroup of patients with a partial response to treatment in whom the decision-making process may be more difficult. Therefore, SAGIT may be a helpful addition to the diagnostic aids available to assist the clinician (5). Furthermore, the SAGIT classification algorithm was consistent with the evaluation of acromegaly based on international guidelines; in fact, 100% of patients with active acromegaly and 100% of patients with controlled acromegaly were correctly classified by CART, with IGF-1 concentration being the only component of this classification.

Although non-biochemical factors (SAGIT components S, A and T) were not retained by the CART analyses for CGE-DC and following international guidelines, they indeed influenced therapeutic decisions. The latter highlights the importance to include variables other than biochemistry for the care of patients with acromegaly and may help to standardize and optimize treatment decision-making. Acromegaly staging using SAGIT examines not only biochemical parameters, but also signs and symptoms, associated comorbidities, and tumor features, and thus provides the possibility to make more informed treatment decisions to better address acromegaly disease control.

Post hoc analyses revealed mean SAGIT global scores at baseline were higher in uncontrolled/active acromegaly, and where treatment was initiated or intensified at baseline. ROC curve analyses performed to determine the SAGIT global score threshold for classification, found that the optimal cut-off score was 6 based on CGE-DC and 7 for both the investigator's therapeutic decision and disease status based on international guidelines. These cut-off scores serve as a threshold between the binary outcomes of each classification (e.g. whether acromegaly is controlled or non-controlled based on CGE-DC) and can be used to aid interpretation of the SAGIT global score, when determining acromegaly disease control status. Therefore, based on this finding it can be inferred that SAGIT may

be particularly helpful with patients in whom there could be uncertainty in determining disease control by having a narrow score threshold (between 6 and 7).

In this regard, there is a requirement to better define pituitary adenomas with aggressive behaviour, in which resistance to treatment is a key component (38-40). Given the complex management challenges faced by these patients, referral to a Pituitary Tumor Center of Excellence (PTCOE) may be desirable for optimal treatment (29,41). SAGIT results could be employed by supporting endocrinologists in non-specialized centers to optimally define patients for referral to PTCOEs for evaluation.

After the initial development of SAGIT another international effort was undertaken to develop ACRODAT® (ACROmegaly Disease Activity Tool) a different practical tool to assist clinicians in the management of acromegaly. Like SAGIT, ACRODAT is a clinician-reported outcome instrument with five components: IGF-1 levels, tumor status, comorbidities, signs and symptoms, and health-related QoL. Recently, a validation study of ACRODAT underlined IGF-1 and tumor status as main determinants for routine clinical decision making (42).

The validation of ACRODAT was based on online opinions of pituitary experts, who categorized 52 hypothetical patient scenarios as having stable, mild or significant disease activity (42). The SAGIT instrument validation study used real-world two-year prospective data, of 227 patients with acromegaly and better reflects routine clinical practice.

An important consideration for managing acromegaly is the patient's perspective on disease activity. Recent evidence suggests that patients with acromegaly place more value than expert endocrinologists on "patient-centered" parameters (i.e. signs/symptoms, comorbid conditions, and QoL) and often weighted them equally with "clinical" parameters (i.e. tumor size and IGF-1 levels) (43). This is in contrast to physicians, who in the ACRODAT validation study, favoured "clinical" parameters (42). In light of this discrepancy, which highlights the need to raise clinicians' awareness of patients' views (43), a limitation of the SAGIT validation study is that it was not designed to evaluate SAGIT from the patients' perspective. Therefore, it is unclear how patients with acromegaly assess the relative importance of the five components of SAGIT, and whether their perspectives differ from those of physicians. Nevertheless, the use of diagnostic aids to assess acromegaly requires integration with patient choice, in clinical decision-making (23).

In conclusion, this study validates the SAGIT instrument for use in clinical practice, exhibiting optimal accuracy when predicting the clinical global evaluation of disease control (CGE-DC). Moreover, SAGIT results were consistent with classifications by international guidelines, and consistent with the investigator's therapeutic decision. The availability of practical tools to assist with the assessment and monitoring of acromegaly disease control and progression is important for clinicians and patients. This study also demonstrated that SAGIT may improve and standardize the application of current guidelines and support objective treatment guidance. SAGIT provides a validated, easy to use and sensitive tool to assist clinicians worldwide in capturing and accurately documenting disease severity for integration in day-to-day acromegaly management.

ACKNOWLEDGEMENTS

The authors thank all patients involved in the study, as well as their caregivers, care team, investigators and research staff in participating institutions. The authors thank Marc Lynch, PhD and Amelia Frizell-Armitage, PhD, of Costello Medical, for providing medical writing support which was sponsored by Ipsen in accordance with Good Publication Practice guidelines.

FUNDING

This study was sponsored by Ipsen.

AUTHORS' CONTRIBUTIONS

Substantial contributions to study conception and design: AG, MB, PC, SP, FC, CS, AH, SM; substantial contributions to analysis and interpretation of the data: AG, MB, PC, SP, FC, CS, AH, SM; drafting the article or revising it critically for important intellectual content: AG, MB, PC, SP, FC, CS, AH, SM; final approval of the version of the article to be published: AG, MB, PC, SP, FC, CS, AH, SM.

DATA AVAILABILITY

Where patient data can be anonymised, Ipsen will share all individual participant data that underlie the results reported in this article with qualified researchers who provide a valid research question. Study documents, such as the study protocol and clinical study report, are not always available. Proposals should be submitted to DataSharing@Ipsen.com and will be assessed by a scientific review board. Data are available beginning 6 months and ending 5 years after publication; after this time, only raw data may be available.

REFERENCES

1. Melmed S. Acromegaly pathogenesis and treatment. *J Clin Invest.* 2009;119(11):3189–3202.
2. Melmed S. Pituitary-Tumor Endocrinopathies. *N Engl J Med.* 2020;382(10):937–950.
3. Lavrentaki A, Paluzzi A, Wass JAH, Karavitaki N. Epidemiology of acromegaly: review of population studies. *Pituitary.* 2017;20(1):4–9.
4. Colao A, Grasso LFS, Giustina A, et al. Acromegaly. *Nat Rev Dis Primers.* 2019;5(1):20.
5. Melmed S, Bronstein MD, Chanson P, et al. A Consensus Statement on acromegaly therapeutic outcomes. *Nat Rev Endocrinol.* 2018;14(9):552–561.
6. Vilar L, Vilar CF, Lyra R, Lyra R, Naves LA. Acromegaly: clinical features at diagnosis. *Pituitary.* 2017;20(1):22–32.
7. Giustina A. Acromegaly and Vertebral Fractures: Facts and Questions. *Trends Endocrinol Metab.* 2020;31(4):274–275.
8. Colao A, Spinelli L, Marzullo P, et al. High prevalence of cardiac valve disease in acromegaly: an observational, analytical, case-control study. *J Clin Endocrinol Metab.* 2003;88(7):3196–3201.
9. Dekkers OM, Biermasz NR, Pereira AM, Romijn JA, Vandenbroucke JP. Mortality in acromegaly: a metaanalysis. *J Clin Endocrinol Metab.* 2008;93(1):61–67.
10. Pivonello R, Auriemma RS, Grasso LFS, et al. Complications of acromegaly: cardiovascular, respiratory and metabolic comorbidities. *Pituitary.* 2017;20(1):46–62.
11. Giustina A, Barkan A, Beckers A, et al. A Consensus on the Diagnosis and Treatment of Acromegaly Comorbidities: An Update. *J Clin Endocrinol Metab.* 2019;105(4):e937–e946.
12. Ben-Shlomo A, Sheppard MC, Stephens JM, Pulgar S, Melmed S. Clinical, quality of life, and economic value of acromegaly disease control. *Pituitary.* 2011;14(3):284–294.
13. Fleseriu M, Biller BMK, Freda PU, et al. A Pituitary Society update to acromegaly management guidelines. *Pituitary.* 2020.
14. Giustina A, Barkhoudarian G, Beckers A, et al. Multidisciplinary management of acromegaly: A consensus. *Rev Endocr Metab Disord.* 2020;21(4):667–678.
15. Agrawal N, Ioachimescu AG. Prognostic factors of biochemical remission after transsphenoidal surgery for acromegaly: a structured review. *Pituitary.* 2020;23(5):582–594.
16. The SAGIT® instrument for the classification of acromegaly. Deposited March 2021. https://doi.org/10.6084/m9.figshare.14315993.v1.
17. Giustina A, Bevan JS, Bronstein MD, et al. SAGIT®: clinician-reported outcome instrument for managing acromegaly in clinical practice--development and results from a pilot study. *Pituitary.* 2016;19(1):39–49.
18. Giustina A, Bronstein MD, Chanson P, et al. Staging and managing patients with acromegaly in clinical practice: baseline data from the SAGIT®validation study. *Pituitary.* 2019;22(5):476–487.
19. International Multicenter Validation Study of the SAGIT® Instrument in Acromegaly: Infographic. Deposited March 2021. https://doi.org/10.6084/m9.figshare.14346023.v1
20. Badia X, Webb SM, Prieto L, Lara N. Acromegaly Quality of Life Questionnaire (AcroQoL). *Health Qual Life Outcomes.* 2004;2:13.
21. Busner J, Targum SD. The clinical global impressions scale: applying a research tool in clinical practice. *Psychiatry (Edgmont (Pa : Township)).* 2007;4(7):28-37.
22. Giustina A, Barkan A, Casanueva FF, et al. Criteria for cure of acromegaly: a consensus statement. *J Clin Endocrinol Metab.* 2000;85(2):526–529.
23. Giustina A, Chanson P, Bronstein MD, et al. A consensus on criteria for cure of acromegaly. *J Clin Endocrinol Metab.* 2010;95(7):3141–3148.
24. Zhou X-H, McClish DK, Obuchowski NA. *Statistical methods in diagnostic medicine*. New York: John Wiley & Sons; 2009
25. Ferri C, Flach PA, Hernández-Orallo J. *Learning Decision Trees Using the Area Under the ROC Curve*: Morgan Kaufmann Publishers Inc.; 2002;139–146.
26. Revicki DA, Erickson PA, Sloan JA, Dueck A, Guess H, Santanello NC. Interpreting and reporting results based on patient-reported outcomes. *Value Health.* 2007;10 Suppl 2:S116–124.

27. Supplementary Materials: International Multicenter Validation Study of the SAGIT® Instrument in Acromegaly. Deposited March 2021; updated June 2021. https://doi.org/10.6084/m9.figshare.14346110.v4
28. Caron PJ, Bevan JS, Petersenn S, Houchard A, Sert C, Webb SM. Effects of lanreotide Autogel primary therapy on symptoms and quality-of-life in acromegaly: data from the PRIMARYS study. *Pituitary.* 2016;19(2):149–157.
29. Casanueva FF, Barkan AL, Buchfelder M, et al. Criteria for the definition of Pituitary Tumor Centers of Excellence (PTCOE): A Pituitary Society Statement. *Pituitary.* 2017;20(5):489–498.
30. Mazziotti G, Giustina A. Effects of lanreotide SR and Autogel on tumor mass in patients with acromegaly: a systematic review. *Pituitary.* 2010;13(1):60–67.
31. Giustina A, Mazziotti G, Torri V, Spinello M, Floriani I, Melmed S. Meta-Analysis on the Effects of Octreotide on Tumor Mass in Acromegaly. *PLOS ONE.* 2012;7(5):e36411.
32. Davi MV, Dalle Carbonare L, Giustina A, et al. Sleep apnoea syndrome is highly prevalent in acromegaly and only partially reversible after biochemical control of the disease. *Eur J Endocrinol.* 2008;159(5):533–540.
33. Mazziotti G, Bianchi A, Porcelli T, et al. Vertebral Fractures in Patients With Acromegaly: A 3-Year Prospective Study. *J Clin Endocrinol Metab.* 2013;98(8):3402–3410.
34. Frara S, Maffezzoni F, Mazziotti G, Giustina A. Current and Emerging Aspects of Diabetes Mellitus in Acromegaly. *Trends Endocrinol Metab.* 2016;27(7):470–483.
35. Pelsma ICM, Biermasz NR, van Furth WR, et al. Progression of acromegalic arthropathy in long-term controlled acromegaly patients: 9 years of longitudinal follow-up. *The Journal of Clinical Endocrinology & Metabolism.* 2020;106(1):188-200.
36. Giustina A. [Acromegaly: reducing diagnostic delay]. *Recenti Prog Med.* 2016;107(8):450–451.
37. Giustina A, Bronstein MD, Casanueva FF, et al. Current management practices for acromegaly: an international survey. *Pituitary.* 2011;14(2):125–133.
38. Raverot G, Burman P, McCormack A, et al. European Society of Endocrinology Clinical Practice Guidelines for the management of aggressive pituitary tumours and carcinomas. *Eur J Endocrinol.* 2018;178(1):G1-g24.
39. Gola M, Bonadonna S, Mazziotti G, Amato G, Giustina A. Resistance to somatostatin analogs in acromegaly: an evolving concept? *J Endocrinol Invest.* 2006;29(1):86-93.
40. Giustina A. Pituitary adenoma…nomen omen? *Endocrine.* 2021.
41. Frara S, Rodriguez-Carnero G, Formenti AM, Martinez-Olmos MA, Giustina A, Casanueva FF. Pituitary Tumors Centers of Excellence. *Endocrinol Metab Clin North Am.* 2020;49(3):553-564.
42. van der Lely AJ, Gomez R, Pleil A, et al. Development of ACRODAT®, a new software medical device to assess disease activity in patients with acromegaly. *Pituitary.* 2017;20(6):692–701.
43. Jackson Y, Flood E, Rhoten S, Janssen EM, Lundie M. AcroVoice: eliciting the patients' perspective on acromegaly disease activity. *Pituitary.* 2019;22(1):62–69.

TABLES AND FIGURE LEGENDS

Table 1. Rate of correct classification from the three CART analyses with CGE-DC (A), investigator's therapeutic decision (B) and evaluation based on international guidelines (C) as the binary outcome. All analyses are based on the prospective population (N=227). CART: Classification and Regression Tree; CGE-DC: Clinical Global Evaluation of Disease Control; ITD: investigator's therapeutic decision.

Table 2. QoL at baseline and end-of-study, and meaningful improvements. *Defined as a change in AcroQoL global score above or equal to the responder threshold (increase of >50% of the baseline standard deviation of the score, which corresponded to 9.9). All analyses are based on the prospective population (N=227). AcroQoL: Acromegaly Quality of Life questionnaire; EOS: End-of-Study; CGE-DC: Clinical Global Evaluation in Disease Control; CI: confidence interval; QoL: quality of life.

Figure 1. Design of the SAGIT validation study. The SAGIT validation study was a multicenter, non-interventional, prospective and retrospective, longitudinal study. However, due to the limited number of patients enrolled retrospectively (N=9), all analyses presented in this report are based on the prospective population (N=227). The follow-up period was two years, including a minimum of three visits and a maximum of six visits per patient, as follows: a baseline visit, a maximum of four follow-up visits (carried out during the period up to 12 months after baseline) and an end-of-study/discontinuation visit (carried out between 12 months and up to 24 months after baseline). M: months.

Figure 2. Overview of SAGIT subscores by CGE-DC disease control status over time. **A**. Overview of SAGIT subscores by CGE-DC disease control status at baseline. **B**. Overview of SAGIT subscores by CGE-DC disease control status change from baseline to the end of study visit. *Significant between-group (controlled vs. non-controlled) difference as determined by non-overlapping 95% CIs. All analyses are based on the prospective population (N=227). CGE-DC: Clinical Global Evaluation in Disease Control; CI: confidence interval; SAGIT-S: signs and symptoms; SAGIT-A: associated comorbidities; SAGIT-G: growth hormone levels; SAGIT-I: insulin-like growth factor-1 levels; SAGIT-T: tumor features. Data shown are means and 95% CIs.

Figure 3. Discriminant accuracy of SAGIT to classify acromegaly according to CGE-DC based on the prospective population (N=227). CART analysis with CGE-DC as the binary outcome. In this analysis, IGF-1 (SAGIT-I) and GH (SAGIT-G) concentrations were retained, with the objective of classifying patients' acromegaly as controlled or non-controlled. CART: classification and regression tree; CGE-DC: Clinical Global Evaluation in Disease Control; GH: growth hormone; IGF-1: insulin-like growth factor-1.

Figure 4. Discriminant accuracy of SAGIT to classify patients according to the investigator's therapeutic decision based on the prospective population (N=227). CART analysis with investigator's therapeutic decision as the binary outcome. In this analysis, all five components of SAGIT were retained, with the objective of classifying patients according to therapeutic decision. CART: classification and regression tree; SAGIT-S: signs and symptoms; SAGIT-A: associated comorbidities; SAGIT-G: growth hormone levels; SAGIT-I: insulin-like growth factor-1 levels; SAGIT-T: tumor features.

Figure 5. Discriminant accuracy of SAGIT to classify acromegaly in the prospective population (N=227) using an evaluation based on international guidelines. CART analysis using an evaluation based on international guidelines as the binary outcome. In this analysis, only IGF-1 (SAGIT-I) concentration was retained, with the objective of classifying acromegaly according to disease activity. CART: classification and regression tree; IGF-1: insulin-like growth factor-1.

Figure 6. ROC curve analyses performed post hoc to determine the SAGIT global score threshold. **A**. Determination of the SAGIT global score threshold by ROC curve analysis according to CGE-DC disease control group. **B**. Determination of the SAGIT global score threshold by ROC curve analysis according to the investigator's therapeutic decision.
C. Determination of the SAGIT global score threshold by ROC curve analysis according to an evaluation based on international guidelines. All analyses are based on the prospective population (N=227). CGE-DC: Clinical Global Evaluation in Disease Control; ROC: receiver operating characteristic.

Table 1.

A. Baseline CGE-DC disease control status and CART classification			
CART classification	**Classified as controlled by CART**	**Classified as non-controlled by CART**	**Total**
CGE-DC controlled, n (%)	80 (93.0%)	6 (7.0%)	86
CGE-DC non-controlled, n (%)	12 (13.0%)	80 (87.0%)	92
			178
B. Baseline investigator's therapeutic decision and CART classification			
CART classification	**Classified as treatment unchanged by CART**	**Classified as treatment intensified/initiated by CART**	**Total**
ITD treatment unchanged, n (%)	111 (88.1%)	15 (11.9%)	126
ITD treatment intensified/initiated, n (%)	15 (25.0%)	45 (75.0%)	60
			186
C. Baseline evaluation based on international guidelines and CART classification			
CART classification	**Classified as controlled by CART**	**Classified as active by CART**	**Total**
Guidelines acromegaly controlled, n (%)	47 (100.0%)	0 (0.0%)	47
Guidelines acromegaly active, n (%)	0 (0.0%)	77 (100.0%)	77
			124

Table 2.

	CGE-DC disease control status	
	Controlled (n=109)	**Non-controlled (n=105)**
Baseline visit		
Physical score		
N	108	105
Missing	1	0
Mean [95% CI]	62.6 [57.9, 67.3]	52.6 [48.1, 57.1]
Psychological score		
N	108	105
Missing	1	0
Mean [95% CI]	68.4 [65.1, 71.8]	57.1 [53.3, 60.9]
Global score		
N	108	105
Missing	1	0
Mean [95% CI]	66.3 [62.7, 69.8]	55.4 [51.7, 59.2]
Change from Baseline at EOS		
Physical score		
N	81	89
Missing	19	11
Mean Change [95% CI]	1.8 [0.0, 4.9]	3.4 [0.0, 7.0]
Psychological score		
N	82	89
Missing	18	11
Mean Change [95% CI]	2.3 [0.0, 5.5]	2.6 [0.0, 6.1]
Global score		
N	81	89
Missing	19	11
Mean Change [95% CI]	2.1 [0.0, 5.1]	2.9 [0.0, 6.1]
QoL meaningful improvement*?		
N	81	89
Missing	28	16
Yes, n (%) [95% CI]	11 (13.6) [7.0, 23.0]	29 (32.6) [23.0, 43.3]
No, n (%) [95% CI]	70 (86.4) [77.0, 93.0]	60 (67.4) [56.7, 77.0]

Figure 1

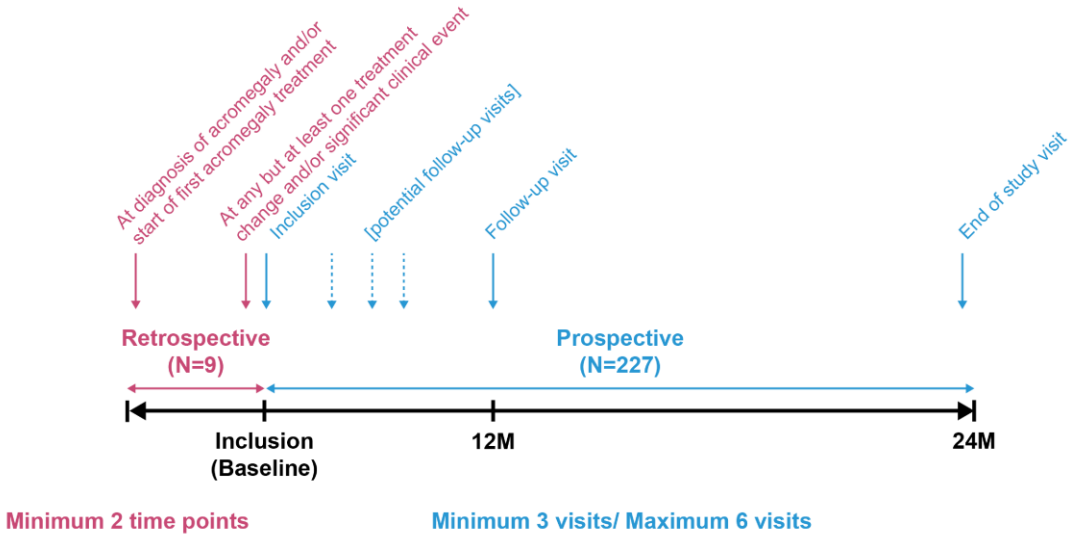

Figure 2A

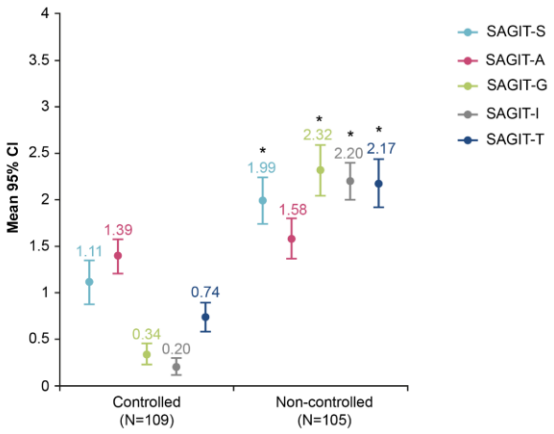

Figure 2B

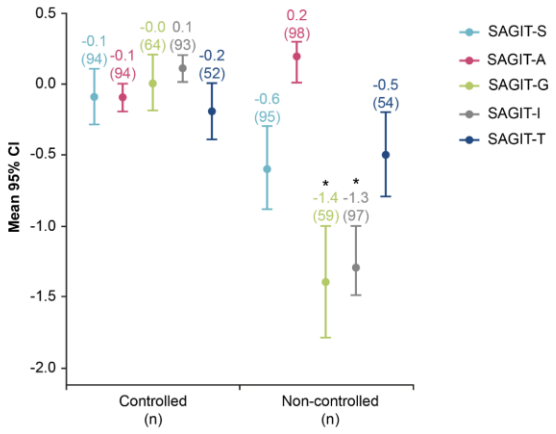

Figure 3

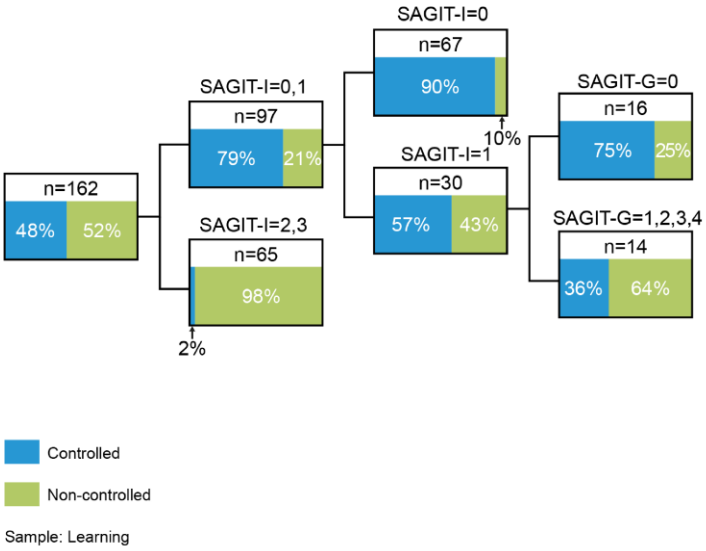

Figure 4

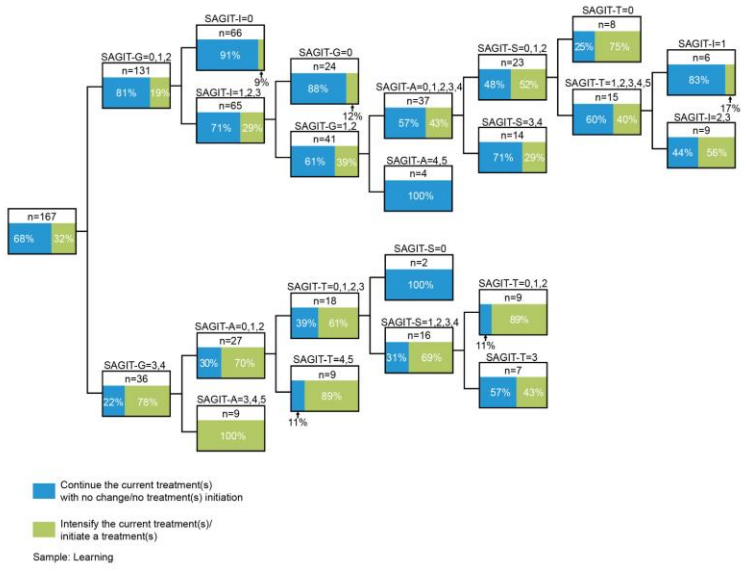

Figure 5

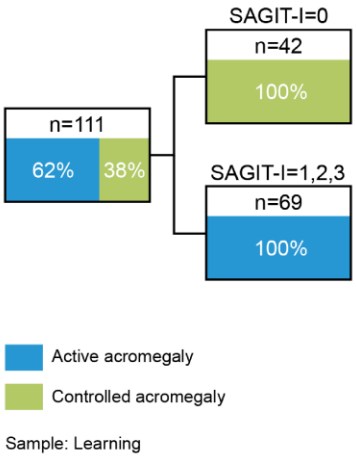

Figure 6A

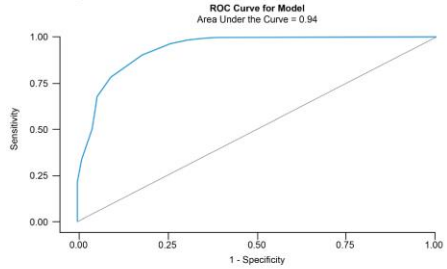

Figure 6B

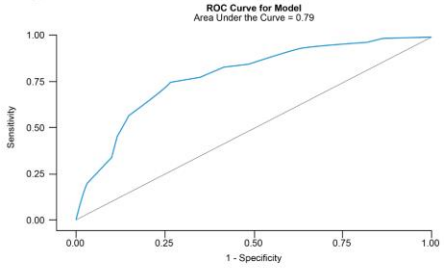

Figure 6C