BMJ Open

Longitudinal association of thyroid-stimulating immunoglobulin levels with clinical characteristics in thyroid eye disease

JaeSang Ko ,[1] Koung Hoon Kook,[2,3] Jin Sook Yoon ,[1] Kyung In Woo ,[4] Jae Wook Yang,[5] Korean Society of Ophthalmic Plastic & Reconstructive Surgery

To cite: Ko J, Kook KH, Yoon JS, *et al*. Longitudinal association of thyroid-stimulating immunoglobulin levels with clinical characteristics in thyroid eye disease. *BMJ Open* 2022;**12**:e050337. doi:10.1136/bmjopen-2021-050337

► Prepublication history for this paper is available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2021-050337).

Received 18 February 2021
Accepted 01 June 2022

For numbered affiliations see end of article.

Correspondence to
Dr Kyung In Woo; eyeminded@skku.edu and Professor Jae Wook Yang; eyeyang@inje.ac.kr

ABSTRACT

Objectives The clinical course of thyroid eye disease (TED) is heterogeneous and predicting patients who may develop the severe sequelae of the disease is difficult. In this study, we evaluated the longitudinal association between changes in serum thyroid-stimulating hormone (TSH) receptor antibody (TRAb) levels and course of disease activity and severity over time.

Design This was a multicentre, prospective, observational study.

Setting Fifteen tertiary care oculoplastic service centres in Korea.

Participants Seventy-six patients with newly diagnosed TED were included and followed up for 12 months.

Methods We evaluated clinical characteristics and serum TRAb levels at baseline, 6 and 12 months of TED diagnosis. Additionally, we analysed longitudinal associations between the serum TRAb levels and clinical activity score (CAS), no signs or symptoms, only signs, soft tissue involvement, proptosis, extraocular muscle involvement, corneal involvement, sight loss (NOSPECS) score and proptosis.

Results Thyroid-stimulating immunoglobulin (TSI) and TSH-binding inhibitory immunoglobulin (TBII) levels decreased during the 1-year follow-up, whereas disease activity measured using CAS decreased mainly in the first 6 months. Disease severity measured using NOSPECS score and proptosis remained unchanged. Moreover, inter-person differences in TBII levels were associated with CAS, NOSPECS score and proptosis over time, whereas inter-person differences in TSI levels were associated with NOSPECS score. Subgroup analysis of patients with a baseline CAS≥4 demonstrated that within-person changes in TSI levels affected the CAS and NOSPECS score.

Conclusions Follow-up measurement of serum TSI and TBII levels may help evaluate TED prognosis and enable accurate clinical decision-making.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This is the first study to investigate longitudinal correlation between thyroid-stimulating immunoglobulin levels and clinical characteristics in thyroid eye disease (TED) in a prospective setting.
⇒ Correlations between TED status and decreases in thyroid-stimulating hormone receptor antibody level were disaggregated to within-person and inter-person effects using a linear mixed model with time-dependent covariates.
⇒ This approach enabled assessment of correlations between autoantibody values at initial presentation and clinical characteristics at follow-ups, as well as correlations between time-dependent changes in autoantibody values and changes in clinical features.
⇒ The findings of this study should be interpreted with caution given the observational nature of the study and lack of standardisation of the baseline clinical characteristics and treatment modalities during follow-up, which may have affected the clinical characteristics over time.

INTRODUCTION

Thyroid eye disease (TED) is a component of autoimmune Graves' hyperthyroidism, wherein the thyroid-stimulating hormone (TSH) receptor antibody (TRAb) stimulates orbital and periorbital tissues.[1 2] The clinical course of TED is heterogeneous, and the natural course is benign in a considerable proportion of TED cases, although some cases may present with substantially severe sequelae.[2–4] Indications that can assist clinicians in determining TED prognosis are limited, such that ascertaining the patients who can develop the potentially blinding sequelae of TED is difficult.

Relationships between TED-related serum TRAb levels and clinical characteristics have been elucidated. Serum TRAb levels correlate directly with the clinical features of TED, with high TRAb level correlating with the disease prevalence and status.[5–11] Additionally, patients with higher baseline TRAb levels demonstrated a higher risk of severe disease course during a 1-year follow-up.[3] However, TRAb levels differ over time and decrease to a large extent in some patients

but not others.[12 13] Most studies have measured serum TRAb level only at the baseline; however, using only baseline values results in the loss of information provided by the serum TRAb level, which changes over time, thereby limiting the assessment of relationships between thyroid autoantibodies and TED.

Currently, two established assays are available for measuring the TRAb levels: the competitive TSH-binding inhibitory immunoglobulin (TBII) assay and functional thyroid-stimulating immunoglobulin (TSI) bioassay.[9 10 14] The former uses the ability of TRAb to inhibit the binding of radiolabelled TSH to TSH receptors,[15] and the latter measures cyclic AMP production after TRAb binds to the TSH receptor, enabling functional TRAb identification.[16 17] Several studies evaluating the correlation between TRAb levels and the clinical course of TED revealed different predictive values between TBII and TSI levels, with most showing that disease activity and TED severity are more closely correlated with the TSI bioassay than the TBII assay.[3 5 18]

Longitudinal studies assessing the association between TRAb levels and the clinical course of TED are limited. The inflammatory phase of TED reportedly correlates with changes in measured TSI levels, indicating that serial measurements of TSI level may be an adjunct in assessing the clinical inflammatory activity of TED[19]; additionally, follow-up measurements of TBII levels also reportedly allow assessment of TED prognosis.[10] However, the association and potential predictive value of TSI levels during the clinical course of TED have not yet been demonstrated in a prospective setting.

Herein, we investigated whether changes in serum TRAb (TBII and TSI) levels over time are associated with TED activity and severity during the disease course. The findings revealed correlations between time-dependent changes in TRAb levels in an individual patient and the clinical course of the disease.

METHODS

Patients

This multicentre, prospective, observational study was conducted by the Korean Society of Ophthalmic Plastic and Reconstructive Surgery from April 2017 to February 2019. During the study period, 91 patients with Graves' disease and TED were enrolled at the 15 participating Korean tertiary care oculoplastic service centres.

Patients with clinically overt TED at <6 months after symptom onset were recruited. Patients with a previous history of TED treatment (except conservative treatments, such as the use of artificial tears), uncertain cases and patients aged<19 years were excluded. Clinical data were recorded, and a blood sample was taken on study inclusion and at 6 and 12 months after inclusion. After enrolment, the study participants received the necessary management for TED at the discretion of each clinician, including conservative management, systemic steroid administration and/or orbital radiotherapy. No surgical procedure was performed on the enrolled patients during the study period.

Disease activity and severity were assessed by an oculoplastic specialist blinded to the laboratory data at each participating institute. The general ophthalmic assessment included an examination of anterior and posterior eye segments, measurement of proptosis using Hertel exophthalmometry and evaluation of ocular movement.[20] The clinical activity score (CAS; range: 0–7) was calculated based on the classic signs of inflammation and comprised seven items.[21] Disease severity was assessed using the modified no signs or symptoms, only signs, soft tissue involvement, proptosis, extraocular muscle involvement, corneal involvement, sight loss (NOSPECS) score,[12] which was calculated as the sum of each class present and ranged from 1 to 18. For each patient, the more severely affected eye (one with more proptosis at enrolment or with a higher CAS in cases of similar extents of proptosis in both eyes) was selected for the study.

TRAb assays

The blood sample collected at each visit was centrifuged at each participating institution and transferred to a central laboratory for analysis. TRAb levels were measured using a third-generation TBII assay with the automated Cobas electrochemiluminescence immunoassay system (Elecsys; Roche Diagnostics GmbH, Penzberg, Germany) according to the manufacturer's instructions. The cut-off value of positivity using this system was 1.75 IU/L. Serum Mc4-TSI level was measured using the Thyretain TSI reporter bioassay (Diagnostic Hybrids, Athens, Ohio, USA) according to the manufacturer's instructions. Results were considered positive at a specimen-to-reference ratio>140% relative to the reference control.

Data analysis and statistical analysis

Data were statistically analysed to determine correlations between changes in clinical indicators of TED, as reflected in the NOSPECS score, CAS and proptosis, and change in measured serum TRAb levels over time in individual patients.[22] We assessed the effect of serum autoantibody levels on clinical TED activity and severity and investigated alterations in serum autoantibody levels and clinical parameters over time using a linear mixed model. We used a linear mixed effects model with an unstructured covariance structure to analyse the data. Additionally, we adopted two approaches for modelling repeatedly measured outcomes. First, we analysed the data using a linear mixed model with time-dependent covariates to show the effect of baseline serum autoantibody levels on clinical characteristics over time. Time-dependent covariates are independent variables that include both within-subject and between-subject variations and can be used to make comparisons across populations and describe time trends and dynamic relationships between covariate and response.[23] Patient age, sex, smoking status and treatment modality (conservative, systemic steroid administration and systemic steroid administration with

orbital radiotherapy) were included as covariates in the model. Second, serum autoantibody levels were used as a covariate, including the baseline level and all repeated measurements.

One particularly salient strength offered by longitudinal data is the ability to disaggregate between-person and within-person effects in the regression of an outcome on a time-varying covariate.[24] This allowed the separation of time-varying within-person changes from time-invariant between-person (inter-person) differences in TRAb levels based on the clinical characteristics of the disease.[25] Moreover, this approach enabled evaluation of correlations between autoantibody levels at initial presentation and clinical characteristics at further follow-ups, as well as correlations between time-dependent changes in autoantibody levels and changes in clinical characteristics. Analyses were performed using SAS software (V.9.4), and a two-sided p<0.05 was considered statistically significant.

Patients with an initial CAS≥4 were included in the active subgroup and subjected to subgroup analysis, which included time-dependent changes in the clinical characteristics and laboratory results and time-dependent correlation analysis.

Patient and public involvement

No patients were involved.

RESULTS

Clinical characteristics and TRAb levels

Of the 91 patients enrolled in this study, 76 completed three clinical and laboratory evaluations. The clinical and biochemical characteristics of the 76 patients with TED are shown in table 1. During the 12-month follow-up period, 24 patients (32%) were conservatively managed, 34 (45%) were prescribed systemic steroid administration (oral or intravenous) and 18 (24%) were managed with combined orbital radiation and systemic steroid administration. Using systemic steroids did not affect the serum TSI or TBII levels of the patients (p=0.4225 and 0.9634, respectively), and TSI and TBII levels decreased significantly during the 12-month follow-up. Regarding clinical characteristics, disease activity measured with CAS decreased during the 12-month follow-up period, although disease severity measured by NOSPECS score and proptosis remained unchanged.

Correlation between baseline TRAb levels and clinical characteristics

Table 2 shows the results of correlation analysis using a linear mixed model with time-dependent covariates between baseline serum TRAb levels and the clinical characteristics of TED. Lower baseline serum TBII level was associated with a future decrease in TED activity, and reduced TED severity (NOSPECS score and proptosis) was associated with lower baseline serum TSI and TBII levels.

Correlation between time-dependent changes in clinical characteristics and TRAb levels

Table 3 shows the effect of the time-dependent changes in TRAb levels on the clinical characteristics of TED. The effect of longitudinally measured TRAb levels over time was disaggregated to inter-person differences and within-person changes using a linear mixed model with time-dependent covariates. TSI levels measured at multiple time points showed inter-person relationships with serial changes in NOSPECS score, indicating that patients with a low average serum TSI levels during the 1-year follow-up showed a greater reduction in their NOSPECS score over

Table 1 Clinical and biochemical characteristics of the study population

Variables	Baseline	6 months	12 months	P value
Age, year	42.4±12.0			
Sex				
Female, n (%)	47 (61.84)			
Male, n (%)	29 (38.16)			
Smoking status				
Non-smoker, n (%)	62 (81.58)			
Smoker, n (%)	14 (18.42)			
TSI, SRR (%)	446.9±185.6	346.9±181.2	302.4±173.5	<0.0001*
TBII, IU/L	11.9±11.7	6.0±8.6	4.6±6.8	<0.0001*
CAS	2.7±1.5	1.4±1.3	1.3±1.5	<0.0001*
NOSPECS score	3.9±2.2	3.4±2.3	3.4±2.4	0.0050*
Proptosis, mm	17.7±3.1	18.0±3.1	17.9±3.2	0.1523*

*P values were calculated between three visits using a linear mixed model.
CAS, clinical activity score; NOSPECS, no signs or symptoms, only signs, soft tissue involvement, proptosis, extraocular muscle involvement, corneal involvement, sight loss; SRR, specimen-to-reference ratio; TBII, thyroid-stimulating hormone receptor-binding inhibitory immunoglobulin; TSI, thyroid-stimulating immunoglobulin.

Table 2 Correlation between baseline serum TRAb levels and future changes in clinical characteristics

	CAS		NOSPECS		Proptosis	
	Estimated change (SE)	P value	Estimated change (SE)	P value	Estimated change (SE)	P value
TSI	0.000 (0.001)	0.999	**0.003 (0.001)**	**0.016**	0.004 (0.002)	0.055
TBII	**0.029 (0.010)**	**0.008**	**0.042 (0.019)**	**0.029**	**0.062 (0.030)**	**0.041**

P values were calculated using a linear mixed model with time-dependent covariates adjusted for age, sex, smoking status, and treatment modality. Numbers in bold denote statistically significant changes.
CAS, clinical activity score; NOSPECS, no signs or symptoms, only signs, soft tissue involvement, proptosis, extraocular muscle involvement, corneal involvement, sight loss; TBII, thyroid-stimulating hormone receptor-binding inhibitory immunoglobulin; TRAb, thyroid-stimulating hormone receptor antibody; TSI, thyroid-stimulating immunoglobulin.

time. Additionally, TBII level showed an inter-person relationship with CAS and the NOSPECS score, indicating that patients with higher TBII levels during the 1-year follow-up demonstrated prolonged disease activity and severity. Within-person changes in TSI and TBII levels did not show a statistically significant correlation with disease activity and severity.

Table 3 Effect of inter-person differences and within-person changes in TRAb levels on clinical characteristics: longitudinal associations during a 1-year follow-up

	Effect of TSI on clinical characteristics	
Variables	Estimated change (SE)	P value*
CAS		
Inter-person	0.0013 (0.0007)	0.0731
Within-person	0.0011 (0.0008)	0.1510
NOSPECS		
Inter-person	**0.0043 (0.0013)**	**0.0013**
Within-person	0.0002 (0.0007)	0.7700
Proptosis		
Inter-person	0.0032 (0.0022)	0.1410
Within-person	−0.0010 (0.0007)	0.1796
	Effect of TBII on clinical characteristics	
Variables	Estimated change (SE)	P value*
CAS		
Inter-person	**0.0392 (0.0109)**	**0.0004**
Within-person	0.0261 (0.0173)	0.1334
NOSPECS		
Inter-person	**0.0520 (0.0242)**	**0.0335**
Within-person	0.0249 (0.0152)	0.1036
Proptosis		
Inter-person	**0.0775 (0.0379)**	**0.0428**
Within-person	0.0017 (0.0167)	0.3095

*P values were calculated using a linear mixed model with time-dependent covariates adjusted for age, sex, smoking status and treatment modality. Numbers in bold denote statistically significant changes.
CAS, clinical activity score; NOSPECS, no signs or symptoms, only signs, soft tissue involvement, proptosis, extraocular muscle involvement, corneal involvement, sight loss; TBII, thyroid-stimulating hormone receptor-binding inhibitory immunoglobulin; TRAb, thyroid-stimulating hormone receptor antibody; TSI, thyroid-stimulating immunoglobulin.

Correlation between time-dependent changes in clinical characteristics and TRAb levels: subgroup analysis

Twenty-five patients with an initial CAS≥4 were included in the subgroup analysis, with this group defined as the active subgroup and other patients included in the inactive subgroup. Subgroup analysis was performed to evaluate the time-dependent correlations between clinical characteristics of TED and TRAb levels in the active subgroup. Table 4 shows that clinical and biochemical characteristics of the active subgroup. This subgroup had more females and smokers than the inactive subgroup (both p values<0.0001). During 12 months of follow-up, 10 patients (40%) were prescribed systemic steroid administration (oral or intravenous) and 15 (60%) were managed with combined orbital radiation and systemic steroid administration. Using systemic steroids did not affect the serum TSI and TBII levels of patients (p=0.9008 and 0.4957, respectively). Additionally, baseline the TBII level was higher in the active than in the inactive subgroups (9.5±9.5 IU/L vs 17.0±14.2 IU/L, p=0.0125), whereas the baseline TSI level of the active subgroup was similar to that of the inactive subgroup (443.9%±210.5% vs 453.2%±123.6%, p=0.8096). Moreover, the active subgroup had a higher CAS (4.4±0.6 vs 1.9±1.0, p<0.0001), higher NOSPECS score (5.6±2.3 vs 3.1±1.6, p<0.0001) and more proptosis (19.2±3.5 mm vs 16.9±2.6 mm, p=0.0020) than the inactive group. Correlation analysis between baseline TRAb levels and clinical characteristics in the active subgroup showed that lower baseline TBII levels correlated with a decrease in CAS and NOSPECS score during the 1-year follow-up (estimated change (CAS and NOSPECS): 0.0523 (p=0.0078) and 0.0566 (p=0.0394), respectively). Furthermore, baseline TSI level did not show a statistically significant correlation with clinical characteristics in the active subgroup.

Table 5 shows the effect of the time-dependent changes in TRAb levels on the clinical characteristics of TED in the active subgroup. The effect of longitudinally measured TRAb level over time was disaggregated to inter-person differences and within-person changes using a linear

Table 4 Clinical and biochemical characteristics: subgroup analysis of patients with active TED at baseline (n=25)

Variables	Baseline	6 months	12 months	P value
Age, years	43.6±11.84			
Sex				
Female, n (%)	10 (40.0)			
Male, n (%)	15 (60.0)			
Smoking status				
Non-smoker, n (%)	17 (68.0)			
Smoker, n (%)	8 (32.0)			
TSI, SRR (%)	453.2±123.6	342.5±139.7	297.9±137.0	<0.0001*
TBII, IU/L	17.0±14.2	9.4±11.6	7.6±10.4	<0.0001*
CAS	4.4±0.6	1.7±1.3	1.7±1.7	<0.0001*
NOSPECS score	5.6±2.3	5.0±2.3	4.9±2.4	0.0877*
Proptosis, mm	19.2±3.5	19.7±3.4	19.6±3.9	0.3414*

*P values were calculated between three visits using a linear mixed model.
CAS, clinical activity score; NOSPECS, no signs or symptoms, only signs, soft tissue involvement, proptosis, extraocular muscle involvement, corneal involvement, sight loss; SRR, specimen-to-reference ratio; TBII, thyroid-stimulating hormone-binding inhibitory immunoglobulin; TED, thyroid eye disease; TSI, thyroid-stimulating immunoglobulin.

mixed model with time-dependent covariates. Within-person changes in TSI level measured at multiple time points affected the CAS and NOSPECS score, whereas inter-person differences in TSI level were unrelated to clinical characteristics. Thus, patients who showed greater reductions in TSI level during follow-up also showed greater reductions in their CAS and NOSPECS score during the 1-year follow-up. By contrast, inter-person differences in TBII level correlated with the CAS and NOSPECS score, whereas within-person changes in TBII level were independent of disease severity and activity.

DISCUSSION

This prospective, observational study involving a large cohort of patients with TED demonstrated the clinical relevance of longitudinal follow-up of both TSI and TBII levels. Both TSI and TBII levels decreased sustainably during the 1-year follow-up, whereas disease activity measured using CAS decreased mainly within the first 6 months and stabilised thereafter. Notably, disease severity measured using NOSPECS score and proptosis remained unchanged. Additionally, multivariate analysis using a linear mixed model with time-dependent covariates revealed that lower TBII level at presentation correlated with future decreases in disease activity and severity, whereas lower TSI level at presentation correlated with a decrease in disease severity but not activity. Although time-dependent changes in TSI and TBII levels did not show significant within-person correlations with clinical characteristics, subgroup analysis of patients with a CAS≥4 at baseline showed that within-person changes in TSI level affected the CAS and NOSPECS score. These findings indicated that patients who demonstrated greater decreases in serum TSI levels over time showed greater reductions in disease severity and activity during the 1-year follow-up.

Multiple cross-sectional studies have demonstrated a relationship between serum TRAb level and TED. Regarding the relationship between TED development and TRAb level, antibody levels showed a positive correlation with TED prevalence according to studies of patients with newly diagnosed, untreated Graves' disease.[3 4] Another study using a patient cohort divided according to TSI quartiles showed that TED prevalence increased with each quartile of TSI levels.[5] Additionally, a previous report indicated that patients with both Graves' disease and TED showed higher TSI levels relative to those without TED, and that TED patients with extraocular muscle enlargement also showed higher TSI levels than patients without myopathy, as well as a positive relationship between TRAb levels and the clinical characteristics of TED.[6] Moreover, both TSI and TBII levels are reportedly associated with disease activity and proptosis,[8] whereas only TSI level correlated significantly with the degree of eyelid swelling, proptosis and extraocular muscle enlargement.[6] In a study involving patients with newly diagnosed and untreated TED, TSI levels were significantly higher in patients with active and severe TED than in those with mild TED and inactive disease, with TSI bioassays showed significantly higher overall positivity among patients with active TED relative to results generated using the TBII assay.[7]

Longitudinal studies on TRAb levels and the clinical characteristics of TED are limited. A retrospective study showed that patients with higher baseline TRAb (TSI and TBII) levels demonstrated a higher risk of severe disease course after 1 year[3]; however, that study did not investigate the relationship between changes in antibody level and clinical characteristics over time, as laboratory

Table 5 Effects of inter-person differences and within-person changes in TRAb levels on clinical characteristics: subgroup analysis of patients with active TED at baseline (n=25)

	Effect of TSI on clinical characteristics	
Variables	**Estimated change (SE)**	**P value***
CAS		
Inter-person	−0.0002 (0.0020)	0.9347
Within-person	**0.0046 (0.0016)**	**0.0064**
NOSPECS		
Inter-person	0.0014 (0.0031)	0.6559
Within-person	**0.0034 (0.0016)**	**0.0385**
Proptosis		
Inter-person	0.0016 (0.0070)	0.8178
Within-person	0.0014 (0.0018)	0.4464
	Effect of TBII on clinical characteristics	
Variables	**Estimated change (SE)**	**P value***
CAS		
Inter-person	**0.0470 (0.0146)**	**0.0023**
Within-person	0.0243 (0.0284)	0.3959
NOSPECS		
Inter-person	**0.0517 (0.0245)**	**0.0402**
Within-person	0.0397 (0.0261)	0.1354
Proptosis		
Inter-person	0.1067 (0.0566)	0.0657
Within-person	−0.0046 (0.0290)	0.8742

*P values were calculated using a linear mixed model with time-dependent covariates adjusted for age, sex, smoking status and treatment modality. Numbers in bold denote statistically significant changes.
CAS, clinical activity score; NOSPECS, no signs or symptoms, only signs, soft tissue involvement, proptosis, extraocular muscle involvement, corneal involvement, sight loss; TBII, thyroid-stimulating hormone receptor-binding inhibitory immunoglobulin; TED, thyroid eye disease; TRAb, thyroid-stimulating hormone receptor antibody; TSI, thyroid-stimulating immunoglobulin.

analysis was performed only once at the initial presentation. Another retrospective chart analysis of 23 patients revealed that changes in the inflammatory phase of TED statistically correlated with changes in measured TSI level, indicating that serial TSI measurements may be adjunct in assessing the clinical inflammatory activity of TED.[19] Similar findings were reported in a prospective study by Eckstein *et al*[12] who demonstrated an association between TBII level and the clinical characteristics of TED, with TBII levels higher in patients with a severe course of TED (CAS≥4, NOSPECS score≥5) relative to those with a mild course (CAS<4, NOSPECS score<5).[12] In this study, patients underwent measurements of TBII levels at multiple time points during follow-up. The results revealed that certain cut-off TBII values could be defined as good or bad prognostic markers at each time point of development after the first 4 months of disease onset, demonstrating the prognostic value of measuring TBII levels at various time points. Similarly, Eckstein *et al*[12] also reported a decrease in TBII level over time but demonstrated no correlation between the extent of this reduction and clinical characteristics.

In terms of the natural course of Graves' disease, the disease may fluctuate in activity, and occasionally patients can spontaneously become euthyroid over time.[26 27] Treatment of the disease affects disease activity. Patients treated with surgery (thyroidectomy) or medication show a gradual decrease in serum TRAb level, and after 1 year, 50% to 60% of patients enter remission of TSH receptor autoimmunity accompanied by the disappearance of TRAb from serum.[26 28 29] Similarly, untreated TED improves spontaneously over time in most patients (described as 'Rundle's curve').[30 31] Moreover, studies report that TRAb levels associated with TED also decrease over time.[12 30] However, most of these studies were performed before the introduction of the bioassay; therefore, little is known about alterations in TSI levels over time. Furthermore, studies analysing the correlation between changes in serum TRAb levels and clinical course over time are limited.

In this study, we assessed correlations between the TED status and decreases in TRAb levels over time. Correlations between these outcomes were disaggregated to within-person and inter-person effects using a linear mixed model with time-dependent covariates, as the measured outcomes over time within each patient are not independent of one another.

Inter-person effects imply that lower average TRAb levels during the disease course correlate a greater reduction in dependent variables (CAS, NOSPECS score and proptosis) over time. Within-person effects suggest that dependent variables (CAS, NOSPECS score and proptosis) tend to decrease over time along with decreases in TRAb levels in individual patients during the disease course. The present analysis revealed that an inter-person difference in TRAb level correlated significantly with disease severity and activity, indicating that patients with low TSI and TBII levels showed a reduction in disease severity and activity during the 1-year follow-up. These results are meaningful, because the mean disease severity (measured using NOSPECS score and proptosis) did not show statistically significant reductions. The findings suggest that low baseline TRAb (both TSI and TBII) levels might be predictive of future decreases in disease severity in these patients.

Within-person analysis indicated that within-person changes in TRAb level in each individual correlated with changes in the disease course over time. No statistically significant results were obtained in this analysis, suggesting that although patients with low TRAb levels showed better prognosis, the decrease in serum TRAb level in one individual patient over time did not affect the clinical course. However, subgroup analysis of patients with a CAS≥4 at presentation demonstrated that within-person changes in

serum TSI level correlated significantly with the CAS and NOSPECS score, whereas baseline TSI level showed no statistically significant correlation with clinical characteristics (CAS and NOSPECS score) in the active subgroup. Thus, we inferred that if TSI levels decreased appropriately over time in patients with active disease at presentation, disease activity and severity will also likely decrease, leading to a better prognosis. These findings suggest that reductions in TSI levels over time have a better prognostic value for evaluating improvements in disease status than a lower baseline TSI level.

This study has limitations. Given the observational nature of the study and the lack of standardisation of the baseline clinical characteristics and treatment modalities during follow-up, differences in management may have affected the clinical characteristics over time. Therefore, variables that could affect the disease course of TED, such as age, sex, treatment modality and smoking, were included as covariates in the linear mixed model, and the adjusted results were presented. Another limitation is that the inclusion criteria were relatively broad, and the subtyping of TED (lipogenic or myogenic subtype orbitopathy) could not be differentiated owing to the lack of imaging data. Furthermore, factors that can affect TRAb levels, such as treatment modalities of Graves' disease and thyroid hormone levels of patients, were not considered in this analysis. Therefore, further interventional prospective studies with standardised management strategies and comprehensive evaluation of Graves' disease status are warranted.

Indicators for evaluating the activity and severity of thyroid ophthalmopathy have always been controversial.[32–34] CAS and the modified NOSPECS score have been challenged as reliable measures of disease status, as they sum the clinical features of the disease with equal weight.[34] For example, some features may worsen, whereas others may improve as the disease progresses. Consequently, a summative score, such as the modified NOSPECS score, might mask the effect of the severity of any one measure of the disease and its change over time. However, these scores have been regarded as reliable measures of disease activity and severity in the clinical setting and widely used as traditional indicators in evaluating TED activity and severity.[35–37] To avoid the limitations associated with the summative scoring system, we separately evaluated proptosis to measure disease severity.

In conclusion, we demonstrated correlated trends associated with reductions in TSI and TBII levels and changes in clinical characteristics of TED over time. Furthermore, in patients with active TED (a CAS≥4 at presentation), time-dependent within-person changes in serum TSI level correlated with disease severity and activity, indicating that active TED patients with smaller reductions in TSI levels over time would likely show prolonged disease activity and more severe disease. These findings offer clinicians insight to inform considerations of more aggressive or prolonged immunosuppressive therapy for TED.

Author affiliations
[1]Department of Ophthalmology, Severance Hospital, Institute of Vision Research, Yonsei University College of Medicine, Seoul, Korea (the Republic of)
[2]Department of Ophthalmology, Ajou University School of Medicine, Suwon, Korea (the Republic of)
[3]Nune Eye Hospital, Seoul, Korea (the Republic of)
[4]Department of Ophthalmology, Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea (the Republic of)
[5]Department of Ophthalmology, Inje University College of Medicine, Inje University Busan Paik Hospital, Busan, Korea (the Republic of)

Acknowledgements We gratefully acknowledge the statistical support of Eun Hwa Kim from the Biostatistics Collaboration Unit, Department of Biomedical Systems Informatics, Yonsei University College of Medicine. Biochemical analyses were supported by Dow Biomedica (Seoul, Republic of Korea).

Collaborators The following investigators contributed their data to this multicentre study: Sang In Khwarg (Seoul National University Hospital, Seoul, Korea); Koung Hoon Kook (Ajou University School of Medicine, Suwon, Korea); Sang Duck Kim (Wonkwang University School of Medicine, Iksan, Korea); Sehyun Baek (Korea University College of Medicine, Seoul, Korea); Jun Hyuk Son (Yeungnam University College of Medicine, Daegu, Korea); Suk-Woo Yang (Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Korea); Jae Wook Yang (Busan Paik Hospital, Inje University College of Medicine, Busan, Korea); Kyung In Woo (Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea); Helen Lew (CHA Bundang Medical Center, CHA University, Seongnam, Korea); Jin Sook Yoon (Yonsei University College of Medicine, Seoul, Korea); Sung Bok Lee (Chungnam National University College of Medicine, Daejeon, Korea); Jeong Kyu Lee (Chung-Ang University Hospital, Chung-Ang University College of Medicine, Seoul, Korea); Jae Woo Jang (Kim's Eye Hospital, Konyang University College of Medicine, Seoul, Korea); Hee Young Choi (Pusan National University College of Medicine, Busan, Korea); Mi Jung Chi (Gachon University Gil Medical Center, Incheon, Korea)

Contributors KIW and JWY conceived of the study. KHK developed and performed the analyses. JK and JSY verified the analytical methods and wrote the main manuscript. KIW is the guarantor of this work and accepts responsibility for the study. All authors provided critical feedback and participated in the research, analysis and review of the manuscript.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval This study involves human participants and was approved by Institutional Review Board of Severance Hospital (approval number 4-2016-1117). Participants gave informed consent to participate in the study before taking part. The complete datasets and written informed consent of each patient were obtained before starting the study. Approval by the institutional ethics review committee of each participating institution was obtained, and the study was performed according to the tenets of the Declaration of Helsinki.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. The datasets generated during and/or analysed during the current study are not publicly available due to the IRB recommendation not to send patient data into the public domain but are available from the corresponding author upon reasonable request.

ORCID iDs
JaeSang Ko http://orcid.org/0000-0002-3011-7213
Jin Sook Yoon http://orcid.org/0000-0002-8751-9467
Kyung In Woo http://orcid.org/0000-0003-4228-6556

REFERENCES

1 Bahn RS. Graves' ophthalmopathy. *N Engl J Med* 2010;362:726–38.
2 Naik VM, Naik MN, Goldberg RA, *et al*. Immunopathogenesis of thyroid eye disease: emerging paradigms. *Surv Ophthalmol* 2010;55:215–26.
3 Jang SY, Shin DY, Lee EJ, *et al*. Relevance of TSH-receptor antibody levels in predicting disease course in Graves' orbitopathy: comparison of the third-generation TBII assay and Mc4-TSI bioassay. *Eye* 2013;27:964–71.
4 Neigel JM, Rootman J, Belkin RI, *et al*. Dysthyroid optic neuropathy. The crowded orbital apex syndrome. *Ophthalmology* 1988;95:1515–21.
5 Khoo DH, Ho SC, Seah LL, *et al*. The combination of absent thyroid peroxidase antibodies and high thyroid-stimulating immunoglobulin levels in Graves' disease identifies a group at markedly increased risk of ophthalmopathy. *Thyroid* 1999;9:1175–80.
6 Noh JY, Hamada N, Inoue Y, *et al*. Thyroid-stimulating antibody is related to Graves' ophthalmopathy, but thyrotropin-binding inhibitor immunoglobulin is related to hyperthyroidism in patients with Graves' disease. *Thyroid* 2000;10:809–13.
7 Ponto KA, Kanitz M, Olivo PD, *et al*. Clinical relevance of thyroid-stimulating immunoglobulins in Graves' ophthalmopathy. *Ophthalmology* 2011;118:2279–85.
8 Gerding MN, van der Meer JW, Broenink M, *et al*. Association of thyrotrophin receptor antibodies with the clinical features of Graves' ophthalmopathy. *Clin Endocrinol* 2000;52:267–71.
9 Woo YJ, Jang SY, Lim THT, *et al*. Clinical association of thyroid stimulating hormone receptor antibody levels with disease severity in the chronic inactive stage of Graves' orbitopathy. *Korean J Ophthalmol* 2015;29:213–9.
10 Eckstein AK, Plicht M, Lax H, *et al*. Clinical results of anti-inflammatory therapy in Graves' ophthalmopathy and association with thyroidal autoantibodies. *Clin Endocrinol* 2004;61:612–8.
11 Kampmann E, Diana T, Kanitz M, *et al*. Thyroid stimulating but not blocking autoantibodies are highly prevalent in severe and active thyroid-associated orbitopathy: a prospective study. *Int J Endocrinol* 2015;2015:1–5.
12 Eckstein AK, Plicht M, Lax H, *et al*. Thyrotropin receptor autoantibodies are independent risk factors for Graves' ophthalmopathy and help to predict severity and outcome of the disease. *J Clin Endocrinol Metab* 2006;91:3464–70.
13 Cappelli C, Gandossi E, Castellano M, *et al*. Prognostic value of thyrotropin receptor antibodies (TRAb) in Graves' disease: a 120 months prospective study. *Endocr J* 2007;54:713–20.
14 Sanders J, Evans M, Premawardhana LDKE, *et al*. Human monoclonal thyroid stimulating autoantibody. *Lancet* 2003;362:126–8.
15 Filetti S, Foti D, Costante G, *et al*. Recombinant human thyrotropin (TSH) receptor in a radioreceptor assay for the measurement of TSH receptor autoantibodies. *J Clin Endocrinol Metab* 1991;72:1096–101.
16 Lytton SD, Li Y, Olivo PD, *et al*. Novel chimeric thyroid-stimulating hormone-receptor bioassay for thyroid-stimulating immunoglobulins. *Clin Exp Immunol* 2010;162:438–46.
17 Lytton SD, Kahaly GJ. Bioassays for TSH-receptor autoantibodies: an update. *Autoimmun Rev* 2010;10:116–22.
18 Jang SY, Shin DY, Lee EJ, *et al*. Clinical characteristics of Graves' orbitopathy in patients showing discrepancy between levels from TBII assays and TSI bioassay. *Clin Endocrinol* 2014;80:591–7.
19 Dragan LR, Seiff SR, Lee DC. Longitudinal correlation of thyroid-stimulating immunoglobulin with clinical activity of disease in thyroid-associated orbitopathy. *Ophthalmic Plast Reconstr Surg* 2006;22:13–19.
20 European Group on Graves' Orbitopathy (EUGOGO), Wiersinga WM, Perros P, *et al*. Clinical assessment of patients with Graves' orbitopathy: the European group on Graves' orbitopathy recommendations to generalists, specialists and clinical researchers. *Eur J Endocrinol* 2006;155:387–9.
21 Bartalena L, Baldeschi L, Dickinson AJ, *et al*. Consensus statement of the European group on Graves' orbitopathy (EUGOGO) on management of Graves' orbitopathy. *Thyroid* 2008;18:333–46.
22 Guerra MW, Shults J, Amsterdam J, *et al*. The analysis of binary longitudinal data with time-dependent covariates. *Stat Med* 2012;31:931–48.
23 Lalonde TL. *Modeling time-dependent covariates in longitudinal data analyses*. Innovative statistical methods for public health data: Springer, 2015: 57–79.
24 Curran PJ, Bauer DJ. The disaggregation of within-person and between-person effects in longitudinal models of change. *Annu Rev Psychol* 2011;62:583–619.
25 Hoffman L, Stawski RS. Persons as contexts: evaluating between-person and within-person effects in longitudinal analysis. *Res Hum Dev* 2009;6:97–120.
26 Laurberg P, Wallin G, Tallstedt L, *et al*. TSH-receptor autoimmunity in Graves' disease after therapy with anti-thyroid drugs, surgery, or radioiodine: a 5-year prospective randomized study. *Eur J Endocrinol* 2008;158:69–75.
27 Codaccioni JL, Orgiazzi J, Blanc P, *et al*. Lasting remissions in patients treated for Graves' hyperthyroidism with propranolol alone: a pattern of spontaneous evolution of the disease. *J Clin Endocrinol Metab* 1988;67:656–62.
28 Pinchera A, Liberti P, Martino E, *et al*. Effects of antithyroid therapy on the long-acting thyroid stimulator and the antithyroglobulin antibodies. *J Clin Endocrinol Metab* 1969;29:231–8.
29 Takamura Y, Nakano K, Uruno T, *et al*. Changes in serum TSH receptor antibody (TRAb) values in patients with Graves' disease after total or subtotal thyroidectomy. *Endocr J* 2003;50:595–601.
30 Menconi F, Profilo MA, Leo M, *et al*. Spontaneous improvement of untreated mild Graves' ophthalmopathy: Rundle's curve revisited. *Thyroid* 2014;24:60–6.
31 Bartley GB. Rundle and his curve. *Arch Ophthalmol* 2011;129:356–8.
32 Kim JW, Woo YJ, Yoon JS. Is modified clinical activity score an accurate indicator of diplopia progression in Graves' orbitopathy patients? *Endocr J* 2016;63:1133–40.
33 Douglas RS, Tsirbas A, Gordon M, *et al*. Development of criteria for evaluating clinical response in thyroid eye disease using a modified Delphi technique. *Arch Ophthalmol* 2009;127:1155–60.
34 Frueh BR. Why the NOSPECS classification of Graves' eye disease should be abandoned, with suggestions for the characterization of this disease. *Thyroid* 1992;2:85–8.
35 Barrio-Barrio J, Sabater AL, Bonet-Farriol E, *et al*. Graves' ophthalmopathy: visa versus EUGOGO classification, assessment, and management. *J Ophthalmol* 2015;2015:1–16.
36 Jang SY, Shin DY, Lee EJ, *et al*. Correlation between TSH receptor antibody assays and clinical manifestations of Graves' orbitopathy. *Yonsei Med J* 2013;54:1033–9.
37 Eckstein AK, Lax H, Lösch C, *et al*. Patients with severe Graves' ophthalmopathy have a higher risk of relapsing hyperthyroidism and are unlikely to remain in remission. *Clin Endocrinol* 2007;67:070630051835001–???.