Aus der I. Medizinischen Klinik und Poliklinik

der Universitätsmedizin der Johannes Gutenberg-Universität Mainz

Multicentric Standardization of a Bioassay for Thyroid Receptor Stimulating Antibodies

Multizentrische Standardisierung eines Bioassays zum Nachweis von Schilddrüsen-stimulierenden Antikörpern

Inauguraldissertation

zur Erlangung des Doktorgrades der Medizin

der Universitätsmedizin der Johannes Gutenberg-Universität Mainz

vorgelegt von

Markus Lehmann

aus Neuwied

Mainz, 2019

Wissenschaftlicher Vorstand: _

1. Gutachter: _

2. Gutachter: _

Tag der Promotion: 30. Juni 2020

Tibi

quia sine te

non potest

“An excuse is what you make after the deed is done, while a justification is what you offer before.” *(Brandon Sanderson, The Way of Kings)*

List of contents:

List of contents 1

List of tables and figures 3

List of abbreviations 7

1. Introduction 9
   - 1.1 Aim of the thesis 10

2. Literature discussion 11
   - 2.1 A short overview of the different bioassays developed for the detection of TSAbs 11
   - 2.2 The discovery of TSAbs and the International Standard for TSAb 15

3. Materials and Methods 17
   - 3.1 Materials 17
     - 3.1.1 TSAb bioassay 17
     - 3.1.2 Biological substances 18
     - 3.1.3 Technical equipment 18
     - 3.1.4 Consumables 19
     - 3.1.5 Software for data analysis 19
   - 3.2 Procedure 20
   - 3.3 Calculation of the results 21
   - 3.4 Methods 22
     - 3.4.1 Patient sera 22
     - 3.4.2 The 2nd International Standard dose-response curve 22
     - 3.4.3 Study design 22
     - 3.4.4 Repetition at the JGU laboratory 24

4. Results 25
4.1 IS dose-response curve 25
4.1.1 Results of the dose-response curves 25
4.1.2 The linear range of the IS dose-response curve 28
4.1.3 Samples above and below the linear range 31
4.2 Measurement of patient serum samples 31
4.2.1 Conversion process for patient samples 32
4.2.2 Overall results 34
4.2.3 Comparison between laboratories 34
4.2.4 Comparison between lots 37
4.2.5 Comparison between users of the same lab 41
4.2.6 Precision study 45
4.2.7 Differences in the measurement of the 2nd IS between laboratories 48
4.2.8 Repetition of the patient measurements at the JGU laboratory 50
4.2.9 Comparison between users 51
4.2.10 Comparison between lots 53
4.2.11 Differences in the measurement of the 2nd IS between the cell lots 56
4.2.12 Comparison between the same user 57

5. Discussion 60
5.1 The International Standard as reference 61
5.1.1 Behavior and availability 61
5.1.2 The linear range 63
5.1.3 Differences in measurements of the 2nd IS in the two laboratories 65
5.1.4 Differences between cell lots 66
5.2 Variability of results reported as SRR% and mIU/L 67
5.3 Conclusion 68

6. Summary 69

7. Zusammenfassung 70

8. Reference list 71

Acknowledgement 78

Curriculum Vitae 79

List of tables and figures

Table 1: Dose-response results for the third cell lot for user 1 at the JGU laboratory 25

Table 2: Mean SRR% values for the dose-response results at the JGU laboratory 27

Table 3: SRR% values for the patient sera and the IS concentrations for cell lot 1 on the first day 32

Table 4: Converted TSAb values for cell lot 1 on the first day 33

Table 5: Mean values for all converted results 34

Table 6: Mean values for converted results at the JGU laboratory 34

Table 7: Mean values for converted results at the DHI laboratory 35

Table 8: Mean values for the moderate TSAb positive samples at the JGU laboratory without outliers 36

Table 9: Comparison between the three cell lots at the JGU laboratory 37

Table 10: Comparison between the three cell lots at the DHI laboratory 39

Table 11:	Comparison between the two users at the JGU laboratory	41
Table 12:	Comparison between the two users at the DHI laboratory	43
Table 13:	Comparison between the two laboratories during precision measurements	45
Table 14:	Intra-assay precision values for the two laboratories	48
Table 15:	Comparison between mean SRR% values of patient samples between the two laboratories	48
Table 16:	Comparison between mean SRR% values of 2nd IS concentrations between the two laboratories	49
Table 17:	Mean values for the repetition of the measurements at the JGU laboratory	50
Table 18:	Comparison between the three cell lots during repetition of the measurements	54
Table 19:	Comparison between mean SRR% values for the three cell lots	56
Table 20:	Comparison between mean SRR% values for the three cell lots for the 2nd IS concentrations	56

Figure 1:	Dose-response curves for the third cell lot for user 1 at the JGU laboratory	26
Figure 2:	Dose-response curves for user 1 at the JGU laboratory	27
Figure 3:	Dose-response curves for user 2 at the JGU laboratory	28
Figure 4:	Display of the linear range	29
Figure 5:	Linear ranges for user 1 at the JGU laboratory	29

Figure 6: Linear ranges for user 2 at the JGU laboratory 30

Figure 7: Comparison between dose-response curve and linear range 31

Figure 8: Comparison between the two laboratories for the low TSAb positive samples 35

Figure 9: Comparison between the two laboratories for the moderate TSAb positive samples 36

Figure 10: Comparison between the three cell lots at the JGU laboratory 38

Figure 11: Comparison between the three cell lots at the DHI laboratory 40

Figure 12: Comparison between the two users at the JGU laboratory 42

Figure 13: Comparison between the two users at the DHI laboratory 44

Figure 14: Comparison between the two laboratories for the low TSAb positive samples during precision measurement 46

Figure 15: Comparison between the two laboratories for the moderate TSAb positive samples during precision measurement 47

Figure 16: Differing behavior between patient serum samples and 2nd IS concentrations at the two laboratories 49

Figure 17: Comparison between converted patient serum samples between original study and repetition 51

Figure 18: Comparison between the two users for the low TSAb positive samples during the repetition measurements 52

Figure 19: Comparison between the two users for the moderate

TSAb positive samples during the repetition measurements 53

Figure 20: Comparison between the three cell lots during the repetition measurements for the low TSAb positive samples 54

Figure 21: Comparison between the three cell lots during the repetition measurements for the moderate TSAb positive samples 55

Figure 22: Comparison between SRR% values of patient samples and IS concentrations for the three cell lots during repetition measurements 57

Figure 23: Comparison between both measurements of user 1 for the low TSAb positive patient serum samples 58

Figure 24: Comparison between both measurements of user 1 for the moderate TSAb positive patient serum samples 59

List of abbreviations:

aa	amino acids
cAMP	3‘,5‘-cyclic adenosine monophosphate
CAS	cell attachment solution
CHO	Chinese hamster ovary
CRE	cAMP response element
bTSH	bovine thyroid stimulating hormone
DBC	dibutyryl cyclic adenosine monophosphate
DHI	Diagnostic Hybrids Inc.
DMSO	dimethylsulfoxid
ECBS	Expert Committee on Biological Standardization
Fab	antigen binding fragment
FBS	fetal bovine serum
FDA	Food and Drug Administration
FRTL5	Fisher-rat thyroid cell line 5
fT3	free triiodothyronine
fT4	free thyroxine
GD	Graves‘ disease
GM	growth medium
GO	Graves‘ orbitopathy
IgG	immunoglobulin G
IU	international units
JGU	Johannes Gutenberg University

LATS	long acting thyroid stimulator
mAb	monoclonal antibody
mIU/L	milli International Units per Liter
MRC	Medical Research Council
NIBSC	National Institute for Biological Standards and Control
RB	reaction buffer
TPO	thyroid peroxidase
TRAb(s)	thyroid receptor antibody(ies)
TSAb(s)	thyroid stimulating antibody(ies)
TBAb(s)	thyroid stimulation blocking antibody(ies)
TSH	thyroid stimulating hormone
TSHR	thyroid stimulating hormone receptor
WHO	World Health Organization

1. Introduction:

Graves' disease (GD) is an autoimmune disorder of the thyroid which is characterized by the production of autoantibodies against the thyroid stimulating hormone (TSH) receptor (TSHR). Stimulation of the TSHR by thyroid stimulating autoantibodies (TSAbs) causes the clinical symptoms of hyperthyroidism in patients with all its appearances, including but not limited to elevated heart rate, poor heat tolerance, sleeping problems, muscle weakness, weight loss and psychological disorders. Furthermore TSAbs may cause Graves' orbitopathy (GO) by binding to orbital fibroblasts in the eye muscles and causing a differentiation into adipocytes with resulting expansion and inflammation as well as edema of the orbital tissue. Another extra-thyroidal manifestation of GD is the pretibial myxedema. Different treatment methods exist for GD. On the one hand there are anti-thyroidal medications like Carbimazole which works by inhibiting the thyroid peroxidase (TPO) and thus suppressing hormone production. On the other hand there are more permanent solutions e.g. surgery or radioiodine therapy which aim to decrease or eradicate the thyroidal tissue. The orbitopathy can be treated with anti-inflammatory drugs like corticoids, decompression-surgery or radiation treatment.

Diagnosis of GD is secured through the clinical appearance with typical symptoms like hyperthyroidism, exophthalmos and goiter combined with laboratory testing. Typically a decrease of the serum concentrations of TSH and an increase of the thyroidal hormones free triiodothyronine (fT3) and free thyroxin (fT4) are found. Pathognomonic for GD is the detection of thyroid receptor autoantibodies (TRAbs). TRAbs can be found in over 98% of untreated patients with GD (Menconi, Marcocci et al. 2014). TRAbs show different characteristics, dependent on their effect on the thyroid. They can be classified in stimulating, blocking or neutral antibodies, depending on their behavior *in vivo* and in cell-based bioassays (Kahaly and Diana 2017).

The activity of TSAbs in patients with GD was shown to correlate positively with the severity of the disease and the development of GO. Conversely, the activity of TSAbs correlates negatively with the effect of an anti-thyroidal therapy as well as permanent euthyroidism after treatment. Thus measurements of TSAbs can help in monitoring disease progression and therefore choosing the appropriate treatment method for the

patient. This is especially important in the light of the risks associated with surgery and other available treatment methods. TSAb are also frequent in patients with Hashimoto's thyroiditis and associated orbitopathy and may help in monitoring this disease (Kahaly, Diana et al. 2016).

Today, the measurement of TSAb is typically done via cell-based bioassays using transfected cell lines expressing a TSHR and a reporter gene (Lytton, Li et al. 2010) and (Leschik, Diana et al. 2013). The activity of the measured patient serum is reported as percentage of Specimen-to-Reference Ratio (SRR%). The result is always in relation to the reference control which is used in the cell-based bioassay. This practice has several disadvantages. Firstly, SRR% represents no absolute measurement but is always dependent on the reference control used, meaning that different non-unified references can lead to different outcomes. Furthermore comparisons among different laboratories or bioassays are only possible within certain limits, which can pose a problem in comprehensive studies. Therefore it would be beneficial for the use of bioassays to report their measurements in an absolute and internationally comparable fashion.

1.1 Aim of this doctoral thesis

The aim of this doctoral thesis is to develop a standardization protocol that enables the conversion of the current SRR% values into International Units per Liter (IU/L) values between different international laboratories using a cell-based bioassay for the measurement of TSAb. Until now the results of the bioassay were always reported as SRR% in relation to the reference control, bovine TSH. It is therefore desirable to use an internationally comparable and absolute unit, similar to the use of the international normalized ratio (INR) instead of Quick% in case of the Prothrombin time.

Utilization of an absolute unit of measurement would simplify comparison between laboratories and enable better studies regarding patients treated in different institutes.

For this purpose, the second International Standard (IS) for thyroid stimulating antibodies from the National Institute for Biological Standards and Control (NIBSC code: 08/204) was utilized.

2. Literature discussion:

2.1 A short overview of the different bioassays developed for the detection of TSAbs

The history of bioassays for TSAbs is closely intertwined with the development of bioassays for TSH. All assays use a TSHR, which is the target for TSH and also TSAbs, to measure how much the thyroid is stimulated. Most assays can be used to measure either TSH or TSAbs.

With regard to thyroid stimulation, the first *in vivo* developed bioassay worked by injecting guinea pigs and mice with thyroxin to suppress the endogenous TSH production as well as feeding them an iodine depleted diet supplemented with ^{131}I, a radioactive iodine isotope. This ensured that the test animals wouldn't produce thyroid hormones themselves and that only radioactive iodine would be present for hormone-synthesis. Later the animals were injected with a certain amount of TSH or patient serum. In the case of TSH, as well as serum containing TSAbs, the thyroid would start producing ^{131}I-loaded thyroid hormones which could then be measured with a scintillation counter (McKenzie 1958).

Working with animal models for the testing of TSH and TSAbs proved quite troublesome. Firstly, the animals had to be bred and cared for. Secondly, the animals were a huge confounder for the results, as each animal poses as an individual with a unique genetic makeup and a high dependency on environmental factors while growing up. This means that different mice or guinea pigs may respond discriminatively to the injected hormones due to variances in upbringing, genetic mutations or food they consumed. As such, researchers aimed to develop *in vitro* bioassays.

The older *in vitro* bioassays used mostly thyroid slices obtained from animals or humans, which were incubated with TSH or patient serum. For example, in 1957, Bakke et al. worked with bovine thyroid slices, which they incubated with TSH and measured the uptake of radioactive iodine into the cells (Bakke and Lawrence 1956)

as well as the increase of their weight as a marker for the stimulation (Bakke, Heideman et al. 1957).

Another assay utilizing thyroid slices derived from mice was developed by Brown and Munro. After feeding mice an iodine depleted diet with addition of thyroxin and ^{131}I, the thyroids were removed, and placed in a buffer solution were they were then incubated with either TSH, Long Acting Thyroid Stimulator (LATS) or the Fab-fragments of LATS. These experiments all showed similar concentrations of ^{131}I-T4 and ^{131}I-T3 in the buffer solution after incubation with the different thyroid stimulators (Brown and Munro 1967).

In the consecutive years, the use of thyroid slices *in vitro* bioassays changed to the use of cells in cell culture. One of the first examples for the measurement of TSH in a bioassay using cell culture was described by Plannels et al. in 1975. He used thyrocytes obtained from adult porcine thyroid glands. The cells were first incubated for three days with TSH in a buffer solution, then radioactive iodine was added and the cells were again incubated. After the incubation period, lysis was induced in the cells and the radioactivity of the suspensions was measured. Although Plannels et al. did not work with serum from patients with GD the use of a cell-culture was a step towards more standardized bioassays (Planells, Fayet et al. 1975).

In the meantime, Wilson et al. discovered that dibutyryl-cAMP (DBC) led to the same stimulatory activity and production of thyroid hormones in thyrocytes as TSH. While cAMP did not show the desired effect when incubated with the thyrocytes, DBC showed a response curve almost identical to TSH. Failure of cAMP to produce the desired effect was attributed to cAMP failing to cross the cell-membrane. This, and further experiments proved that cAMP is the intracellular mediator of TSHR stimulation. (Wilson, Raghupathy et al. 1968)

The discovery that cAMP is the intracellular mediator of TSH binding to the TSHR led to the idea to instead use cAMP as the endpoint in bioassays. Many assays measured an increase of cAMP in thyrocytes. This worked for assays using human and guinea pig thyroid slices (Onaya, Kotani et al. 1973). Another assay which utilized dog thyroid slices was developed by Rapoport (Rapoport 1976) and was then adapted to use a monolayer cell-culture (Rapoport and Adams 1978).

Around the same time, different researchers and laboratories began experimenting with human thyrocyte cultures. One assay described by Toccafondi et al. for the measurement of thyroid stimulation used human thyrocytes in cell monolayers which were incubated with bTSH or TSAb positive serum (Toccafondi, Aterini et al. 1980).

A positive assay improvement was made by Vitti et al. in a bioassay which detected cAMP in a continuous cell line of rat thyroid cells after incubation with TSH, sera from patients with GD and IgG derived from lymphocytes of GD patients. This assay utilized FRTL-5 (Fisher-rat thyroid cell line-5) cells and was more sensitive to detect the stimulatory activity of TSAb than the previously used human thyrocytes (Vitti, Rotella et al. 1983).

A few years later, a new cell line derived from the widely used Chinese hamster was developed by Perret et al. transfecting the Chinese hamster ovary (CHO) cells with human TSHR-DNA. This cell line was tested with TSH and TSAbs. A good correlation was shown between the stimulatory activity and intracellular cAMP increase (Perret, Ludgate et al. 1990).

Several bioassays were developed that used a CHO cell line called JPO9 CHO with a radioactive cAMP endpoint. One of the main advantages of the CHO cell line were that the cell culture was less cumbersome and decreased incubation times in comparison with the earlier bioassays using other cultured cell lines (Michelangeli, Munro et al. 1994).

The next chapter of TSAb bioassays was opened by the transfection of CHO cells with a cAMP response element (CRE) and a luciferase reporter gene (Watson, Ajjan et al. 1998).

In this cell line, stimulation of the human TSHR leads to an intracellular increase in cAMP and therefore an activation of the CRE and transcription of the luciferase gene. In a second step, lysis is induced in the cells and luciferin is added which leads to the emission of light, which can then be quantified with an ELISA reader. A bioassay using an improved cell line called JP26 CHO showed very promising results for the detection of TSAbs and proved to be highly efficient (Wallaschofski and Paschke 1999).

A new bioassay was developed by Araki et al. in 2015 which uses a CHO cell line expressing aequorin and cAMP gated calcium channels alongside the TSHR. When

the TSHR is stimulated, intracellular cAMP is increased through the G-protein coupled adenylate cyclase. This leads to the opening of cAMP gated calcium channels, through which extracellular Ca^{2+} enters the cells. Subsequently the aequorin is activated, emitting light. The light emission is then measured through an illuminometer. The main advantage of this bioassay is its short preparation time, as the bioassay does not depend on cell monolayers and the assay performance takes only about four hours (Araki, Iida et al. 2015).

In 2012, Pierce et al. published a paper about a newly generated CHO cell line which uses cAMP as an endpoint. For this assay, Pierce and colleagues constructed a CHO cell line, which expresses a TSHR-mutant ectopically. The human TSHR was fused to its respective GalphaS subunit with a GTPase-inactivation mutation for this TSHR-mutant. This assay was developed to be more sensitive to TSAb and showed a high specificity (Pierce, Sandrock et al. 2012).

The cell-based bioassay in this paper applied a genetically modified CHO-K1 cell line. It consists of a CHO cell line transformed with a plasmid comprising the firefly luciferase reporter gene under control of an alpha 4 glycoprotein promoter with tandemly repeated CREs (Lytton, Li et al. 2010). This assay utilizes a chimeric human TSHR where the amino acids (aa) 261-370 within the N-terminal extracellular domain were replaced by aa from the rat luteinizing and chorionic gonadotrophin hormone receptor. The use of the chimeric TSHR showed that it responded less to thyroid blocking antibodies (TBAbs) possibly present in patients with hypothyroidism and is therefore more specific to TSAbs (Tahara, Ishikawa et al. 1997) and (Grasso, Kim et al. 1999).

2.2 The discovery of TSAbs and the International Standard for TSAb

When Adams et al. used mice *in vivo* to measure the effect of TSH and serum of patients with GD, they discovered that the serum from patients with Graves' hyperthyroidism led to a delayed but longer acting response of the thyroid. They supposed to call this agent "Long Acting Thyroid Stimulator" (LATS) (Adams 1958) and (McKenzie 1958). LATS was detected in many patients with the clinical appearance of GD and in almost all patients with exophthalmos and localized pretibial myxedema (Meek, Jones et al. 1964).

To facilitate and simplify the research with patient serum, in 1964 the first standard for LATS was described. It was isolated from one patient with thyrotoxicosis through GD and labeled the Medical Research Council (MRC) Research Standard A for LATS (LATS A) (Dorrington and Munro 1964).

This substance was later replaced by the MRC Research Standard B for LATS. This standard consists of a serum sample from a patient with LATS activity and was labeled with the NIBSC code 65/122 (LATS-Standard 2013).

After the discovery, that the LATS is in fact a stimulating antibody to the TSHR, the name LATS subsequently changed to Thyroid Stimulating Antibody (TSAb) (Smith and Munro 1970).

The first International Standard (1st IS) for TSAb was later introduced by the NIBSC and the World Health Organization (WHO) in 1995. The preparation, labeled with the NIBSC code 90/672, was obtained by freeze-drying plasma proteins from a pregnant woman with GD whose plasma needed to be regularly exchanged via plasmapheresis during pregnancy. Calibration of the 1st IS was obtained by comparing the activity against the MRC Research Standard B for LATS, code 65/122. This preparation was used, among other applications, for the calibration of several competitive binding assays (Hata 2010).

Due to exhaustion of the original 90/672 stocks, the WHO Expert Committee on Biological Standardization (ECBS) recognized the need for a replacement of the 1st IS in the year 2006. As a follow-up material for 90/672 a human monoclonal thyroid stimulating antibody was put forth. This autoantibody was first isolated in the year 2003

from the lymphocytes of a 19-year-old patient with hyperthyroidism due to GD. For this purpose the patients' lymphocytes were infected with Epstein Barr virus and fused with a mouse/human hybrid cell line (K6H6/B5). After cloning, a single colony producing high concentrations of the TSHR autoantibody was isolated and it could be shown that the produced antibody, now named M22, was monoclonal (Sanders, Evans et al. 2003). The antigen-binding fragment (Fab) of the monoclonal antibody was shown to stimulate cAMP production in TSHR-transfected CHO cells on a similar level to porcine TSH, while higher doses of the intact M22 antibody were necessary to cause the same level of stimulation. Furthermore the stimulation of TSHR by M22 Fab and IgG was inhibited by patient sera with TSHR blocking activity similar to the stimulation by TSH (Sanders, Jeffreys et al. 2004).

An international collaborative study was undertaken by 13 laboratories in six countries to calibrate the M22 antibody preparation, which was now labeled 08/204, relative to the 1st IS, examine its behavior in receptor binding assays and bioassay systems and to determine the stability of the antibody (Burns 2010). The study concluded that a pool of serum with a high autoantibody titer from patients with GD might be a better replacement for the 1st IS, as it would resemble the original composition of the preparation for the 1st IS but since such a preparation was not available, 08/204 would serve as a replacement. Regarding the characteristics of the preparation, the study showed that there were no consistent slope differences in dose-response curves of 08/204 and 90/672, indicating that the assay response is comparable for both preparations. Measurements of the potency of 08/204 revealed that one ampoule contains 0.113 IU in comparison to the 1st IS for receptor binding assays, while it is 0.242 IU for bioassays. The between-laboratory variability for bioassays was measured with a percentage of coefficient variation (CV%) of 26.8%. This is higher than the between-laboratory variability for receptor binding assays with a CV% of 14.5%. This difference was ascribed to the 08/204 preparation being only stimulatory in nature compared to 09/672 containing both stimulatory and blocking antibodies. 08/204 was then introduced as the 2nd International Standard at the 61st meeting of the WHO ECBS in the year 2010. One vial of the 2nd IS 08/204 (0.113 IU) is equal to 1 µg of the thyroid stimulating monoclonal antibody (mAb) M22.

3. Materials and Methods

3.1 Materials

3.1.1 TSAb bioassay:

The experimental work for this doctoral thesis was conducted with the FDA-cleared Thyretain® bioassay (Quidel, San Diego, CA, USA). The Thyretain® kit contains the following contents (compare Thyretain® kit instructions):

1. CHO Mc4 FreshFrozenCells®: Cryovials containing CHO Mc4 cells cryogenically preserved in medium containing DMSO. Reagent is stored at -70° C or lower.
2. Cell Attachment Solution, 200 mL: A proprietary reagent that promotes rapid cell attachment is used to treat the wells of a 96-well plate prior to planting the cells. Reagent is stored at 2° to 30° C.
3. Growth Medium, 200 mL: Ham's F-12 cell culture medium containing 10% fetal bovine serum (FBS). Reagent is stored at 2° to 8° C.
4. Reaction Buffer, 500 mL: A proprietary buffer that augments the reaction of TSI with the TSHR. Reagent is stored at 2° to 8° C.
5. Control Set:
    a) Positive Control, 0.5 mL: TSI-containing human serum which yields a value that is ≥140% of the Reference Control. Reagent is stored at -70° C or lower.
    b) Reference Control, 0.5 mL: A bovine TSH (bTSH) containing solution against which controls and test specimens are compared. Reagent is stored at -70° C or lower.
    c) Normal Control, 0.5 mL: Human serum that is negative for the presence of TSI which yields a value that is <140% of the reference control. Reagent is stored at -70° C or lower.

6. Luciferase Assay Reagent Set:

   a) Luciferase Substrate, 1 vial: A lyophilized beetle luciferin substrate which is converted by luciferase to oxyluciferin and light. Reagent is stored at -20° C or lower.

   b) Luciferase Assay Buffer Solution, 1 vial, 10 mL: A cell culture lysis buffer.

3.1.2 Biological Substances:

1. Healthy Donor Serum: Serum taken from control donors without thyroidal or autoimmune disorders and no known chronic conditions.
2. Patient Serum: Serum taken from patients with Graves' hyperthyroidism
3. WHO 2nd International Standard for Thyroid Stimulating Antibody, NIBSC code 08/204, 0.113mIU per ampoule

3.1.3 Technical equipment:

1. ELISA Reader:	Tecan Infinite® M200
2. CO_2 -Incubator:	Sanyo MCO-18AIC
3. Laminar hood:	Holten LaminAir
4. Microscope:	Hund Wetzlar Wilovert S
5. Water bath:	Köttermann Labortechnik
6. Vortex mixer stirrer:	IKA Labortechnik VF2

3.1.4 Consumables:

96 well-plates	Corning Inc. Costar #3603, 96 well assay plate, black, clear bottom with lid, tissue culture treated, polystyrene
Small falcons	Eppendorf Conical Tubes, 15 mL, sterile, pyrogen-, DNase-, RNase- and DNA-free, colorless
Large falcons	Eppendorf Conical Tubes, 50 mL, sterile, pyrogen-, DNase-, RNase- and DNA-free, colorless
Small tubes	Eppendorf Tubes® 3810X, 1.5 mL, g-Safe®, colorless
Disposable reagent reservoirs	VistaLab Technologies disposable reagent reservoirs, 100 mL, sterile, pyrogen-, DNase-, RNase- and DNA-free, white

3.1.5 Software for data analysis:

1. Microsoft Excel 2010
2. GraphPad Prism 8.1.0
3. IBM SPSS Statistics 22

3.2 Procedure

The measurements of serum TSAb activity were performed according to the manufacturer's instructions:

Step 1: Carried out aseptically under the laminar hood:

1. The inner 48 wells of the 96 well-plates were treated with 100 µl Cell Attachment Solution (CAS) for 10 min at room temperature. Afterwards the CAS solution was discarded.
2. One CHO-Mc4 cell vial was thawed in a 37°C water bath and mixed with 5ml Growth Medium (GM) heated to 37°C and 100 µl were pipetted in each of the 48 wells.
3. The plates were incubated for 15-18h at 37°C, 5% CO_2.

Step 2: Carried out on the benchtop:

4. The plates were examined under the microscope regarding confluence of the cell monolayers. Plates exhibiting signs of microbial contamination, sub-confluence or over-confluence were disqualified.
5. The patient serum, components of the control-set and reaction buffer (RB) were heated to 37°C
6. Both patient serum samples and the three kit controls (positive, reference and normal control) were diluted 1:11 in RB.
7. The remaining GM in the plates was discarded and each well was rinsed with 100 µl RB, which was again decanted.
8. The wells were filled with 100 µl RB and subsequently 100 µl of the diluted patient serum or kit controls. Each serum and control sample was measured in triplicate.
9. The plates were again incubated at 37°C, 5% CO_2 for three hours.

Step 3: Carried out on the benchtop:

10. The lyophilized luciferase substrate and luciferase assay buffer solution were thawed and mixed before use.
11. The RB and samples were discarded from the wells and 75 µl of the mixed luciferase solution was added in each well.
12. After 10 min the luminescence was measured with the Elisa plate reader and the raw data were reported as relative light units (RLU).

3.3 Calculation of the results

Results of the bioassay were reported as percentage of specimen-to-reference ratio (SRR%). The first step is to calculate for each sample the mean value of the triplicate measurements:

Average RLU = (RLU well 1 + RLU well 2 + RLU well 3) / 3.

Afterwards, the patient RLU value is compared to the reference RLU:

*SRR% = ([Average RLU patient] / [Average RLU Reference Control]) * 100%.*

For each specimen the CV% value was determined using the following formula:

*CV% = ([Standard deviation RLU] / [Average RLU])*100%.*

If the CV% was above 15% the results were discarded and the assay would be re-run.

The cutoff of the assay is at 140%, which means that results with SRR% ≥ 140% are considered as TSAb positive.

3.4 Methods

3.4.1 Patient sera:

The patient serum samples that were measured during the study were taken from patients with GD which were monitored and treated at the JGU laboratory. There were nine patient serum samples in total, from patients whose blood was frequently taken during diagnosis and treatment of GD hyperthyroidism. The samples were chosen such that three serum samples were measured between SRR% of 140 - 279% in the original bioassay, those were called the low TSAb positive patient serum samples. Three serum samples were measured between SRR% of 280 – 419% and were called the moderate TSAb positive patient serum samples. The last three serum samples, the high TSAb positive patient serum samples, were measured above SRR% of 420%.

For comparison, serum from four healthy donors without thyroid or other autoimmune disorders was used. The healthy donors were associates at the JGU laboratory.

3.4.2 The 2nd International Standard dose-response curve:

To evaluate whether the 2nd IS could be used to convert the SRR% respectively the RLU obtained by the original bioassay into International Units (IU), it was necessary to obtain a complete dose-response curve for the mAb. For the preparation of the IS dose-response curve, one vial of the 2nd IS (0.113 IU) was dissolved in two ml phosphate buffered saline. This stock solution was diluted in a 1:11 mixture of healthy donor serum and RB to generate the final IS concentrations (0.3125, 0.625, 1.25, 2.5, 5, 10, 20, 40, 60, 80, 100, 120, 160 and 200 mIU/L) used for the dose-response curve.

For conversion, the range of the IS dose-response curve was chosen, that could best be approximated with a linear function of the type $y = a*x + b$.

3.4.3 Study design:

Performance of the bioassay and the conversion of SRR% into mIU/L were analyzed in a study comprising three different laboratories, each with two different users to run

the bioassay. The following three laboratories participated in performing the experiments: The Diagnostic Hybrids Inc. (DHI) laboratory in Ohio, USA; the Mayo clinic in Rochester, USA and the Molecular Thyroid Research Laboratory of the Johannes Gutenberg University (JGU) Medical Center in Mainz, Germany.

All dilutions and samples were prepared in one batch by the DHI laboratory and then distributed as frozen aliquoted samples to the remaining participants. Unfortunately, the Mayo clinic in Rochester lost their samples due to internal difficulties and thus could not present their data. Due to these circumstances, only the results of the other two participants will be reported and discussed.

The study was subdivided into three different parts:

1. In the first part of the study, the concentrations of the IS dose-response curve were measured by each user with three different cell lots over three consecutive days, producing nine dose-response curves per user.
2. For the second part of the study, the linear range of the dose-response curve of the 2nd IS (5, 10, 20, 40 and 80 mIU/L) and 13 patient serum samples were tested on one plate. Again, each user measured all samples three times with three different cell lots each, for a total of nine measurements per user per sample. The 13 patient serum samples consisted of the following:
    a. Three high TSAb positive samples previously measured with SRR% > 420%.
    b. Three moderate TSAb positive samples previously measured with SRR% 280 - 419%.
    c. Three low TSAb positive samples previously measured with SRR% 140 - 279%.
    d. Four normal sera from healthy subjects.
3. In the final part of the study, one user per laboratory measured the same patient serum sample set as in the second part, but this time with only one cell lot over 20 consecutive days.

3.4.4 Repetition at the JGU laboratory:

After the initial study, enough material was left at the JGU laboratory in order to repeat the second part of the study. Similar to the original study, two users carried out the measurements of the patient serum samples including the linear conversion curve with each of the three cell lots on three consecutive days. Again, this produced nine samples for each user and 18 samples in total.

4. Results

4.1 IS dose-response curve

4.1.1 Results of the dose-response curves

During the first part of the study, the aim was to obtain the dose-response curve of the 2nd IS for TSAb. When plotting the common logarithm of the IS concentrations against the SRR%, the resulting curve shows a concentration-dependent increase resembling a sigmoidal curve. As an example for the results of the dose-response curve measurements, the data for three curves, measured by one user over three days and the associated dose-response curves are shown below:

2nd IS (mIU/L)	**Mean (SRR%)**	**SD**	**CV%**
0.3125	36	2	5.7%
0.625	42	5	13.0%
1.25	48	5	9.4%
2.5	66	6	8.6%
5	105	5	5.0%
10	167	43	25.6%
20	258	17	6.7%
40	331	40	12.0%
60	397	26	6.6%
80	426	42	9.9%
100	458	67	14.5%
120	494	43	8.7%
160	524	33	6.3%
200	507	72	14.1%

Table 1: *2nd IS concentrations measured with the third cell lot by user 1 at the JGU laboratory over three consecutive days with their respective mean values, standard deviations and coefficient of variation.*

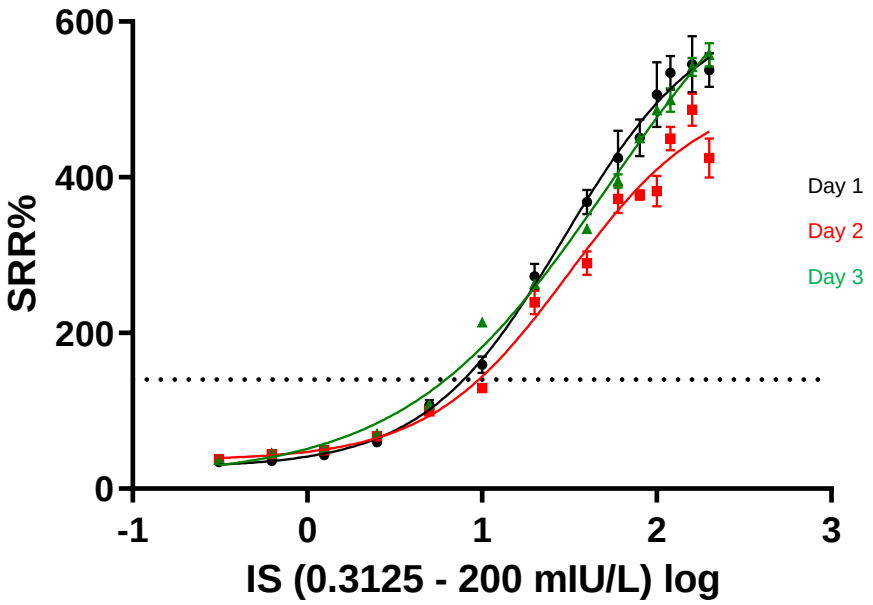

Figure 1: *The dose-response curves for the 2nd IS measured by user 1 with the third cell lot on three consecutive days.*

At the JGU laboratory, the two users generated 18 dose-response curves: Each user measured each of the three lots on three days. The mean values for these 18 curves, as well as their standard deviations and coefficients of variation are shown in the following table. The associated dose-response curves are shown below.

2nd IS (mIU/L)	Mean (SRR%)	SD	CV%
0,3125	44.7	11.0	24.6%
0,625	49.5	11.4	22.9%
1,25	58.2	14.5	25.0%
2,5	74.9	15.4	20.5%
5	124.1	32.5	26.2%
10	186.3	45.6	24.5%
20	280.1	51.7	18.5%
40	392.3	94.2	24.0%
60	427.6	82.4	19.3%
80	464.6	99.0	21.3%
100	494.9	116.1	23.5%
120	533.1	126.8	23.8%
160	557.9	127.3	22.8%
200	614.4	182.7	29.7%

Table 2: *The mean SRR% values for the dose-response measurement for the two users at the JGU laboratory with the respective standard deviations and coefficient of variation.*

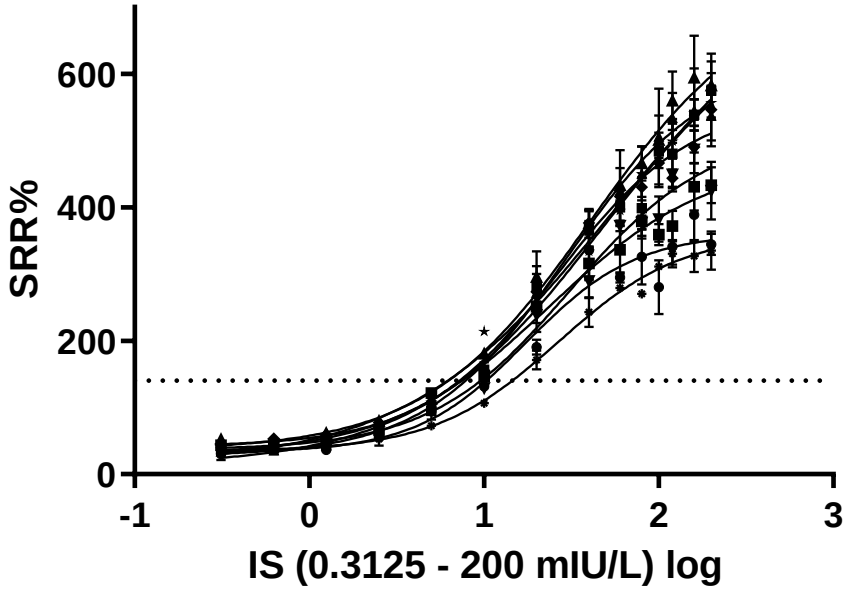

Figure 2: *The nine dose-response curves obtained by user 1 at the JGU laboratory. The dotted line indicates the assay cut-off at SRR of 140%.*

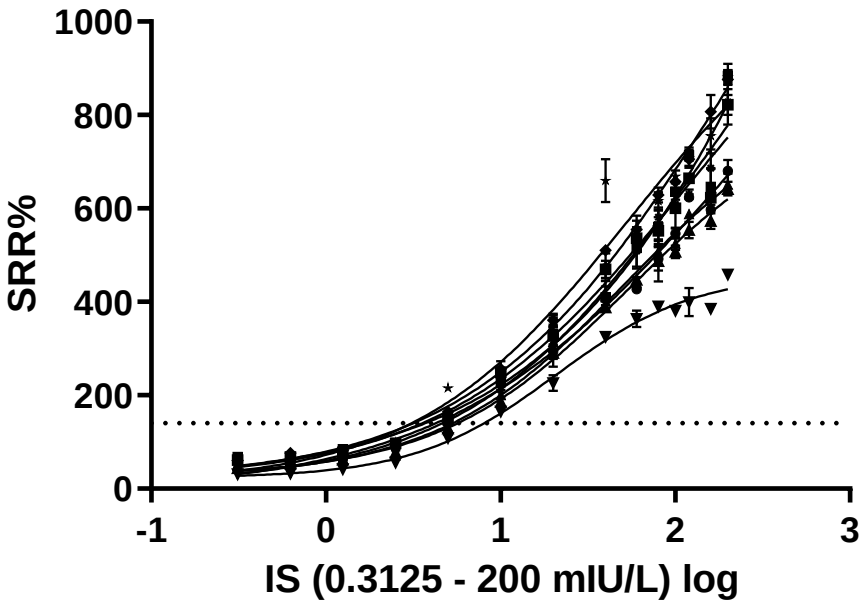

Figure 3: *The nine dose response curves obtained by user 2 at the JGU laboratory. Again, the dotted line shows the assay cut-off of the bioassay at 140%.*

4.1.2 The linear range of the IS dose-response curve:

For the conversion of patient samples, a linear range of the dose-response curve had to be figured out. As can be seen from the figure below, the linear range for the dose-response curve for the 2nd IS is between 5 to 80 mIU/L. Below 5 mIU/L and above 80 mIU/L the dose-response curve shows a plateau and therefore these concentrations cannot be considered for the linear equation. The SRR% value of patient samples within the linear range can be converted into mIU/L. The linear curves derived from the concentrations 5 – 80 mIU/L showed consistently a high coefficient of determination (R^2-value). The coefficient of determination was always above 95%, thus proving the linear approximation to be a good fit.

Still, as there is a lot of variance between the standard curves, the IS concentrations in the linear range were measured on each plate.

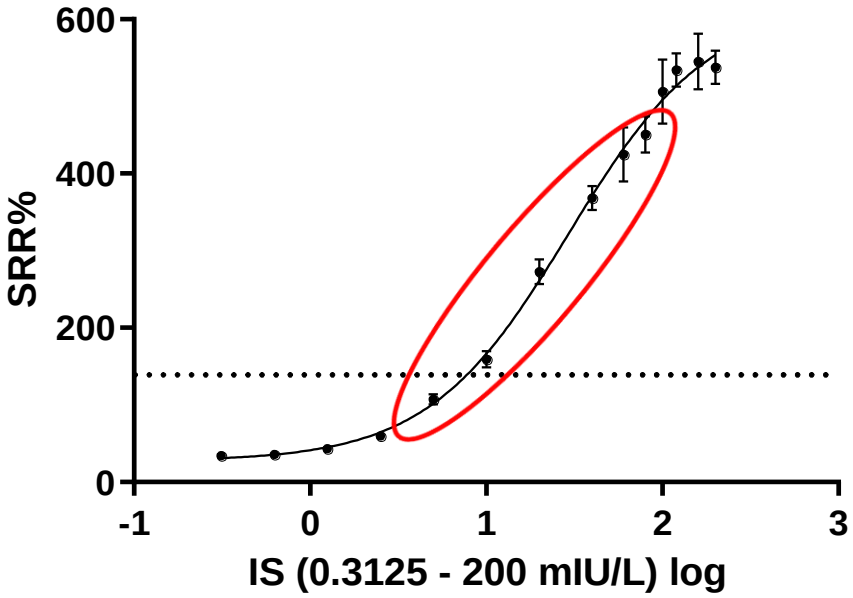

Figure 4: *The red circle shows the linear range between 5 and 80 mIU/L of the sigmoidal dose-response curve.*

The linear ranges of the IS are shown in the following figures:

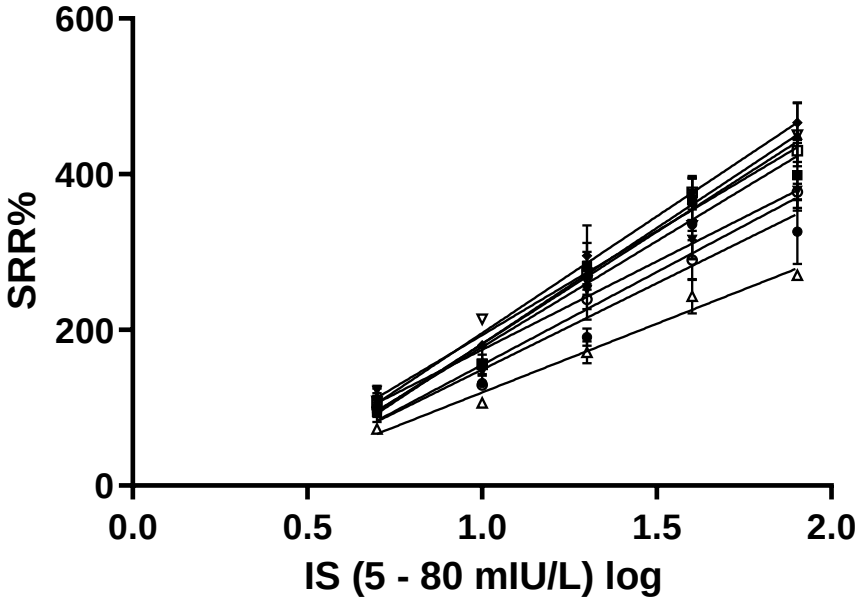

Figure 5: *Linear ranges of the nine dose-response curves acquired by user 1 at the JGU laboratory.*

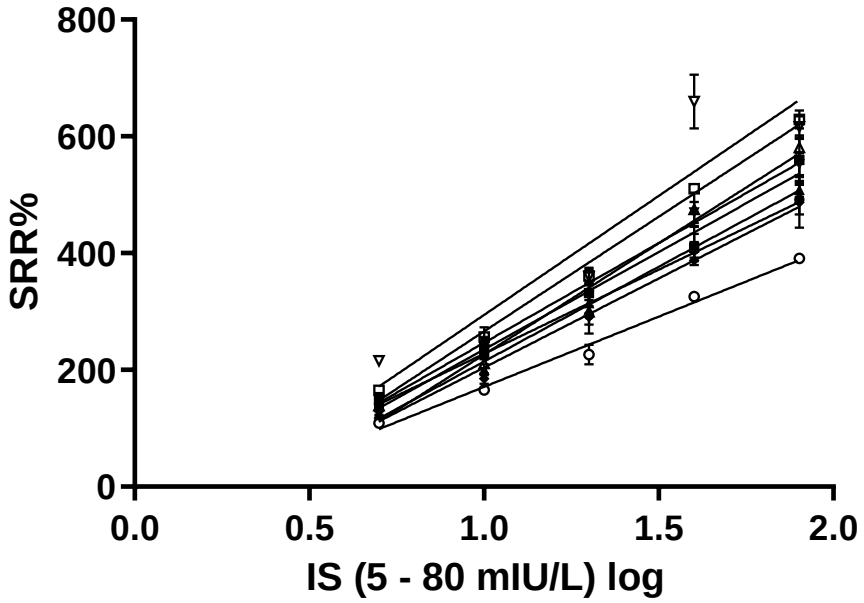

Figure 6: *Linear ranges of the nine dose-response curves acquired by user 2 at the JGU laboratory.*

The linear curve used to approximate the increase in SRR% for the IS concentrations from 5 – 80 mIU/L can be described by a linear equation of the type:

$$a * x + b = c$$

Herein the parameters “a” and “b” are determined by the curves slope and interception of the y-axis. “c” is the resulting SRR% and “x” is the equivalent of log(mIU/L). To derive the mIU/L value, the equation has to be solved for “x” and the tenth potency is to be taken:

$$log(x) = log((c-b)/a)$$

Solving this equation for all acquired patient SRR% values, the equivalent mIU/L can be determined.

The calculation of the linear equations were done with the linear regression tool of Microsoft Excel.

4.1.3 Samples above and below the linear range:

As shown before, below 5 mIU/L and above 80 mIU/L the sigmoidal standard-curve, from which the linear range was derived, showed plateaus. Therefore a linear equation cannot be used in order to convert patient samples that are below 5 mIU/L and above 80 mIU/L.

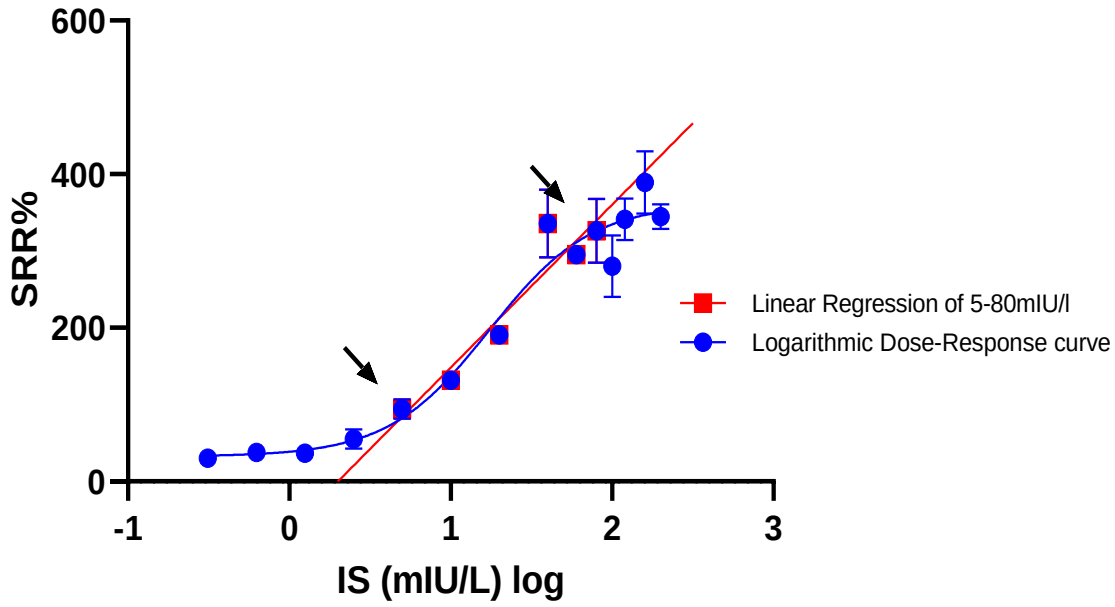

Figure 7: *This figure shows a dose-response curve and the linear regression for the six data points of the linear range. The blue and red lines align quite well for this range. Above 80 mIU/L and below 5 mIU/L (marked with the arrows), the dose-response curve deviates from the linear regression line.*

4.2 Measurement of patient serum samples

During the study, 13 different patient serum samples were measured. Of these 13 samples, four were from healthy control subjects. These samples could not be converted as their SRR% was permanently below the aforementioned linear range (< 5 mIU/L). The remaining nine patient serum samples contained three low positive, three moderate positive and three high TSAb positive serum samples. While the low positive and moderate TSAb positive serum samples were within the linear range, the

three high TSAb positive serum samples were always measured above the linear range and thus could not be converted, too. Hence, only the results of the low and moderate TSAb positive samples will be reported and discussed in the following chapters.

4.2.1 Conversion process for patient samples:

First, the conversion process when measuring the patient samples will be shown:

For cell lot 1 (#022514), on the first day of measurement, the following SRR% values were obtained by one of the users:

Patient Samples	Mean TSAb SRR%
Low TSAb 1	194
Low TSAb 2	196
Low TSAb 3	199
Mod TSAb 1	350
Mod TSAb 2	306
Mod TSAb 3	349
High TSAb 1	592
High TSAb 2	594
High TSAb 3	614
Healthy control 1	39
Healthy control 2	41
Healthy control 3	37
Healthy control 4	19

IS (mIU/L)	Mean TSAb SRR%
80	349
40	294
20	222
10	144
5	111

Table 3: *SRR% values obtained by user 1 on the first day of measurement with cell lot 1. Top: the low, moderate and high TSAb positive patient samples as well as the healthy control sera. Bottom: the concentrations of the 2nd IS for the conversion curve. In red the high TSAb positive serum samples above the linear range.*

The first three samples are the low TSAb positive patient sera, the following three samples are the moderate TSAb positive patient sera, the three samples highlighted in red are the high TSAb positive patient sera and the last four samples are the control sera from healthy donors. In the table below, the IS concentrations for the standard curve, measured in mIU/L, are shown. For each patient serum the SRR% value is reported. As shown, the high positive sera have SRR% values above the SRR% of the 80 mIU/L standard sample and are above the linear range. Samples above the linear range cannot be converted, as due to the plateaus of the standard curve, the linear equation deviates from the standard curve.

Linear regression for the IS concentrations gives the following equation:

$$y = 208x - 47 \ (R^2 = 0.99)$$

The coefficient of determination for the linear regression of the standard concentrations is very good with an R^2 value of 0.99. This shows that the linear equation is a good fit for the IS concentrations. For conversion, the equation is solved for *x* and SRR% values are inserted for *y*. Subsequently, the tenth potency is taken. The results are shown in the table below:

Patient samples	**TSAb (mIU/L)**
Low TSAb 1	14.3
Low TSAb 2	14.6
Low TSAb 3	15.2
Mod TSAb 1	80.3
Mod TSAb 2	49.8
Mod TSAb 3	79.5

Table 4: *Converted patient serum samples for the first cell lot on the first day of measurement for user 1 at the JGU laboratory*

As the healthy donor samples and the high positive samples could not be converted, they were left out. The results of the second part of the study will encompass the low and medium TSAb positive samples converted to mIU/L.

4.2.2 Overall results:

For the first examination of the results, the mean from all measurements done by the four users of both laboratories was taken. The overall mean values for the low TSAb positive samples are 13.2, 12.0 and 13.7 mIU/L respectively with CV% values between 14 and 23%. The moderate TSAb positive samples show mean values of 33.7, 48.0 and 50.3 mIU/L with higher CV% values of 35-40%.

Patient samples	Mean (mIU/L)	SD	CV%
Low TSAb 1	13.2	2.9	21.7%
Low TSAb 2	12.0	2.7	23.0%
Low TSAb 3	13.7	2.0	14.3%
Mod TSAb 1	50.3	19.8	39.4%
Mod TSAb 2	33.7	12.5	37.0%
Mod TSAb 3	48.0	17.2	35.8%

Table 5: *Mean values for the patient measurements by all four users during the second part of the study. Included are standard deviation and coefficient of variation for the patient samples.*

The results for the bioassay show quite high variances, especially in the moderate TSAb positive samples, as can be seen by the high CV% values. To evaluate where these variances stem from, different influencing factors are shown in the following paragraphs.

4.2.3 Comparison between laboratories:

Comparison between the two laboratories show lower mean values for all patient serum samples at the JGU laboratory, as can reported in the following tables:

JGU	Mean (mIU/L)	SD	CV%
Low TSAb 1	11.5	2.0	17.5%
Low TSAb 2	10.6	2.3	21.4%
Low TSAb 3	12.5	1.3	10.7%
Mod TSAb 1	38.5	19.5	50.8%
Mod TSAb 2	25.5	9.6	37.5%
Mod TSAb 3	40.3	19.7	48.8%

DHI	Mean (mIU/L)	SD	CV%
Low TSAb 1	14.9	2.6	17.5%
Low TSAb 2	13.3	2.6	19.3%
Low TSAb 3	15.0	1.7	11.1%
Mod TSAb 1	62.2	11.4	18.4%
Mod TSAb 2	41.9	9.3	22.2%
Mod TSAb 3	55.8	9.7	17.5%

Tables 6-7: *The mean values of the converted patient samples measured at the JGU and the DHI laboratories with the respective standard deviations and coefficient of variation.*

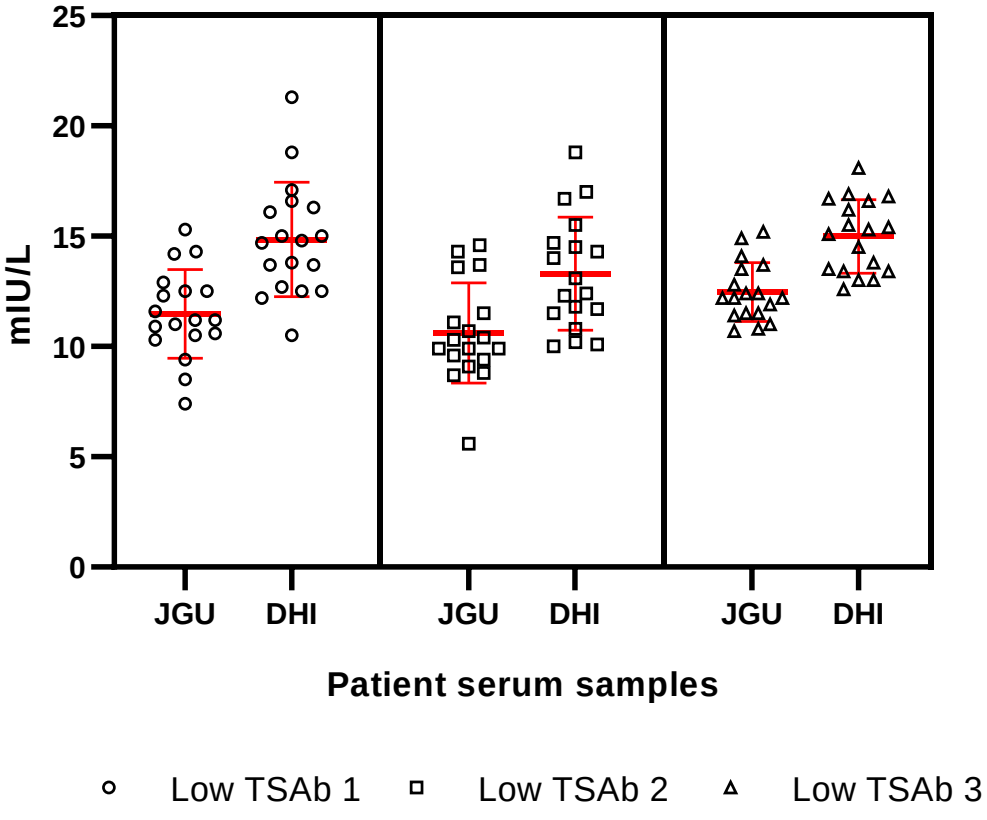

Figure 8: *Comparison between the JGU and DHI laboratories for the three different low TSAb positive patient serum samples.*

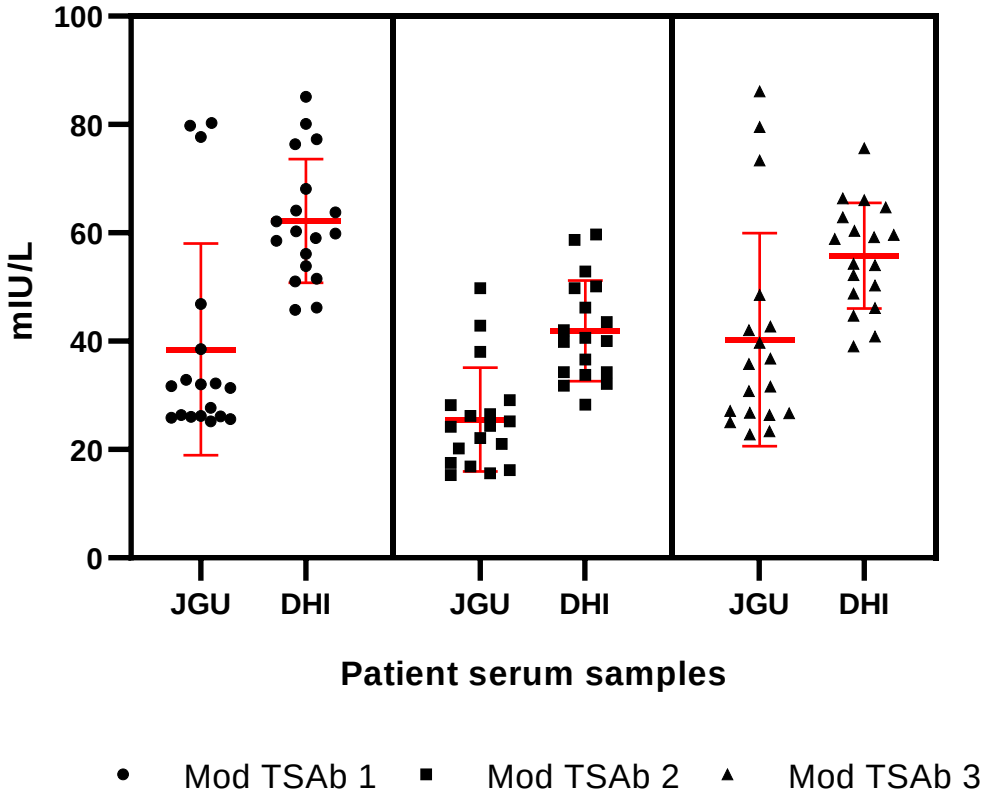

Figure 9: *Comparison between the JGU and DHI laboratories for the three moderate TSAb positive patient serum samples.*

In comparison to the DHI laboratory, the measurements at the JGU laboratory show a very high variance for the moderate TSAb positive values. The high CV% values can be explained by three outlier measurements. Omission of the outliers gives mean values of 30.3, 21.9 and 32.4 for the moderate TSAb positive patient serum measurements and a CV% of 20-25%.

JGU	Mean (mIU/L)	SD	CV%
Mod TSAb 1	30.3	6.0	19.7%
Mod TSAb 2	21.9	4.8	21.7%
Mod TSAb 3	32.4	8.0	24.8%

Table 8: *The mean values for the moderate TSAb positive samples measured at the JGU laboratory without outliers.*

In summary, each patient measured at the two laboratories showed higher results at the DHI laboratory than at the JGU laboratory. The low TSAb positive patient serum samples were measured at approximately 20-30% higher at the DHI laboratory, while the moderate TSAb positive patient serum samples were measured 40-65% higher. With omission of the outlier measurements at the JGU, the results were 70-100% higher at the DHI laboratory for the moderate TSAb positive samples.

4.2.4 Comparison between lots:

The comparison between the three cell lots were made for each laboratory separately to prevent differences between the laboratories from overshadowing the comparison between the cell lots.

JGU	**Mean (mIU/L)**		
	Cell lot 1	**Cell lot 2**	**Cell lot 3**
Low TSAb 1	11.4	11.3	11.7
Low TSAb 2	11.1	10.7	10.1
Low TSAb 3	12.5	11.9	13.0
Mod TSAb 1	45.1	40.1	30.2
Mod TSAb 2	30.2	24.8	21.5
Mod TSAb 3	45.5	42.1	33.2

Table 9: *Comparison between the three cell lots for the low and moderate TSAb positive patient serum samples at the JGU laboratory. The CV% values were between 8 and 24% for the low TSAb positive patient serum samples and between 10 and 61% for the moderate TSAb positive patient serum samples*

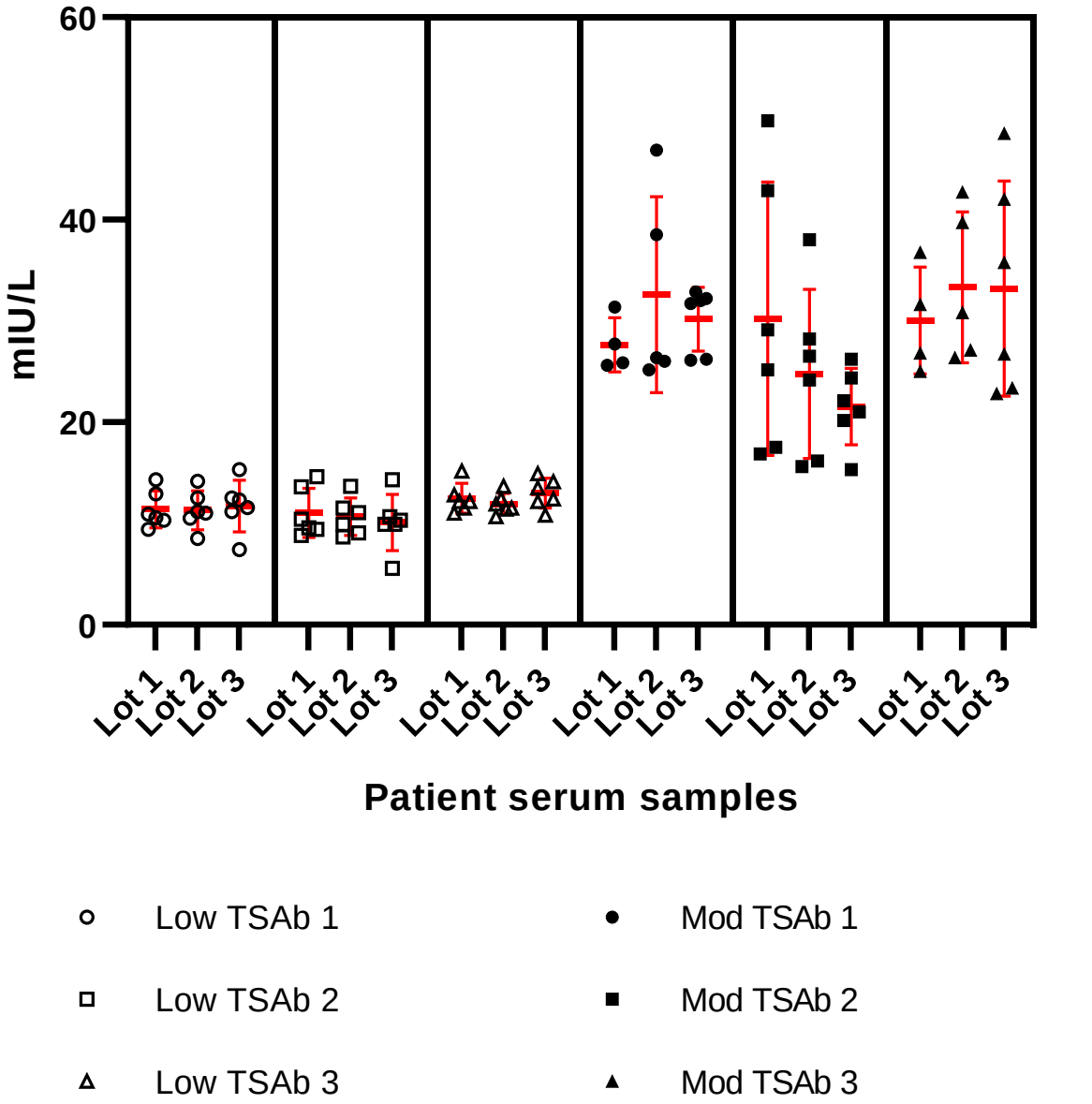

Figure 10: *Comparison between the three cell lots for the six patient serum samples at the JGU laboratory.*

For the low TSAb positive samples, the three lots show no great differences. Although the moderate TSAb positive samples show greater differences between the lots, they also exhibit quite high variances. No significant difference can be found between the three different cell lots that were used.

DHI	Mean (mIU/L)		
	Cell lot 1	**Cell lot 2**	**Cell lot 3**
Low TSAb 1	14.8	13.8	16.0
Low TSAb 2	13.4	12.4	14.1
Low TSAb 3	15.3	14.0	15.7
Mod TSAb 1	64.2	60.2	62.2
Mod TSAb 2	44.6	37.2	44.0
Mod TSAb 3	56.1	52.3	59.0

Table 10: *Comparison between the three cell lots for the low and moderate TSAb positive patient serum samples at the DHI laboratory. The CV% values were between 8 and 24% for the low TSAb positive patient serum samples and between 8 and 29% for the moderate TSAb positive patient serum samples*

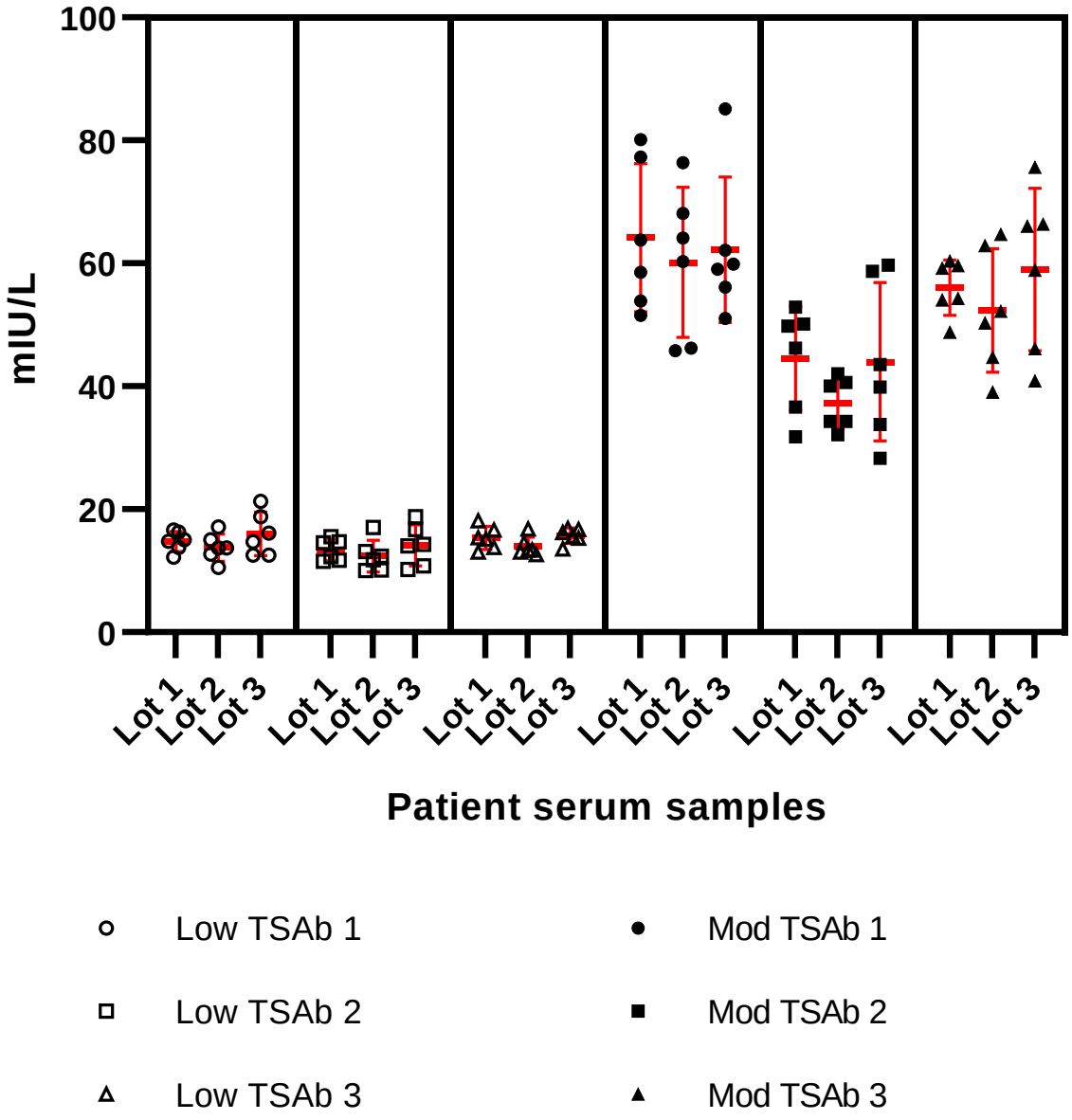

Figure 11: *Comparison between the three cell lots for the six patient serum samples at the DHI laboratory*

Similar to the measurements at the JGU laboratory, the differences between the three lots are quite low for the low TSAb positive patient serum samples. While the measurements at the JGU laboratory showed higher variances for the moderate TSAb positive serum samples, results from the DHI laboratory showed smaller discrepancies between lots and also lower CV% values.

Recapulatory, there was no significant difference between the results obtained from each of the cell lots. Neither the measurements from the JGU laboratory nor the measurements from the DHI laboratory allowed the conclusion that the three different cell lots are responsible for the fluctuations of the results.

4.2.5 Comparison between users of the same lab:

The last comparable influencing factor for the results is posed by the two users in each laboratory that performed the experiments. Although the procedure of the bioassay is exactly defined, human error and small differences in the approach between two users cannot be eliminated.

JGU	**Mean (mIU/L)**		**SD**		**CV%**	
	User 1	**User 2**	**User 1**	**User 2**	**User 1**	**User 2**
Low TSAb 1	12.1	10.8	2.4	1.3	19.9%	12.3%
Low TSAb 2	11.5	9.7	2.9	0.9	24.9%	8.9%
Low TSAb 3	12.6	12.3	1.3	1.4	10.6%	11.4%
Mod TSAb 1	49.5	27.4	23.0	2.6	46.4%	9.5%
Mod TSAb 2	31.5	19.5	9.8	4.1	31.2%	21.0%
Mod TSAb 3	52.4	28.2	21.8	3.8	41.7%	13.6%

Table 11: *Comparison between the two users for the low and moderate TSAb positive patient serum samples at the JGU laboratory.*

As can be seen from the following figure, the high variance and higher mean values for the moderate TSAb positive samples for the first user stem mostly from the three outlier measurements. The remaining results show quite low differences between the two users, although user 1 had consistently higher measurements compared to user 2.

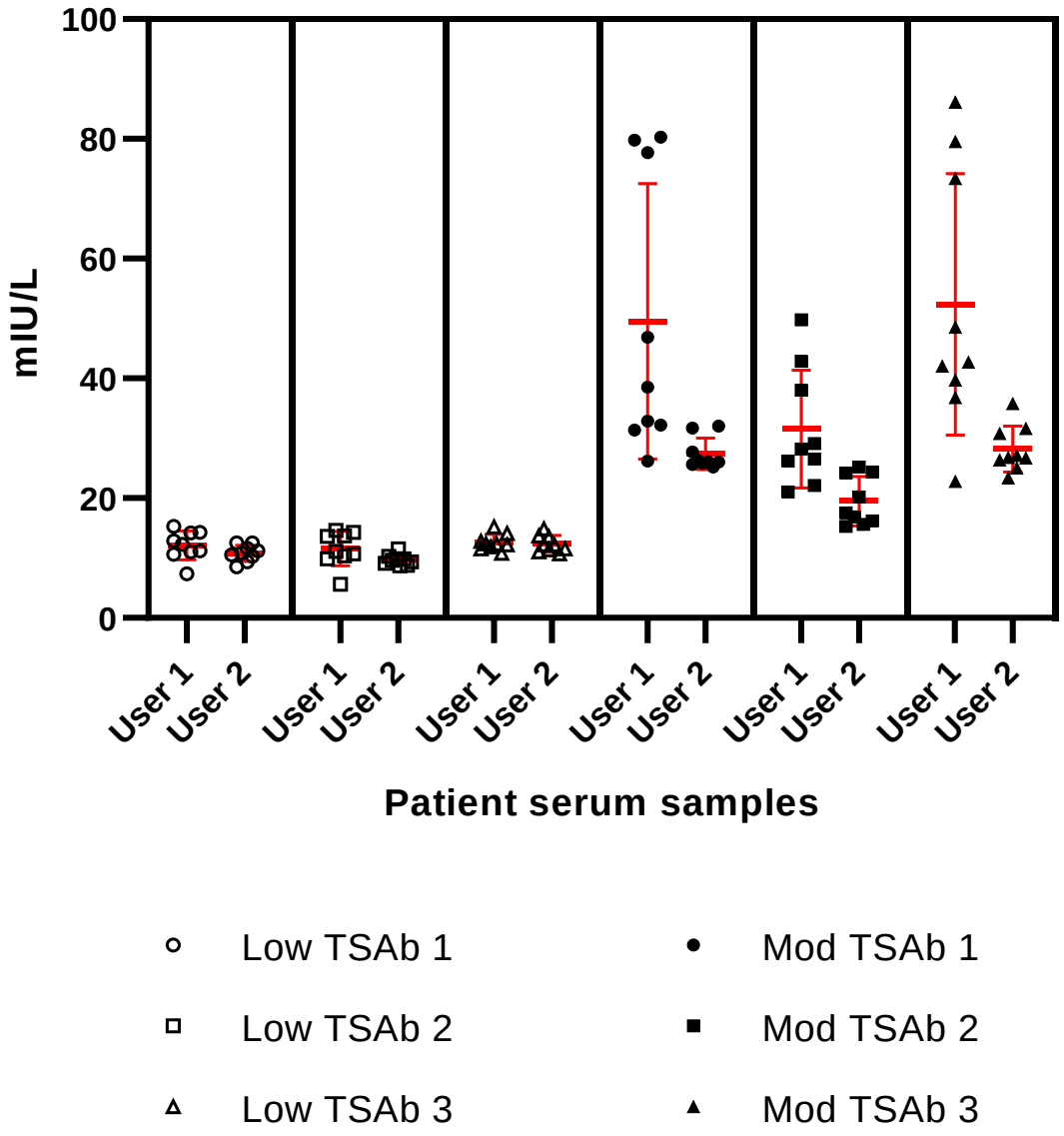

Figure 12: *Comparison between the two users at the JGU laboratory.*

DHI	Mean (mIU/L)		SD		CV%	
	User 1	User 2	User 1	User 2	User 1	User 2
Low TSAb 1	15.1	14.6	2.0	3.2	13.4%	21.8%
Low TSAb 2	13.4	13.2	3.0	2.2	22.5%	16.8%
Low TSAb 3	14.9	15.1	1.9	1.5	12.8%	9.8%
Mod TSAb 1	61.2	63.2	7.7	14.7	12.5%	23.3%
Mod TSAb 2	37.7	46.2	5.9	10.4	15.7%	22.5%
Mod TSAb 3	55.6	56.0	9.8	10.3	17.6%	18.4%

Table 12: *Comparison between the two users for the low and moderate TSAb positive patient serum samples at the DHI laboratory.*

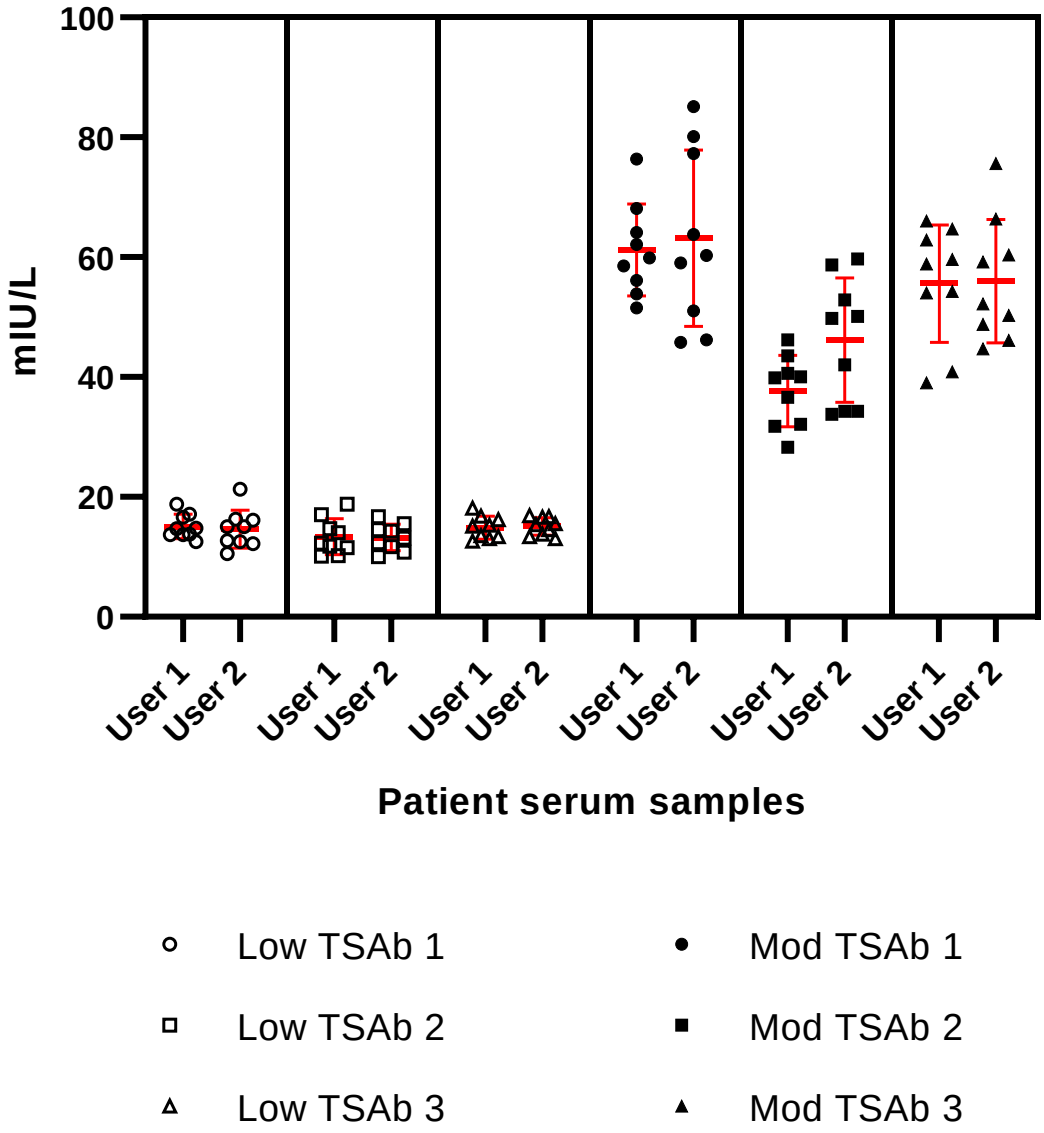

Figure 13: *Comparison between the two users at the DHI laboratory.*

At the DHI laboratory, the variance between the two users was quite low, the mean values measured by each user were for each serum sample within one standard deviation. CV% values are comparable as well.

In summary, there was no significant differences between the measurements of the two users at either the DHI or the JGU laboratory regarding the results of the study.

4.2.6 Precision study:

For the investigation of the assay precision, the patient serum samples were measured by one user over 20 consecutive days. For each measurement, the same cell lot was used and only one plate was tested.

As can be seen from the tables below, the results for the measurements were noticeably higher at the DHI laboratory.

Precision	**Mean (mIU/L)**		**SD**		**CV%**	
	JGU	**DHI**	**JGU**	**DHI**	**JGU**	**DHI**
Low TSAb 1	12.7	16.1	1.9	3.5	15.1%	21.6%
Low TSAb 2	11.8	13.8	2.0	2.6	16.5%	18.6%
Low TSAb 3	13.2	14.7	2.0	2.3	14.8%	15.9%
Mod TSAb 1	39.8	63.7	8.7	15.3	21.8%	24.0%
Mod TSAb 2	22.7	35.7	4.1	6.6	17.9%	18.6%
Mod TSAb 3	39.5	66.4	14.3	13.3	36.3%	20.1%

Table 13: *The mean values for the low and moderate TSAb positive patient serum samples during the precision measurements at the two laboratories.*

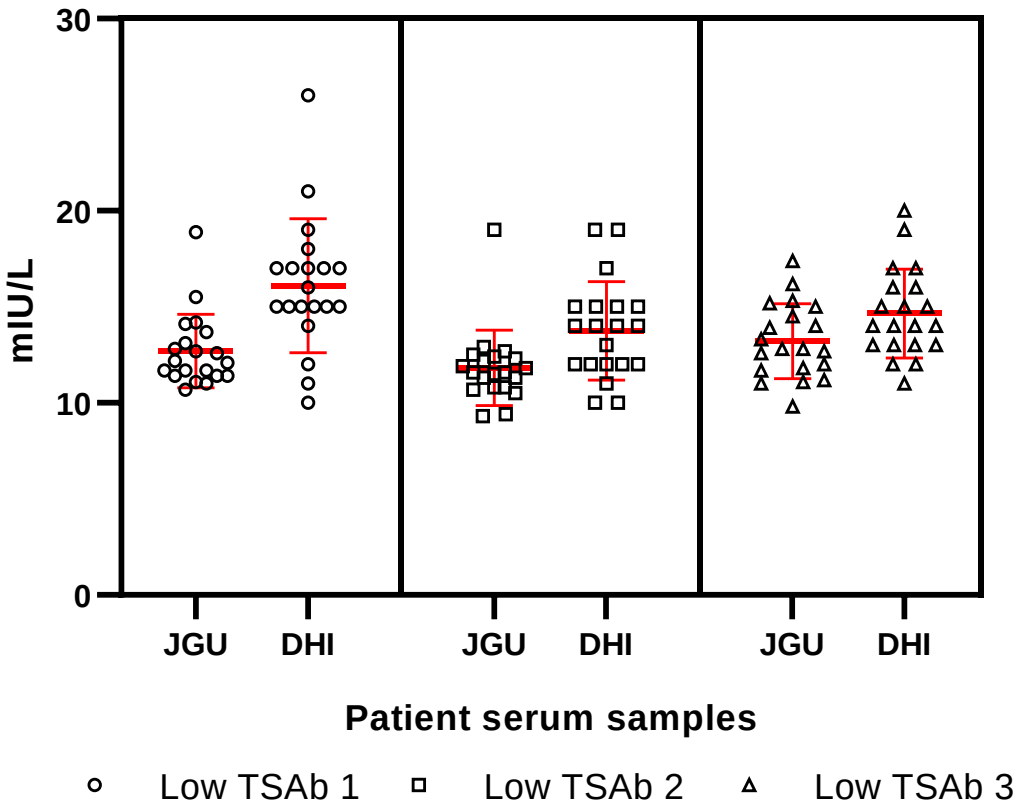

Figure 14: *Inter-assay precision for the low TSAb positive patient serum samples, comparison between the JGU and DHI laboratories.*

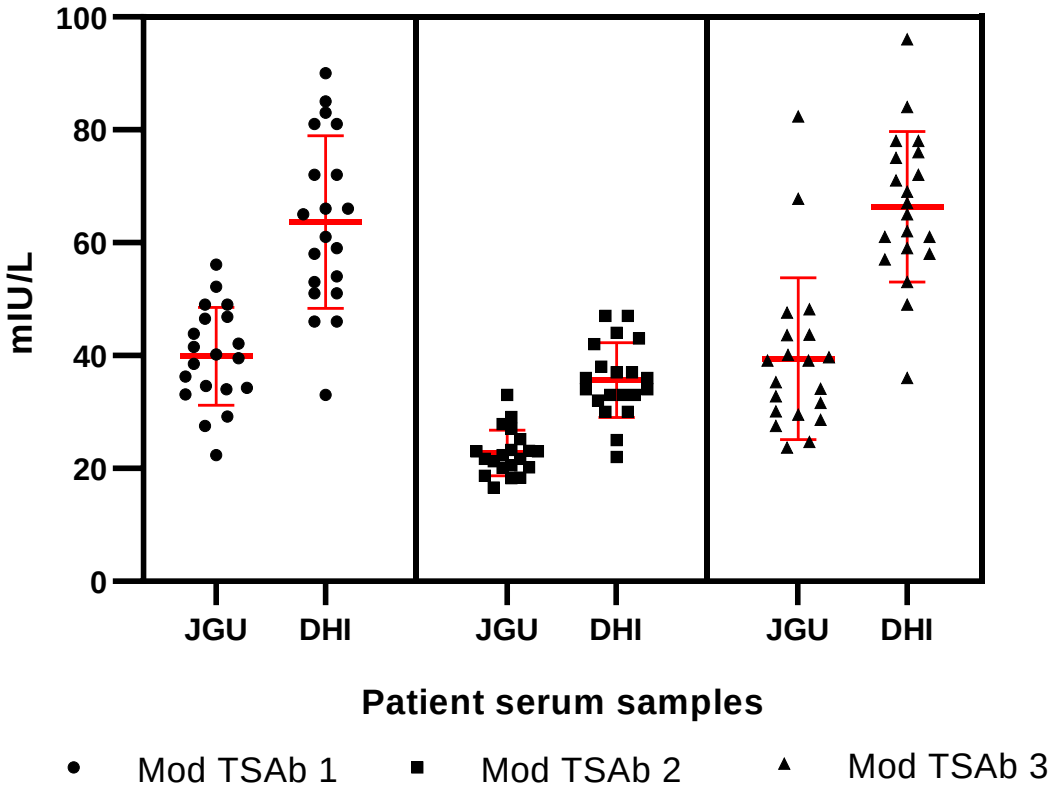

Figure 15: *Inter-assay precision for the moderate TSAb positive patient serum samples, comparison between the JGU and DHI laboratories.*

Overall, the inter-assay precision had a CV% value of 20.4% at the JGU laboratory and of 19.8% at the DHI laboratory and was thus quite well comparable.

For the intra-assay precision, the mean CV% values for the patient serum samples can be seen in the following table. Intra-assay precision was very good with mean CV% values mostly under 5%.

Intra-Assay	Mean CV% values	
	JGU	**DHI**
Low TSAb 1	4.9%	2.8%
Low TSAb 2	4.7%	3.1%
Low TSAb 3	3.8%	4.1%
Mod TSAb 1	3.9%	4.2%
Mod TSAb 2	3.4%	5.3%
Mod TSAb 3	4.1%	4.0%

Table 14: *The intra-assay precision values for the low and moderate TSAb positive patient samples.*

4.2.7 Differences in the measurement of the 2nd IS between laboratories:

The main reason for the higher results measured at the DHI laboratory is a differing behavior of the 2nd IS during measurements at both laboratories. As can be seen from the following tables, while the patient serum samples were measured about 5 to 16% higher at the JGU laboratory, 2nd IS concentrations were measured 19 to 27% higher.

Patient Samples	JGU	DHI	JGU/DHI
Low TSAb 1	553%	514%	108%
Low TSAb 2	641%	562%	114%
Low TSAb 3	622%	557%	112%
Mod TSAb 1	363%	336%	108%
Mod TSAb 2	299%	285%	105%
Mod TSAb 3	357%	339%	105%
High TSAb 1	233%	209%	111%
High TSAb 2	224%	194%	115%
High TSAb 3	236%	203%	116%
Normal Serum 1	36%	46%	78%
Normal Serum 2	38%	48%	80%
Normal Serum 3	37%	45%	83%
Normal Serum 4	20%	25%	78%

Table 15: *Comparison between mean SRR% values of the patient samples during the precision study.*

IS concentrations	JGU	DHI	JGU/DHI
5 mIU/l	433%	359%	120%
10 mIU/l	386%	304%	127%
20 mIU/l	277%	227%	122%
40 mIU/l	195%	164%	119%
80 mIU/l	131%	110%	119%

Table 16: *Comparison between mean SRR% values of the 2nd IS concentrations during the precision study.*

As can be seen from the following graph, patient serum and 2nd IS show different behavior during measurement in the two laboratories. When comparing the relation between the patient serum samples measured at both laboratories and the 2nd IS concentrations measured at both laboratories, the two graphs diverge.

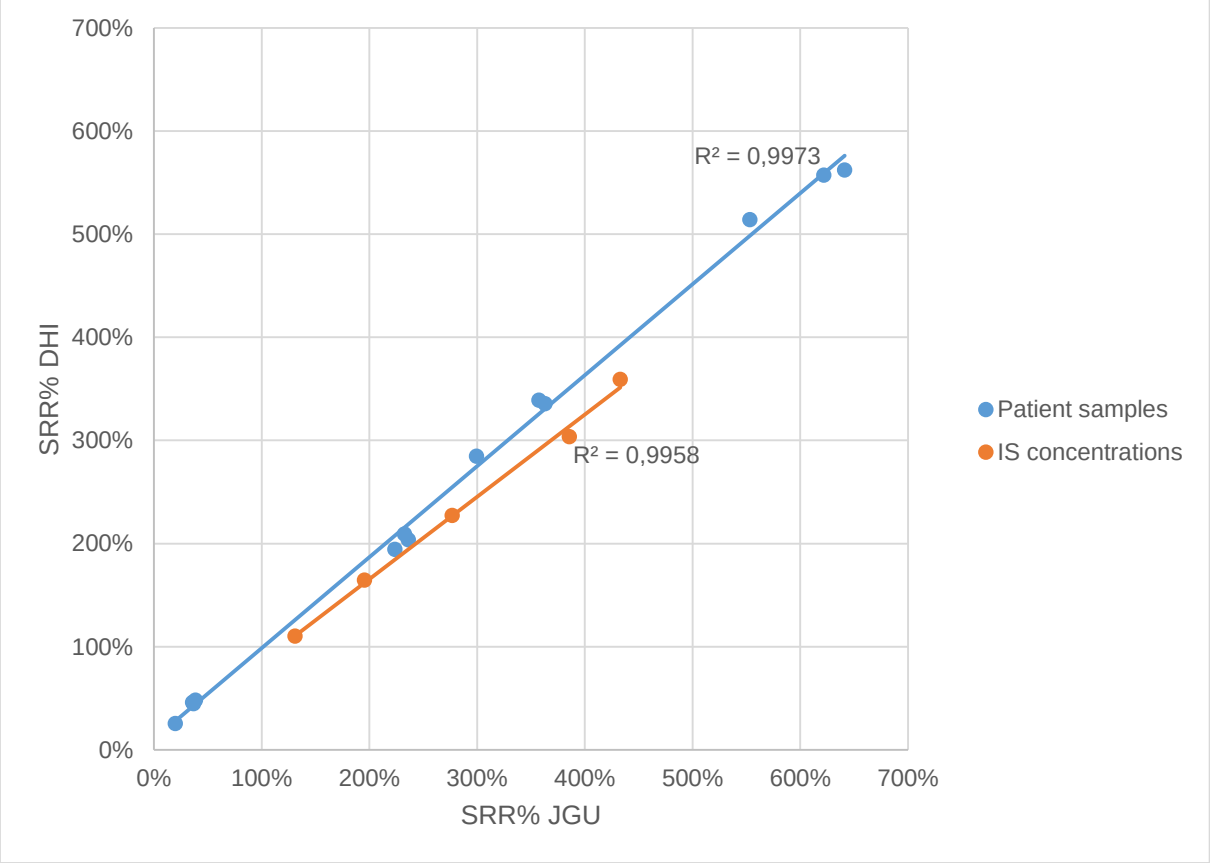

Figure 16: *Expectation would be that IS concentrations and patient samples would lie on the same line, but they differ in their measurements between the laboratories*.

The higher the 2nd IS concentrations and the higher the TSAb levels, the more the two graphs diverge. The diverging graphs prove that there is a fundamental difference in the measurements between the patient serum samples and the IS concentrations of the linear range.

4.2.8 Repetition of the patient measurements at the JGU laboratory:

After analysis of the data gathered from the three parts of the study, there was enough material left at the JGU laboratory to repeat the second part of the study. This time, two users, the first being the same as during the original iteration of the study, measured the patient samples on three consecutive days with three cell lots each day and converted the measurements via the standard curve.

In the following table, the overall mean value of the patient samples converted to mIU/L measured by both users with three different cell lots will be shown. CV% values were between 18-48% for the low TSAb positive samples and 37-62% for the moderate TSAb positive samples. As before, the high TSAb positive samples could not be converted as they were consistently above the linear range. During the repetition, some moderate TSAb positive samples had to be discarded (only of the first moderate TSAb positive patient serum sample, labelled Mod TSAb 1) as they were outside the linear range.

Patient samples	**Mean (mIU/L)**	**SD**	**CV%**
Low TSAb 1	10.8	3.3	30.7%
Low TSAb 2	5.8	1.1	18.7%
Low TSAb 3	18.2	8.8	48.4%
Mod TSAb 1	56.7	20.8	36.6%
Mod TSAb 2	18.6	7.7	41.2%
Mod TSAb 3	26.3	16.2	61.7%

Table 17: *Results for the repetition of the experiments at the JGU laboratory.*

When comparing the results of the patient measurements from the original part of the study to the repetition, some differences were noted. Firstly, for the moderate TSAb positive samples, the mean values were even lower compared to the first measurement. Secondly, the results of the repetition measurements were more

spread, especially for the third low TSAb positive sample. To allow further examination, the cell lots and users will be examined separately.

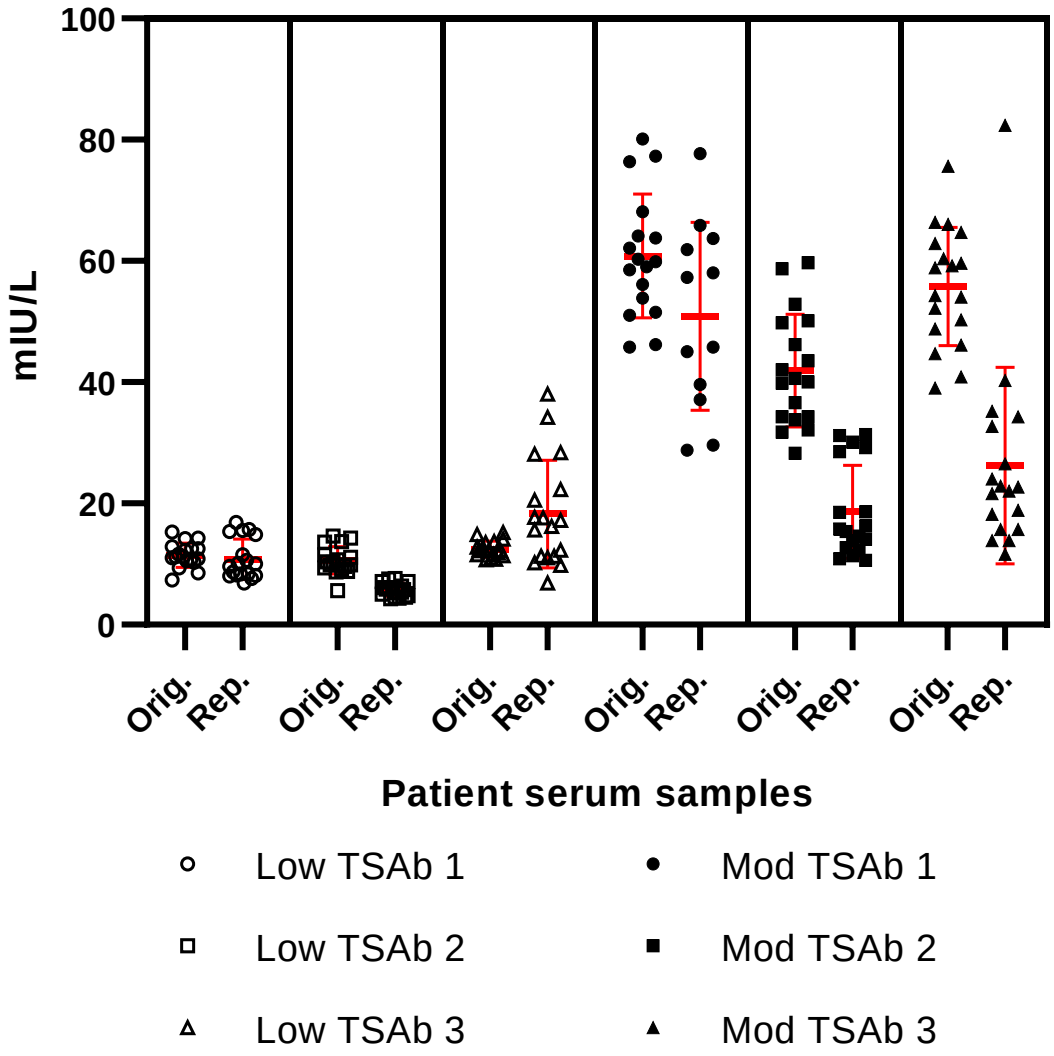

Figure 17: *Comparison between all patient samples measured during the original part of the study and the repetition measurements.*

4.2.9 Comparison between users:

When comparing the results of the two users that performed the repetition measurements, it can be seen that user 1 gained overall higher results than user 2. While the mean measurements for the low TSAb positive samples were 11.8, 6.4, and 23.0 mIU/L for the first user, they were 9.8, 5.1, and 13.5 mIU/L for the second user.

During the measurements of the moderate TSAb positive samples, user 1 provided mean values of 57.6, 20.7, and 31.7 mIU/L while user 2 provided 46.0, 16.6, and 20.8 mIU/L for the respective patient samples. For the first moderate TSAb positive sample, four of nine measurements for user 1 were above the linear range and thus didn't count for the mean value, while only two of the nine measurements had to be dropped for user 2. Supposing these measurements had been converted, the first moderate TSAb positive patient sample would probably have had an even higher mean value for user 1.

The CV% values were quite high for both users, ranging from 12 - 50% for the low TSAb positive samples and from 33 - 62% for the moderate TSAb positive samples.

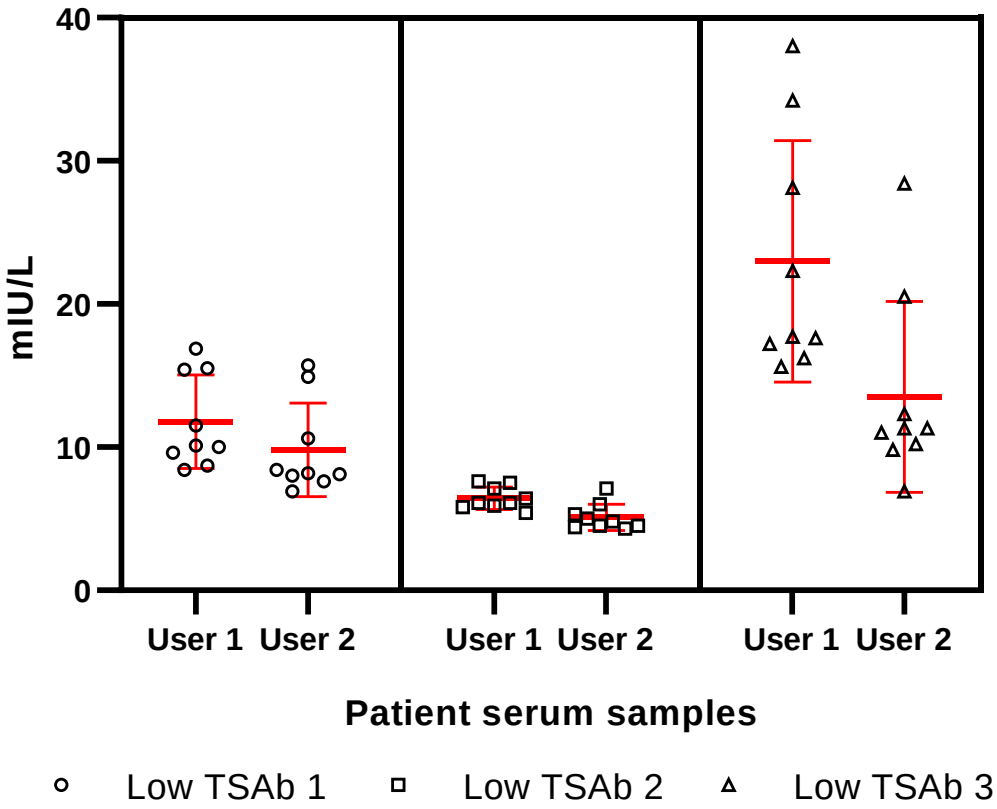

Figure 18: *Comparison between the two users for the low TSAb positive samples during the repetition.*

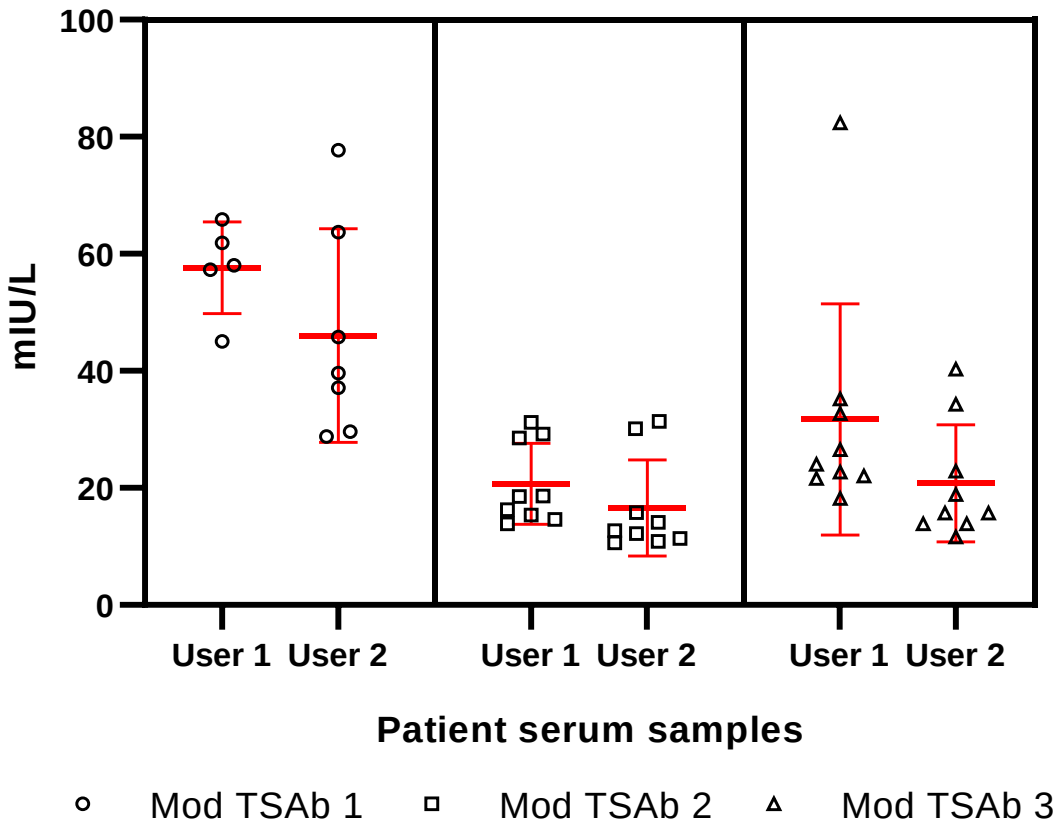

Figure 19: *Comparison between the two users for the moderate TSAb positive patient samples during the repetition.*

4.2.10 Comparison between lots:

Comparison between the three cell lots used during the repetition of the measurements showed some interesting data. While the mean values of the first and the second cell lot were quite similar, the third cell lot showed much higher mean values for the patient samples. This is an increase of around 44 - 77% for the respective patient samples. For the first of the moderate TSAb positive samples, most measurements had to be removed, as they were above the linear range, and only two samples could be retained.

Repetition JGU	Mean (mIU/L)		
	Cell lot 1	Cell lot 2	Cell lot 3
Low TSAb 1	8.6	8.9	14.8
Low TSAb 2	5.2	5.4	6.7
Low TSAb 3	14.4	13.6	26.7
Mod TSAb 1	48.6	48.3	82.1
Mod TSAb 2	13.7	14.5	27.7
Mod TSAb 3	18.1	19.4	41.3

Table 18: *Comparison between the three cell lots during the repetition of the measurement at the JGU laboratory.*

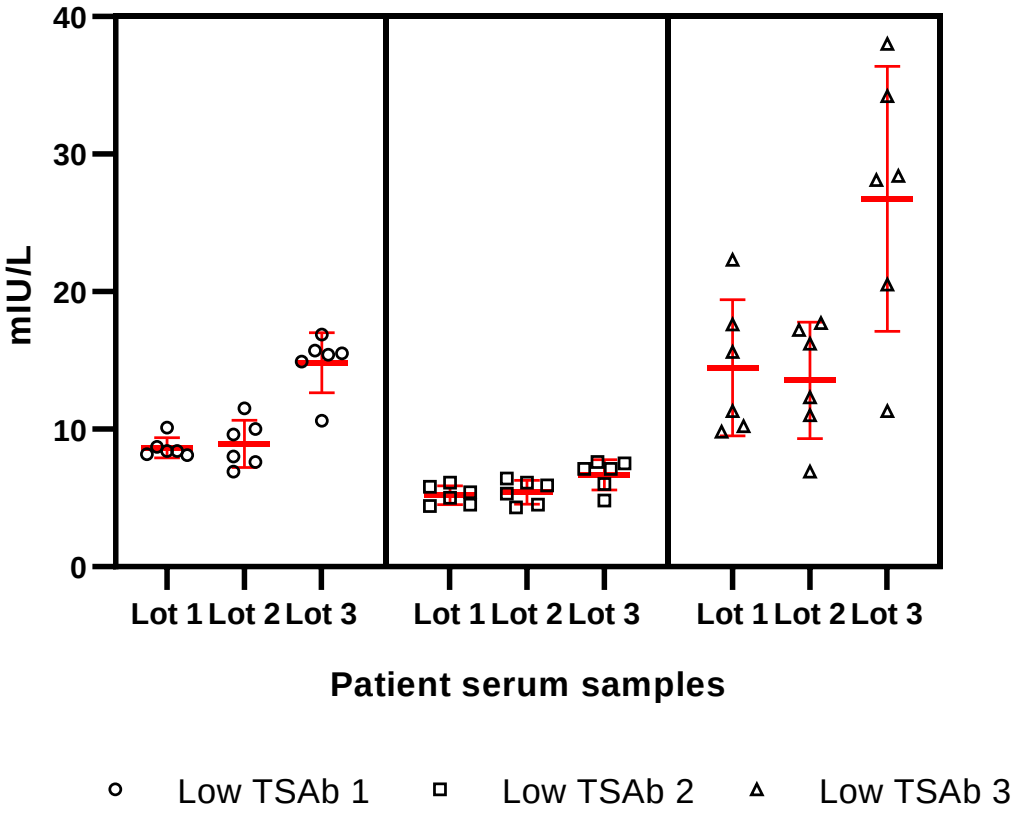

Figure 20: *Comparison between the three lots for the low TSAb positive patient samples during the repetition of the measurements.*

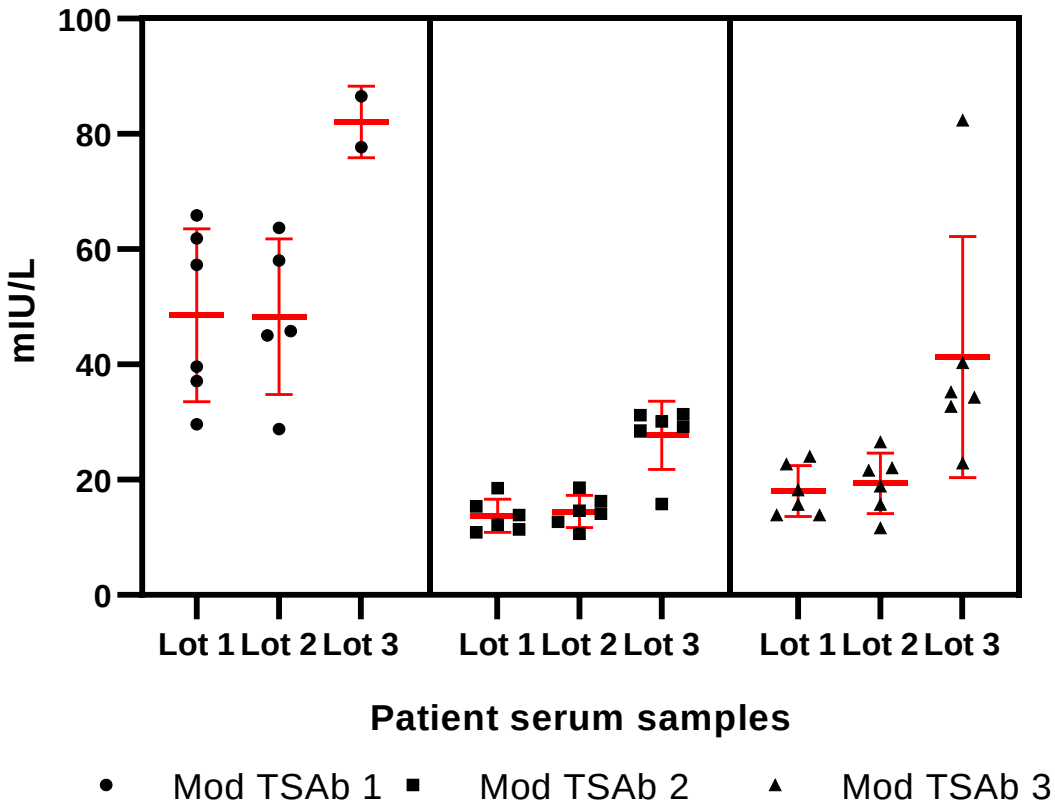

Figure 21: *Comparison between the three lots for the moderate TSAb positive samples during the repetition of the measurements.*

This deviation of the third cell lot could not be observed during the measurements of the original study, where the cell lots proved to be no real confusing factor.

4.2.11 Differences in the measurement of the 2nd IS between the cell lots:

As during comparison between the laboratories, there is a big underlying difference between the first two cell lots and the third cell lot during the repetition of the measurements.

Patient Samples	Mean of Lot 1 + 2	Lot 3	Mean of Lot 1 + 2 / Lot 3
Low TSAb 1	225%	233%	97%
Low TSAb 2	169%	160%	105%
Low TSAb 3	270%	277%	97%
Mod TSAb 1	413%	413%	100%
Mod TSAb 2	275%	288%	95%
Mod TSAb 3	304%	315%	96%
High TSAb 1	471%	428%	110%
High TSAb 2	626%	550%	114%
High TSAb 3	586%	546%	107%

Table 19: *The mean SRR% values for the patient serum samples for the first two lots are quite similar to the SRR% values measured with lot 3.*

IS concentrations (mIU/L)	Mean of Lot 1 + 2	Lot 3	Mean of Lot 1 + 2 / Lot 3
5	449%	378%	119%
10	400%	336%	119%
20	313%	269%	116%
40	232%	201%	115%
80	161%	130%	124%

Table 20: *The SRR% values for the IS concentrations of the linear range are about 15 to 24% higher for the first two cell lots compared to the third cell lot.*

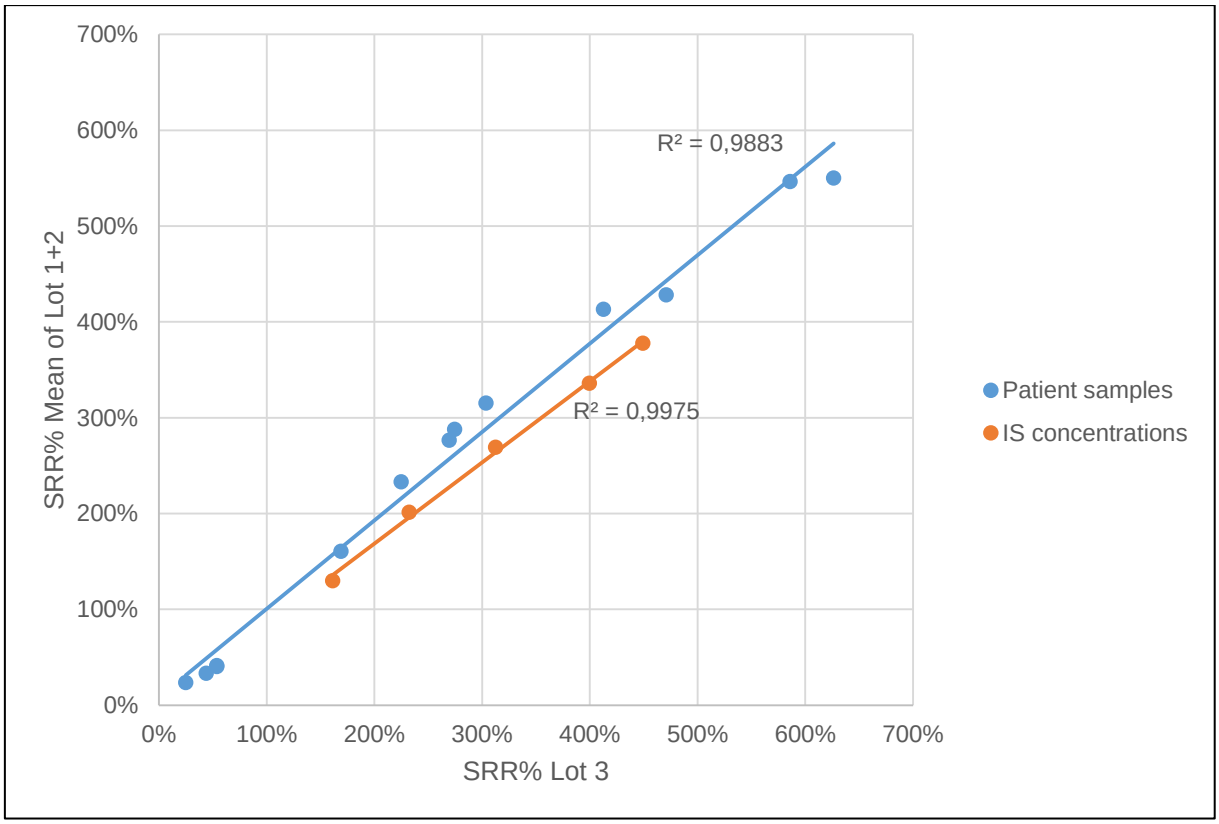

Figure 22: *Comparison between SRR% values of patient serum samples and 2nd IS concentrations between the different lots used during the repetition of the measurements.*

The difference between the higher mIU/L values measured with the third cell lot compared to the other two cell lots stems mostly from the fact that the IS concentrations of the standard curve produced much higher results in comparison to the patient serum samples. Analogous to the earlier comparison between the two laboratories, again the IS that shows a differing behavior in comparison to the patient serum samples.

4.2.12 Comparison between the same user:

During the measurement repetition, two users at the JGU laboratory measured the samples. While the first user was the same as during the original study, the second user differed from the original study. Due to this, measurements of user 1 during the measurement repetition were opposed to measurements during the original study.

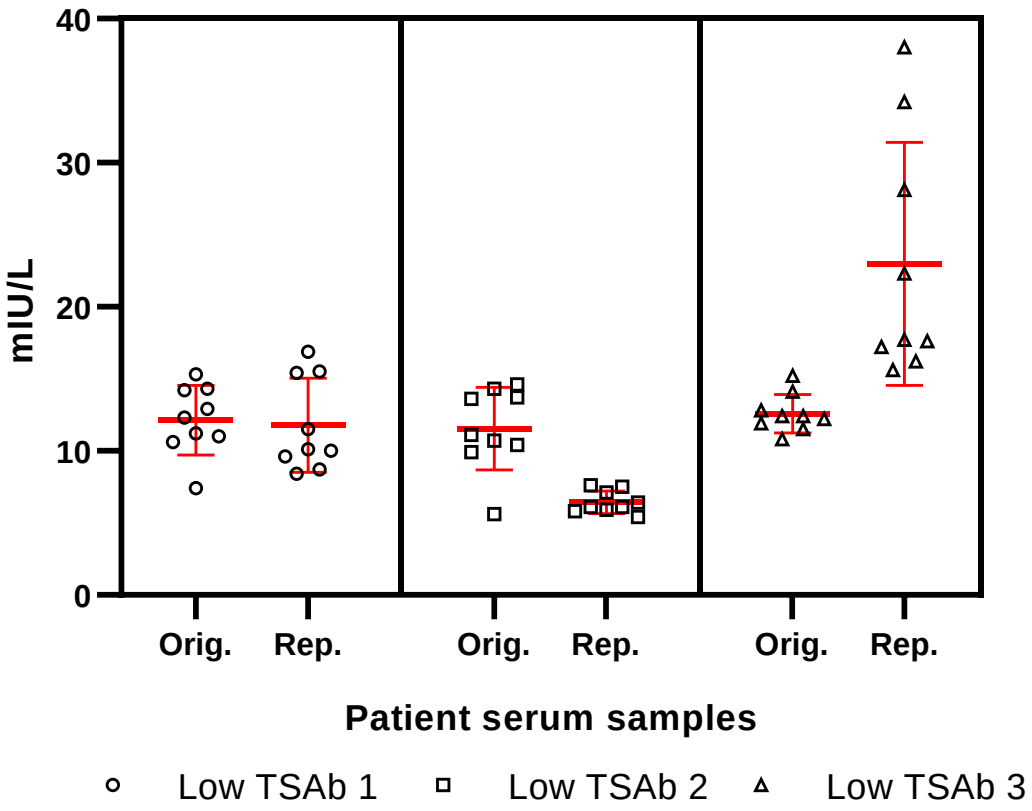

Figure 23: *Comparison between original and repetition measurements for the low TSAb positive patient samples for user 1.*

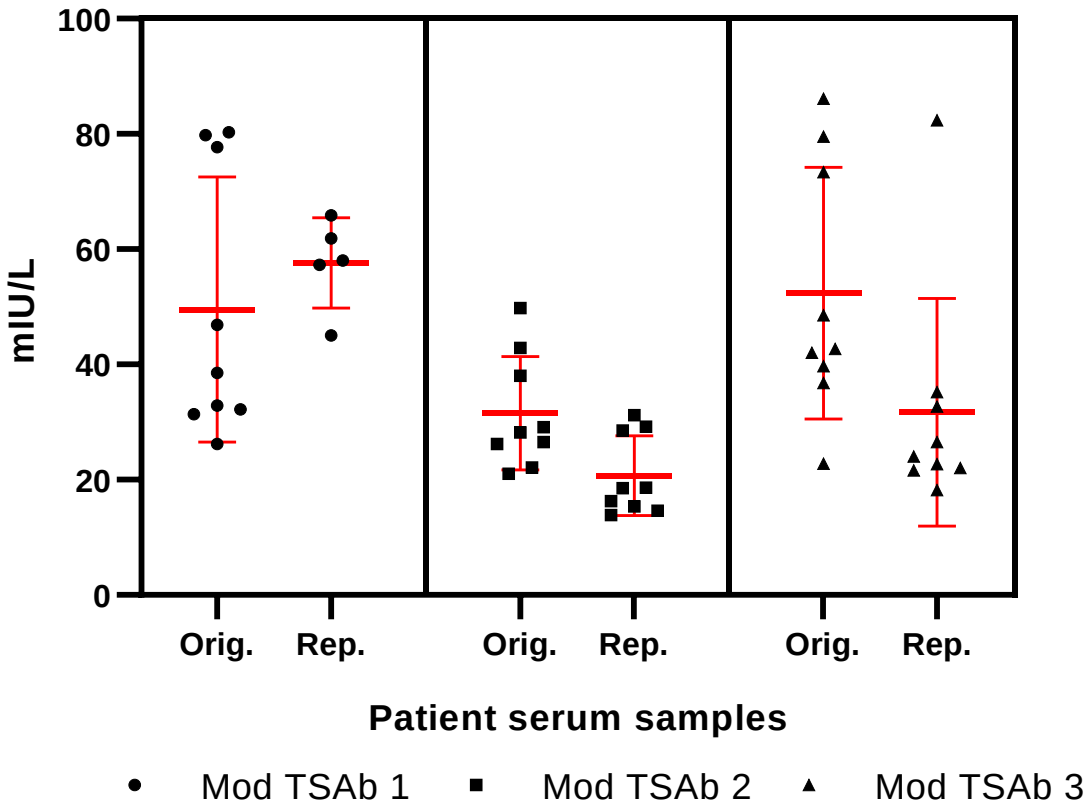

Figure 24: *Comparison between original and repetition measurements for the moderate TSAb positive patient samples for user 1.*

As can be seen from the figures, while some results were quite similar (e.g. for low positive sample 1), most samples differed in varying amounts. The differences between the measurements were not systematic, as some samples were higher in the repetition and some samples provided lower results than during the original study. Most of the higher results were measured with cell lot 3, which provided higher results during the repetition of the measurements.

5. Discussion

Up to date there is no standardized TSAb bioassay available. The rationale of this doctoral thesis was to standardize a cell-based bioassay for the measurement of TSAb by using international reference material and subsequently to compare the converted mIU/L results with other laboratories that measured the same set of samples. The multicenter standardization of the test results among different laboratories that run the TSAb bioassay will enable a more accurate comparison of the TSAb results.

This TSAb bioassay was used to show that serum TSAb levels in pediatric patients with GD and/ or GO correlate strongly with clinical severity and activity. (Diana, Brown et al. 2014) Another study including adult patients with GD and GO revealed that both symptoms (retrobulbar pain) and signs (redness and swelling of the eyelids as well as swelling of the caruncle) of GO strongly correlate with TSAb (Ponto, Kanitz et al. 2011) and (Kampmann, Diana et al. 2015). TSAb levels also show the onset of dysthyroid optic neuropathy in patients with GD and GO (Ponto, Diana et al. 2015).

In its original state, the bioassay uses a reference control whereby the measured samples are compared to, and thus reports the results as a specimen-to-reference ratio (SRR%). Further distinction of the positive patient serum samples can only be achieved by dividing the samples through arbitrary limits. Standardization of this bioassay means, that the results can be reported not as a percentage of the used reference, but instead in physical units. Standardization of the bioassay was achieved by using the 2nd IS to first create a calibration curve and then to convert the SRR% values of the bioassay into mIU/L values (Diana, Kanitz et al. 2015). In this doctoral thesis, the standardization process and conversion of defined patient serum samples was compared between two different laboratories.

5.1 The International standard as reference:

5.1.1 Behavior and Availability:

In order to achieve standardization, instead of the previously used reference, a reference with a precisely defined quantity had to be used. For this, the 2nd International Standard (IS) for Thyroid Stimulating Antibodies was chosen. The 2nd IS is a mAb, called M22 (Sanders, Jeffreys et al. 2004) which was isolated from the lymphocytes of a patient with GD (TSAb-Standard 2013). This antibody binds to the TSHR and has a stimulating effect on the thyrocytes or other cells transfected with the TSHR as it triggers intracellular cyclic AMP production and thus acts similar to TSH. It furthermore inhibits binding of TSH to the TSHR (Sanders, Evans et al. 2003).

After preparing dilution series of the 2nd IS, the dose-response curve was measured in the bioassay. The dose-response curve showed a linear range between 5 and 80mIU/L of the IS. A linear curve regression of the measurements could then provide a linear equation. This linear equation was used in the experiments to convert the patient samples to the standard units (mIU/L).

In order for the 2nd IS to be a suitable reference, the monoclonal antibody M22 has to behave the same way in the bioassay as the patient samples. This proved to be a vulnerable point of the standardization process and quite possibly one of the sources of error.

Firstly, the 2nd IS already showed a quite high variance during the establishment. In the studies constituting this mAb as the next IS for TSAb, the usage in bioassays had a coefficient of variation of 26.8% (TSAb-Standard 2013). This shows that patient samples which are measured via the bioassay and then converted to mIU/L through a standard curve generated by measurements of the 2nd IS must also show variances around 25% in inter-laboratory comparisons. This variability might be mirrored by variances during the measurement of patient samples, meaning that during the run of one plate, patient samples as well as the 2nd IS provide higher or lower results and thus cancel each other out, but this cannot always be assumed to happen.

Secondly, the 2nd IS differs in its composition from the patient serum that is measured and thus might provide different results during measurement. As previously mentioned, the 2nd IS is a mAb, which is of pure stimulatory nature. Serum from patients with autoimmune thyroid disorders might contain not only one monoclonal antibody but a mixture of different antibodies all targeting the TSHR. (Evans, Sanders et al. 2010) These antibodies might also have different effects on the thyroid. It is proven, that some autoantibodies occurring in GD patients are of stimulatory nature, while other antibodies occurring might be of blocking nature as they inhibit the binding of TSH to the receptor and might be weak agonists (Endo, Kasagi et al. 1978). Other antibodies are neutral in their interaction with the receptor, meaning they bind to the receptor without stimulating intracellular cAMP production or inhibiting binding of TSH to the receptor. Instead they seem to be able to induce apoptosis in the thyroid cells (Morshed, Ando et al. 2010). There are even case files of patients exhibiting both stimulatory and blocking antibodies against the TSHR, which switch during antithyroid treatment between hyper- and hypothyroidism (McLachlan and Rapoport 2013). In the light of these findings, M22 might not be the best reference for the magnitude of possible antibody variations occurring in patients with GD. This concern was already raised when the antibody was considered as a replacement for the 1st International Standard for Thyroid Stimulating Antibodies, but as there was no suitable alternative available, the committee approved the antibody as the replacement (Burns 2010).

In favor for the 2nd IS as a reference control is mostly its availability. The antibody is produced by an immortal, cloned cell-line of B-lymphocytes (Sanders, Evans et al. 2003). It was also heavily researched in its properties and molecular structure (Sanders, Jeffreys et al. 2004). Another study describing the binding of the antibody to the TSHR through X-ray diffraction analysis showed that the M22 antibody has very similar binding features to TSH in its interaction with the TSHR (Sanders, Chirgadze et al. 2007). Nowadays, the antibody is readily available from the NIBSC of the WHO in ampoules of 113 mIU and can be prepared and stored in every suitable laboratory.

In conclusion, although the 2nd IS might not be the best conceivable reference to imitate the behavior of TSAb positive patient sera, it still serves the purpose and has many advantages due to its availability and standardized distribution by the National Institute for Biological Standards and Control of the World Health Organization. It is thus

suitable to fulfill the role for the standardization process and might be distributed together with the bioassay kit if the standardized bioassay should be commercialized.

5.1.2 The linear range:

The dose-response curve of the 2nd IS shows a sigmoidal course. As previously shown, the linear range of the SRR% values plotted against the common logarithm of the 2nd IS concentrations extends from 5 to 80mIU/L. Linear regression for these concentrations always gave a very good curve fit with coefficients of determination above 95%.

It was shown, that the cut-off point of the original bioassay, set at SRR of 140%, corresponds to a concentration of 9.54 mIU/L IS with a standard deviation of 1.68 mIU/L. (Diana, Kanitz et al. 2015) As such, the previous cut-off point can be approximated by the second concentration of the standard curve, 10 mIU/L.

For the lower limit of 5 mIU/L, the previous patient measurements have shown, that low TSAb positive serum samples, in the original bioassay measured from 140-280%, are well within the linear range, most starting, as could be expected, at about 10 mIU/L. For the moderate TSAb positive serum samples ranging from 280-420%, the linear range was also a good fit, with most samples ranging between 20 and 60mIU/L. Unfortunately, all experiments showed that high TSAb positive patient samples, which were measured above 420% in the original bioassay, could not be converted, as their results were well above the SRR% values for the 80 mIU/L IS concentration and thus outside the upper limit of the linear range.

High TSAb positive serum samples above the linear range (>80 mIU/L) can be converted by dilution. As the linear range encompasses more than a ten-fold increase from the minimal concentration (5 mIU/L) to the maximal concentration (80 mIU/L), a possible dilution could be 1:10 of the original serum. This should ensure that each patient serum sample reaches the conversion range after a couple of dilution steps. For example a patient serum sample containing e.g. 500 mIU/L TSAb could be diluted 1:10 and would then be measured at 50 mIU/L within the linear range.

The ideal standard for conversion would have a linear range that extends as far as possible to allow for the majority of the patient samples to be converted. The 2nd IS

proves to be quite well suited in this regard, as only very high TSAb positive patient serum samples cannot be converted and would have to be diluted before measurement to allow for conversion. It is known, that these very high TSAb positive samples correlate with a much higher probability for the development of GO (Ponto, Kanitz et al. 2011), as well as relapse after pharmaceutical therapy (Giuliani, Cerrone et al. 2012).

During clinical application this could mean that patients with results in the bioassay above the linear range benefit more from either higher doses of thyroid blocking medication or even from therapies aimed to reduce the amount of thyroidal tissue. Therapies to reduce the amount of thyroidal tissue are primarily surgical (total or subtotal thyroidectomy) or radiotherapeutic (radioiodine therapy). These therapies constitute a permanent dependency on hormonal replacement therapy with thyroxine for the patients, as the thyroidal tissue is removed from the body. Additionally, surgery and radioiodine therapy pose additional risks. Radioiodine therapy is not possible during pregnancy, and might increases the risk for hematopoietic illnesses and even leukemia (Schroeder, Kuendgen et al. 2012). Thyroidectomy carries the risk of nerve palsy, especially recurrent laryngeal nerve palsy. Unilateral or even bilateral recurrent laryngeal nerve damage can lead to vocal cord palsy, respiratory distress and even the need for tracheotomy (Padur, Kumar et al. 2016) and (Pardal-Refoyo and Ochoa-Sangrador 2016). Another potent risk is transient and even permanent hypoparathyroidism (Chadwick 2017), which can be quite debilitating for the patient.

While the high TSAb positive patients are at a higher risk for relapse after medical therapy and extrathyroidal complications through GD, they might benefit more from surgery or radioiodine therapy despite the greater risks involved. On the other hand the low and moderate TSAb positive patients might benefit more from medical therapy and the possibility of periodically measuring TSAb serum levels in an internationally comparable unit.

5.1.3 Differences in measurements of the 2nd IS in the two laboratories:

During the standardization study, the patient serum samples and the different IS concentrations were prepared at the DHI laboratory, frozen and then send to the other laboratories. This approach should ensure that there is no systematical error during the preparation of the samples by different persons. The process for the measurement of the samples was also standardized between the laboratories.

Nonetheless, evaluation of the results showed that the DHI laboratory reported continuously higher results for the patient serum measurements than the JGU laboratory. A possible reason for this can be found in the differing behavior of the 2nd IS during measurement. SRR% values for the TSAb positive patient sera were measured continuously between 5 and 15% higher at the JGU laboratory. After conversion of the patient serum samples, the results were always lower at the JGU laboratory in comparison to the DHI laboratory. The reason is that the IS concentrations of the standard curve were measured at even higher SRR% values at the JGU laboratory, namely between 19 and 27% higher than at the DHI laboratory. Therefore, this discrepancy is the main cause for the higher results at the DHI laboratory in comparison to the JGU laboratory.

There are a couple of possible reasons for the differences. The first influencing factor is the material and equipment used in the two laboratories. The ELISA reader, incubators and other laboratory equipment at the DHI laboratory differed from the equipment that was used at the JGU laboratory. Any part of this might have led to the 2nd IS being measured differently to the patient samples. Another factor might be that the 2nd IS concentrations could have reacted differently than the patient samples in regard to freezing, transportation and storage, and might have degraded slower or faster depending on the environment. Thirdly, although all involved persons followed protocol, there could have been some fundamental difference between the users at the American laboratory and the German laboratory while carrying out the measurements.

In this case it is very unfortunate that the laboratory at the Mayo clinic could not perform the experiments and deliver their data. With a third data set, it might have been possible to show which source caused the error in the measurement or it might have been possible to rule out at least some possible sources of error.

5.1.4 Differences between cell lots:

During the repetition of the patient serum sample measurements and conversions at the JGU laboratory, another very interesting development was visible. While during the original study the measurements showed no differences between the three different cell lots that were used, this was not the case during the repetition of the measurements. Here, the third cell lot showed significantly higher mIU/L values compared to the other two cell lots, ranging from 44 to 77% higher. This discrepancy is even more baffling, as the material used was exactly the same as during the original experiments. Again, the reason for this discrepancy cannot be deduced clearly. The cell lots were all stored at -80 degree Celsius in adjoining boxes in the same freezer. During the preparation of the plates, the cell lots and the plates used to measure the patient samples were all treated equally. One of the possible explanations could be degradation of either the one or the other two cell lots and thus a different behavior during the measurements, although the expiration date was not reached for the cell lots.

When comparing the behavior of patient serum samples and IS concentrations against each other, the same abnormalities in the measurement of the 2nd IS can be seen, as the IS concentrations were measured lower with cell lot 3 than with the other cell lots, while the patient serum samples showed no differences.

This differing behavior of the 2nd IS is most obvious during the comparison between the two laboratories during the original part of the study and during the comparison between the three cell lots during the repetition of the measurements.

Which exact mechanism leads to the IS standard curve being measured differently, cannot be reconstructed from the data gathered during the measurements. Still, the 2nd IS seems to be the biggest confounder during the standardization process.

5.2 Variability of results reported as SRR% and mIU/L:

The SRR% values gathered for the patient serum samples at both laboratories during the precision measurements showed a CV of about 10 to 15%. The variability of the SRR% values was not dependent on the level of TSAb and there was no discernible difference in the variability between the low, moderate or high TSAb positive patient samples. For the calibrated samples reported as mIU/L this was different. After standardization, the variability of the results depended highly upon the TSAb level. While the low TSAb positive samples had only slightly higher CV% values, in the case of the precision measurements between 16 and 22%, the moderate TSAb positive patient samples showed a much higher variance, 29-36% for the precision measurements.

The higher variability for the moderate TSAb positive samples can be derived from deviating measurement of the 2nd IS. As could be seen previously, the higher the concentrations measured, the more the 2nd IS deviates from the patient samples. This is not only the case in the samples shown above, the 2nd IS deviates also in the cases where it is not as apparent from the behavior of the patient samples.

Optimally, standardization of the bioassays results should have little to no impact on the variability of the results. While this is true for the low TSAb positive samples, this cannot be said for the moderate TSAb positive samples, where the higher variability should not be ignored.

This shows that indeed, the 2nd IS might not be the definitive answer as a reference control for the standardization of the bioassay and other references with similar properties but without the flaws of the 2nd IS for TSAb should be considered.

5.3 Conclusion:

The standardized bioassay showed a quite high variability, especially when comparing the results acquired by the different international laboratories. The differences in measurement were also amplified by higher TSAb levels, as the moderate TSAb positive patient serum samples showed higher variability in comparison to the low TSAb positive patient serum samples.

Focusing on the reasons for the differences, no consistent error produced by the different users or the different cell lots used was observed. Bias seemed to be produced by the two different laboratories carrying out the measurements, as well as the substance used as reference material, namely the 2nd IS for TSAbs. This mAb showed a differing behavior from the patient serum samples utilized during the measurements. This deviation is revealed especially in the inter-assay comparison and during the repetition of the measurements when comparing the different cell lots.

The reasons for the deviation of the 2nd IS from the patient serum samples could be multitudes and could not be identified during this work. Possible explanations could be the monoclonal nature of the antibody in contrast to the diverse mixture of antibodies existing in the patient serum samples. Other possible factors could be additives present in the dry matter of the 2nd IS ampoules distributed by the NIBSC or the phosphate buffered saline which was used to dissolve the 2nd IS.

For future prospects, a different reference for standardization should be considered. Possible references for standardization as a replacement for the 2nd IS could include a mixture of patient sera, which should be guaranteed to act similar to the patient serum which is measured. Another possible reference for standardization could be the bovine TSH, which was used as the reference in the original assay to report the results as SRR%. Provided these references show the same behavior in the bioassay as the patient serum, the favorable reference would be the reference with the wider linear range and maybe the better availability.

6. Summary:

In this doctoral thesis, the standardization process across several international laboratories for a bioassay detecting thyroid stimulating antibodies (TSAbs) is presented. This process consisted of using the 2nd IS for TSAbs to create a dose-response curve, where the linear range was identified. This linear range comprised 2nd IS concentrations between 5 mIU/L and 80 mIU/L. The linear range was then used to convert several previously selected patient serum samples from SRR% to mIU/L. The results of this process showed a high variability, especially when comparing the results of the two conducting laboratories, which was amplified by higher levels of TSAb respectively 2nd IS.

A reason for the high variability could be identified in a differing behavior of the 2nd IS in comparison to the TSAb positive patient serum samples during measurement.

This doctoral thesis shows the feasibility of standardizing a TSAb bioassay but at the same time demonstrates that the 2nd IS increases the variability of the bioassay's results considerably. Conclusively, the 2nd IS might not be the best choice as a reference for the standardization process. This means that a different reference for standardization should be considered.

Thus, although this paper does not provide a definitive answer for a standardized bioassay for TSAb, it may pave the way for such an endeavor.

7. Zusammenfassung:

In dieser Doktorarbeit wird die Standardisierung eines Bioassays für Schilddrüsen-stimulierende Antikörper (TSAbs) durch zwei internationale Laboratorien beschrieben. Zur Standardisierung wurde der zweite Internationale Standard (IS) für TSAbs verwendet um eine Dosis-Wirkung Kurve zu erzeugen, bei der ein linearer Bereich definiert werden konnte. Der lineare Bereich des zweiten IS reichte von 5 mIU/L bis 80 mIU/L. Mithilfe dieses linearen Bereiches wurden im Anschluss mehrere zuvor ausgewählte Serumproben von Patienten mit Schilddrüsenüberfunktion von SRR% in mIU/L konvertiert. Die Ergebnisse dieses Prozesses zeigten besonders beim Vergleich zwischen den zwei Laboratorien eine hohe Variabilität, welche durch höhere Konzentrationen von TSAb beziehungsweise zweiter IS erhöht wurde.

Als Hauptgrund für die hohe Variabilität konnte ein abweichendes Verhalten des zweiten IS im Vergleich zu den TSAb positiven Patientenproben identifiziert werden.

Diese Doktorarbeit zeigt somit die Machbarkeit der Standardisierung eines Bioassays für TSAb, zeigt aber zugleich, dass der zweite IS, der als Referenzmedium verwendet wurde, die Variabilität der Ergebnisse durch das Standardisierungsverfahren deutlich erhöht. Zusammenfassend scheint der zweite IS nicht die beste Wahl als Referenzmedium für das Standardisierungsverfahren zu sein und andere Referenzmedien sollten in Erwägung gezogen werden.

Obwohl diese Doktorarbeit keine endgültige Antwort für einen standardisierten Bioassay für TSAb liefert, ebnet sie doch den Weg für dieses Unterfangen.

Reference List:

Adams, D. D. (1958). "The presence of an abnormal thyroid-stimulating hormone in the serum of some thyrotoxic patients." J Clin Endocrinol Metab **18**(7): 699-712.

Araki, N., M. Iida, N. Amino, S. Morita, A. Ide, E. Nishihara, M. Ito, J. Saito, T. Nishikawa, K. Katsuragi and A. Miyauchi (2015). "Rapid bioassay for detection of thyroid-stimulating antibodies using cyclic adenosine monophosphate-gated calcium channel and aequorin." Eur Thyroid J **4**(1): 14-19.

Bakke, J. L., L. Heideman, Jr., N. L. Lawrence and C. Wiberg (1957). "Bioassay of thyrotropic hormone by weight response of bovine thyroid slices." Endocrinology **61**(4): 352-367.

Bakke, J. L. and N. Lawrence (1956). "The effect of thyroid stimulating hormone upon the iodide collecting mechanism of thyroid tissue slices." Endocrinology **58**(5): 531-545.

Brown, J. and D. S. Munro (1967). "A new in vitro assay for thyroid-stimulating hormone." J Endocrinol **38**(4): 439-449.

Burns, C. T., R.; Rigsby, P.; (2010). "WHO International Collaborative Study of the proposed 2nd International Standard for Thyroid Stimulating Antibody."

Chadwick, D. R. (2017). "Hypocalcaemia and permanent hypoparathyroidism after total/bilateral thyroidectomy in the BAETS Registry." Gland Surg **6**(Suppl 1): S69-S74.

Diana, T., R. S. Brown, A. Bossowski, M. Segni, M. Niedziela, J. Konig, A. Bossowska, K. Ziora, A. Hale, J. Smith, S. Pitz, M. Kanitz and G. J. Kahaly (2014). "Clinical relevance of thyroid-stimulating autoantibodies in pediatric graves' disease-a multicenter study." J Clin Endocrinol Metab **99**(5): 1648-1655.

Diana, T., M. Kanitz, M. Lehmann, Y. Li, P. D. Olivo and G. J. Kahaly (2015). "Standardization of a bioassay for thyrotropin receptor stimulating autoantibodies." Thyroid **25**(2): 169-175.

Dorrington, K. J. and D. S. Munro (1964). "A Standard for the Long-Acting Thyroid Stimulator (Lats) in the Serum of Patients with Thyrotoxicosis." J Endocrinol **31**: 21-28.

Endo, K., K. Kasagi, J. Konishi, K. Ikekubo, T. Okuno, Y. Takeda, T. Mori and K. Torizuka (1978). "Detection and properties of TSH-binding inhibitor immunoglobulins in patients with Graves' disease and Hashimoto's thyroiditis." J Clin Endocrinol Metab **46**(5): 734-739.

Evans, M., J. Sanders, T. Tagami, P. Sanders, S. Young, E. Roberts, J. Wilmot, X. Hu, K. Kabelis, J. Clark, S. Holl, T. Richards, A. Collyer, J. Furmaniak and B. R. Smith (2010). "Monoclonal autoantibodies to the TSH receptor, one with stimulating activity and one with blocking activity, obtained from the same blood sample." Clin Endocrinol (Oxf) **73**(3): 404-412.

Giuliani, C., D. Cerrone, N. Harii, M. Thornton, L. D. Kohn, N. M. Dagia, I. Bucci, M. Carpentieri, B. Di Nenno, A. Di Blasio, P. Vitti, F. Monaco and G. Napolitano (2012). "A TSHR-LH/CGR chimera that measures functional thyroid-stimulating autoantibodies (TSAb) can predict remission or recurrence in Graves' patients undergoing antithyroid drug (ATD) treatment." J Clin Endocrinol Metab **97**(7): E1080-1087.

Grasso, Y. Z., M. R. Kim, C. Faiman, L. D. Kohn, K. Tahara and M. K. Gupta (1999). "Epitope heterogeneity of thyrotropin receptor-blocking antibodies in Graves' patients as detected with wild-type versus chimeric thyrotropin receptors." Thyroid **9**(6): 531-537.

Hata, M. (2010). "Comparison of a Novel Cell-based Reporter Assay and a Competitive Binding ELISA for the Detection of Thyrotropin-Receptor (TSHR) Autoantibodies (TRAb) in Graves' Disease Patients."

Kahaly, G. J. and T. Diana (2017). "TSH Receptor Antibody Functionality and Nomenclature." Front Endocrinol (Lausanne) **8**: 28.

Kahaly, G. J., T. Diana, J. Glang, M. Kanitz, S. Pitz and J. Konig (2016). "Thyroid Stimulating Antibodies Are Highly Prevalent in Hashimoto's Thyroiditis and Associated Orbitopathy." J Clin Endocrinol Metab **101**(5): 1998-2004.

Kampmann, E., T. Diana, M. Kanitz, D. Hoppe and G. J. Kahaly (2015). "Thyroid Stimulating but Not Blocking Autoantibodies Are Highly Prevalent in Severe and Active Thyroid-Associated Orbitopathy: A Prospective Study." Int J Endocrinol **2015**: 678194.

LATS-Standard (2013). "Non WHO Reference Material Long-Acting Thyroid Stimulator (LATS) NIBSC code: 65/122 Instructions for use."

Leschik, J. J., T. Diana, P. D. Olivo, J. Konig, U. Krahn, Y. Li, M. Kanitz and G. J. Kahaly (2013). "Analytical performance and clinical utility of a bioassay for thyroid-stimulating immunoglobulins." Am J Clin Pathol **139**(2): 192-200.

Lytton, S. D., Y. Li, P. D. Olivo, L. D. Kohn and G. J. Kahaly (2010). "Novel chimeric thyroid-stimulating hormone-receptor bioassay for thyroid-stimulating immunoglobulins." Clin Exp Immunol **162**(3): 438-446.

McKenzie, J. (1958). "The bioassay of thyrotropin in serum." Endocrinology **63**(3): 372-382.

McKenzie, J. (1958). "Delayed thyroid response to serum from thyrotoxic patients." Endocrinology **62**(6): 865-868.

McLachlan, S. M. and B. Rapoport (2013). "Thyrotropin-blocking autoantibodies and thyroid-stimulating autoantibodies: potential mechanisms involved in the pendulum swinging from hypothyroidism to hyperthyroidism or vice versa." Thyroid **23**(1): 14-24.

Meek, J. C., A. E. Jones, U. J. Lewis and W. P. Vanderlaan (1964). "Characterization of the Long-Acting Thyroid Stimulator of Graves' Disease." Proc Natl Acad Sci U S A **52**: 342-349.

Menconi, F., C. Marcocci and M. Marino (2014). "Diagnosis and classification of Graves' disease." Autoimmun Rev **13**(4-5): 398-402.

Michelangeli, V. P., D. S. Munro, C. W. Poon, A. G. Frauman and P. G. Colman (1994). "Measurement of thyroid stimulating immunoglobulins in a new cell line transfected with a functional human TSH receptor (JPO9 cells), compared with an assay using FRTL-5 cells." Clin Endocrinol (Oxf) **40**(5): 645-652.

Morshed, S. A., T. Ando, R. Latif and T. F. Davies (2010). "Neutral antibodies to the TSH receptor are present in Graves' disease and regulate selective signaling cascades." Endocrinology **151**(11): 5537-5549.

Onaya, T., M. Kotani, T. Yamada and Y. Ochi (1973). "New in vitro tests to detect the thyroid stimulator in sera from hyperthyroid patients by measuring colloid droplet formation and cyclic AMP in human thyroid slices." J Clin Endocrinol Metab **36**(5): 859-866.

Padur, A. A., N. Kumar, A. Guru, S. N. Badagabettu, S. R. Shanthakumar, M. B. Virupakshamurthy and J. Patil (2016). "Safety and Effectiveness of Total Thyroidectomy and Its Comparison with Subtotal Thyroidectomy and Other Thyroid Surgeries: A Systematic Review." J Thyroid Res **2016**: 7594615.

Pardal-Refoyo, J. L. and C. Ochoa-Sangrador (2016). "Bilateral recurrent laryngeal nerve injury in total thyroidectomy with or without intraoperative neuromonitoring. Systematic review and meta-analysis." Acta Otorrinolaringol Esp **67**(2): 66-74.

Perret, J., M. Ludgate, F. Libert, C. Gerard, J. E. Dumont, G. Vassart and M. Parmentier (1990). "Stable expression of the human TSH receptor in CHO cells and

characterization of differentially expressing clones." Biochem Biophys Res Commun **171**(3): 1044-1050.

Pierce, M., R. Sandrock, G. Gillespie and A. W. Meikle (2012). "Measurement of thyroid stimulating immunoglobulins using a novel thyroid stimulating hormone receptor-guanine nucleotide-binding protein, (GNAS) fusion bioassay." Clin Exp Immunol **170**(2): 115-121.

Planells, R., G. Fayet, S. Lissitzky, G. Hennen and J. Closset (1975). "Bioassay of thyrotropin using isolated porcine thyroid cells." FEBS Lett **53**(1): 87-91.

Ponto, K. A., T. Diana, H. Binder, N. Matheis, S. Pitz, N. Pfeiffer and G. J. Kahaly (2015). "Thyroid-stimulating immunoglobulins indicate the onset of dysthyroid optic neuropathy." J Endocrinol Invest **38**(7): 769-777.

Ponto, K. A., M. Kanitz, P. D. Olivo, S. Pitz, N. Pfeiffer and G. J. Kahaly (2011). "Clinical relevance of thyroid-stimulating immunoglobulins in graves' ophthalmopathy." Ophthalmology **118**(11): 2279-2285.

Rapoport, B. (1976). "Dog thyroid cells in monolayer tissue culture: adenosine 3', 5'-cyclic monophosphate response to thyrotropic hormone." Endocrinology **98**(5): 1189-1197.

Rapoport, B. and R. J. Adams (1978). "Bioassay of TSH using dog thyroid cells in monolayer culture." Metabolism **27**(12): 1732-1742.

Sanders, J., D. Y. Chirgadze, P. Sanders, S. Baker, A. Sullivan, A. Bhardwaja, J. Bolton, M. Reeve, N. Nakatake, M. Evans, T. Richards, M. Powell, R. N. Miguel, T. L. Blundell, J. Furmaniak and B. R. Smith (2007). "Crystal structure of the TSH receptor in complex with a thyroid-stimulating autoantibody." Thyroid **17**(5): 395-410.

Sanders, J., M. Evans, L. D. Premawardhana, H. Depraetere, J. Jeffreys, T. Richards, J. Furmaniak and B. Rees Smith (2003). "Human monoclonal thyroid stimulating autoantibody." Lancet **362**(9378): 126-128.

Sanders, J., J. Jeffreys, H. Depraetere, M. Evans, T. Richards, A. Kiddie, K. Brereton, L. D. Premawardhana, D. Y. Chirgadze, R. Nunez Miguel, T. L. Blundell, J. Furmaniak and B. Rees Smith (2004). "Characteristics of a human monoclonal autoantibody to the thyrotropin receptor: sequence structure and function." Thyroid **14**(8): 560-570.

Schroeder, T., A. Kuendgen, S. Kayser, N. Kroger, F. Braulke, U. Platzbecker, V. Klarner, F. Zohren, D. Haase, M. Stadler, R. Schlenk, A. G. Czibere, I. Bruns, R. Fenk, N. Gattermann, R. Haas, G. Kobbe and U. Germing (2012). "Therapy-related myeloid neoplasms following treatment with radioiodine." Haematologica **97**(2): 206-212.

Smith, B. R. and D. S. Munro (1970). "The nature of the interaction between thyroid stimulating gamma G-globulin (long-acting thyroid stimulator) and thyroid tissue." Biochim Biophys Acta **208**(2): 285-293.

Tahara, K., N. Ishikawa, K. Yamamoto, A. Hirai, K. Ito, Y. Tamura, S. Yoshida, Y. Saito and L. D. Kohn (1997). "Epitopes for thyroid stimulating and blocking autoantibodies on the extracellular domain of the human thyrotropin receptor." Thyroid **7**(6): 867-877.

Toccafondi, R. S., S. Aterini, M. A. Medici, C. M. Rotella, A. Tanini and R. Zonefrati (1980). "Thyroid-stimulating antibody (TSab) detected in sera of Graves' patients using human thyroid cell cultures." Clin Exp Immunol **40**(3): 532-539.

TSAb-Standard (2013). "WHO International Standard 2nd International Standard for Thyroid Stimulating Antibody NIBSC code: 08/204 Instructions for use."

Vitti, P., C. M. Rotella, W. A. Valente, J. Cohen, S. M. Aloj, P. Laccetti, F. S. Ambesi-Impiombato, E. F. Grollman, A. Pinchera, R. Toccafondi and L. D. Kohn (1983). "Characterization of the optimal stimulatory effects of graves' monoclonal and serum

immunoglobulin G on adenosine 3',5'-monophosphate production in fRTL-5 thyroid cells: a potential clinical assay." J Clin Endocrinol Metab **57**(4): 782-791.

Wallaschofski, H. and R. Paschke (1999). "Detection of thyroid stimulating (TSAB)- and thyrotropin stimulation blocking (TSBAB) antibodies with CHO cell lines expressing different TSH-receptor numbers." Clin Endocrinol (Oxf) **50**(3): 365-372.

Watson, P. F., R. A. Ajjan, J. Phipps, R. Metcalfe and A. P. Weetman (1998). "A new chemiluminescent assay for the rapid detection of thyroid stimulating antibodies in Graves' disease." Clin Endocrinol (Oxf) **49**(5): 577-581.

Wilson, B., E. Raghupathy, T. Tonoue and W. Tong (1968). "TSH-like actions of dibutyryl-cAMP on isolated bovine thyroid cells." Endocrinology **83**(4): 877-884.

Acknowledgment

Ohne meinen Doktorvater _ wäre diese Arbeit nicht möglich gewesen. Ich danke ihm vielmals für das Überlassen des Themas, die fachliche Unterstützung und die Möglichkeit während meines Studiums als HiWi in seinem Labor zu arbeiten.

Mein besonderer Dank geht an _ für die tolle Zusammenarbeit während meiner Zeit im Labor an der JGU Mainz, die vielen Stunde und Tage, die wir zusammen an Experimenten oder Auswertungen gearbeitet haben und natürlich auch die tollen Korrekturhinweise beim Schreiben der Doktorarbeit.

Ich danke auch _, an die ich mich immer mit organisatorischen Fragen wenden konnte, und die für viele schöne Erlebnisse im Labor gesorgt hat.

Ganz besonders möchte ich mich auch bei _ für die Einweisung in das Labor, die Einführung in die Experimentierarbeit und die tolle Zusammenarbeit bei der Durchführung der Studien bedanken.

Curriculum Vitae

Name:	Markus Christopher Lehmann
Geburtsdatum:	18.08.1990
Geburtsort:	Neuwied
Eltern:	_
	_

Schulausbildung:

1997-2000	Grundschule Heddesdorfer Berg, Neuwied
2000-2009	Rhein-Wied Gymnasium, Neuwied, Abschluss mit Abitur
2009-2010	FSJ an der Ganztags-Grundschule Weißenthurm

Medizinstudium

2010-2016	Studium der Humanmedizin in Mainz
	2012: 1. Abschnitt der ärztlichen Prüfung (Physikum)
	2015: 2. Abschnitt der ärztlichen Prüfung
	2016: Staatsexamen für Humanmedizin

Berufslaufbahn:

Seit 2017	Assistenzarzt in der Klinik für Allgemein-, Viszeral- und Gefäßchirurgie des Marienhaus Klinikums St. Elisabeth Neuwied