Applying Prediction Models and AI to Optimize Growth Hormone Therapy in Children with Short Stature

Helena-Jamin Ly

Department of Pediatrics
Institute of Clinical Sciences
Sahlgrenska Academy, University of Gothenburg

UNIVERSITY OF GOTHENBURG

Gothenburg 2025

Cover illustration: Nurturing Growth, by Helena-Jamin Ly

The cover design was conceptualized and designed by the author using ChatGPT software as a tool for its creation.

Applying Prediction Models and AI to Optimize Growth Hormone Therapy in Children with Short Stature

© Helena-Jamin Ly 2025
helena-jamin.ly@gu.se or ly.helena@gmail.com

ISBN 978-91-8115-082-7 (PRINT)
ISBN 978-91-8115-083-4 (PDF)
http://hdl.handle.net/2077/84689

Printed in Borås, Sweden 2025
Printed by Stema Specialtryck AB

“Somewhere something incredible is waiting to be known”

— Carl Sagan

To My Family

Applying Prediction Models and AI to Optimize Growth Hormone Therapy in Children with Short Stature

Helena-Jamin Ly

Department of Pediatrics, Institute of Clinical Sciences
Sahlgrenska Academy, University of Gothenburg
Gothenburg, Sweden

ABSTRACT

Background: Growth hormone (GH) therapy has revolutionized the management of short stature in children, offering improved growth outcomes for those with GH deficiency and other related conditions. This thesis focuses on optimizing GH therapy through predictive tools and advanced machine learning models, with an emphasis on individualizing care to enhance patient outcomes and cost-effectiveness.

Studies summaries: Study 1 evaluated the Gothenburg prediction model, which accurately identified poor responders to GH therapy, reducing the likelihood of unnecessary treatment. In contrast to a Nordic study, where 28% of children with GH deficiency failed to achieve a sufficient growth response on treatment, 98% of patients selected using the Gothenburg model achieved adequate growth outcomes with GH therapy. Study 2 compared the Gothenburg and KIGS (Pfizer International Growth Database) prediction models, both of which demonstrated exceptional accuracy in predicting first-year growth ($r = 0.99$ for both models). This comparison highlighted the flexibility of clinical choice, as either model could effectively guide GH therapy planning.

Insulin-like growth factor 1 (IGF-1) is a potential biomarker that reflects the body´s response to GH. Study 3 explored the complexities of IGF-1 interpretation during GH therapy, particularly in early puberty, where discrepancies between clinical pubertal characteristics and biochemical hormone levels led to overestimated IGF-1 standard deviation scores (SDS). This study underscored the need for improved tools to enhance IGF-1 interpretation for better clinical decision-making.

Study 4 introduced machine learning approaches to predict IGF-1 SDS and optimize GH dosing strategies. Symbolic regression emerged as a powerful tool, providing accurate predictions with minimal input variables, while Explainable Boosting Machine (EBM) excelled in identifying complex feature relationships. These models not only predicted IGF-1 SDS with high accuracy (R^2 = 0.47 - 0.51) but also demonstrated clinical applicability by balancing precision with interpretability.

Conclusion: This thesis establishes a strong foundation for incorporating predictive models into routine clinical practice, facilitating more personalized and effective GH therapies. By utilizing advanced technologies and addressing key clinical challenges, this work contributes to optimizing treatment strategies, improving patient outcomes, and advancing precision healthcare.

Keywords: Explainable Boosting Machine, Growth Hormone, Growth Hormone Therapy, Insulin-like Growth Factor 1, Machine Learning, Prediction Models, Sex Steroids, Symbolic Regression.

ISBN 978-91-8069-978-91-8115-082-7 (PRINT)
ISBN 978-91-8069-978-91-8115-083-4 (PDF)

SAMMANFATTNING PÅ SVENSKA

BAKGRUND

Barn med tillväxthormonbrist riskerar att bli kortvuxna, och tillväxthormon-behandling är en central del av pediatrisk endokrinologi för att hjälpa dessa barn att nå sin fulla tillväxtpotential. Det är dock en utmaning att identifiera vilka barn som verkligen behöver behandling, eftersom det kan vara svårt att ställa en korrekt diagnos. Svår tillväxthormonbrist är relativt enkel att diagnosticera, men mildare former kan vara betydligt svårare att identifiera. Idag saknas en standardiserad metod för diagnostik, vilket gör att bedömningen istället baseras på en helhetsvärdering av flera undersökningar- en process som ofta är subjektiv.

Utöver diagnostiken är korrekt dosering av tillväxthormon en stor utmaning. Barns känslighet för behandlingen varierar, och fel dosering kan ha konsekvenser — för höga doser innebär risker, medan för låga doser ger otillräcklig tillväxt. I denna avhandling undersöks hur prediktionsmodeller, artificiell intelligens (AI) och maskininlärning kan optimera behandlingen för barn som får tillväxthormon.

RESULTAT

De två första studierna fokuserade på att identifiera de barn som har störst nytta av tillväxthormonbehandling. Prediktionsmodeller, det vill säga matematiska verktyg som kan förutsäga tillväxtrespons innan behandlingen påbörjas, utvärderades. Resultaten visade att dessa modeller fungerar väl i klinisk praxis och kan vara värdefulla verktyg för att identifiera rätt patienter. Genom att använda prediktionsmodeller kan man undvika onödig behandling för barn som sannolikt inte kommer att svara på behandlingen, vilket samtidigt bidrar till en mer kostnadseffektiv vård.

När rätt barn har identifierats blir nästa utmaning att dosera tillväxthormon korrekt. Den senare delen av avhandlingen fokuserade på att använda modern teknik, som AI och maskininlärning, för att optimera doseringen. Ett viktigt verktyg för att utvärdera rätt dos är att mäta nivåerna av insulin-like growth factor 1 (IGF-1), ett protein som teoretiskt speglar effekten av tillväxthormon. Dock påverkas IGF-1 nivåer av flera faktorer, såsom ålder, kön, könshormoner, näringsstatus och inflammation, vilket gör tolkningen komplex. Studier på vuxna har visat att låga nivåer av IGF-1 är kopplade till ökad risk för hjärtkärlsjukdom, medan höga nivåer är associerade med en ökad

risk för vissa cancerformer. Behandlingsmålet är därför att hålla IGF-1 nivåerna inom normalintervallet för friska barn.

För att uppnå detta användes maskininlärningsmodeller såsom Symbolic Regression och Explainable Boosting Machine (EBM) för att förutsäga IGF-1-nivåer hos barn som behandlas med tillväxthormon. Dessa modeller möjliggör en individanpassad dosering baserad på det önskade IGF-1 svaret, vilket säkerställer en mer exakt och säker behandling för varje barn.

SLUTSATS

Denna avhandling lägger grunden för en mer individanpassad vård och visar hur avancerad teknologi kan integreras i klinisk praxis för att förbättra resultaten för barn med tillväxthormonbehandling. Prediktionsmodeller och maskininlärning har potential att förbättra både diagnostik och behandling, vilket kan leda till bättre tillväxtresultat och ökad säkerhet för patienterna.

LIST OF PAPERS

This thesis is based on the following studies:

Study 1: **Ly HJ**, Fors H, Nilsson S, Dahlgren J

A Prediction Model Could Foresee Adequate Height Response in Children Eligible for Growth Hormone Treatment

Acta Paediatrica 2022 Feb; 111(2): 346-353

Study 2: **Ly HJ**, Lindberg A, Fors H, Dahlgren J

Comparison of Two Prediction Models in a Clinical Setting to Predict Growth in Prepubertal Children on Recombinant Growth Hormone

Growth Hormone & IGF Research 2023 Feb; 68:101523

Study 3: **Ly HJ**, Ankarberg-Lindgren C, Fors H, Nilsson S, Dahlgren J

Interpreting IGF-1 in Children Treated with Recombinant Growth Hormone: Challenges During Early Puberty

Frontiers of Endocrinology 2025 Jan 21;15:1514935

Study 4: **Ly HJ**, Suvilehto J, Skyman B, Ankarberg-Lindgren C, Fors H, Dahlgren J

Applying Artificial Intelligence to Predict IGF-1 in Boys Treated with Growth Hormone: A Focus on Maintenance and Pubertal Phase Predictors

Manuscript submitted

CONTENTS

ABBREVIATIONS

ALS	Acid-labile Subunit
BSA	Body Surface Area
EBM	Explainable Boosting Machine
EMA	European Medicines Agency
FDA	Food and Drug Administration
GAM	Generalized Additive Models
GH	Growth Hormone
GHD	Growth Hormone Deficiency
GHBP	Growth Hormone Binding Protein
GHR	Growth Hormone Receptor
GHRH	Growth Hormone-Releasing Hormone
IGF-1	Insulin-like Growth Factor 1
IGFBP	Insulin-like Growth Factor Binding Protein
iGRO	Individualized Growth Response Optimization
ISS	Idiopathic Short Stature
ITT	Insulin Tolerance Test
JAK	Janus Kinase
KIGS	Pfizer International Growth Database
MAPK	Mitogen-Activated Protein Kinase
MAPE	Mean Absolute Percentage Error

MAE	Mean Absolute Error
MPH	Mid-Parental Height
MRI	Magnetic Resonance Imaging
PI3K	Phosphoinositide 3-Kinase
PAPP-A2	Pregnancy-Associated Plasma Protein-A2
SDS	Standard Deviation Score
SGA	Small for Gestational Age
SHOX	Short Stature Homeobox-Containing Gene
STAT	Signal Transducer and Activator of Transcription
STC	Stanniocalcin
TSH	Thyroid-Stimulating Hormone

1 INTRODUCTION

1.1 HISTORY

The origins of endocrinology trace back to 1849, when Adolph Berthold of Göttingen introduced the concept of internal secretions. He observed that transplanting the testis of a rooster to a different part of its body prevented atrophy of the comb, establishing a foundational principle in the field (1).

June 1, 1889, 72-year-old Brown-Séquard presented his self-experimentation in Paris. Over two weeks, he injected himself with testicular fluid daily and claimed rejuvenation, suggesting that the testes produce a substance, then referred to as *dynamogenic*, that could be extracted and used to restore vitality in aging organisms (2).

Investigations into the role of the pituitary gland were initiated through clinical studies and also anatomical analyses of individuals with gigantism and adults displaying acromegalic characteristics. Frietsche, a Swiss physician, provided in 1884 the earliest detailed clinical description of a 44-year-old male presenting with features characteristic of acromegaly, including enlarged hands and feet, coarse facial characteristics, and prominent jaw. These features were later described with the term *acromegaly* as per Pierre Marie in 1886 (3).

In 1909, American neurosurgeon Cushing carried out a transsphenoidal procedure on a farmer from South Dakota with acromegaly, who subsequently lived for at least 21 more years. Cushing later introduced the term *growth hormone*, hereafter referred to as GH, and advocated for further research to identify specific pituitary hormones (4).

A landmark study influencing the development of GH therapy was conducted by Evans and Long in 1922. A substance from the anterior lobe of oxen pituitaries was extracted and injected into rats, effectively demonstrating its growth-stimulating properties (5).

Houssay and Biassotti advanced understanding of the metabolic roles of pituitary hormones by demonstrating that diabetic dogs experienced a reduction in blood sugar levels following pituitary removal in 1930 (6). For years thereafter, there was debate over whether the regulation of carbohydrate metabolism and the diabetogenic effect involved multiple pituitary factors. By the 1950s, consensus emerged that these effects were attributable to a single

substance—*growth hormone*. Extensive research by Houssay on the pituitary gland and carbohydrate metabolism earned the Nobel Prize in 1947.

In the meantime, Li and Evans at the University of California, Berkeley, achieved the first successful isolation of bovine GH in 1944 (7). Their team later purified human GH and characterized its primary structure as a protein consisting of 191 amino acids with two disulfide bonds. In 1969, they successfully achieved the first chemical synthesis of human GH (8). The purification and optimization of human GH marked the beginning of a new era in therapeutic use. A 1958 report by Raben is often cited as the first documented case demonstrating a highly favorable growth response to GH extracted from human pituitaries in a boy with GH deficiency (9). Westphal published methodological studies in the 1960s on measuring human GH and its clinical applications in diagnosing and evaluating human GH levels in various patient populations (10).

Following the success of pituitary-derived GH therapy, the demand for human pituitaries surged. In response, the National Pituitary Agency (NPA) was established in 1960 in the U.S. to oversee the collection, purification, and distribution of GH. Between 1963 and 1985, the NPA facilitated treatment for approximately 7,700 children in the U.S. and 27,000 globally with GH deficiency (11).

The scarcity of pituitary-derived GH significantly hindered progress in the field until 1979, when the biotechnology company Genentech successfully inserted the human GH gene into bacteria; enabling the production of recombinant GH (12). Large-scale manufacturing of recombinant GH began in 1985, coinciding with the abrupt discontinuation of pituitary-derived GH therapy after several patients were diagnosed with Creutzfeldt-Jakob disease. Fortunately, the development of recombinant human GH came at an opportune time, offering a safe, uncontaminated product combined with potential for unlimited supply.

See the history timeline summarized in Figure 1.

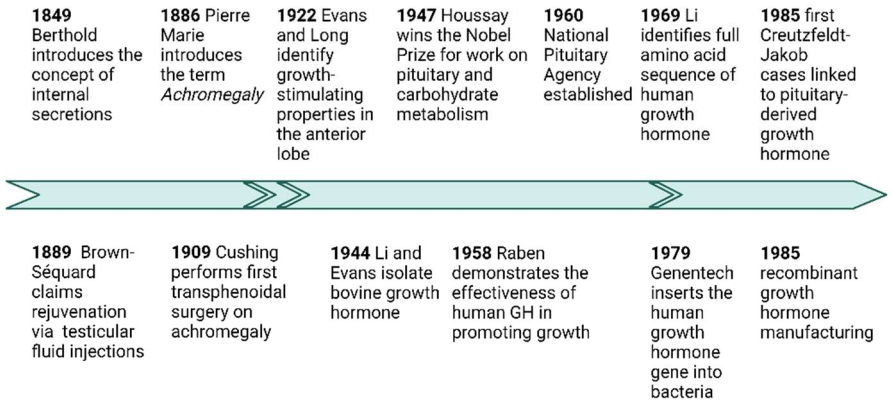

Figure 1. History Timeline for Growth Hormone. Created with BioRender.com.

1.2 GROWTH HORMONE

1.2.1 HYPOTHALAMUS AND PITUITARY

The pituitary gland, or hypophysis, consists of the anterior lobe (the adenohypophysis) and the posterior lobe (the neurohypophysis) and is situated at the base of the brain within a cavity of the sphenoid bone, a butterfly-shaped bone at the base of the skull. This cavity is known as the sella turcica and provides a protective enclosure for the pituitary gland. The hypophysis interacts closely with hypothalamus at the base of the brain, a neural structure that plays a crucial role in maintaining homeostasis by integrating signals from the periphery with signals from other brain regions and relaying them to the pituitary gland and autonomic nervous system.

Natural GH secretory pulses, result from the interaction of regulatory peptides, primarily the hypothalamic peptides growth hormone-releasing hormone (GHRH) and somatostatin (13).

From the arcuate nucleus of the hypothalamus, GHRH is released to stimulate the production of GH in the pituitary. Interestingly, GHRH was initially isolated not from human hypothalamic tissue but from non-hypothalamic pancreatic tumor cells in individuals with acromegaly (14). Somatostatin, also produced by hypothalamic neurons, serves as an inhibitory regulator of GH secretion while also suppressing thyroid-stimulating hormone (TSH) and prolactin release (15). Once released, somatostatin travels through the hypothalamo-portal vascular system to the anterior pituitary, where it suppresses GH secretion by targeting somatotroph cells. At this site, its inhibitory role outweighs the stimulatory effects of GHRH on GH release (16). Furthermore, GH regulation operates through a classic negative feedback loop, wherein increased concentrations of GH and insulin-like growth factor-1 (IGF-1) trigger somatostatin-producing neurons to enhance somatostatin secretion and activity, thereby reducing GH production.

1.2.2 ISOFORMS AND SECRETION MECHANISMS

GH is present in multiple isoforms, with the 22 kiloDalton (kDa) variant being the most prevalent, accounting for approximately 50% of circulating GH. This primary isoform, composed of 191 amino acids and organized into a three-dimensional fold of a four-helix bundle protein, is the main form responsible for promoting growth. The second most abundant isoform, the 20 kDa variant, constitutes 5-10% of circulating GH. This smaller form, consisting of 176 amino acids, exhibits reduced growth-promoting activity compared to the 22 kDa isoform (17). Some conditions such as Turner syndrome are found to have more non-22 kDa isoforms diminishing the effect on the GH-receptor (18).

GH is released from the somatotroph cells in the anterior lobe in a distinctive pulsatile pattern with intermittent bursts of hormone release rather than a continuous steady flow throughout the day. The majority of pulses occur during the first few hours of nocturnal sleep, particularly during slow-wave deep sleep. The pulsatile pattern, established during infancy after the infant develops diurnal sleep, is characterized by episodes of increased GH secretion interspersed with periods of baseline levels of secretion which seems to be required for optimal intracellular signaling (19). This rhythmic release allows the body to respond dynamically to changing internal and external stimuli, ensuring that the GH is released in appropriate amount and at the right times to support various physiological functions.

The pulsatile secretion is orchestrated by a complex interplay between signals from the hypothalamus and the pituitary gland. The regulation involves both stimulatory signals through GHRH, ghrelin and inhibitory signals via somatostatin (20). See Figure 2. GHRH appears to play a significant role, as nighttime GH secretion decreases following the administration of GHRH antagonists (21). Ghrelin, a peptide mostly produced in the gastrointestinal tract, acts on the pituitary and hypothalamus by modulating activity through the vagus nerve and stimulates the secretion of GH acting via GHRH from the hypothalamus, which activates somatotroph cells in the pituitary gland to secrete GH (22). GH acts on the liver to stimulate IGF-1 production and promotes bone growth directly on chondrocyte progenitor cells (23).

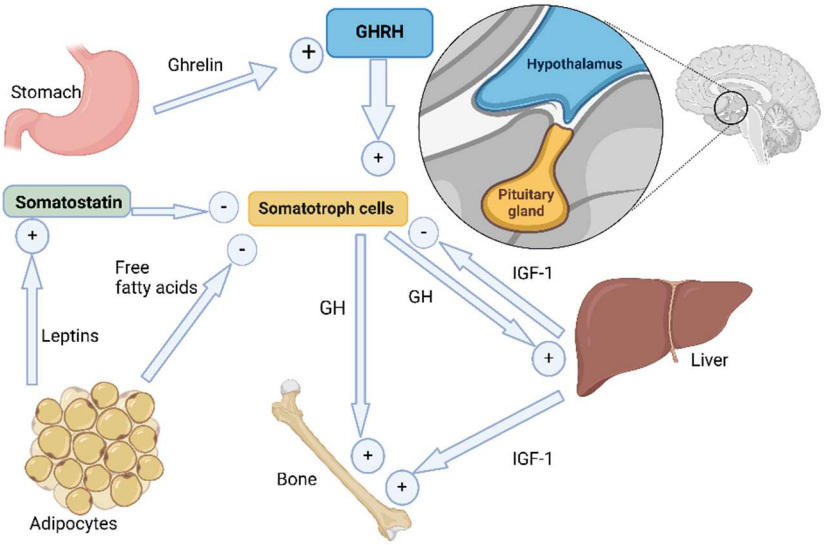

Figure 2. The regulatory network of growth hormone secretion and its effects. Positive and negative regulatory interactions are denoted by "+" and "-", respectively. Created with BioRender.com.

It was previously believed that GH acted exclusively through IGF-1, as described by the somatomedin hypothesis. However, Isaksson challenged this by demonstrating that local GH injections in hypophysectomized rats stimulated longitudinal bone growth directly (23). See Figure 3.

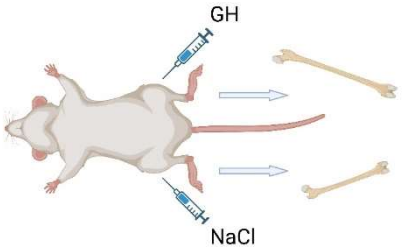

Figure 3. Mouse experiment done by Isaksson et al. showing the direct effect of growth hormone on longitudinal bone growth. Created with BioRender.com.

Adipocytes suppress GH secretion through leptin-mediated stimulation of somatostatin release and direct inhibition by free fatty acids on somatotroph cells (24). In addition to somatostatin, the pulsatile secretion of GH is regulated by inhibitory factors such as feedback suppression by GH and IGF-1 at the level of the hypothalamus and pituitary gland. Elevated blood glucose level can also inhibit GH release, although the underlying mechanisms are not yet fully understood.

Furthermore, emerging evidence suggests that insulin delivered to the liver via the portal system may enhance the liver's sensitivity to GH, facilitating IGF-1 production (25, 26). During fasting, reduced insulin levels in the liver decrease its responsiveness to GH, resulting in lower IGF-1 levels despite compensatory increases in GH secretion. A recent review by Yuen et al. (27) emphasized the critical role of portal insulin delivery in regulating the GH/IGF-1 axis by affecting hepatic GH receptor synthesis, thereby altering GH sensitivity and IGF-1 production. They underscored the importance of interpreting GH and IGF-1 levels in conditions where insulin levels are either elevated (e.g., obesity, Cushing's syndrome, or glucocorticoid treatment) or reduced (e.g., malnutrition, type 1 diabetes), as these changes can disrupt hepatic GH sensitivity and lead to a misalignment in the GH/IGF-1 axis.

1.2.3 GROWTH HORMONE SIGNALING

GH acts by binding to the growth hormone receptor (GHR) on the cell surface, triggering an intracellular signaling cascade. The GHR is composed of three main regions: an extracellular domain of 246 amino acids, a 24-amino acid helical transmembrane segment, and an intracellular domain of 350 amino acids. The close-up structure of the GH and its receptor from De Vos et al. can be seen in Figure 4 with two binding sites (28). The GHR exists primarily as a dimer, stabilized by interactions between its transmembrane helices (29). While GHRs are most abundantly expressed on hepatocytes, they are also present on various other cell types throughout the body.

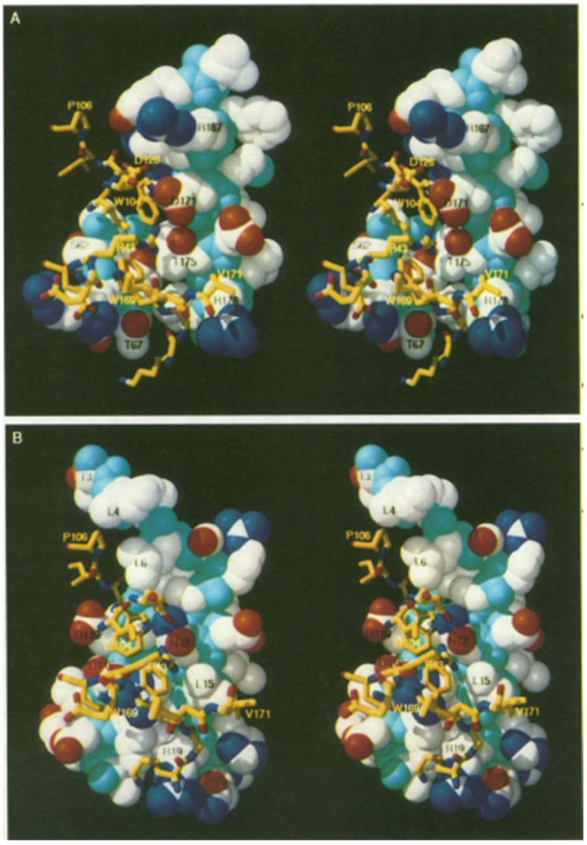

Figure 4. Close-up of growth hormone (space filling model) and receptor (stick model) with two binding sites in A and B. Adapted with permission from De Vos, Published in Science 1992:306-12.

Upon GH binding to the GHR, conformational changes occur, activating intracellular signaling pathways, primarily the Janus kinase (JAK)-signal transducer and activator of transcription (STAT) pathway. Specifically, GH first binds site 1 to one receptor ligand, then binding site 2 to the other ligand causing a dimerization of the receptor (30), this brings the two JAK2 molecules into close proximity allowing a transphosphorylation to occur thus activating JAK2. This leads to the phosphorylation of specific tyrosine residues on the intracellular domain of GHR. These phosphorylated residues create docking sites for signaling molecules, triggering the activation of STAT proteins, particularly STAT5b. STAT5b dimerizes and translocate from the cytoplasm to the nucleus, where it initiates the transcription of target genes, including IGF-1, IGF-2, IGFBP-3, and ALS, which play key roles in mediating the effects of growth hormone (31, 32). See Figure 5.

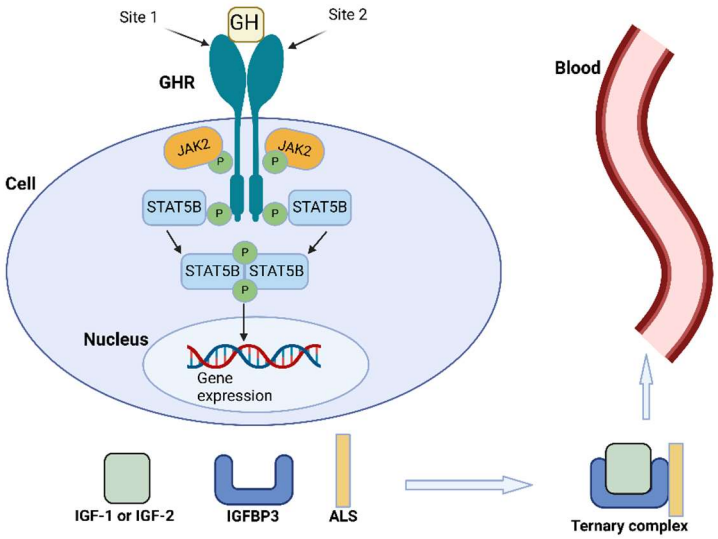

Figure 5. Growth hormone binding to the receptor leading to activation of Janus kinase 2 (JAK2) and signal transducer and activator of transcription 5B (STAT5B), which promotes the expression of genes encoding e.g. insulin-like growth factors 1 and 2 (IGF-1 and 2), IGF-binding protein 3 (IGFBP3), and the acid labile subunit (ALS). These components form the ternary complex that enters the circulation. Created with BioRender.com.

In addition to the JAK-STAT pathway, GH also activates other signaling pathways, including the mitogen-activated protein kinase (MAPK)-extracellular signal-regulated kinase (ERK) pathway. This pathway contributes to broader cellular responses related to development and growth (33). The extracellular domain of the GHR can be cleaved and released into the bloodstream as growth hormone binding protein (GHBP). GHBP binds approximately half of the circulating GH, extending its half-life. However, its precise physiological role remains incompletely understood (34).

1.3 INSULIN-LIKE GROWTH FACTOR 1

IGF-1 functions in autocrine, paracrine, and endocrine manners and is expressed in various tissues, including muscle, fat and the growth plate, where a key role is played in growth, metabolic effects, and cellular repair (35). Additionally, IGF-1 also influences aging, cancer, insulin sensitivity, and cardiovascular health (36). IGF-1 is a single-chain polypeptide composed of 70 amino acids, stabilized by three disulfide bridges. A complete amino acid sequence was first identified in 1978 by Rinderknecht and Humbel (37). The structure of IGF-1 is similar to insulin, sharing 43% of its amino acid sequence with human proinsulin. Both peptides can trigger metabolic and growth-promoting effects through their specific receptors, which share structural and functional similarities (38). Despite their structural resemblance, binding remains specific to their respective receptors. Soos and Siddle found that the shared functional effects are likely due to hybrid receptors in certain cells, such as myocytes, which consist of one half insulin receptor polypeptide and one half IGF-1 receptor polypeptide (39). IGF-1 and insulin both can activate overlapping signaling pathways like the Phosphoinositide 3-kinase-Proteine Kinas B (PI3K-AKT) for regulating various biological processes, including cell survival, metabolism, growth, and migration, and the MAPK/ERK pathway for growth and proliferation.

IGF-1 is bound to a family of six specific IGF binding proteins (IGFBP) in extracellular fluid. The IGFBPs are labeled 1 to 6 with its own characteristics (40). Approximately 75% of IGF-1 is bound to IGFBP-3 in ternary complexes with acid-labile subunit (ALS), which stabilizes IGF-1 and facilitates its transport across body compartments. The production and formation of all three units in the liver is induced by the pulsatile action of GH as previously mentioned.

The activity of IGF-1 relies on the modification of IGFBP-3, which releases free IGF-1 to bind to the receptor. Modifications, such as glycosylation or phosphorylation of IGFBP-3, enhances its binding to IGFs, while proteolysis, dephosphorylation, or deglycosylation reduce IGF binding (41). The cleavage of IGFBP-3 is mainly by the specific protease pregnancy-associated plasma protein (PAPP-A2) (42) while modification of protease activity is through stanniocalcin (STC). IGFBPs can also interact with other cellular structures, which decreases affinity for bound IGF-1 making it more readily available (43).

1.3.1 NUTRITION AND INFLAMMATION

A caloric or protein-restricted diet has been shown to decrease IGF-1 levels in both children and adults (44). Fasting significantly reduces IGF-1 levels, as shown by Isley et al. a 36% decrease after five days of fasting, which partially recovers with refeeding (45). Both proteins and energy are critical for IGF-1 levels. IGFBP-3 or its proteolytic activity does not, conversely, seem to be affected by caloric restrictions, which may explain the theory that the decreased IGF-1 during fasting is caused by a lower production of IGF-1 rather than dependent on IGFBP-3 activity (44).

While obese adults generally show decreased GH secretion and low to low-normal total IGF-1 levels, obese children often display normal, or even elevated, IGF-1 levels as reviewed by A Juul (46). The largest study reviewed reported IGF-1 and IGFBP-3 levels increased by +0.41 SDS and +1.61 SDS, respectively, in obese children (47).

Patients with untreated inflammatory disease like coeliac disease and Crohn's disease show lower IGF-1 and IGFBP-3 levels probably due to malnutrition and inflammation (48, 49).

1.3.2 IGF-1 AND PUBERTY

Extensive research has shown that during childhood, serum levels gradually increase from low levels at birth, with no significant differences between boys and girls in larger healthy cohorts. Juul et al. showed that IGF-1 levels typically rise throughout childhood, with a sharp increase during puberty, peaking at around 14.5 years in girls and 15.5 years in boys (50).

IGF-1 correlates with height velocity in prepubertal children, but not in pubertal children as a decline in height velocity in late puberty showed a

remaining elevation of IGF-1 (50-53). IGF-1 levels gradually decline with age in adults (50, 54). In cases of precocious puberty, IGF-1 and IGFBP-3 levels are elevated and correlate with growth rate in both boys and girls (55, 56).

The pubertal increase in IGF-1 and IGFBP-3 is driven by elevated GH secretion, which is stimulated by sex steroids (57, 58). Administered estradiol, in particular, plays a key role in GH stimulation but has a biphasic effect on IGF-1, where higher doses have been shown to reduce IGF-1 levels (59). Administration of exogenous androgens in boys with delayed sexual maturation increases IGF-1 concentrations by enhancing the pulse amplitude of GH secretion, without altering pulse frequency (60). However, studies in men have shown decreased GH secretion with the use of estrogen receptor blockers but not androgen receptor blockers, again highlighting the critical role of estrogen (61, 62). Additionally, in prepubertal boys treated with dihydrotestosterone, which cannot be aromatized to estrogen, IGF-1 levels did not increase but instead decreased (63). These findings suggest that the aromatization of androgens to estrogen is essential for the rise in GH and IGF-1 levels during puberty in males.

In addition to the effects mentioned above, estradiol also promotes growth independently of GH and its receptor. Venken et al. demonstrated in GHR-disrupted mice that estradiol stimulates longitudinal bone growth by increasing hepatic IGF-1 production (64).

1.3.3 IGF-1 ASSAYS AND REFERENCE RANGES

A significant challenge in the clinical use of IGF-1 measurements is the substantial variability in results produced by different assays (65). A consensus established guidelines for assay validation and the development of robust normative data in order to address this challenge (66). The use of recombinant international Standard 02/254 (67) for all assays is recommended and emphasizes the need for IGF-1 assays to be insensitive to interference from IGFBPs.

The correct interpretation of IGF-1 levels requires reliable reference ranges. Several authors have published reference data adjusted for age and sex derived from large cohorts of children, adolescents, or adults, utilizing various immunoassay techniques (50, 52, 68).

Reference ranges including gender, age, and pubertal status are recommended (69). Löfqvist et al. developed a model that simultaneously relates serum

IGF-1 levels to age, pubertal stage, and gender in children. This model is the first to accurately convert serum IGF-1 concentrations into SDS from infancy to adulthood (53).

When diagnosing GHD, Inoue-Lima et al. demonstrated that IGF-1 reference ranges adjusted for pubertal stage provided the highest positive predictive value compared to reference ranges based solely on chronological age or bone age (70).

1.3.4 IGF-1 AND SAFETY

Both elevated IGF-1 levels, as seen in acromegalic patients (71), and low IGF-1 levels, as in adults with GHD, are associated with increased mortality (72). Major meta-analysis by Burgers et al., including 14,906 participants, found a non-linear relationship between IGF-1 levels and mortality, with both high and low levels linked to an elevated risk of all-cause mortality (73). Epidemiological studies in adults have demonstrated an association between IGF-1 levels within the top quartile of the normal reference range and an increased risk of colorectal, breast, and prostate cancer (74). A nested case-control study also found that individuals with low IGF-1 levels and high IGFBP-3 levels faced a significantly increased risk of developing ischemic heart disease over a 15-year period (75).

The mechanisms underlying the relationship between serum IGF-1 levels and the risks for cancer and ischaemic heart disease remain unclear. Confounding factors may contribute but are challenging to exclude. It is important to note that the studies mentioned above are epidemiological studies conducted in adults, and their relevance or applicability to children remains uncertain, highlighting the need for further evaluation of IGF-1 levels in pediatric populations, particularly when assessing safety in GH therapy.

The assessment of bioactive IGF-1 levels offers additional insights, especially in cases like short patients born small for gestational age (SGA), where achieving adequate growth may require supraphysiological IGF-1 levels. In a study by Wegmann et al. involving a cohort of short SGA children treated with high-dose GH, 68% of patients reached total IGF-1 levels above +2 SD, while only 15% exceeded this threshold for bioactive IGF-1 (76). When examining cumulative lifetime IGF-1 exposure in an SGA cohort, Kjaer et al. demonstrated that GH therapy during childhood resulted in only a slight increase compared to levels observed in healthy individuals (77).

Despite the uncertainties, the potential risks associated with both low and high IGF-1 levels underscore the importance of maintaining IGF-1 levels within the normal range during GH therapy, as recommended by international guidelines (78). While the use of IGF-1 and IGFBP-3 as biomarkers of GH action has been proposed (79), their relevance has not yet been fully explored in-depth.

For now, IGF-1 is primarily recommended as a marker for long-term safety and treatment adherence, rather than being solely used as a measure of GH therapy efficacy. International guidelines caution against basing GH dosing exclusively on IGF-1 levels, emphasizing instead a weight-based dosing approach, despite known individual variations in sensitivity to GH (78). Further research is needed to refine the role of IGF-1 in guiding treatment and ensuring both efficacy and safety.

1.4 GROWTH HORMONE DEFICIENCY

Growth Hormone Deficiency (GHD) in children can be isolated, or occur alongside other pituitary hormone deficiencies, and may be either congenital or acquired. While many causes of GHD in children have been identified, the majority of cases remain idiopathic.

The definition of short stature in children is a height below 2 SDS from the population mean. GHD is often considered in a child with short stature, although it is rare as an underlying cause. In children with a height above -2 SDS for gender and age, the estimated prevalence of GHD is 2% (80). Mameli et al. found the overall prevalence of GHD in children and adolescent to range between 1/1,107 and 1/8,646 (81).

Diagnosing GHD in childhood is a complex process that involves thorough clinical and auxological evaluations, along with biochemical testing and radiological assessments (82). If left untreated, childhood-onset GHD results in permanent short stature. Before considering GHD in children, other causes for short stature, such as genetic, organic, hormonal, metabolic, or psychogenic causes, need to be ruled out.

1.4.1 DIAGNOSING GHD

Diagnosing GHD in children is a multifaceted process including the following aspects (82-84):

1.4.1.1 MEDICAL HISTORY AND PHYSICAL EXAMINATION

Children with GHD may develop characteristics of normal body proportions, relatively large head, sarcopenia, alterations in body composition for example increased body fat, and reduced bone mineral density. Consistent growth below the normal range for age and gender may indicate a potential GHD.

A detailed medical history and physical examination are essential to assess growth and body proportions, as abnormal proportions strongly indicate skeletal dysplasia rather than GHD. Facial and body dysmorphic features should also be evaluated to exclude syndromic causes. An analysis of auxology and growth charts is essential, focusing on how the child's growth compares to appropriate growth standards, parental heights, and trends over time. Several growth disorders are distinguished by their characteristic growth patterns.

A thorough family history is crucial including low birth weight or length, early-life failure to thrive, developmental delays, intellectual disabilities, or behavioral issues for primary growth disorders such as dysregulation of the epiphyseal growth plate. Secondary growth disorders, including GHD, caused by external influences on the epiphyseal growth plate, may present with indicators such as abnormal weight changes, anorexia, fatigue, abdominal symptoms, signs of increased intracranial pressure, or the effects of certain medications (84).

1.4.1.2 BIOCHEMICAL TESTING

Biochemical testing often includes IGF-1 and IGFBP-3 levels, an indirect measurement of GH secretion. These levels are assessed in blood despite IGF-1 also being produced in peripheral tissues, where it contributes to both local and systemic effects. Low levels of IGF-1 and/or IGFBP-3 should prompt further evaluation for GHD, although there is no consensus on the exact cut-off values for these markers, as their levels can be influenced by various confounding factors (70, 85, 86). See chapter 1.3.

Rikken et al. analyzed a Dutch cohort of children with GHD or idiopathic short stature (ISS) and found when IGF-1 was below -2 SDS, the likelihood of a GHD diagnosis was 65%, with a specificity of 78%. Similarly, for IGFBP-3 levels below -2 SDS, the likelihood of a GHD diagnosis was 53%, with a specificity of 81% (87). A low IGFBP-3 is less sensitive but more specific than a low IGF-1. The additional value of IGFBP-3 as a complement IGF-1 is, however, unclear, except in young children under three years of age (88). Cianfarani et al. reported that an IGF-1 cut-off below -1.9 z-score yielded a sensitivity of 73% and a specificity of 95% for diagnosing GHD. For IGFBP-

3, the sensitivity was 30% and the specificity was 98% (85). Ibba et al. found a sensitivity of 68% and specificity of 63% using an IGF-1 cut-off of -1.5 SDS (89). These findings highlight the diagnostic variability and limitations of relying solely on IGF-1 and IGFBP-3.

Furthermore, IGF-1 and IGFBP-3 levels can indicate other abnormalities in the GH-IGF-1 axis. Low levels combined with an elevated GH peak suggest bioinactive GH or GH insensitivity, warranting further evaluation through IGF-1 generation tests.

Wit et al. developed recommendations for laboratory testing in children with growth failure by combining clinical insights and evidence-based approaches (84). The basic screening for short stature includes a complete blood count, IGF-1, thyroid-stimulating hormone (TSH), and free T4 to rule out hypothyroidism; transglutaminase IgA antibodies to screen for celiac disease; and measurements of sodium, potassium, creatinine, calcium, phosphate, and alkaline phosphatase to evaluate for kidney insufficiency or metabolic bone disorders. Genetic testing on short girls to rule out Turner syndrome. For children under the age of 3, blood gas analysis and serum IGFBP-3 are recommended to rule out renal tubular acidosis and GHD. Children 10 years or older with affected height and weight, inflammatory markers are recommended to screen for inflammatory bowel disease. This comprehensive approach highlights the need to consider less obvious causes of growth failure. For example, Van Rijn et al. demonstrated in a systematic review that 2-8% of children with short stature may have celiac disease, even in the absence of gastrointestinal symptoms, underscoring the importance of excluding this as a potential underlying cause (90).

1.4.1.3 RADIOLOGIC ASSESSMENTS

X-ray of the left hand/wrist to evaluate both bone age, but also adult height prediction can be useful. The extent of bone age delay assists in differentiating between various types of growth disorders. Bone age estimation should thus be part of a routine growth failure evaluation (82). Most primary growth disorders show a bone age close to chronological age. A hand/wrist X-ray can also reveal signs of abnormalities related to short stature homeobox-containing (SHOX) haploinsufficiency (91). For a more detailed assessment, MRI of the pituitary is recommended to identify organic causes of GHD. In cases where GHD is diagnosed, performing an MRI is essential to evaluate the structure and positioning of the pituitary gland, aiding in the identification of potential underlying abnormalities (82).

1.4.1.4 GENETIC EVALUATION

Height is a multigenic trait influenced by multiple genes, with 80-90% of adult height determined by heritability (92). Genome-wide association studies have identified numerous genetic loci linked to height, many of which are involved in cellular or metabolic processes, growth-regulating systems, or critical functions within the growth plate (93, 94).

Advances in genetic tools have facilitated the identification of new genetic disorders responsible for GHD in short children previously classified as SGA or ISS (95). Homma et al. demonstrated that 34% of children born SGA with persistent short stature, along with features such as dysmorphic characteristics, major malformations, developmental delays, and/or intellectual disabilities, carried pathogenic or likely pathogenic variants in genes linked to growth disturbances (96).

Genetic testing should be conducted if a primary growth disorder is suspected, particularly in cases where the child exhibits disproportionate growth and/or dysmorphic features (84). Flechtner et al. reported a prevalence of skeletal dysplasia of 21.8% in patients with ISS and 20.9% in those born SGA (97). Furthermore, the prevalence increased to 33% when a parent was also short. Children with GHD who experience significant growth impairment and have a confirmed family history should undergo genetic mutation screening. The progressive endocrine disorders observed in some of these cases highlight the need for ongoing long-term monitoring (98).

1.4.1.5 GROWTH HORMONE STIMULATION TEST

A single random measurement of GH in plasma provides limited information about overall GH secretion over time. To address this, provocative tests were developed based on the assumption that the maximum GH peak following a stimulus injection could distinguish individuals with normal GH secretion from those with GHD.

The first GH stimuli in use was insulin-induced hypoglycemia in the insulin tolerance test (ITT) (99). Various other test combinations, such as arginine, glucagon, propranolol, and L-Dopa, have been reported, but these act through different mechanisms at the pituitary or hypothalamic levels and are associated with varying risk profiles, response variability, and limited reproducibility.

Extremely low GH levels are indicative of severe GHD; however, no definitive cut-off exists to differentiate normal GH levels from partial GHD. Since GHD exists on a continuum from partial deficiency to complete deficiency, there is

no consensus on the optimal test cut-off, as controlled or evidence-based studies linking these thresholds to adult height are lacking.

Early reports defined a GH response above 5 µg/L as sufficient but as GH availability increased, most centers raised the cut-off to 7 µg/L to expand treatment eligibility. Following the introduction of recombinant GH in the 1990s, the cut-off was further increased to 10 µg/L (100). 10 µg/L is marginally below the mean response for normally growing children tested with provocation tests, hence the cut off (101). Studies have shown that a GH peak cut-off of < 10 µg/L can effectively predict growth response during the first few years of GH therapy (102, 103). In recent years, many centers have, however, readjusted the cut-off to 7 µg/L, reflecting updates in modern methods and reference standards (104).

Rosenfeld et al. outlined several limitations associated with provocative GH testing (100):

1. Provocative GH tests are non-physiological and fail to replicate the natural GH secretion pattern due to differences in dosage, administration, and interactions,
2. The definition of a normal GH rise after stimulation is arbitrary,
3. Age dependency and the role of sex steroid priming has not fully been established. To avoid false positive results, administration of sex steroids prior to GH provocative tests has been proposed as GH secretion is generally lower prior to puberty (78),
4. GH assays, on which the tests are reliant, are variable in accuracy,
5. Cost, risk of tests and discomfort for the child. The pharmacological agents used to stimulate GH secretion may cause side effects, with insulin-induced hypoglycemia potentially leading to seizures, particularly in children with severe GHD,
6. Poor reproducibility of results,
7. Severe GHD can be identified but partial GHD among normal short children is difficult to identify. There is a possibility that a partial GHD can be masked by potent provocative agents.

Bright et al. showed that provocative testing has a high false positive rate and the probability of true-positive result in a child with short stature is approximately 1 in 36 cases (105). International guidelines recommend against relying solely on GH provocative test results for diagnosing GHD, as these tests are associated with numerous pitfalls, as outlined above (78).

1.4.1.6 GROWTH HORMONE SPONTANEOUS TEST

Measurements of spontaneous GH secretion are arguably superior to provocation tests as it reflects the physiological secretion of GH in the individual (106). Screening requires multiple blood sampling every 20 to 30 minutes to capture the pulsatile nature of GH secretion, to construct a GH profile. Interpreting the GH profiles include analyzing the amplitude and frequency of the GH peaks, the baseline between the peaks, and calculation of the total amount of GH secretion during the testing period. See Figure 6 for a normal growth hormone spontaneous night profile, and Figure 7 for a pathological example. The duration of GH measurements has been reduced from 24 hours to 12 hours overnight, as nighttime peaks are the most informative (107). Many of the pitfalls highlighted above for provocative testing also apply to spontaneous testing. Spontaneous measurements are expensive and requires hospitalization and therefore the use have been limited to only a few centers worldwide.

Divergent results between provocative testing and nocturnal spontaneous GH tests have shown to be highly variable from 6-42% depending on cut-off levels (3-10 µg/L) (108).

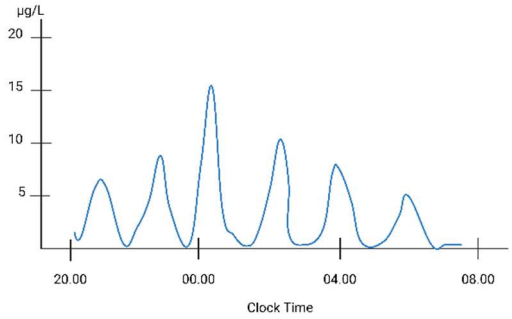

Figure 6. Normal growth hormone spontaneous night profile with several peaks. Created with BioRender.com.

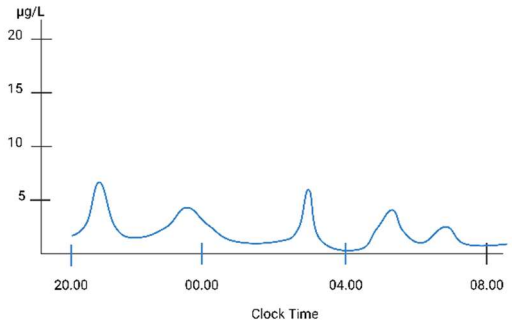

Figure 7. Pathologic growth hormone spontaneous night profile with stunted low peaks. Created with BioRender.com.

1.5 GROWTH HORMONE TREATMENT

In Sweden, as well as other countries in the European Union, the approved indications for GH therapy in children by EMA, European Medicines Agency, currently include GHD, SGA without catch-up growth, Turner syndrome, renal insufficiency, growth disorders caused by SHOX gene mutations, Prader-Willi syndrome, and Noonan syndrome.

The Food and Drug Administration (FDA) approved in 2003 GH therapy for children with ISS in the U.S. In Sweden ISS alone is not an approved indication for GH therapy.

The recommended GH dose for children with GHD range from 0.025-0.035 mg/kg/day. International guidelines recommend dosing by body weight based on previous positive effects on growth seen in multiple studies (78). Dosing based on body surface area (BSA) is not recommended, as it may lead to lower total GH doses in older children compared to weight-based dosing (109). This recommendation is further supported by the lack of rigorous studies directly comparing the two approaches and insufficient evidence favoring the superiority of BSA-based dosing.

Children undergoing GH therapy represent a diverse range of underlying causes and varying degrees of sensitivity to treatment. Despite receiving the same GH dose, they often exhibit a wide spectrum of responsiveness, highlighting the need for individualized dosing strategies (110-112).

1.5.1 IGF-1 DOSE TITRATION

Given that GH levels regulate the production and release of IGF-1 in the liver, a potential approach to optimize GH therapy is to use IGF-1 as a biomarker for dosing. For other hormonal deficiencies, such as hypothyroidism, treatment aims to titrate the dose to achieve hormone levels comparable to those in a healthy population. Applying a similar strategy for GH therapy, with dose titration based on IGF-1 levels corresponding to a healthy cohort, would be a reasonable approach. By titrating GH doses based on IGF-1 levels, treatment could be tailored to the individual, improving efficacy while minimizing the risks associated with elevated IGF-1 levels.

Cohen and colleagues conducted several studies investigating IGF-1 based dosing of GH therapy in short children (113-115). A randomized controlled trial compared IGF-1-based dosing with conventional weight-based dosing for children with GHD. The study found that adjusting GH doses to achieve higher IGF-1 levels led to improved growth responses, although this approach required higher average GH doses (113). A subsequent study suggested potential dose-reducing and safety-improving effects of IGF-1-based dosing in children with GHD and ISS (114).

Jensen et al. demonstrated, however, in the North European Small-for-Gestational-Age Study that titrating GH doses based on IGF-1 levels in children born SGA was less effective than conventional dosing strategies, resulting in reduced height gain (116).

The GH dose effect on IGF-1 levels is complex. Studies have shown that during the first years of GH therapy, IGF-1 levels increase depending on GH dose (117, 118). Cohen et al. showed that GH therapy induces a positive correlation between IGF-1 levels and a two-year adjusted change in height SDS. The study population indicated, however, a plateau in the GH dose-response curve at an intermediate dose of approximately 50 μg/kg per day; higher doses providing no additional growth benefit but with elevated IGF-1 levels (119). Lundberg et al. demonstrated that a higher GH dose resulted in a greater increase in IGF-1 levels (delta IGF-1) during both prepubertal and pubertal growth, ultimately leading to a larger overall height gain (120). The so called LG growth study, a long-term, observational cohort study conducted in South Korea, reports that IGF-1 changes were only weakly associated with treatment effects during the first years of treatment, and were no longer significant from the third year (121). Oberle et al. found no significant

association between IGF-1 z-scores and GH dose in their cohort of prepubertal children receiving GH therapy (122).

1.5.2 GROWTH HORMONE AND SAFETY

The Safety and Appropriateness of Growth Hormone Treatments in Europe (SAGhE) study, a retrospective investigation into the long-term safety of GH therapy, studied individuals from eight European countries that received GH therapy during childhood, focusing on cancer incidence and mortality due to specific causes. Initial findings from the French cohort indicated a heightened risk of all-cause mortality and stroke in children receiving GH therapy (123, 124). Although the SAGhE study raised potential concerns, its findings should be viewed cautiously due to methodological limitations and the absence of supporting evidence from other cohorts. The latest comprehensive analysis, combining data from all cohorts, reported, conversely, no increase in all-cause mortality (125).

1.6 PREDICTION MODELS

A prediction model is a mathematical or computational tool designed to forecast future outcomes or trends based on existing data and predefined variables.

Based on the complexity in diagnosing GHD, identifying a child with severe GHD may be obvious due to several signs, but the intermediate group with partial GHD or ISS can be challenging. Prediction models, which use the characteristics of the child in contribution with predictors to forecast future height gain can assist in diagnosing and may help in various ways:

1. Identifying patients - healthcare resources can be allocated more efficiently to children most likely to benefit from GH therapy and minimize the risk of unnecessary treatment,
2. Personalizing treatment - GH dosing can be tailored based on estimated responsiveness, using baseline auxological and biochemical factors to determine the appropriate dose, considering the unique characteristics of children,
3. Treatment monitoring - assist in tracking the progress of GH therapy and evaluating adherence to treatment.

A clinically relevant prediction model should include the following according to Wit et al. (126):

1. **Validation**: model should be tested on an independent cohort different from the one used to develop it,
2. **Accuracy**: model should demonstrate minimal prediction error,
3. **Comprehensiveness**: model should account for wide variability in treatment response,
4. **Practicality**: model should rely on easily accessible and standardized variables,
5. **Integration**: model should incorporate treatment modalities as variables,
6. **Biological Relevance**: model should be grounded in biological principles,
7. **Usability**: model should be straightforward to apply in clinical practice.

1.6.1 KIGS PREDICTION MODELS

KIGS (Pfizer International Growth Database) prediction models for GH therapy is a comprehensive tool used to predict and analyze the response of children to GH therapy (127). An extensive global database of clinical data from children undergoing GH therapy is used as a basis. The prediction for first-year growth response was based on data from 593 GHD patients (148 girls and 455 boys), with fewer data points available for subsequent years. Using this dataset, the models were developed through multiple linear regression analysis and can generate growth predictions for up to four years. The prediction variables in the models include: 1) Age at onset, 2) Birth weight, 3) GH dose, 4) Parental heights, 5) Body weight, 6) Height, and 7) Maximum GH peak. The best model explained 61% of the variability of the response and the most important variable was the maximum GH peak. The more severe GHD, the better first year growth prediction. Furthermore, there was an inverse relationship between the growth response and chronological age, in addition to the deficit between child's growth and target height. Younger and smaller children exhibited, consequently, a more substantial growth response to GH therapy in the first year.

A separate model, excluding the maximum GH peak, explained 45% of the variance in treatment response. The most significant predictor in this model

was the height deficit relative to target height. Both models produced comparable results overall; the model without the maximum GH peak tended, however, to underestimate growth response in individuals with severely reduced GH secretory capacity (127).

Other prediction models for children with Turner syndrome, SGA, ISS and chronic renal insufficiency are also available from data included in KIGS (127-130).

The KIGS prediction models are integrated into a user-friendly cloud-based tool called Individualized Growth Response Optimization (iGRO) (131). First-year growth serves as the most significant predictor for subsequent prepubertal growth velocity. Height velocity becomes, however, more complex after the onset of puberty, making annualized height velocity challenging to assess. Predictive algorithms have been developed to address this by modelling total pubertal growth, from puberty onset to near-adult height, for conditions such as GHD, Turner syndrome, and SGA.

1.6.2 GOTHENBURG PREDICTION MODELS

The Gothenburg prediction models focus on a broad group of prepubertal short children for whom the diagnosis of GHD is not clearly defined. In 2000, Albertsson-Wikland et al. published several prediction models incorporating different variables. These models were developed using data from 269 short prepubertal children (45 girls and 224 boys) and validated with a separate cohort of 149 children (132). The modeling technique involved nonlinear data fitting and empirical testing; chosen because the relationship between growth response and the variables was inherently nonlinear.

The model with the best prediction accuracy was the 24-h GH profile data added to the Basic + early growth model including data of: 1) Parental heights, 2) Birth weight and length, 3) Weight and height at 1 year, 2 years, 1 year prior to prediction and at prediction, and 4) maximum spontaneous GH peak. This model could predict the first-year growth on GH therapy with an accuracy of ± 0.37 SDS (± 1.96 standard deviation of residuals (SD_{res})) using a fixed standard dose of 33µg/kg/day.

The spontaneous maximum GH peak over 24 hours was found to be the most informative variable. The prediction model including early growth variables and GHmax24h has been used in prospective studies on personalized GH

regimens to estimate responsiveness (111). Later models with broader inclusion criteria as preterm children and SGA were also developed (133).

1.6.3 MACHINE LEARNING PREDICTION MODELS

Machine learning with origins in statistics is a powerful tool for developing artificial intelligent (AI) applications that can uncover patterns and relationships within data (134). These insights can be used to enhance understanding of a phenomenon or to predict future outcomes.

The two primary categories of machine learning scenarios are unsupervised learning and supervised learning. Unsupervised learning involves training a model on data without labeled outcomes or explicit guidance, aiming to reveal hidden structures or pattern in the data. Supervised learning, in contrast, uses training data consisting of input observations and their corresponding known outputs to create a model that can predict outputs for new datasets. The goal is to develop a model that produces predicted values closely matching the observed ones in the training set, avoiding underfitting and associated high bias, typically a result from an overly simple model with insufficiently informative inputs. The model must generalize well to new data to prevent overfitting, which occurs when the model is overly complex or relies on too many features relative to a small training dataset. The key to building an effective model lies in striking a balance between underfitting and overfitting (135).

Complex deep learning models like Neural Networks mimics the human brain by using multiple layers of interconnected nodes to learn complex patterns (136), while Random Forests combine numerous decision trees to produce highly accurate and reliable predictions (137). The decision-making processes of these algorithms are often highly complex or hidden, making them difficult to interpret, and are commonly referred to as black-box models. White-box models are based on internal decision-making processes that can be traced and easily understood (138). Black-box models often outperform in predictive power, whereas white-box models prioritize transparency and interpretability, sometimes at the cost of predictive accuracy. This trade-off is, however, not always apparent when the data is well-structured and the features are highly informative (139). Use of interpretable models in healthcare is of crucial significance. Despite a growing reliance on black-box models, particularly in fields like radiology and automation, their trustworthiness depends heavily on the quality and reliability of the training data.

1.6.3.1 MACHINE LEARNING LINEAR REGRESSION

Linear regression is one of the most widely used machine learning techniques and the simplest of the models presented here (140). The relationship between a dependent variable (Y) and one or more independent variables (X), is represented by approximating them with a linear equation.

The key difference between traditional regression and machine learning-based regression lies in their goals and approaches. Traditional regression focuses on understanding the relationship between dependent and independent variables, emphasizing interpretability and hypothesis testing. In contrast, machine learning-based regression prioritizes prediction accuracy, using methods like cross-validation to avoid overfitting. Machine learning models are more flexible, allowing for the use of larger datasets and hyperparameter tuning to enhance performance, though often at the expense of interpretability (141).

Simple linear regression can be enhanced using regularization techniques. Ridge regression reduces the size of the coefficients to prevent overfitting, while Lasso regression selects a smaller set of predictors to minimize prediction errors for the response variable. Both methods simplify the model and help prevent overfitting.

1.6.3.2 SYMBOLIC REGRESSION

Symbolic regression is a form of genetic programming that begins by generating random mathematical expressions to fit the data (142). The best-fitting expressions are then selected, combined, and modified through random changes to produce better equations that minimize errors. This iterative process continues over multiple generations to refine the models further. This approach does not require a predefined functional form, allowing it to discover new and complex equations that capture underlying data patterns. Unlike classic regression, symbolic regression can automatically select features from raw input variables and combine them into new features that capture underlying patterns in the data (143). The resulting model is a mathematical expression that is often interpretable, providing valuable insights into the relationships between variables (144).

1.6.3.3 EXPLAINABLE BOOSTING MACHINE

Explainable Boosting Machine (EBM) is a tree-based, cyclic gradient-boosting model that builds on Generalized Additive Models (GAMs) with pairwise interactions. GAMs are a type of statistical model that extends linear regression by allowing for greater flexibility in the relationship between the dependent variable and each independent variable. Unlike traditional linear models, which assume a linear relationship, GAMs enable the relationship between the

target and each predictor to be modeled as a smooth, non-linear function. This flexibility allows GAMs to capture more intricate patterns in the data without sacrificing interpretability. GAM helps us understand how different features affect the outcome by looking at the effect of each feature separately and then adding them together in a prediction (145).

Tree-based means the model uses decision trees, which are non-linear predictive models capable of capturing complex relationships between variables. Cyclic gradient boosting refers to the iterative process where, instead of building new decision trees for the entire dataset at each step, the model focuses on individual features in a cyclic manner. Each subsequent tree corrects the errors of the previous iteration, refining the contribution of specific features to build a more accurate and interpretable model (146).

Within EBM the most important features can be visualized with the implementation of the InterpretML package and a global features plot, where the most important features affecting the prediction is ranked. The higher importance score, the stronger impact the feature has on the model's prediction, see Figure 8.

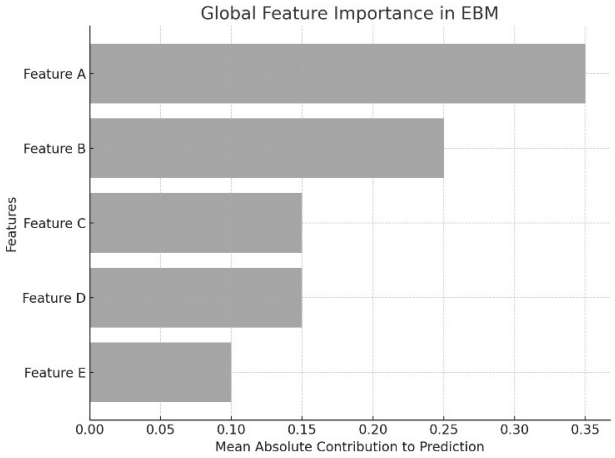

Figure 8. Global feature importance plot for an Explainable Boosting Machine (EBM). The horizontal bars show the average impact of each feature on the model's predictions, with larger values reflecting greater influence. Created with *OpenAI. (2024). ChatGPT. URL: https://openai.com/*

Feature importance plots illustrate the relationship between each single feature and the predicted outcome when all other variables are held on average. See Figure 9. The appearance of the plot can reveal whether a feature has a linear or non-linear relationship with the target variable. The top graph illustrates the relationship between the feature values and their contribution to the prediction, with the blue shaded area representing uncertainty. The bottom graph shows the distribution of the feature values as a histogram, providing context for the range and frequency of the feature within the dataset.

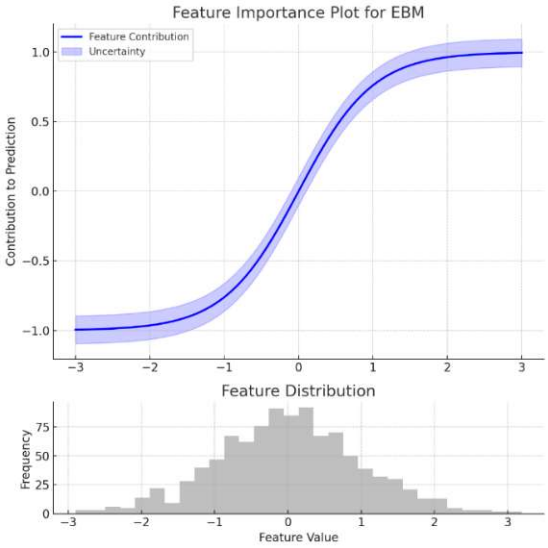

Figure 9. Feature importance plot for an Explainable Boosting Machine (EBM). Created with *OpenAI. (2024). ChatGPT. URL: https://openai.com/*

1.6.4 COMPARISON OF PREDICTION MODELS

Quality of a prediction model can be expressed in the amount of variance where the coefficient of determination (R^2) close to 1 is ideal. A high R^2 indicates a strong fit between the model and the observed data; but does not imply that the model generalizes well to new data.

Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) measure the accuracy of the models by quantifying the error. MAE measures the average error between the predicted and observed values expressed in the same units. Small errors implicate better accuracy for the model. MAPE measures the average percentage error and highlights the proportional accuracy and allows comparison across datasets with different scales or units.

Another way to express the accuracy of prediction models as used for the Gothenburg prediction models is the SD_{res}, measuring the average distance between observed data points and the predicted values from a model. A value close to zero indicates a better prediction.

2 AIMS AND HYPOTHESES

The overall aim of this thesis is to optimize GH therapy for short children in a clinical setting, focusing on identifying patients who would benefit from treatment and to improve strategies for individualized care.

The following hypotheses is the basis of this thesis:

- Diagnosing GHD in children, particularly milder cases, is challenging, often leading to overtreatment. Unnecessary treatment poses potential risks, including side effects and undue burden on the child.
- Use of IGF-1 as a biomarker for GH dose titration has the potential to improve treatment accuracy by reducing the risk of overtreatment and its associated adverse effects. Significant pitfalls remain in its clinical application.
- Sophisticated AI machine learning models provide an accurate prediction of IGF compared to traditional statistical approaches. Important features affecting the prediction can also be identified.

2.1 SPECIFIC AIMS

Study 1

- Assess the clinical accuracy of the Gothenburg prediction model in real-world settings,
- Identify children who will benefit from GH therapy in terms of growth outcomes, minimizing unnecessary interventions.

Study 2

- Validate and compare the accuracy of the Gothenburg prediction model and the KIGS models in predicting growth outcomes during GH therapy using clinical data,
- Enhance decision-making prior to initiating GH therapy in children by improving access to validated prediction models.

Study 3

- Evaluate a well-defined group of children in early puberty with short stature undergoing GH therapy, focusing on how pubertal maturation, and sex steroid levels influence the interpretation of IGF-1 levels,
- Describe the characteristics of children with high or low serum IGF-1 SDS.

Study 4

- Compare the performance of three AI machine learning models (Linear Regression, Symbolic Regression, and Explainable Boosting Machine) to demonstrate the potential advantages of more advanced machine learning techniques in facilitating the clinical use of IGF-1,
- Identify the most important factors influencing IGF-1 prediction in both short and long term.

3 PATIENTS AND METHODS

3.1 STUDY SUBJECTS

Study 1

All short children who started GH therapy based on a growth prediction at the Queen Silvia Children's Hospital in Gothenburg between 2004 and 2016 were considered for inclusion in the study. The cohort comprised 121 children, 44 girls and 77 boys, all of whom had received at least one year of GH treatment initiated based on a favourable first-year growth prediction of ≥ 0.7 SDS. Exclusion criteria included chronic diseases, endocrinopathies other than GHD, oncological conditions, poor adherence, or evident dysmorphic features suggestive of an underlying syndrome. Pubertal children with breast Tanner stage > 1 for girls and testicular volume ≥ 4 ml for boys were also excluded.

Study 2

The inclusion and exclusion criteria were the same as those in study 1, except that a first-year prediction of ≥ 0.7 SDS was not required, which resulted in the inclusion of 5 additional patients. However, three patients initially included in study 1 were excluded from study 2 due to newly discovered suspicions of a syndrome, rheumatologic disease, and one case of late-identified low treatment adherence. This resulted in a cohort of 123 prepubertal children, consisting of 47 girls and 76 boys.

Study 3 and 4

Study 3 and Study 4 included participants from a prior clinical dose-response trial conducted in Sweden, which consisted of two parts: the Catch-Up Study (111) and the subsequent Maintenance Study (147). Patients from the Catch-Up Study who met the eligibility criteria for the second phase were included in the second part.

The Catch-up Study included prepubertal children, with a height SDS of ≤ -2 or a growth velocity SDS of ≤ -1. Additionally, eligible children needed to be ≤1 SDS below their mid-parental height (MPH) SDS, born with a weight or length above -2.5 SDS, and born at a gestational age of ≥ 30 weeks. Accessible growth data was required, including measurements of length/height and weight at birth, at 1 year (± 3 months) and 2 years (± 6 months) of age, as well as 1

year (± 3 months) before starting treatment. At least two stadiometer measurements during the pretreatment year were also necessary.

Exclusion criteria included chronic illness, clinical syndromes, disproportionate body measurements, catch-up growth during the pretreatment year, or an MPH SDS > 1.5. Diagnostic classification was determined using the arginine-insulin tolerance test (AITT) but 24-hour spontaneous profiles were also measured. Children were randomized into standard or individualized GH dose groups in a 1:2 ratio. 128 patients (38 girls) completed the protocol for the first Catch-up phase.

In the following Maintenance Study, patients were selected from the individualized-dose group of the previous study based on specific eligibility criteria. To participate, children had to remain prepubertal and had obtained a height at least 1.5 SDS from MPH. For the weight-based dose control group, no auxological criteria were applied. A total of 98 children (72 boys, 26 girls) formed the intention-to-treat (ITT) population, which included 33 children with GHD and 65 without GHD. GHD diagnosis in this study was based on both the arginine-insulin tolerance test (AITT) and the 24-hour spontaneous profile. Participants were randomized into three groups receiving half their previous individualized GH dose, continuation of their unchanged individualized dose or receiving a standard GH dose of 43 µg/kg/day.

Study 3 includes patients from the Maintenance Study, excluding five boys due to incomplete data. Study 4 focuses exclusively on boys from the same trial, excluding seven due to incomplete data. See Figure 10.

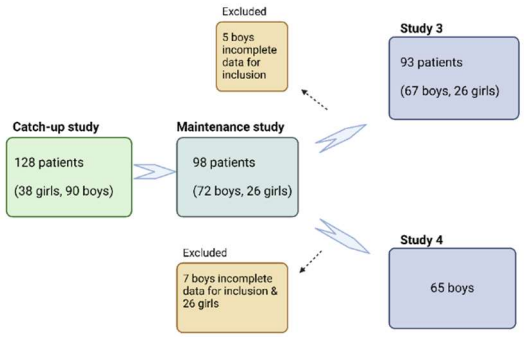

Figure 10. Patient cohorts for study 3 and 4 from previous clinical trials. Created with BioRender.com.

3.2 AUXOLOGICAL MEASUREMENTS

All height measurements at standing position at the Pediatric endocrine clinic at Queen Silvias Children's hospital were conducted using an Ulmer stadiometer that is wall-mounted. This device, equipped with an incremental scale and an optoelectronic measuring system, provides a precision of 1 mm. Body weight was recorded with an accuracy of 0.1 kg using a step scale (SECA 701, Germany).

Early growth data for the children were collected from child welfare centers, where various measurement methods were used. Birth weight and birth length were adjusted for gestational age and converted into SDS according to Niklasson et al. (148). The one- and two-year height outcomes from the Gothenburg prediction model are expressed in height SDS according to Karlberg et al. (149). Therefore, all height measurements in Studies 1 and 2 were converted accordingly to ensure comparability.

3.3 HORMONAL ANALYSES

3.3.1 IGF-1 AND IGFBP-3 ANALYSES

A radioimmunoassay (RIA) (Mediagnost GmbH, Tübingen, Germany), a technique that quantifies substances using specific antibodies and radioactively labelled antigens, was employed to measure IGF-1 and IGFBP-3. This method provided a total coefficient of variation (CV) of 16% at low IGF-1 concentrations (40 µg/L) and 9% at higher concentrations (≥ 225 µg/L). For IGFBP-3, the CV remained constant at 10% across concentrations. Additionally, the imprecision for the IGF-1/IGFBP-3 ratio was 17% at low values (2.1 µg/L) and 15% at higher values (≥ 5.0 µg/L).

From March 2018, the analysis method was updated to a chemiluminescent immunoassay (IDS iSYS, Immunodiagnostic Systems Holdings, Tyne and Wear, United Kingdom). This method, unlike RIA, utilizes light-emitting chemicals rather than radioactive isotopes for detection, being safer and more environmentally friendly. A high degree of agreement with the previous RIA method was demonstrated, correlation coefficient, $r = 0.98$, negating the need for a conversion factor between methods.

3.3.2 GROWTH HORMONE ANALYSES

Serum GH concentrations were measured using a fluoroimmunometric assay with monoclonal antibodies, following the WHO IRP 80/505 Standard AutoDELFIA hGH method (PerkinElmer Life and Analytical Sciences, Turku, Finland). This method is an immunoassay that uses fluorescent labels for detection. On June 16, 2014, this method was replaced by a chemiluminescent immunoassay (IDS-iSYS, Immunodiagnostic Systems Holdings, Tyne and Wear, United Kingdom). The AutoDELFIA method reported GH levels in mU/L, while the IDS-iSYS expressed them in µg/L. To ensure consistency between methods, the transformation formula IDS-iSYS = AutoDELFIA / 2.6 was applied.

3.3.3 SEX STEROID ANALYSES

Serum estradiol and testosterone levels were measured simultaneously using high-sensitivity gas chromatography-tandem mass spectrometry (GC-MS/MS) (150). The CV for estradiol was 19% at 8 pmol/L and 6% at concentrations ≥ 36 pmol/L. For testosterone, the CV was 16% at 0.3 nmol/L and < 10% at levels > 1.5 nmol/L. Previously established biological reference ranges (151) for estradiol and testosterone in children were adopted due to the high agreement between the GC-MS/MS method and earlier methods, such as extraction-based RIAs.

MS/MS is an advanced analytical technique commonly used for the quantification of substances such as hormones. The technique involves isolating ions of interest and generating a mass spectrum containing only those specific ions. The process operates in two steps: first, the sample is ionized and sorted by their size and weight in the first mass analyzer; second, the selected ions are fragmented into smaller pieces and analyzed in the second mass analyzer to produce a unique fragmentation pattern, known as the *mass spectrum*. When combined with gas chromatography, the accuracy is further enhanced, as a separation step is included to isolate individual compounds before analysis. This method is particularly effective for measuring low-abundance analytes, such as hormones, with high precision. The specific GC–MS/MS method developed by Ankarberg-Lindgren et al., with its high sensitivity and specificity, has shown to be well-suited for measuring androgens and estrogens (150).

3.4 STATISTICAL AND MACHINE LEARNING METHODS

A p-value of less than 0.05 was regarded as statistically significant for all studies in this thesis. When comparing observed versus predicted growth in Study 1, a modified version of Bland-Altman plot was used, including the 95% upper and lower limits of agreement, representing ± 2 SD from the mean. Correlations in Studies 1 and 2 were analyzed using Pearson correlation, which calculates a coefficient to express the strength and direction of the relationship. Additionally, Study 2 used studentized residual plots to identify outliers and evaluate the accuracy of the prediction models used in the study.

Study 3 employed linear mixed-effects models to compare sample groups, accounting for measurements that were repeatedly taken from the same subjects.

Machine learning models were used for study 4. To avoid overfitting, a hold-out validation approach was used, reserving 30% of the data as a separate subset to evaluate the performance of models. To optimize the hyperparameters, which are predefined settings that control the learning process for each model type, grid search with cross-validation was used on the training data. Once the optimal hyperparameters were determined, the final model was trained on the entire training dataset using the selected hyperparameters.

For the linear regression analyses, the scikit-learn implementation, an open-source library in Python was used. For symbolic regression, the Gene-Pool Optimal Mixing Evolutionary Algorithm (GP-GOMEA), developed by Virgolin et al. (152), implemented through the GPG Python package for symbolic regression was used. The algorithm is available at https://github.com/marcovirgolin/gpg. For the Explainable Boosting Machine, the implementation provided by the InterpretML package was used (146).

3.5 ETHICAL CONSIDERATIONS

The studies included in this thesis were conducted in accordance with the Declaration of Helsinki and received approval from the Research Ethics Committee in Gothenburg, Sweden.

Studies 1 and 2: Approval number 171-18.

Studies 3 and 4: Approval number 320-03, 449-16 and approval for the trial by the Medical Product Agency of Sweden with study number NRA 6280003.

The first two studies were retrospective, with data collected from databases and medical records. Their retrospective nature and use of de-identified data did not require informed consent, as specified by the approval from the ethics committee.

Data for Studies 3 and 4 were collected prospectively; however, the data analysis plan for the pubertal years was retrospective in nature. For these studies informed consent was obtained from the parents, and assent was gathered from the child when age appropriate.

The cohort for Study 3 and 4 was originally part of a clinical trial in which examinations and follow-ups were conducted every three months. Blood samples were collected and stored at -80 degrees Celsius for future use. Following updated ethical approval (449-16), selected samples were analyzed for estradiol, testosterone, IGF-1 and IGFBP-3. These additional analyses had no adverse impact on the patients.

4 RESULTS

4.1 STUDY 1

The primary objective of the first study was to evaluate the accuracy of the Gothenburg prediction model in identifying children who would benefit from GH therapy in a clinical setting. The cohort included 121 patients with a median age of 5.32 years (range: 3.0–11.8 years). After one and two years of GH therapy, the mean height gains were 1.00 ± 0.32 SDS and 1.50 ± 0.43 SDS, respectively.

Of the 121 patients, 13 did not achieve the expected growth of ≥ 0.7 SDS after one year of treatment. Among these, two patients were classified as poor responders, having grown less than 0.5 SDS.

The correlation between first-year predicted and observed growth in SDS revealed a regression line with an R^2 of 0.17, indicating an underestimation of low predictions and an overestimation of high predictions. The residual deviation, measured as the standard deviation of the differences between observed and predicted growth, was 0.31 SDS. See Figure 11.

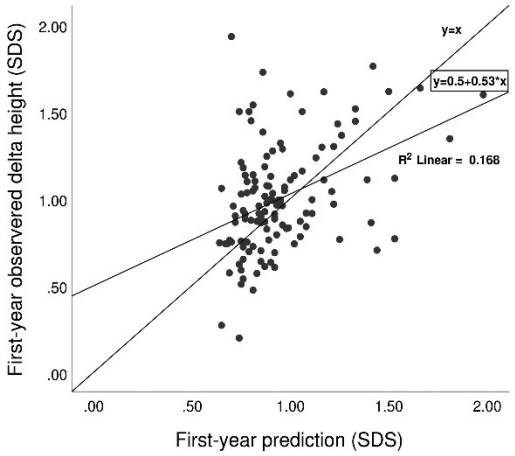

Figure 11. The correlation between the first-year predicted and observed growth is shown. Reproduced with permission from Ly et al., Acta Paediatrica, 2022 Feb;111(2):346-353.

The predicted and observed height outcomes for the first and second years of GH therapy were compared, as shown in Figure 12. For the first year, the 95% limits were 0.65 SDS and -0.57 SDS, with a standard deviation of 0.31. For the two-year prediction, the limits were 0.93 SDS and -0.80 SDS, with a standard deviation of 0.44.

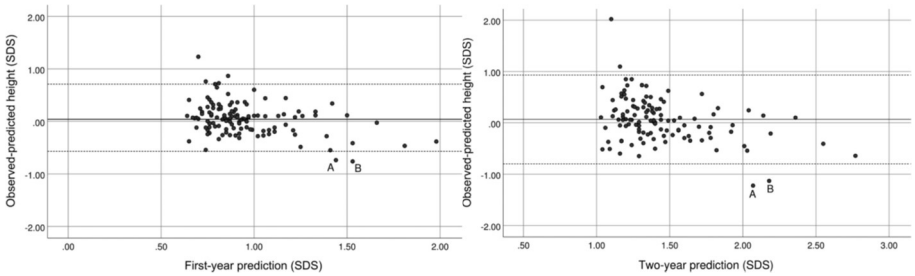

Figure 12. Graphs illustrating the discrepancy between observed and predicted first-year (left) and two-year (right) growth responses relative to the predicted values. The dashed lines indicate the 95% agreement limits, with patients A and B showing considerably less growth than anticipated. Reproduced with permission from Ly et al., Acta Paediatrica, 2022 Feb;111(2):346-353.

4.2 STUDY 2

Study 2 compared the accuracy of two prediction models by comparing the growth outcome after one year on GH therapy.

The growth outcome was assessed by comparing the actual height after one year of treatment, expressed as mean (SD); 113.0 cm (10.3), with the predicted heights from the Gothenburg model showing 112.9 cm (10.5), and the KIGS models, 112.7 cm (9.9).

Figures 13 and 14 demonstrate the correlation between predicted one year growth of GH therapy versus prediction based on the Gothenburg model and KIGS models respectively.

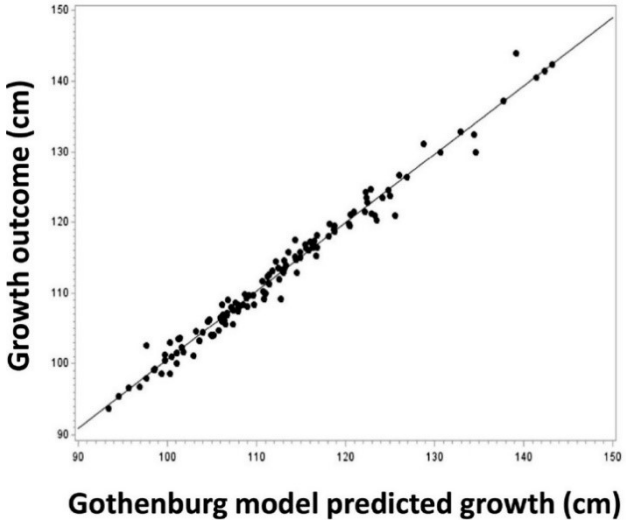

Figure 13. Observed growth outcomes compared to predicted growth by the Gothenburg model after one year of GH therapy. Pearson correlation coefficient r = *0.990,* p ≤ *0.0001. Reproduced with permission from Ly et al., Growth Horm IGF Res. 2023 Feb:68:101523.*

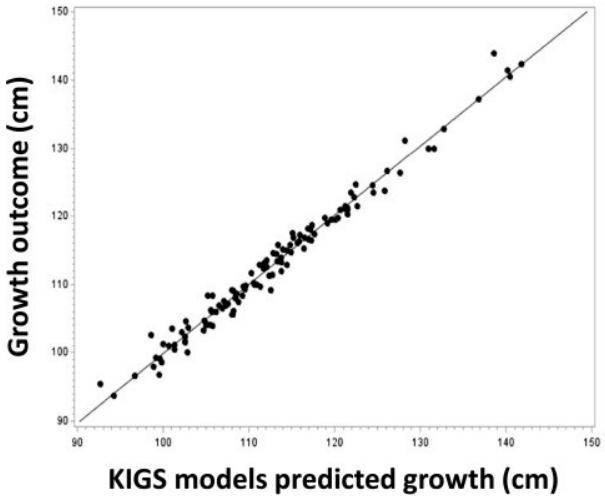

Figure 14. Observed growth outcomes compared to predicted growth by the KIGS models after one year of GH therapy. Pearson correlation coefficient r = *0.991,* p ≤ *0.001. Reproduced with permission from Ly et al., Growth Horm IGF Res. 2023 Feb:68:101523.*

Figure 15 displays the mean and (SD) studentized residuals for the predicted growth response according to the Gothenburg model, which was found to be 0.10 (0.81). Similarly Figure 16 presents the corresponding values for the KIGS models, which was 0.03 (0.96).

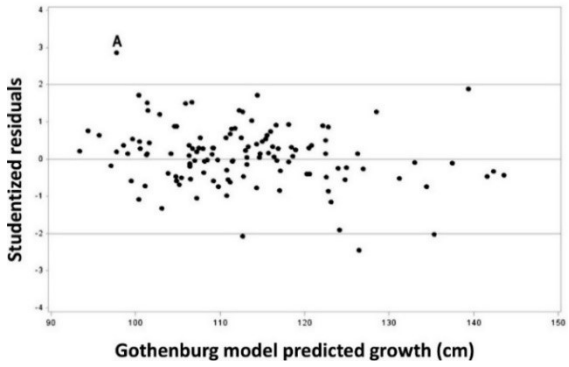

Figure 15. Studentized residuals for each patient, based on the first-year predictions from the Gothenburg model, are displayed. Patient A's growth was underestimated, resulting in a high studentized residual. Reproduced with permission from Ly et al., Growth Horm IGF Res. 2023 Feb:68:101523.

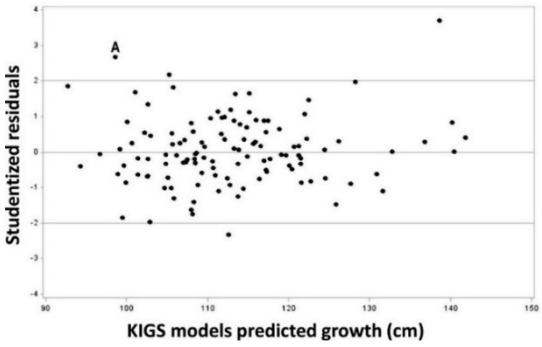

Figure 16. Studentized residuals for each patient are presented relative to the first-year predictions using the KIGS models. Similarly, patient A exhibits a high studentized residual, indicating underestimated growth. Reproduced with permission from Ly et al., Growth Horm IGF Res. 2023 Feb:68:101523.

When comparing both models, a high correlation of 0.995 with p≤0.0001 was found. See Figure 17.

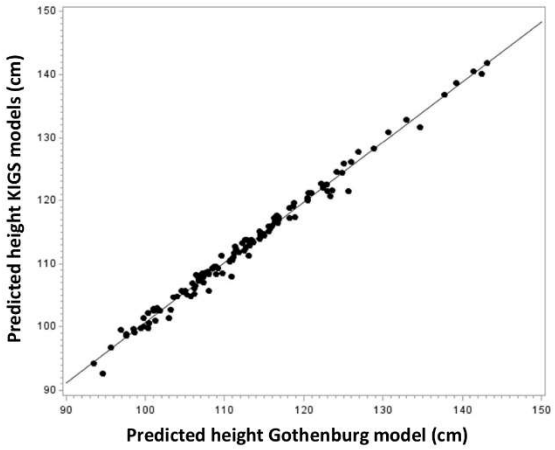

Figure 17. The correlation between predictions made by the KIGS models and the Gothenburg model is shown. Pearsson correlation coefficient r = 0.995, p ≤ 0.0001. Reproduced with permission from Ly et al., Growth Horm IGF Res. 2023 Feb:68:101523.

4.3 STUDY 3

Study 3 is a descriptive analysis addressing the clinical challenges of interpreting IGF-1 levels during early puberty. Pubertal thresholds for estradiol and testosterone were set at ≥ 25 pmol/L and ≥ 0.47 nmol/L, respectively.

Among 58 samples from girls at Tanner breast stage 1, 15.5% (9/58) exceeded this threshold. Adjusting IGF-1 SDS using the early puberty reference range instead of the prepubertal range lowered the median IGF-1 SDS from 2.04 to 1.44, $p < 0.05$, with 4 out of 9 values falling below the guideline-recommended upper limit of 2 SDS (Figure 18).

For boys with testes <4 mL, 15.7% (24/153) had testosterone levels above the pubertal threshold of ≥0.47 nmol/L. Using the early puberty reference range

reduced the median IGF-1 SDS from 1.96 to 1.29, p=0.02, and 5 out of 24 values dropped below the 2 SDS upper limit (Figure 19).

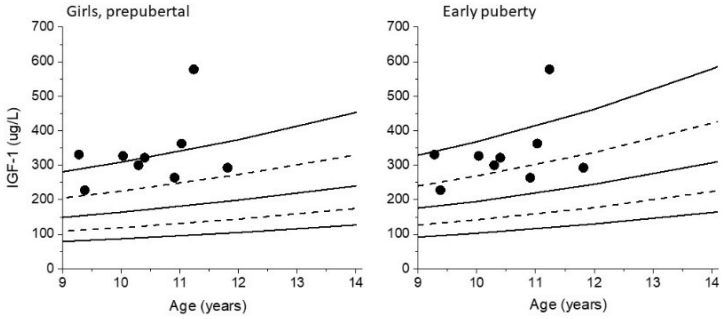

Figure 18. IGF-1 levels for girls at Tanner breast stage 1 with pubertal estradiol concentrations are shown in comparison to the prepubertal reference range for girls (left). The same IGF-1 levels are also compared to the early puberty reference range (right). Dashed lines +/- 1 SDS, solid lines mean and +/- 2 SDS. Reproduced with permission from Ly et al. Front Endocrinol,15:1514935.

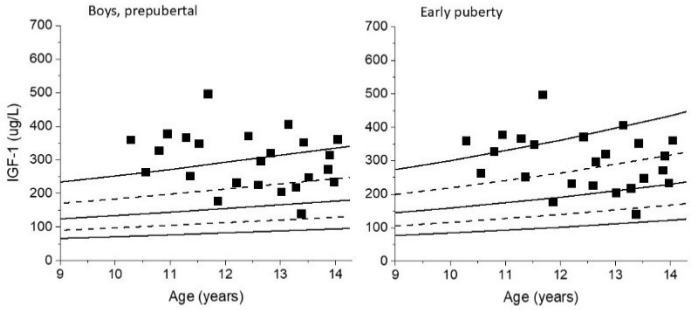

Figure 19. IGF-1 levels for prepubertal boys with pubertal testosterone levels are displayed relative to the prepubertal reference range for boys (left) and to the early puberty reference range (right). Dashed lines +/- 1 SDS, solid lines mean and +/- 2 SDS. Reproduced with permission from Ly et al. Front Endocrinol,15:1514935.

When examining the differences between the groups with IGF-1 SDS above and below 2 SDS, the following observations were made, see Table 1. The dataset included 475 samples of IGF-1, regardless of pubertal staging, of which 166 (35%) had levels ≥ 2 SDS. Comparing mean GH dose in mg/kg/day, BMI in kg/m^2, estradiol in girls and testosterone in boys showed significant differences between the groups of higher or lower IGF-1 SDS. The group with high IGF-1 SDS ≥ 2 had higher mean GH dose, lower BMI kg/m^2 and much lower estradiol and testosterone.

Table 1. Differences in low and high IGF-1 SDS groups.

Variables	Units		IGF-1 <2 SDS		IGF-1 ≥2 SDS	P-value
		n=	Median (range)	n=	Median (range)	
Mean growth hormone dose	mg/kg/day	283	0.038 (0.01-0.10)	140	0.042 (0.02-0.10)	p<0.001
Mean growth hormone dose	mg/m^2	283	1.15 (0.42-3.22)	140	1.24 (0.49-2.94)	p=0.08
BMI	kg/m^2	319	17.9 (13.0-27.7)	153	17.5 (13.5-26.1)	p<0.001
BMI	SDS	318	0.10 (-2.65-2.26)	151	-0.18 (-4.16-2.49)	p=0.23
Estradiol (girls)	pmol/L	83	102 (1-1070)	42	13 (3-214)	p<0.001
Testosterone (boys)	nmol/L	112	6.9 (0.04-31.2)	236	0.35 (0.11-27.2)	p<0.001

4.4 STUDY 4

The machine learning models demonstrated varying performance in predicting IGF-1 SDS at 3 and 12 months. The more advanced models; symbolic regression showed R^2 values of 0.47 and 0.51 for 3- and 12-month predictions, respectively, with minimal predictors, while EBM achieved R^2 values of 0.47 and 0.42. Mean absolute error (MAE) and Mean absolute percentage error (MAPE) were calculated for each model, and the number of predictors required varied between methods. Full details are provided in Table 2.

Table 2. The Performance of each model for the 3-month and 12-month IGF-1 SDS predictions.

		Linear Regression	Symbolic Regression	Explainable Boosting Machine
3-Month	R^2	0.07	0.47	0.47
	MAE	0.78	0.55	0.55
	MAPE	0.68	0.47	0.32
	Predictors (n)	33	7	63
12-Month	R^2	0.22	0.51	0.42
	MAE	0.67	0.54	0.55
	MAPE	1.16	1.37	1.22
	Predictors (n)	33	6	40

The 3-month prediction made by symbolic regression used 7 variables; Birth length, IGF-1 SDS, height velocity one year before treatment, weight, missing injections past year, delta IGF-1 at 1 year on treatment, and change in GH dose past 3 months. For the 12-month prediction the following 6 key features were found; Birth length, IGF-1 SDS, change in weight last 12 months, age, IGF-1/IGFBP-3 ratio, and decrease in testosterone.

Table 3 lists the most important features influencing IGF-1 SDS predictions at 3 and 12 months, as determined by the EBM model, ranked by their level of importance.

Table 3. Key features identified by Explainable Boosting Machine for the 3-month and 12-month predictions.

3-month predictions	12-month predictions
IGF-1 SDS	IGF-1 SDS
Height velocity one year before treatment	Age
GH dose	Target height difference
Target height difference	Birth length
IGF-1/IGFBP3	Testosterone
IGF-1 increase	Testicle size
Age	Height velocity one year before treatment
Growth second year of life	Change in IGF-1 first 3 months on treatment
IGF-1 SDS & GH dose	GH dose
Change in weight past 12 months	Estradiol
Age at treatment start	Change in target height difference past 12 months
Birth weight	Stimulated GH peak
weight	IGF-1/IGFBP3 & Age at treatment start
Estradiol increase	Weight
IGF-1 SDS & delta IGF-1 at 1 year on treatment	IGF-1 SDS & Age

5 DISCUSSION

5.1 MAIN FINDINGS

The overarching aim of this thesis was to optimize GH therapy for short children in a clinical setting, with a focus on identifying patients who were most likely to benefit from therapy and enhancing approaches for personalized treatment strategies.

Study 1 validated the Gothenburg prediction model and found that the model successfully identifies short children who would benefit from GH therapy while avoiding unnecessary treatment with potential risks. In comparison to a Nordic study where 28% of children with GHD were classified as poor responders, achieving a catch-up growth of less than 0.5 SDS after the first year of GH therapy (153), our findings demonstrated that using the Gothenburg prediction model in a clinical setting resulted in only 2 out of 121 children (less than 2%) being classified as poor responders.

Study 2 further advanced the use of predictive tools to avoid unnecessary treatments and guide clinical decisions by comparing the available models to facilitate their application. The study demonstrated that both the Gothenburg and KIGS models achieved remarkable accuracy in predicting first-year growth, with correlation coefficients of $r = 0.99$ and $r = 0.991$, respectively. Furthermore, a direct comparison between the two models revealed an impressive correlation of $r = 0.995$, underscoring their reliability and consistency in clinical prediction. Given that both models demonstrate exceptional accuracy, their use in clinical settings is highly recommended.

Study 3 highlighted the complexity of monitoring IGF-1 during early puberty due to biological variability and identified the need of refined interpretation tools to be able to optimize GH therapy. 15.5% of the prepubertal girls and 15.7% of the prepubertal boys were found to have pubertal levels of sex steroids making the IGF-1 SDS levels to be difficult to interpret. This misalignment between clinical pubertal status and biochemical markers can lead to overestimation or misclassification of IGF-1 levels, potentially resulting in inappropriate adjustments to GH dosing or misdiagnosis.

Further comparisons were made to characterize the differences between low and high IGF-1 SDS levels, aiming to identify key factors influencing IGF-1. Significant differences were observed in GH doses, BMI (kg/m²), and sex steroid levels between the groups. Given the complexity of IGF-1 and the multitude of factors that affect its levels, more advanced and sophisticated techniques are, however, required to further enhance its clinical utility of IGF-1 as a reliable biomarker for GH therapy monitoring and decision-making.

Study 4 introduces cutting-edge machine learning tools for personalized treatment optimization. To demonstrate the superiority of more advanced machine learning models, we included linear regression as a baseline because the close resemblance to traditional statistical regression, offering a familiar reference point for comparison. Symbolic regression emerged as the most effective model, delivering the highest predictive accuracy (R^2 values of 0.47 and 0.51 for three and 12 months, respectively) while utilizing a minimal number of predictors, underscoring its clinical practicality. EBM excelled, however, in providing detailed insights into feature contributions and relationships, making it particularly valuable for interpretability and understanding the underlying factors influencing IGF-1 SDS predictions.

5.2 OPTIMIZING DIAGNOSIS WITH PREDICTION MODELS

Diagnosing GHD is a complex process that still heavily relies on the subjective judgment of clinicians and lacks alignment with evidence-based, universally accepted practices. A systematic review, predominantly including data from European countries, estimated the prevalence of GHD in children and adolescents to range between 1 in 1,107 and 1 in 8,646 (81). This review highlighted significant heterogeneity in diagnostic criteria across centers, emphasizing the need for the scientific community to refine methodologies and develop standardized, harmonized protocols for diagnosing and managing GHD across regions.

Currently, the diagnostic approach includes a combination of medical history, physical examination, biochemical testing, radiological assessments, genetic evaluation, and some form of GH testing (82-84). However, the reliance on clinician interpretation at multiple stages underscores the need for objective, reproducible diagnostic tools to reduce variability and improve diagnostic accuracy.

Interpreting IGF-1 levels in the diagnosis of GHD has proven to be complex. While the influence of puberty on IGF-1 levels has been extensively studied (50, 154), it does not fully account for the nuanced effects of sex steroids on IGF-1 levels, as highlighted in Study 3.

GH stimulation testing, considered the "gold standard" in many centers, often results in a high rate of false-positive diagnoses of idiopathic GHD. In 35 of 36 short children, a GHD diagnosis may be false positive if relied upon solely on a provocative test (105). Highlighting the issue of over-diagnosis is crucial, as under-diagnosis, particularly of severe GHD requiring urgent treatment, is relatively uncommon. Adding to the difficulty in interpretation, Henry discusses in a review the arbitrary nature of the cut-offs for GH peak, which have shifted between 5, 7, and 10 µg/L over time, further complicating their reliability (155).

To facilitate the complex diagnostic process, our findings suggest that prediction models could aid in identifying patients who are most likely to benefit from GH therapy in terms of growth outcomes. Using the Gothenburg prediction model to select children eligible for GH therapy reduced the proportion of poor responders—defined as first-year catch-up growth below 0.5 SDS—to just 2%, compared to 28% in a Nordic study (153).

The initial clinical validation of the Gothenburg model resulted in a low R^2 of 0.17, characterized by an underestimation of low predictions and an overestimation of high predictions. This pattern of under- and overestimation aligns with findings by de Ridder et al. (156), who observed similar results when clinically validating the KIGS model. The low R^2 is likely primarily due to a time delay of several months between the prediction date and the start of GH therapy, which significantly impacted the model's performance. Additionally, the omission of patients with predictions below 0.7 SDS from the calculations may have further influenced the results.

To address this limitation, study 2 aligned the prediction date with the actual GH start date, eliminating the time discrepancy. In this study, Pearson correlation was applied to assess the linear relationship between observed and predicted growth for both the Gothenburg and KIGS models. The results showed excellent correlations ($r = 0.990$, $p \leq 0.0001$ for the Gothenburg model and $r = 0.991$, $p \leq 0.0001$ for the KIGS model), underscoring the predictive accuracy of both models when proper account is made for timing.

The two prediction models were compared in Study 2, demonstrating a strong correlation of $r = 0.995$, $p \leq 0.0001$ indicating that both models are equally capable of predicting first-year growth on GH. The Gothenburg model bases its predictions on a standard dose of 33 µg/kg/day, with broader inclusion criteria and theoretically no upper limits for maximum GH peaks or parental heights. However, a notable limitation of the model is its reliance on the spontaneous maximum GH peak, a parameter not widely used. Additionally, the Gothenburg model requires more detailed early growth auxology compared to the KIGS model, which may limit its applicability in some clinical settings.

In contrast, the KIGS models, accessible through the online iGRO software (131), require fewer input variables and offer dose-dependent predictions, making them more user-friendly than the Gothenburg model. The KIGS models were developed using data from a diverse cohort of investigators and countries worldwide. However, the KIGS models do not allow predictions for children over 10 years of age or for those with a birth weight below -2 SDS, whereas the Gothenburg model includes children with birth weights as low as -2.5 SDS, offering broader applicability in this regard.

The Gothenburg model had a narrower range of residuals (0.10 ± 0.81 SD) compared to the KIGS models (0.03 ± 0.96 SD), indicating that the Gothenburg model's predictions were slightly more consistent and showed less variability relative to the observed data. This suggests that the Gothenburg model may provide more stable predictions in a clinical setting.

A cost-effectiveness analysis based on the KIGS prediction models assessed the economic benefits of using prediction tools in GH therapy. The study included 5,333 GHD patients and 1,173 Turner syndrome patients from the KIGS database, applying results to Germany (157). Patients were categorized as high, average, or low responders, allowing tailored dosages or discontinuation for ineffective therapy. The findings suggested that this approach could significantly reduce both GH usage and overall costs while maintaining comparable growth outcomes in the treated population.

As both the Gothenburg and KIGS models demonstrate exceptional accuracy, their implementation in clinical practice is highly recommended. Choosing between the two models depends on specific clinical needs and settings. The choice of model should be guided by the availability of required input data, the clinical setting, and the specific needs of the patient population.

5.3 OPTIMIZING DOSING STRATEGIES WITH PREDICTION MODELS

As outlined in the study's objectives, individualizing treatment strategies is a key focus. Once patients have been accurately identified, the next challenge lies in determining the appropriate GH dosage. The primary goals of GH therapy for patients with GHD are to enhance growth velocity, facilitate the normalization of growth during childhood, and achieve an adult height that aligns with the child's genetic potential (78).

Despite the well-documented variability in children's responsiveness to GH therapy, current international guidelines continue to recommend weight-based dosing due to insufficient evidence supporting alternative strategies (78). One proposed approach is dose titration based on IGF-1 levels. This method is, however, challenging due to the complexity and variability of IGF-1. Factors such as age, gender, pubertal hormones, nutritional status, and inflammation significantly influence IGF-1 levels. Interpretation is further complicated by differences in assay methodologies and the use of varying reference ranges.

Study 3, with its descriptive focus, highlights the challenges in interpreting IGF-1 levels during the early pubertal years, where clinical pubertal characteristics may not align with biochemical levels of pubertal hormones. The study revealed a discrepancy between pubertal hormones and pubertal status in early puberty, with 9 out of 48 (15.5%) samples from girls and 24 out of 153 (15.7%) samples from boys displaying prepubertal characteristics despite having hormone levels indicative of puberty. This misalignment complicates the interpretation of IGF-1 SDS, as the reference ranges are based on clinical pubertal characteristics. If the reference range were adjusted to early puberty instead of prepubertal reference ranges, the IGF-1 levels in these children would fall below the +2 SDS cut-off, thereby avoiding the need to adjust the GH dose, which could potentially lead to suboptimal treatment.

To further explore the complexities of IGF-1 interpretation, we evaluated all IGF-1 SDS data independently of pubertal staging, to assess different characteristics of children with IGF-1 $\geq$ 2 SDS or below. The group with higher IGF-1 SDS showed significantly lower estradiol levels in girls and lower testosterone levels in boys, and vice versa, aligning with our suspicion that IGF-1 SDS is overestimated in early puberty due to the misalignment between clinical pubertal characteristics and sex steroid levels.

These findings underscore the need for tools that can account for the complexities and variability in IGF-1 levels, especially during early puberty, where overestimations are common. Prediction models integrating IGF-1, alongside other variables, may provide a more accurate framework for optimizing GH therapy.

To demonstrate the potential of more advanced prediction models constructed by machine learning methods, we compared several models ranging from more simple linear regression representing classical statistics to more complex models like symbolic regression and EBM.

Symbolic regression and EBM are examples of white-box models, where the internal decision-making process is fully transparent and interpretable. This transparency offers significant advantages over black-box models, particularly in fields like medicine, where understanding the reasoning behind predictions is crucial. Despite the growing reliance on black-box models, such as neural networks, model trustworthiness is strongly dependent on the quality and reliability of the training data. For instance, a study by Zech et al. (158) demonstrated that a neural network trained to interpret pulmonary X-rays based its predictions on the presence of the word "portable" in the image metadata rather than the actual medical content. This type of error highlights the risks of black-box models, which can inadvertently latch onto spurious correlations.

White-box models, like symbolic regression and EBM, inherently avoid these pitfalls by providing clear, interpretable insights into how predictions are made.

The success of a machine learning solution largely depends on the quality of the data and the effectiveness of the chosen algorithms. When working with real-world data, it is crucial to thoroughly investigate data collection methods. Historical data often includes ambiguities, missing values, outliers, and irrelevant information, highlighting the need for high-quality data to build an effective model. A wide range of machine learning algorithms is available for analyzing data and generating insights. However, selecting the right algorithm for a specific application can be challenging, as the performance of each algorithm varies depending on the characteristics of the data (159). The success of a machine learning solution and its applications depends significantly on both the data and the learning algorithms.

Symbolic regression and EBM managed to predict IGF-1 SDS during maintenance growth phase and puberty with an accuracy of R^2 up to 0.47 and 0.51, while also identifying key features. Artificial intelligence and machine learning have previously been used to predict GH deficiency, growth outcome or adherence to treatment (160-162), but not for IGF-1 prediction.

Studies in non-GH-treated populations have demonstrated that both low, and high serum IGF-1 levels, even within the normal range, are associated with increased cancer risk and overall mortality, emphasizing the critical role of IGF-1 regulation for both treatment safety and efficacy (73). While models designed to predict growth response may optimize growth outcomes, they do not inherently address safety considerations, which can be managed by focusing on IGF-1 as the predictive target.

The strong performance of symbolic regression models, coupled with their use of a small set of easily accessible variables for both 3-month and 12-month predictions, underscores their practical applicability and potential to improve clinical decision-making. This aligns with previous studies where symbolic regression has demonstrated value in medical applications (138) and effectiveness with small datasets (163). A high-performing prediction model requiring minimal input variables is, additionally, more practical for clinical implementation, as it increases the likelihood that necessary data will be readily available in routine practice.

Although EBM demonstrated marginally lower prediction accuracy compared to symbolic regression, distinct advantages exist, particularly in illustrating the relationships between features in fine detail. As such EBM excels in identifying complex patterns, uncovering a broader range of contributing features than symbolic regression. These strengths make EBM a valuable tool for exploring and understanding intricate feature interactions within clinical data.

EBM effectively captured the complex relationship between GH dose and the 3-month IGF-1 SDS prediction, a pattern not detected by linear regression. The model revealed that doses around 0.039 mg/kg/day marked a transition from a negative to a positive correlation with IGF-1 SDS, but at doses above approximately 0.05 mg/kg/day, the relationship plateaued. This is consistent with findings from Cohen et al. (119), who observed that higher doses beyond 0.05 mg/kg/day did not result in better growth outcomes. Notably, their study, conducted during the catch-up phase of treatment, showed that boys still

exhibited increases in IGF-1 levels at higher doses, in contrast to girls that did not exhibit such increases. Our findings also align with these observations despite focusing on a different phase of treatment, further supporting the dose-response plateau at similar thresholds.

For IGF-1 predictions to be clinically applicable, a limit to the number of predictors is crucial to ensure that the necessary data is readily available. We thus believe that the IGF-1 prediction algorithms developed using symbolic regression represent valuable clinical tools. These models have the potential to facilitate more personalized GH dosing by accurately forecasting IGF-1 responses based on individual patient characteristics and parameters.

The advancements presented in this thesis underscore the transformative potential of integrating prediction models into clinical practice. By harnessing the interpretability of white-box models, such as symbolic regression, and EBM, clinicians can make data-driven decisions with confidence; ensuring that treatment strategies are not only effective but also safe. As these models predict IGF-1 SDS with reasonable accuracy while also identifying key contributing factors, they highlight their value in addressing both the variability in patient response and the complexity of dosing strategies.

As the field evolves, the focus must shift toward refining these models for broader clinical adoption. This includes improving data quality, standardizing IGF-1 reference ranges, and ensuring the accessibility of these tools in everyday practice. With these advancements, the path toward truly personalized GH therapy becomes clearer — ushering in a new era of precision medicine that directly benefits patients, optimizing outcomes while simultaneously minimizing risks. This work not only contributes to scientific knowledge but also sets a foundation for the future of individualized treatment in pediatric endocrinology.

5.4 STRENGTHS

Clinical Impact and Practical Applicability

This thesis, particularly Study 1, demonstrates direct clinical impact by showcasing the potential of prediction models to significantly reduce the number of poor responders and unnecessary GH treatments. This minimizes risks for patients and also enhances cost-effectiveness. Study 2 compared the Gothenburg and KIGS prediction models, showing that both are highly

accurate. This finding increases access to prediction tools by providing clinicians with multiple effective options, enhancing their practical availability and flexibility in clinical settings.

Study 3, with the descriptive character provides valuable insights into the limitations of current IGF-1 reference ranges. The study underscores the need for improved tools and prediction models that integrate these factors to enhance the reliability of IGF-1 interpretation, ultimately aiding in the optimization of GH therapy during critical developmental stages.

Development of machine learning prediction models in study 4, especially through symbolic regression, highlights strong performance with minimal input variables, making these models practical and ready for implementation in routine clinical settings.

Novelty

This thesis represents several pioneering efforts.

- First clinical validation of the Gothenburg prediction model and also the first clinical comparison of the Gothenburg and KIGS models.
- Address the challenges of interpreting IGF-1 levels during puberty, identifying discrepancies between clinical and biochemical markers, and proposing innovative prediction models to improve reliability. A prediction model for IGF-1 SDS in children treated with GH after the catch-up growth phase has not previously been developed using advanced machine learning.
- Incorporation of state-of-the-art methods such as symbolic regression and EBM introduces groundbreaking tools for clinical decision-making and emphasizes personalized treatment approaches.

Cohort Strength

A key strength of this thesis is the use of a patient cohort in study 3 and 4 that originates from a previous clinical trial. This provides highly structured, reliable, and robust data, an uncommon advantage in studies relying on real-world clinical data. Furthermore, the availability of longitudinal data within this cohort adds an additional layer of robustness, allowing for the evaluation of changes over time and enhancing the validity of the findings. This well-curated and longitudinally rich dataset ensures the accuracy of the machine

learning models and supports more comprehensive insights into the effectiveness of GH treatment strategies.

Focus on Personalized Treatment

The thesis places a strong emphasis on tailoring GH treatment for short children by utilizing both traditional prediction models and advanced machine learning techniques to optimize individual outcomes.

Safety Considerations

By incorporating IGF-1 as a predictive target, the thesis balances the optimization of growth outcomes with safety, addressing the risks associated with abnormal IGF-1 levels.

Cost-Effectiveness

The ability of prediction models to tailor GH dosages improves economic efficiency and aligns with evidence-based practices, ensuring resource optimization while maintaining high standards of patient care.

5.5 LIMITATIONS

Limited Long-Term Data

Studies 1 and 2 focus on short-term outcomes, leaving the long-term impact of prediction models unexplored.

Follow-Up Limitations

Study 1 effectively identified poor responders to GH therapy. However, a significant limitation was the lack of follow-up for patients deemed ineligible for treatment. Due to the retrospective nature of the study, monitoring the untreated growth of these patients was not feasible, which might have otherwise provided valuable data as a control group.

Assumption of Data Availability

Prediction models depend on detailed patient data, which may not always be accessible in all clinical environments. For example, the Gothenburg model requires early growth data, which is typically available in Swedish cohorts due to the widespread use of child welfare centers but not always available in

regions without such infrastructure. Obtaining the necessary data could pose a significant challenge, potentially limiting model applicability.

Cohort size

While the machine learning models in study 4 demonstrated promising accuracy, several limitations must be considered. A key challenge is the reliance on the quality and structure of the input data. Although the dataset used was highly structured due to its origin in a previous clinical trial, the relatively small cohort size limits the generalizability of the findings to broader populations. Machine learning models are inherently data-driven, where performance and reliability heavily depend on the size and diversity of the training data. A larger cohort could potentially enhance the accuracy and robustness of these models.

Additionally, the relatively small cohort size increases the risk of overfitting, where the model might perform exceptionally well on the training data but struggles to replicate this performance on external or unseen datasets.

EBM identified key features influencing IGF-1 SDS predictions. While these features provide valuable insights, the reliance on them is limited due to the significant uncertainty highlighted in the EBM plots. The presence of these grey areas suggests that while the models captured key trends, the predictions for some data points may be less reliable, particularly in regions with sparse or highly variable data.

These considerations highlight the critical need for future studies with larger and more diverse cohorts to validate and refine the findings of Study 4. A significant limitation of Study 4 is the insufficient number of girls, which precluded meaningful statistical analysis. Given potential differences in how girls respond compared to boys, further research is essential to enhance our understanding of IGF-1 prediction and optimize GH therapy for both sexes.

6 CONCLUSION

This thesis highlights the potential of integrating AI and prediction models into clinical practice to optimize GH therapy for short children. By addressing the complexities of correct GH diagnosis and GH dosing, enhancement of clinical decision making with prediction models can be demonstrated in diagnostic accuracy and treatment.

The validation of the Gothenburg prediction model and its comparison with the KIGS models illustrate the significant progress that can be made in identifying patients most likely to benefit from GH therapy. These models have proven effective in reducing the number of poor responders, minimizing unnecessary treatments, and hence leading to cost-effectiveness.

The thesis also shed light on the challenges of interpreting IGF-1 levels during early puberty, underscoring the need for refined tools that account for biological variability in IGF-1. The use of advanced machine learning methods, such as symbolic regression and EBM, further enhances the ability to predict IGF-1 responses and offering tailored dosing strategies. Symbolic regression combines high predictive accuracy with transparency, while being clinically practical with the low number of clinical accessible features.

The findings presented here emphasize the importance of balancing efficacy with safety. While growth optimization remains a key goal, the regulation of IGF-1 levels is critical to reduce long-term risks, such as those associated with abnormal IGF-1 concentrations. By focusing on personalized approaches to both diagnosis and treatment, this work demonstrates the feasibility of precision medicine in pediatric endocrinology.

This thesis provides a robust foundation for integrating prediction models into routine clinical practice, paving the way for more personalized and effective GH therapies. By leveraging advanced technologies and addressing key clinical challenges, it outlines a pathway to improving patient outcomes, ensuring accurate treatment delivery, and contributing to the broader vision of precision healthcare.

7 FUTURE PERSPECTIVES

Long-Term Impact Studies

Future research should focus on conducting longitudinal studies to evaluate the long-term outcomes of using the Gothenburg and KIGS prediction models. These studies would not only assess the achievement of adult height, but also explore additional markers of sustained efficacy, such as long-term metabolic health, quality of life, and psychosocial outcomes.

By following patients into adulthood, researchers can validate whether the initial predictions and treatment strategies guided by these models continue to benefit patients over time.

Improved IGF-1 Reference Ranges

Future research should focus on developing IGF-1 reference ranges based on sex steroid levels rather than clinical assessments of pubertal status. As sex steroid levels provide a more objective measure of pubertal progression compared to clinical staging, such references would better align with the biological variability in IGF-1 regulation. This approach could mitigate discrepancies observed during early puberty, where clinical and biochemical markers often misalign, as demonstrated in Study 3.

A potential approach would be developing individualized IGF-1 reference ranges using multivariate regression models that incorporate key variables such as puberty, BMI, and sex. These tailored reference ranges would address the variability in IGF-1 levels influenced by these factors, thereby enhancing diagnostic precision and improving clinical decision-making in GH therapy.

Study 4 in a Female Cohort

A significant limitation of Study 4 was the lack of a sufficient number of female participants, leaving potential sex-specific differences unexplored. Future research should replicate Study 4 with a focus on a female cohort to address this gap.

Clinical Validation of Machine Learning Models

An important future step is the clinical validation of the machine learning models developed in Study 4, including symbolic regression and EBM. While

these models showed strong predictive accuracy in a structured dataset, testing their performance on independent and diverse real-world cohorts is essential to confirm their generalizability and reliability.

Clinical validation would assess their integration into routine workflows, including ease of use, data availability, and practical application in real-time clinical decisions. Comparing model predictions with actual patient outcomes would ensure their accuracy and effectiveness in guiding GH therapy. This step is crucial to translate the promising findings of Study 4 into practical tools for personalized GH therapy, improving safety, dosing strategies, and patient outcomes.

ACKNOWLEDGEMENT

I would like to express my deepest gratitude to all those who have supported me during this research and the completion of my thesis. Your encouragement and belief in me made this journey truly rewarding.

I would like to express my heartfelt gratitude to my supervisor, **Jovanna Dahlgren**, for encouraging me to push my boundaries and guiding me toward discovering my full potential. I would also like to thank you for your patience, encouragement, and expert guidance, which have been instrumental in shaping this thesis.

I sincerely thank my co-supervisor, **Hans Fors**, for your unwavering support and for always being there to help, whether in research or clinical work. You have taught me everything I know about pediatric endocrinology, and your dedication, expertise, and passion for the field have been truly inspiring. You have been more than a mentor – your guidance and the example you set have not only shaped my journey but have also inspired me to follow in your footsteps.

I am deeply thankful to my co-authors for their dedication and contributions to this thesis. Special thanks to **Carina Ankarberg-Lindgren** for her meticulous attention to detail, always catching nuances that others might miss; to **Juulia Suvilehto** for her excellent collaboration, consistently prompt responses, and for opening the world of AI to me; and to **Staffan Nilsson**, a true lifesaver when last-minute statistical challenges arose. Your expertise and support have been essential to completion of this work.

Thank you, **Annika Reims**, a mentor and truly role model. Your dedication for the patients and the clinic is admirable, and I am deeply grateful for your support in always allowing us to prioritize research through your thoughtful scheduling.

Sólveig Óskarsdóttir, my mentor in my pediatric training, now a cherished friend and colleague, thank you for your invaluable guidance and encouragement. You brought me to my first international conference in Florida, supported me in delivering my first oral presentation, and opened the door to the amazing world of research. Your passion for the children with 22q11-deletion syndrome are beyond words. Thank you for introducing me to the inspiring world of 22qties.

Everyone at the GP-GRC, Department of Pediatrics, University of Gothenburg, thank you for providing exceptional laboratory support. Your expertise and dedication have been vital to the success of this research.

Lillemor Ljungberg for your invaluable support and assistance with data retrieval. Your contribution has been greatly appreciated.

To all my colleagues—doctors, nurses, and medical secretaries—at the Pediatric Endocrine Clinic, Queen Silvia Children's Hospital: thank you for your constant support, for all the happy times we've shared, and for stepping up to take on extra work while I dedicated myself to this thesis. Your help and friendship have been truly invaluable.

All my sister by bond, your friendship has been invaluable. I am truly blessed to have such incredible women in my life. To **Kathy Vu**, always available 24/7 with your unwavering support and care. Your presence has been a constant source of comfort and strength. To **Annika Eriksson**, thank you for always offering cheerful words and much-needed confidence boosts. Your positivity and encouragement have been a constant source of motivation for me. To **Layla Mirzaei** despite the distance, we share a special bond that nothing can break. Your encouraging words and remarkable ambition have been a constant inspiration to me. To **Anna Svedlund** we share so much in common, and I am deeply grateful for all dinners and walks that have kept me sane not only through this process but through life itself. Thank you for your guidance both in navigating work and family life – it has been a source of immense support and inspiration!

My sisters by blood, **Lilly**, **Susanna** and **Amy.** Thank you for your love, support, and for always being my safe haven. Your presence in my life is something I deeply treasure.

My parents, thank you for your unconditional love. 感謝你們一直以來無條件的支持和愛. 你們教會了我努力會有回報，這一直是我前進的動力.

My children, **Edward, Emilia** and **Nicholas,** thank you for your patience and understanding as I spent countless evenings working. You have been my greatest source of joy and motivation, and I hope this work inspires you to believe that hard work and determination pay off – only the sky is the limit.

My husband, **Dave Trinh**—thank you for your endless love and support, for taking care of our home and the children, allowing me to focus fully on my work. A special thanks as well for your help in editing the text of this thesis—it truly made a difference. Your patience, encouragement, and unwavering belief in me have meant the world. I couldn't have achieved this without you, and to show just how grateful I am, I'm dedicating this entire page to you.

"Appreciation is a wonderful thing. It makes what is excellent in others belong to us as well." — Voltaire

REFERENCES

1. Berthold AA. Transplantation der hoden. Arch Anat Physiol. 1849:42-6.
2. Celestin L-C. The Father of Hormonal Therapy: 1889–1893. Springer; 2014.
3. de Herder WW. Acromegaly and Gigantism in the Medical Literature. Case Descriptions in the Era Before and the Early Years After the Initial Publication of Pierre Marie (1886). Pituitary. 2009;12(3):236-44.
4. Cushing H. The Pituitary Body and its Disorders: Clinical States Produced by Disorders of the Hypophysis Cerebri: JB Lippincott; 1912.
5. Evans HM, Long JA. Characteristic Effects Upon Growth, Oestrus and Ovulation Induced by the Intraperitoneal Administration of Fresh Anterior Hypophyseal Substance. Proceedings of the National Academy of Sciences of the United States of America. 1922:38-9.
6. Houssay B, Biasotti A. La Diabetes Pancreatica de los Perros Hipofisoprivos. Rev Soc Argent Biol. 1930;6:251-96.
7. Li CH, Evans HM. The Isolation of Pituitary Growth Hormone. Science. 1944;99(2566):183-4.
8. Li CH, Dixon JS, Liu W-K. Human Pituitary Growth Hormone: XIX. The Primary Structure of the Hormone. Archives of Biochemistry and Biophysics. 1969;133(1):70-91.
9. Raben M. Treatment of a Pituitary Dwarf with Human Growth Hormone. The Journal of Clinical Endocrinology & Metabolism. 1958;18(8):901-3.
10. Westphal O. Human Growth Hormone. A Methodological and Clinical Study. Acta Paediatr Scand. 1968(182):1-81.
11. Franklin SL, Geffner ME. Growth Hormone: the Expansion of Available Products and Indications. Pediatric Clinics. 2011;58(5):1141-65.
12. Cronin MJ, From the Division of Research Technology G. Pioneering Recombinant Growth Hormone Manufacturing: Pounds Produced Per Mile of Height. The Journal of pediatrics. 1997;131(1):S5-S7.
13. Goldenberg N, Barkan A. Factors Regulating Growth Hormone Secretion in Humans. Endocrinology and Metabolism Clinics. 2007;36(1):37-55.
14. Rivier J, Spiess J, Thorner M, Vale W. Characterization of a Growth Hormone-releasing Factor from a Human Pancreatic Islet Tumour. Nature. 1982;300(5889):276-8.

15. Thorner M, Vance ML, Hartman M, Holl R, Evans W, Veldhuis J, et al. Physiological Role of Somatostatin on Growth Hormone Regulation in Humans. Metabolism. 1990;39(9):40-2.
16. Ben-Shlomo A, Melmed S. Pituitary Somatostatin Receptor Signaling. Trends in Endocrinology & Metabolism. 2010;21(3):123-33.
17. Lewis UJ, Dunn JT, Bonewald LF, Seavey BK, Vanderlaan WP. A Naturally Occurring Structural Variant of Human Growth Hormone. 1978.
18. Boguszewski CL, Jansson C, Boguszewski MC, Rosberg S, Carlsson Br, Albertsson-Wikland K, et al. Increased Proportion of Circulating Non-22-kilodalton Growth Hormone Isoforms in Short Children: A Possible Mechanism for Growth Failure. J Clin Endocrinol Metab. 1997;82(9):2944-9.
19. Veldhuis JD, Keenan DM, Pincus SM. Motivations and Methods for Analyzing Pulsatile Hormone Secretion. Endocrine Reviews. 2008;29(7):823-64.
20. Bonnefont X, Lacampagne A, Sanchez-Hormigo A, Fino E, Creff A, Mathieu M-N, et al. Revealing the Large-scale Network Organization of Growth Hormone-secreting Cells. Proceedings of the National Academy of Sciences. 2005;102(46):16880-5.
21. Jessup SK, Malow BA, Symons KV, Barkan AL. Blockade of Endogenous Growth Hormone-releasing Hormone Receptors Dissociates Nocturnal Growth Hormone Secretion and Slow-wave Sleep. European journal of endocrinology. 2004;151(5):561-6.
22. Muccioli G, Tschöp M, Papotti M, Deghenghi R, Heiman M, Ghigo E. Neuroendocrine and Peripheral Activities of Ghrelin: Implications in Metabolism and Obesity. European Journal of pharmacology. 2002;440(2-3):235-54.
23. Isaksson OG, Jansson J-O, Gause IA. Growth Hormone Stimulates Longitudinal Bone Growth Directly. Science. 1982;216(4551):1237-9.
24. Tannenbaum GS, Gurd W, Lapointe M. Leptin Is a Potent Stimulator of Spontaneous Pulsatile Growth Hormone (GH) Secretion and the GH Response to GH-Releasing Hormone. Endocrinology. 1998;139(9):3871-5.
25. Nijenhuis-Noort EC, Berk KA, Neggers SJ, van der Lely AJ. The Fascinating Interplay between Growth Hormone, Insulin-Like Growth Factor-1, and Insulin. Endocrinology and Metabolism. 2024;39(1):83-9.
26. Jones JI, Clemmons DR. Insulin-like Growth Factors and their Binding Proteins: Biological Actions. Endocrine Reviews. 1995;16(1):3-34.
27. Yuen KCJ, Hjortebjerg R, Ganeshalingam AA, Clemmons DR, Frystyk J. GrowthHhormone/insulin-like Growth Factor I Axis in

Health and Disease States: An Update on the Role of Intra-portal Insulin. Front Endocrinol (Lausanne). 2024;15:1456195.
28. De Vos AM, Ultsch M, Kossiakoff AA. Human Growth Hormone and Extracellular Domain of its Receptor: Crystal Structure of the Complex. Science. 1992;255(5042):306-12.
29. Brooks AJ, Dai W, O'Mara ML, Abankwa D, Chhabra Y, Pelekanos RA, et al. Mechanism of Activation of Protein Kinase JAK2 by the Growth Hormone Receptor. Science. 2014;344(6185):1249783.
30. Cunningham BC, Ultsch M, de Vos AM, Mulkerrin MG, Clauser KR, Wells JA. Dimerization of the Extracellular Domain of the Human Growth Hormone Receptor by a Single Hormone Molecule. Science. 1991;254(5033):821-5.
31. Waters MJ, Rowlinson SW, Clarkson RW, Chen C-M, Lobie PE, Norstedt G, et al. Signal Transduction by the Growth Hormone receptor. Proceedings of the Society for Experimental Biology and Medicine. 1994;206(3):216-20.
32. Brooks AJ, Waters MJ. The Growth Hormone Receptor: Mechanism of Activation and Clinical Implications. Nature Reviews Endocrinology. 2010;6(9):515-25.
33. Lanning NJ, Carter-Su C. Recent Advances in Growth Hormone Signaling. Reviews in Endocrine and Metabolic Disorders. 2006;7:225-35.
34. Baumann G. Growth Hormone Binding Protein. The Soluble Growth Hormone Receptor. Minerva Endocrinologica. 2002;27(4):265-76.
35. Firth SM, Baxter RC. Cellular Actions of the Insulin-like Growth Factor Binding Proteins. Endocrine reviews. 2002;23(6):824-54.
36. Frystyk J, Teran E, Gude MF, Bjerre M, Hjortebjerg R. Pregnancy-associated Plasma Proteins and Stanniocalcin-2 - Novel Players Controlling IGF-I Physiology. Growth Horm IGF Res. 2020;53-54:101330.
37. Rinderknecht E, Humbel RE. The Amino Acid Sequence of Human Insulin-like Growth Factor I and its Structural Homology with Proinsulin. Journal of Biological Chemistry. 1978;253(8):2769-76.
38. Dunger DB, Dunger DB, Acerini CL. Does Recombinant Human Insulin-like Growth Factor-1 Have a Role in the Treatment of Diabetes? Diabetic Medicine. 1997;14.
39. Soos M, Siddle K. Immunological Relationships between Receptors for Insulin and Insulin-like Growth Factor I. Evidence for Structural Heterogeneity of Insulin-like Growth Factor I Receptors Involving Hybrids with Insulin Receptors. Biochemical Journal. 1989;263(2):553-63.
40. Baxter RC, Binoux MA, Clemmons DR, Conover CA, Drop SL, Holly JM, et al. Recommendations for Nomenclature of the Insulin-like

Growth Factor Binding Protein Superfamily. J Clin Endocrinol Metab. 1998;83(9):3213.
41. Ranke M, Elmlinger M. Functional Role of Insulin-like Growth Factor Binding Proteins. Hormone Research in Paediatrics. 1997;48(Suppl. 4):9-15.
42. Dauber A, Muñoz-Calvo MT, Barrios V, Domené HM, Kloverpris S, Serra-Juhé C, et al. Mutations in Pregnancy-associated Plasma Protein A2 Cause Short Stature due to Low IGF-I Availability. EMBO molecular medicine. 2016;8(4):363-74.
43. Liu B, Weinzimer SA, Gibson TB, Mascarenhas D, Cohen P. Type Iα Collagen is an IGFBP-3 Binding Protein. Growth hormone & IGF research. 2003;13(2-3):89-97.
44. Smith WJ, Underwood LE, Clemmons DR. Effects of Caloric or Protein Restriction on Insulin-like Growth Factor-I (IGF-I) and IGF-binding Proteins in Children and Adults. J Clin Endocrinol Metab. 1995;80(2):443-9.
45. Isley WL, Underwood LE, Clemmons DR. Dietary Components That Regulate Serum Somatomedin-C Concentrations in Humans. J Clin Invest. 1983;71(2):175-82.
46. Juul A. Serum Levels of Insulin-like Growth Factor I and its Binding Proteins in Health and Disease. Growth Horm IGF Res. 2003;13(4):113-70.
47. Wabitsch M, Blum W, Muche R, Heinze E, Haug C, Mayer H, et al. Insulin-like Growth Factors and their Binding Proteins before and after Weight Loss and their Associations with Hormonal and Metabolic Parameters in Obese Adolescent Girls. International Journal of Obesity and Related Metabolic Disorders: Journal of the International Association for the Study of Obesity. 1996;20(12):1073-80.
48. Jansson UH, Kristiansson B, Magnusson P, Larsson L, Albertsson-Wikland K, Bjarnason R. The Decrease of IGF-I, IGF-binding Protein-3 and Bone Alkaline Phosphatase Isoforms during Gluten Challenge Correlates with Small Intestinal Inflammation in Children with Coeliac disease. Eur J Endocrinol. 2001;144(4):417-23.
49. Kirschner BS, Sutton MM. Somatomedin-C levels in Growth-impaired Children and Adolescents with Chronic Inflammatory Bowel Disease. Gastroenterology. 1986;91(4):830-6.
50. Juul A, Bang P, Hertel NT, Main K, Dalgaard P, Jorgensen K, et al. Serum Insulin-like Growth Factor-I in 1030 Healthy Children, Adolescents, and Adults: Relation to Age, Sex, Stage of Puberty, Testicular Size, and Body Mass Index. J Clin Endocrinol Metab. 1994;78(3):744-52.
51. Rosenfield RI, Furlanetto R, Bock D. Relationship of Somatomedin-C Concentrations to Pubertal Changes. J Pediatr. 1983;103(5):723-8.

52. Brabant G, Wallaschofski H. Normal Levels of Serum IGF-I: Determinants and Validity of Current Reference Ranges. Pituitary. 2007;10(2):129-33.
53. Lofqvist C, Andersson E, Gelander L, Rosberg S, Blum WF, Albertsson Wikland K. Reference Values for IGF-I throughout Childhood and Adolescence: a Model that Accounts Simultaneously for the Effect of Gender, Age, and Puberty. J Clin Endocrinol Metab. 2001;86(12):5870-6.
54. Goodman-Gruen D, Barrett-Connor E. Epidemiology of Insulin-like Growth Factor-I in Elderly Men and Women. The Rancho Bernardo Study. Am J Epidemiol. 1997;145(11):970-6.
55. Juul A, Flyvbjerg A, Frystyk J, Muller J, Skakkebaek NE. Serum Concentrations of Free and Total Insulin-like Growth Factor-I, IGF Binding Proteins -1 and -3 and IGFBP-3 Protease Activity in Boys with normal or Precocious Puberty. Clin Endocrinol (Oxf). 1996;44(5):515-23.
56. Juul A, Scheike T, Nielsen CT, Krabbe S, Muller J, Skakkebaek NE. Serum Insulin-like Growth Factor I (IGF-I) and IGF-binding Protein 3 Levels are Increased in Central Precocious Puberty: Effects of Two Different Treatment Regimens with Gonadotropin-releasing Hormone Agonists, without or in Combination with an Antiandrogen (Cyproterone Acetate). J Clin Endocrinol Metab. 1995;80(10):3059-67.
57. Rose SR, Municchi G, Barnes KM, Kamp GA, Uriarte MM, Ross JL, et al. Spontaneous Growth Hormone Secretion Increases during Puberty in Normal Girls and Boys. J Clin Endocrinol Metab. 1991;73(2):428-35.
58. Blum WF, Albertsson-Wikland K, Rosberg S, Ranke MB. Serum Levels of Insulin-like Growth Factor I (IGF-I) and IGF Binding Protein 3 Reflect Spontaneous Growth Hormone Secretion. J Clin Endocrinol Metab. 1993;76(6):1610-6.
59. Copeland KC. Effects of Acute High Dose and Chronic Low Dose Estrogen on Plasma Somatomedin-C and Growth in Patients with Turner's Syndrome. J Clin Endocrinol Metab. 1988;66(6):1278-82.
60. Martha PM, ROGOL JR AD, Veldhuis JD, Kerrigan JR, Goodman DW, Blizzard RM. Alterations in the Pulsatile Properties of Circulating Growth Hormone Concentrations during Puberty in Boys. The Journal of Clinical Endocrinology & Metabolism. 1989;69(3):563-70.
61. Metzger DL, Kerrigan JR. Androgen Receptor Blockade with Flutamide Enhances Growth Hormone Secretion in Late Pubertal Males: Evidence for Independent Actions of Estrogen and Androgen. J Clin Endocrinol Metab. 1993;76(5):1147-52.

62. Metzger DL, Kerrigan JR. Estrogen Receptor Blockade with Tamoxifen Diminishes Growth Hormone Secretion in Boys: Evidence for a Stimulatory Role of Endogenous Estrogens during Male Adolescence. J Clin Endocrinol Metab. 1994;79(2):513-8.
63. Keenan BS, Richards GE, Ponder SW, Dallas JS, Nagamani M, Smith ER. Androgen-stimulated Pubertal Growth: the Effects of Testosterone and Dihydrotestosterone on Growth Hormone and Insulin-like Growth factor-I in the Treatment of Short Stature and Delayed Puberty. J Clin Endocrinol Metab. 1993;76(4):996-1001.
64. Venken K, Schuit F, Van Lommel L, Tsukamoto K, Kopchick JJ, Coschigano K, et al. Growth without Growth Hormone Receptor: Estradiol is a Major Growth Hormone-independent Regulator of Hepatic IGF-I Synthesis. J Bone Miner Res. 2005;20(12):2138-49.
65. Frystyk J, Freda P, Clemmons DR. The Current Status of IGF-I Assays - a 2009 Update. Growth Horm IGF Res. 2010;20(1):8-18.
66. Clemmons DR. Consensus Statement on the Standardization and Evaluation of Growth Hormone and Insulin-like Growth Factor Assays. Clin Chem. 2011;57(4):555-9.
67. Burns C, Rigsby P, Moore M, Rafferty B. The First International Standard For Insulin-like Growth Factor-1 (IGF-1) for Immunoassay: Preparation and Calibration in an International Collaborative Study. Growth Hormone & IGF Research. 2009;19(5):457-62.
68. Bereket A, Turan S, Omar A, Berber M, Ozen A, Akbenlioglu C, et al. Serum IGF-I and IGFBP-3 Levels of Turkish Children during Childhood and Adolescence: Establishment of Reference Ranges with Emphasis on Puberty. Horm Res. 2006;65(2):96-105.
69. Collett-Solberg PF, Ambler G, Backeljauw PF, Bidlingmaier M, Biller BMK, Boguszewski MCS, et al. Diagnosis, Genetics, and Therapy of Short Stature in Children: A Growth Hormone Research Society International Perspective. Horm Res Paediatr. 2019;92(1):1-14.
70. Inoue-Lima TH, Vasques GA, Scalco RC, Nakaguma M, Mendonca BB, Arnhold IJP, et al. IGF-1 Assessed by Pubertal Status has the Best Positive Predictive Power for GH Deficiency Diagnosis in Peripubertal Children. J Pediatr Endocrinol Metab. 2019;32(2):173-9.
71. Dekkers O, Biermasz N, Pereira A, Romijn J, Vandenbroucke J. Mortality in Acromegaly: a Metaanalysis. J Clin Endocrinol Metab. 2008;93(1):61-7.
72. Stochholm K, Gravholt CH, Laursen T, Laurberg P, Andersen M, Kristensen LØ, et al. Mortality and GH Deficiency: a Nationwide Study. European Journal of Endocrinology. 2007;157(1):9-18.
73. Burgers AMG, Biermasz NR, Schoones JW, Pereira AM, Renehan AG, Zwahlen M, et al. Meta-analysis and Dose-response Metaregression: Circulating Insulin-like Growth Factor I (IGF-I) and Mortality. The

Journal of Clinical Endocrinology & Metabolism. 2011;96(9):2912-20.
74. Clayton PE, Banerjee I, Murray PG, Renehan AG. Growth Hormone, the Insulin-like Growth Factor Axis, Insulin and Cancer Risk. Nature Reviews Endocrinology. 2011;7(1):11-24.
75. Juul A, Scheike T, Davidsen M, Gyllenborg J, Jørgensen T. Low Serum Insulin-like Growth Factor I is Associated with Increased Risk of Ischemic Heart Disease: a Population-based Case-control Study. Circulation. 2002;106(8):939-44.
76. Wegmann MG, Jensen RB, Thankamony A, Frystyk J, Roche E, Hoey H, et al. Increases in Bioactive IGF do not Parallel Increases in Total IGF-I during Growth Hormone Treatment of Children Born SGA. J Clin Endocrinol Metab. 2020;105(4):e1291-e8.
77. Kjaer ASL, Jensen RB, Petersen JH, Linneberg A, Kårhus LL, Henriksen LS, et al. Tracking and Cumulative Lifetime Exposure to IGF-I in 6459 Healthy Individuals and in SGA Children Treated With GH. J Clin Endocrinol Metab. 2022;108(3):642-52.
78. Grimberg A, DiVall SA, Polychronakos C, Allen DB, Cohen LE, Quintos JB, et al. Guidelines for Growth Hormone and Insulin-Like Growth Factor-I Treatment in Children and Adolescents: Growth Hormone Deficiency, Idiopathic Short Stature, and Primary Insulin-Like Growth Factor-I Deficiency. Horm Res Paediatr. 2016;86(6):361-97.
79. Tillmann V, Patel L, Gill MS, Whatmore AJ, Price DA, Kibirige MS, et al. Monitoring Serum Insulin-like Growth Factor-I (IGF-I), IGF Binding Protein-3 (IGFBP-3), IGF-I/IGFBP-3 Molar Ratio and Leptin during Growth Hormone Treatment for Disordered Growth. Clin Endocrinol (Oxf). 2000;53(3):329-36.
80. Fayter D, Nixon J, Hartley S, Rithalia A, Butler G, Rudolf M, et al. Effectiveness and Cost-effectiveness of Height-screening Programmes during the Primary School Years: a Systematic Review. Archives of disease in childhood. 2008;93(4):278-84.
81. Mameli C, Guadagni L, Orso M, Calcaterra V, Wasniewska MG, Aversa T, et al. Epidemiology of Growth Hormone Deficiency in Children and Adolescents: a Systematic Review. Endocrine. 2024:1-8.
82. Growth Hormone Research S. Consensus guidelines for the Diagnosis and Treatment of Growth Hormone (GH) Deficiency in Childhood and Adolescence: Summary Statement of the GH Research Society. GH Research Society. J Clin Endocrinol Metab. 2000;85(11):3990-3.
83. Oostdijk W, Grote FK, de Muinck Keizer-Schrama SM, Wit JM. Diagnostic Approach in Children with Short Stature. Hormone Research in Paediatrics. 2009;72(4):206-17.

84. Wit JM, Kamp GA, Oostdijk W, Triage DWGo, Children DoGDi. Towards a Rational and Efficient Diagnostic Approach in Children Referred for Growth Failure to the General Paediatrician. Hormone Research in Paediatrics. 2019;91(4):223-40.
85. Cianfarani S, Tondinelli T, Spadoni GL, Scirè G, Boemi S, Boscherini B. Height Velocity and IGF-I Assessment in the Diagnosis of Childhood Onset GH Insufficiency: Do We Still Need a Second GH Stimulation Test? Clinical Endocrinology. 2002;57(2):161-7.
86. Wit JM, Bidlingmaier M, de Bruin C, Oostdijk W. A Proposal for the Interpretation of Serum IGF-I Concentration as Part of Laboratory Screening in Children with Growth Failure. J Clin Res Pediatr Endocrinol. 2020;12(2):130-9.
87. Rikken B, Van Doorn J, Ringeling A, Van den Brande JL, Massa G, Wit JM. Plasma Levels of Insulin-like Growth Factor (IGF)-I, IGF-II and IGF-binding Protein-3 in the Evaluation of Childhood Growth Hormone Deficiency. Hormone Research in Paediatrics. 1998;50(3):166-76.
88. Cianfarani S, Liguori A, Boemi S, Maghnie M, Iughetti L, Wasniewska M, et al. Inaccuracy of Insulin-like Growth Factor (IGF) Binding Protein (IGFBP)-3 Assessment in the Diagnosis of Growth Hormone (GH) Deficiency from Childhood to Young Adulthood: Association to Low GH Dependency of IGF-II and Presence of Circulating IGFBP-3 18-kilodalton Fragment. J Clin Endocrinol Metab. 2005;90(11):6028-34.
89. Ibba A, Corrias F, Guzzetti C, Casula L, Salerno M, Di Iorgi N, et al. IGF1 for the Diagnosis of Growth Hormone Deficiency in Children and Adolescents: a Reappraisal. Endocrine Connections. 2020;9(11):1095-102.
90. Van Rijn J, Grote F, Oostdijk W, Wit J. Short Stature and the Probability of Coeliac Disease, in the Absence of Gastrointestinal Symptoms. Archives of Disease in Childhood. 2004;89(9):882-3.
91. Rappold G, Blum WF, Shavrikova EP, Crowe BJ, Roeth R, Quigley CA, et al. Genotypes and Phenotypes in Children with Short Stature: Clinical Indicators of SHOX Haploinsufficiency. Journal of medical genetics. 2007;44(5):306-13.
92. Zahnleiter D, Uebe S, Ekici AB, Hoyer J, Wiesener A, Wieczorek D, et al. Rare Copy Number Variants are a Common Cause of Short Stature. PLoS Genetics. 2013;9(3):e1003365.
93. Ranke MB. Treatment of Children and Adolescents with Idiopathic Short Stature. Nature Reviews Endocrinology. 2013;9(6):325-34.
94. Marouli E, Graff M, Medina-Gomez C, Lo KS, Wood AR, Kjaer TR, et al. Rare and Low-frequency Coding Variants Alter Human Adult Height. Nature. 2017;542(7640):186-90.

95. Wit JM, Oostdijk W, Losekoot M, van Duyvenvoorde HA, Ruivenkamp CA, Kant SG. Mechanisms in Endocrinology: Novel Genetic Causes of Short Stature. European Journal of Endocrinology. 2016;174(4):R145-R73.
96. Homma TK, Freire BL, Kawahira RSH, Dauber A, de Assis Funari MF, Lerario AM, et al. Genetic Disorders in Prenatal Onset Syndromic Short Stature Identified by Exome Sequencing. The Journal of Pediatrics. 2019;215:192-8.
97. Flechtner I, Lambot-Juhan K, Teissier R, Colmenares A, Baujat G, Beltrand J, et al. Unexpected High Frequency of Skeletal Dysplasia in Idiopathic Short Stature and Small for Gestational Age Patients. European Journal of Endocrinology. 2014;170(5):677-84.
98. Alatzoglou KS, Turton JP, Kelberman D, Clayton PE, Mehta A, Buchanan C, et al. Expanding the Spectrum of Mutations in GH1 and GHRHR: Genetic Screening in a Large Cohort of Patients with Congenital Isolated Growth Hormone Deficiency. J Clin Endocrinol Metab. 2009;94(9):3191-9.
99. Roth J, Glick SM, Yalow RS, Berson SA. Hypoglycemia: A Potent Stimulus to Secretion of Growth Hormone. Science. 1963;140(3570):987-8.
100. Rosenfeld RG, Albertsson-Wikland K, Cassorla F, Frasier SD, Hasegawa Y, Hintz RL, et al. Diagnostic Controversy: the Diagnosis of Childhood Growth Hormone Deficiency Revisited. J Clin Endocrinol Metab. 1995;80(5):1532-40.
101. Ghigo E, Bellone J, Aimaretti G, Bellone S, Loche S, Cappa M, et al. Reliability of Provocative Tests to Assess Growth Hormone Secretory Status. Study in 472 Normally Growing Children. J Clin Endocrinol Metab. 1996;81(9):3323-7.
102. Cole TJ, Hindmarsh PC, Dunger DB. Growth hormone (GH) Provocation Tests and the Response to GH Treatment in GH Deficiency. Archives of Disease in Childhood. 2004;89(11):1024-7.
103. Wilson DM, Frane J. A Brief Review of the Use and Utility of Growth Hormone Stimulation Testing in the NCGS: Do We Need to Do Provocative GH Testing? Growth Hormone & IGF Research. 2005;15:21-5.
104. Ranke MB, Mullis P-E. Diagnostics of Endocrine Function in Children and Adolescents: Karger Medical and Scientific Publishers; 2011.
105. Bright GM, Morris PA, Rosenfeld RG. When Is a Positive Test for Pediatric Growth Hormone Deficiency a True-positive Test? Hormone Research in Paediatrics. 2022;94(11-12):399-405.
106. Albertsson-Wikland K, Rosberg S. Analyses of 24-hour Growth Hormone Profiles in Children: Relation to Growth. J Clin Endocrinol Metab. 1988;67(3):493-500.

107. Kristrom B, Lofqvist C, Rosberg S, Albertsson Wikland K. Effect of Spontaneous GH Secretion and the GH Sampling Period on the Accuracy of Models for Predicting Growth Responses to GH Treatment. J Clin Endocrinol Metab. 2001;86(10):4963-4.
108. Lennartsson O, Nilsson O, Lodefalk M. Discordance between Stimulated and Spontaneous Growth Hormone Levels in Short Children is Dependent on Cut-off Level and Partly Explained by Refractoriness. Frontiers in Endocrinology. 2020;11:584906.
109. Hughes IP, Harris M, Cotterill A, Ambler G, Cowell CT, Cutfield WS, et al. Comparison of Weight-vs Body Surface Area-based Growth Hormone Dosing for Children: Implications for Response. Clinical Endocrinology. 2014;80(3):384-94.
110. Wit JM, Rekers-Mombarg LT, Cutler GB, Crowe B, Beck TJ, Roberts K, et al. Growth hormone (GH) Treatment to Final Height in Children with Idiopathic Short Stature: Evidence for a Dose Effect. J Pediatr. 2005;146(1):45-53.
111. Kristrom B, Aronson AS, Dahlgren J, Gustafsson J, Halldin M, Ivarsson SA, et al. Growth Hormone (GH) Dosing during Catch-up Growth Guided by Individual Responsiveness Decreases Growth Response Variability in Prepubertal Children with GH Deficiency or Idiopathic Short Stature. J Clin Endocrinol Metab. 2009;94(2):483-90.
112. Ranke MB, Lindberg A, Chatelain P, Wilton P, Price DA, Albertsson-Wikland K. The Potential of Prediction Models Based on Data from KIGS as Tools to Measure Responsiveness to Growth Hormone. Hormone Research in Paediatrics. 2004;55:44 - 8.
113. Cohen P, Rogol AD, Howard CP, Bright GM, Kappelgaard AM, Rosenfeld RG, et al. Insulin Growth Factor-based Dosing of Growth Hormone Therapy in Children: a Randomized, Controlled Study. J Clin Endocrinol Metab. 2007;92(7):2480-6.
114. Cohen P, Weng W, Rogol AD, Rosenfeld RG, Kappelgaard AM, Germak J. Dose-sparing and Safety-enhancing Effects of an IGF-I-based Dosing Regimen in Short Children Treated with Growth Hormone in a 2-year Randomized Controlled Trial: Therapeutic and Pharmacoeconomic Considerations. Clin Endocrinol (Oxf). 2014;81(1):71-6.
115. Cohen P, Germak J, Rogol AD, Weng W, Kappelgaard AM, Rosenfeld RG, et al. Variable Degree of Growth Hormone (GH) and Insulin-like Growth Factor (IGF) Sensitivity in Children with Idiopathic Short Stature Compared with GH-deficient Patients: Evidence from an IGF-based Dosing Study of Short Children. J Clin Endocrinol Metab. 2010;95(5):2089-98.
116. Jensen RB, Thankamony A, O'Connell SM, Kirk J, Donaldson M, Ivarsson SA, et al. A Randomised Controlled Trial Evaluating IGF1

Titration in Contrast to Current GH Dosing Strategies in Children Born Small for Gestational age: the North European Small-for-Gestational-Age Study. Eur J Endocrinol. 2014;171(4):509-18.
117. Liu H-J, Wang L-H, Chen L. Evaluation of Safety and Efficacy of Growth Hormone Therapy by IGF-1 Z score in Children with Short Stature. Advances in therapy. 2019;36:2374-83.
118. Ranke M, Schweizer R, Elmlinger M, Weber K, Binder G, Schwarze C, et al. Relevance of IGF-I, IGFBP-3, and IGFBP-2 Measurements during GH treatment of GH-deficient and Non-GH-deficient Children and Adolescents. Hormone Research in Paediatrics. 2001;55(3):115-24.
119. Cohen P, Bright GM, Rogol AD, Kappelgaard AM, Rosenfeld RG, American Norditropin Clinical Trials G. Effects of Dose and Gender on the Growth and Growth Factor Response to GH in GH-deficient Children: Implications for Efficacy and Safety. J Clin Endocrinol Metab. 2002;87(1):90-8.
120. Lundberg E, Kristrom B, Jonsson B, Albertsson-Wikland K, study g. Growth Hormone (GH) Dose-dependent IGF-I Response Relates to Pubertal Height Gain. BMC Endocr Disord. 2015;15:84.
121. Kim JH, Kim SJ, Lee J, Shin CH, Seo JY. Factors Affecting IGF-I Level and Correlation with Growth Response during Growth Hormone Treatment in LG Growth Study. PLoS One. 2021;16(7):e0252283.
122. Oberle M, Grimberg A, Bamba V. Treatment of Pre-pubertal Patients with Growth Hormone Deficiency: Patterns in Growth Hormone Dosage and Insulin-like Growth Factor-I Z-scores. J Clin Res Pediatr Endocrinol. 2017;9(3):208-15.
123. Poidvin A, Touze E, Ecosse E, Landier F, Bejot Y, Giroud M, et al. Growth Hormone Treatment for Childhood Short Stature and Risk of Stroke in Early Adulthood. Neurology. 2014;83(9):780-6.
124. Carel J-C, Ecosse E, Landier F, Meguellati-Hakkas D, Kaguelidou F, Rey G, et al. Long-term Mortality after Recombinant Growth Hormone Treatment for Isolated Growth Hormone Deficiency or Childhood Short Stature: Preliminary Report of the French SAGhE Study. J Clin Endocrinol Metab. 2012;97(2):416-25.
125. Sävendahl L, Cooke R, Tidblad A, Beckers D, Butler G, Cianfarani S, et al. Long-term Mortality After Childhood Growth Hormone Treatment: the SAGhE Cohort Study. The Lancet Diabetes & Endocrinology. 2020;8(8):683-92.
126. Wit J, Ranke M, Albertsson-Wikland K, Carrascosa A, Rosenfeld R, Van Buuren S, et al. Personalized Approach to Growth Hormone Treatment: Clinical Use of Growth Prediction Models. Hormone Research in Paediatrics. 2013;79(5):257-70.

127. Ranke MB, Lindberg A, Chatelain P, Wilton P, Cutfield W, Albertsson-Wikland K, et al. Derivation and Validation of a Mathematical Model for Predicting the Response to Exogenous Recombinant Human Growth Hormone (GH) in Prepubertal Children with Idiopathic GH Deficiency. KIGS International Board. Kabi Pharmacia International Growth Study. J Clin Endocrinol Metab. 1999;84(4):1174-83.
128. Ranke M, Lindberg A, Chatelain P, Wilton P, Cutfield W, Albertsson-Wikland K, et al. Predicting the Response to Recombinant Human Growth Hormone in Turner Syndrome: KIGS Models. Acta Pædiatrica. 1999;88:122-5.
129. Ranke MB, Lindberg A, KIGS International Board. Prediction Models for Short Children Born Small for Gestational age (SGA) Covering the Total Growth Phase. Analyses Based on Data from KIGS (Pfizer International Growth Database). BMC Medical Informatics and Decision Making. 2011;11:1-8.
130. Mehls O, Lindberg A, Nissel R, Haffner D, Hokken-Koelega A, Ranke MB. Predicting the Response to Growth Hormone Treatment in Short Children with Chronic Kidney Disease. J Clin Endocrinol Metab. 2010;95(2):686-92.
131. Loftus J, Lindberg A, Aydin F, Gomez R, Maghnie M, Rooman R, et al. Individualised Growth Response Optimisation (iGRO) Tool: an Accessible and Easy-to-use Growth Prediction System to Enable Treatment Optimisation for Children Treated with Growth Hormone. J Pediatr Endocrinol Metab. 2017;30(10):1019-26.
132. Wikland KA, Kristrom B, Rosberg S, Svensson B, Nierop AF. Validated Multivariate Models Predicting the Growth Response to GH Treatment in Individual Short Children with a Broad Range in GH Secretion Capacities. Pediatr Res. 2000;48(4):475-84.
133. Dahlgren J, Kristrom B, Niklasson A, Nierop AF, Rosberg S, Albertsson-Wikland K. Models Predicting the Growth Response to Growth Hormone Treatment in Short Children Independent of GH Status, Birth Size and Gestational Age. BMC Med Inform Decis Mak. 2007;7:40.
134. Frank M, Drikakis D, Charissis V. Machine-learning Methods for Computational Science and Engineering. Computation. 2020;8(1):15.
135. Badillo S, Banfai B, Birzele F, Davydov, II, Hutchinson L, Kam-Thong T, et al. An Introduction to Machine Learning. Clin Pharmacol Ther. 2020;107(4):871-85.
136. Schmidhuber J. Deep Learning in Neural Networks: An Overview. Neural Networks. 2015;61:85-117.
137. Breiman L. Random Forests. Machine Learning. 2001;45:5-32.
138. Affenzeller M, Burlacu B, Dorfer V, Dorl S, Halmerbauer G, Königswieser T, et al., Editors. White Box vs. Black Box Modeling:

On the Performance of Deep Learning, Random Forests, and Symbolic Regression in Solving Regression Problems. Computer Aided Systems Theory–EUROCAST 2019: 17th International Conference, Las Palmas de Gran Canaria, Spain, February 17–22, 2019, Revised Selected Papers, Part I 17; 2020: Springer.
139. Rudin C. Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead. Nature Machine Intelligence. 2019;1(5):206-15.
140. Maulud D, Abdulazeez AM, Editors. A Review on Linear Regression Comprehensive in Machine Learning2020.
141. Han J, Pei J, Tong H. Data Mining: Concepts and Techniques: Morgan Kaufmann; 2022.
142. Koza JR. Genetic Programming as a Means for Programming Computers by Natural Selection. Statistics and Computing. 1994;4:87-112.
143. Makke N, Chawla S. Interpretable Scientific Discovery with Symbolic Regression: A Review. Artificial Intelligence Review. 2024;57(1):2.
144. Schmidt M, Lipson H. Distilling Free-form Natural Laws from Experimental data. Science. 2009;324(5923):81-5.
145. Hastie TJ. Generalized Additive Models. Statistical Models in S: Routledge; 2017. p. 249-307.
146. Nori H, Jenkins S, Koch P, Caruana R. InterpretML: A Unified Framework for Machine Learning Interpretability. ArXiv Preprint ArXiv:190909223. 2019.
147. Decker R, Albertsson-Wikland K, Kristrom B, Halldin M, Gustafsson J, Nilsson NO, et al. GH Dose Reduction Maintains Normal Prepubertal Height Velocity After Initial Catch-Up Growth in Short Children. J Clin Endocrinol Metab. 2019;104(3):835-44.
148. Niklasson A, Ericson A, Fryer JG, Karlberg J, Lawrence C, Karlberg P. An Update of the Swedish Reference Standards for Weight, Length and Head Circumference at Birth for Given Gestational Age (1977-1981). Acta Paediatr Scand. 1991;80(8-9):756-62.
149. Karlberg J, Engstrom I, Karlberg P, Fryer JG. Analysis of Linear Growth Using a Mathematical Model. I. From Birth to Three Years. Acta Paediatr Scand. 1987;76(3):478-88.
150. Ankarberg-Lindgren C, Dahlgren J, Andersson MX. High-sensitivity Quantification of Serum Androstenedione, Testosterone, Dihydrotestosterone, Estrone and Estradiol by Gas Chromatography-tandem Mass Spectrometry with Sex- and Puberty-specific Reference Intervals. J Steroid Biochem Mol Biol. 2018;183:116-24.
151. Ankarberg-Lindgren C, Mahler T, Norjavaara E. Biological Reference Intervals for Serum Estradiol and Testosterone in Children. Horm Res Paediatr. 2013;80(suppl 1):179

152. Virgolin M, Alderliesten T, Witteveen C, Bosman PA. Improving model-based Genetic Programming for Symbolic Regression of Small Expressions. Evolutionary Computation. 2021;29(2):211-37.
153. Bang P, Bjerknes R, Dahlgren J, Dunkel L, Gustafsson J, Juul A, et al. A Comparison of Different Definitions of Growth Response in Short Prepubertal Children Treated with Growth Hormone. Horm Res Paediatr. 2011;75(5):335-45.
154. Alberti C, Chevenne D, Mercat I, Josserand E, Armoogum-Boizeau P, Tichet J, et al. Serum Concentrations of Insulin-like Growth Factor (IGF)-1 and IGF Binding Protein-3 (IGFBP-3), IGF-1/IGFBP-3 ratio, and Markers of Bone Turnover: Reference Values for French Children and Adolescents and Z-score Comparability with Other References. Clin Chem. 2011;57(10):1424-35.
155. Henry RK. Childhood Growth Hormone Deficiency, a Diagnosis in Evolution: the Intersection of Growth Hormone History and Ethics. Growth Hormone & IGF Research. 2020;55:101358.
156. de Ridder MA, Stijnen T, Hokken-Koelega AC. Validation and Calibration of the Kabi Pharmacia International Growth Study Prediction Model for Children with Idiopathic Growth hormone deficiency. J Clin Endocrinol Metab. 2003;88(3):1223-7.
157. Kaspers S, Ranke MB, Han D, Loftus J, Wollmann H, Lindberg A, et al. Implications of a Data-driven Approach to Treatment with Growth Hormone in Children with Growth Hormone Deficiency and Turner Syndrome. Applied Health Economics and Health Policy. 2013;11(3):237-49.
158. Zech JR, Badgeley MA, Liu M, Costa AB, Titano JJ, Oermann EK. Variable Generalization Performance of a Deep Learning Model to Detect Pneumonia in Chest Radiographs: a Cross-sectional Study. PLoS Medicine. 2018;15(11):e1002683.
159. Sarker IH, Kayes A, Watters P. Effectiveness Analysis of Machine Learning Classification Models for Predicting Personalized Context-aware Smartphone Usage. Journal of Big Data. 2019;6(1):1-28.
160. Murray PG, Stevens A, De Leonibus C, Koledova E, Chatelain P, Clayton PE. Transcriptomics and Machine Learning Predict Diagnosis and Severity of Growth Hormone Deficiency. JCI Insight. 2018;3(7).
161. Smyczyńska U, Smyczyńska J, Hilczer M, Stawerska R, Tadeusiewicz R, Lewiński A. Pre-treatment Growth and IGF-I Deficiency as Main Predictors of Response to Growth Hormone Therapy in Neural Models. Endocrine Connections. 2018;7(1):239-49.
162. Spataru A, van Dommelen P, Arnaud L, Le Masne Q, Quarteroni S, Koledova E. Use of Machine Learning to Identify Patients at Risk of Sub-optimal Adherence: Study Based on Real-world Data from 10,929

Children Using a Connected Auto-injector Device. BMC Medical Informatics and Decision Making. 2022;22(1):179.
163. Wilstrup C, Kasak J. Symbolic Regression Outperforms Other Models for Small Data Sets. ArXiv Preprint ArXiv:210315147. 2021.