Pubertal development in Norwegian girls

Ingvild Særvold Bruserud

Thesis for the degree of Philosophiae Doctor (PhD)
University of Bergen, Norway
2020

UNIVERSITY OF BERGEN

Pubertal development in Norwegian girls

Ingvild Særvold Bruserud

Thesis for the degree of Philosophiae Doctor (PhD)
at the University of Bergen

Date of defense: 23.10.2020

Year: 2020

Title: Pubertal development in Norwegian girls

Name: Ingvild Særvold Bruserud

Print: Skipnes Kommunikasjon / University of Bergen

Scientific environment

The thesis originates from the PhD program of the Department of Clinical Science, Faculty of Medicine at the University of Bergen based on data from the Bergen Growth Studies 1 and 2 (BGS1 and BGS2). The current work was carried out from June 2017 to June 2020 within the environment in relation to the Department of Paediatrics and Youth Medicine at Haukeland University Hospital, WestPaed Research. Financial support was provided by the Western Norway Health Authorities.

My main supervisor was Professor Pétur B. Júlíusson, principal investigator of the BGS1 and BGS2, and affiliated with the Department of Paediatrics and Youth Medicine at Haukeland University Hospital, the Department of Clinical Science at the University of Bergen, and the Department of Health Registry Research and Development at the National Institute of Public Health. My co-supervisors were PhD Mathieu Roelants affiliated to Environment and Health, Department of Public Health and Primary Care, Katolieke University of Leuven; Professor Karen Rosendahl, affiliated to the Department of Radiology, University Hospital of North Norway and the Department of Clinical Medicine, University of Tromsø; and Professor Robert Bjerknes, affiliated to the Department of Paediatrics and Youth Medicine at Haukeland University Hospital and Department of Clinical Science, University of Bergen. Professors Júlíusson and Bjerknes are paediatricians. PhD Roelants is a nutritionist and biostatistician. Professor Rosendahl is a radiologist. Professor Geir Egil Eide, a biostatistician from the Centre for Clinical Research, Haukeland University Hospital, has contributed to and supervised the statistical analysis and interpretation in the first two studies in this dissertation.

Acknowledgements

I appreciate the financial support that I have received from the Western Norwegian Regional Health Authority (Helse-Vest). This PhD project could not have been done without it. There are many people that deserve thanks for their contribution to this thesis, each in their own way. First of all, the children who participated deserve thanks for volunteering. To encounter children and adolescents is always rewarding, and to meet them in a research setting was no exception.

My main supervisor Pétur B. Júlíusson deserves many thanks for introducing me to research on paediatric growth and development; first to the practical and logistical parts of research, such as recruitment and data sampling, then to handle the data and write papers. First in the field of paediatric overweight and obesity, then to puberty. I am truly grateful for his generosity with time, knowledge and patience, and that he has given me the opportunity to be a part of the research group and to work with the data from the Bergen Growth Studies - it has been a privilege. My co-supervisor Mathieu Roelants has without doubt lifted the quality of this work, and his contribution has undoubtedly been important for the final results. I have very much appreciated his thorough, kind help and support in statistics and in writing. Thanks also to my co-supervisor Karen Rosendahl, for teaching a (total) novice how breast ultrasound can be performed in pediatric patients, and for increasing my understanding of how to be humble in research. Thank you, co-supervisor Robert Bjerknes for your overall perspectives and for providing to-the-point comments that give a fresh view on things. Geir Egil Eide also deserves mentioning for his contributions to the statistics, thank you.

I would like to thank the Department of Pediatric and Adolescent Medicine for being my formal employer during the data collection and during the PhD period. Thank you, Janicke and Ida Kari, for showing flexibility, and for challenging me to do ‘my own piece of research’.

I am grateful for my office mates, previous and current co-PhD-students, and others that have been present during these years, making this PhD-period everything

but lonely: Maria, Ida, Ester, Hanna, Arne, Kristine, Thomas, Gottfried, Tonje and Merete. Thank you for nice lunch and coffee breaks including academic and not-so-academic discussions.

Magnus Rise Sveen deserves many thanks for always keeping his good mood while being my only colleague for 5 months during the data collection in 2016. You made recruitment and examinations of almost 1200 children fun! André Madsen, thank you for your helpful advices and support and for asking questions that challenged my thoughts or brought new perspectives. Ninnie Hélen Bakken Oehme, this PhD-journey would definitely not have been the same without you (literally) by my side. I have truly appreciated to work with you and to have someone to share this 'academic life' with, which is filled with both triumphs and frustrations.

My friends in all parts of the country, no one mentioned no one forgotten, thank you for being there!

I would finally like to thank my family. May and Øystein, thank you for your encouragements, and for being there for us. My big sisters, thank you for still laghing at my jokes and cheering for me. And mamma and pappa, thank you for your unconditional love and for supporting me and my projects, as a 'stipendiat' in the mountains, or in academia. My children, Solveig and Trygve, you are my greatest sources of joy, laughter, distraction and pride. Lastly and most important: Øyvind, the love of my life and best friend, you are my rock.

Ingvild Særvold Bruserud, June 2020

Summary of the thesis

Background

Evidence suggests that girls worldwide are experiencing a decrease in age at pubertal onset, but data from Norwegian girls have not been available. Early pubertal timing is associated with adverse health outcomes. Pubertal breast development is currently categorized using the Tanner B scale and relies on visual inspection and/or palpation. Breast ultrasound appears to be a good alternative to this Tanner B assessment because it can separate adipose tissue from breast tissue and facilitates both staging and volumetric measurements.

Objectives

The objectives were to explore ultrasound as an alternative method to assess breast development throughout puberty, to investigate the timing of puberty in Norwegian girls and construct pubertal references, and to shed light on the relationship between anthropometric indicators of body composition and the timing of puberty.

Methods

This thesis is based on data from the Bergen Growth Study 1 (BGS1; 2003–2006, n = 1485, aged 8–15.5 years) and the Bergen Growth Study 2 (BGS2; 2016 and 2017, n = 703, aged 6–16 years). Data on menarche were recorded in both studies, while breast development was assessed only in the BGS2. Six distinct ultrasound-based breast stages were described, and the performance of these were evaluated in terms of observer agreement and in relation to the standardized Tanner B scale. Descriptive pubertal references were estimated from the BGS2. The associations between indirect (body mass index, BMI) and direct (skinfolds and waist circumference (WC)) measurements of fat with the timing of menarche were explored based on data from the BGS1. The ages at menarche between the studies were compared.

Results

The ultrasound staging system performed well in terms of the agreement between one and two observers; however, direct measurements of size and volume lacked the necessary precision. When comparing ultrasound staging and Tanner B staging, an overall good agreement was found. A modest decline in the age at menarche was observed during the last decade, although pubertal references suggest that pubertal timing in Norway is similar to that of neighbouring countries. After adjustment for age and sex, both high and low levels of BMI, WC, subscapular skinfold (SSF) and triceps skinfold were all associated with an earlier and later menarche. In a fully adjusted analysis of all the measurements together, only a high BMI was related to earlier menarche, while a low BMI and a low SSF were associated with later menarche.

Conclusions

Ultrasound is a feasible method to determine the breast developmental stage. This research has clinical implications for the evaluation of puberty in paediatric and adolescent female patients in Norway by providing descriptive references based on a healthy sample of girls. The decline in age at menarche was not explained by BMI and therefore warrants further investigations in population-based studies. BMI was the strongest anthropometric predictor of early and late menarche.

Contents

Scientific environment **2**

Acknowledgements **3**

Summary of the thesis **5**

Contents **7**

Abbreviations **10**

List of Publications **11**

Related publications **12**

1. Introduction **13**

1.1 Normal pubertal development in females *13*

1.1.1 Physical changes during puberty 13

1.2 Assessing breast development *15*

1.2.1 Breast ultrasound 18

1.3 Secular trends in the timing of puberty *21*

1.3.1 Historical trends in the timing of menarche 21

1.3.2 Recent trends in pubertal timing 21

1.4 Health outcomes associated with early pubertal onset *23*

1.5 Factors that influence pubertal timing *23*

1.5.1 Perinatal factors 23

1.5.2 Endocrine-disrupting chemicals 23

1.5.3 Stress and socioeconomic factors 24

1.5.4 Body mass index, body composition and the timing of puberty 24

1.5.5 Weight-related anthropometric measures 25

2. Aims and objectives **28**

3. Methods **30**

3.1 Bergen Growth Studies: Setting and design *30*

3.1.1 The Bergen Growth Study 1 30

3.1.2 The Bergen Growth Study 2 30

3.2 Participants *32*

3.3 Ultrasound technique and assessments *32*
3.3.1 Ultrasound procedure 33
3.3.2 Reliability of the ultrasound examination of the breast 36
3.4 Tanner assessments of breast and pubic hair development *36*
3.5 Menarche *37*
3.6 Anthropometric measurements *37*
3.7 Questionnaire *38*
3.8 Quality control *38*
3.9 Statistical methods *39*
3.9.1 Observer agreement 40
3.9.2 Logistic regression and Cox proportional hazards 42
3.9.3 References for the age at entry of pubertal markers 43
3.9.4 Power calculations 44
3.10 Ethics *44*
4. Summary of results 46
4.1 Paper I *46*
4.2 Paper II *47*
4.3 Paper III *50*
5. Discussion 51
5.1 Methodological considerations *51*
5.1.1 Participants and study design 51
5.1.2 Ultrasound procedure 55
5.1.3 Tanner staging 56
5.1.4 Menarche 57
5.1.5 Anthropometric measurements 57
5.1.6 Questionnaire data 58
5.1.7 Age range 58
5.1.8 Statistical considerations 58
5.1.9 Ethical considerations 59
5.2 Discussion of results *59*
5.2.1 Observer agreement 59
5.2.2 Method comparison 61
5.2.3 Ultrasound and Tanner staging and adipose tissue 63

5.2.4 Pubertal references compared to current references 64
5.2.5 Secular trend in the age at menarche 65
5.2.6 Tempo of maturation versus adiposity 68

6. Conclusions 69

7. Future perspectives 71

References 73

Appendix 86

Abbreviations

BGS1	Bergen Growth Study 1
BGS2	Bergen Growth Study 2
BMI	Body mass index (kg/m2)
CI	Confidence interval
E2	Estradiol
GAM	Generalized additive model
GLM	Generalized linear model
HPG	Hypothalamic-pituitary-gonadal axis
ISB	Ingvild Særvold Bruserud
IOTF	International Obesity Task Force
KR	Karen Rosendahl
NHANES III	Third National Health and Nutrition Examination Survey
PROS	Pediatric Research in Office Settings
SD	Standard deviation
SDS	Standard deviation score
SSF	Subscapular skinfold
Tanner B	Tanner Breast Stage
Tanner PH	Tanner Pubic Hair Stage
TSF	Triceps skinfold
US B	Ultrasound-based Breast Stage
WC	Waist circumference

List of Publications

Paper I

Ingvild Særvold Bruserud, Mathieu Roelants, Ninnie Hélen Bakken Oehme, Geir Egil Eide, Robert Bjerknes, Karen Rosendahl, Pétur B. Juliusson (2018). Ultrasound assessment of pubertal breast development in girls: intra- and interobserver agreement, *Pediatric Radiology* 48, 1576–1583.

Paper II

Ingvild Særvold Bruserud, Mathieu Roelants, Ninnie Helén Bakken Oehme, Andre Madsen, Geir Egil Eide, Robert Bjerknes, Karen Rosendahl and Pétur B. Júlíusson (2020). References for Ultrasound Staging of Breast Maturation, Tanner Breast Staging, Pubic Hair, and Menarche in Norwegian Girls*, The Journal of Clinical Endocrinology & Metabolism*, Volume 105, Issue 5, dgaa107.

Paper III

Heiko Bratke, Ingvild Særvold Bruserud, Bente Brannsether, Jörg Aßmus, Robert Bjerknes, Mathieu Roelants, and Pétur B. Júlíusson (2017). Timing of menarche in Norwegian girls: associations with body mass index, waist circumference and skinfold thickness. *BioMed Central Pediatrics*, 17(1), 138.

Related publications

Ninnie Helén Bakken Oehme, Mathieu Roelants, Ingvild Særvold Bruserud, Geir Egil Eide, Robert Bjerknes, Karen Rosendahl, Pétur B. Júlíusson. Ultrasound-based measurements of testicular volume in 6- to 16-year-old boys — intra- and interobserver agreement and comparison with Prader orchidometry (2018). *Pediatric Radiology* 48, 1771–1778.

André Madsen, Ninnie Helén Bakken Oehme, Mathieu Roelants, Ingvild Særvold Bruserud, Geir Egil Eide, Kristin Viste, Robert Bjerknes, Bjørg Almås, Karen Rosendahl, Jørn V. Sagen, Gunnar Mellgren, Pétur B. Júlíusson. Testicular ultrasound to stratify hormone references in a cross-sectional Norwegian study of male puberty (2020). The Journal of Clinical Endocrinology & Metabolism, Volume 105, Issue 6, dgz094.

Ninnie Helén Bakken Oehme, Mathieu Roelants, Ingvild Særvold Bruserud, André Madsen, Geir Egil Eide, Robert Bjerknes, Karen Rosendahl, and Pétur B. Júlíusson. Reference data for testicular volume measured with ultrasound and pubic hair in Norwegian boys are comparable with Northern European populations (2020). Acta Paediatrica. 2020; 00: 1– 8.

André Madsen, Ingvild Særvold Bruserud and Bjørn-Fredrik Bertelsen, Mathieu Roelants, Ninnie Helén Bakken Oehme, Kristin Viste, Robert Bjerknes, Bjørg Almås, Karen Rosendahl, Gunnar Mellgren, and Jørn V. Sagen and Pétur B. Júlíusson. Hormone references for ultrasound breast staging and endocrine profiling to predict female onset of puberty. Submitted June 2020.

1. Introduction

1.1 Normal pubertal development in females

1.1.1 Physical changes during puberty

Puberty is the period of physical transition from being a child to becoming an adult. Physical changes that take place during female puberty include breast growth and maturation, pubic hair development, maturation of the internal genitalia, occurrence of the first menstruation (menarche), a pubertal growth spurt until final height is reached, and changes in body composition and fat distribution (1, 2). These pubertal events usually occur in a predictable sequence. Accelerated height growth has been recognized as an early manifestation of puberty together with the breast development and the appearance of pubic hair (1). Height growth during puberty accounts for approximately 17% of the final height (3).

Body composition comprises the proportion and amount of fat mass, muscle mass, water and skeletal mass. Lean mass includes all body components except fat tissue. During puberty, both the amount and proportion of body fat increase (4, 5), and the fat distribution relocates more towards the hips (6). The bone mass increases during puberty as a result of an increased size of the bones and also due to an increased bone density (7). The pubertal gains in height and the accretion of bone tissue in puberty precede the peak bone mineralization (8).

Menarche is the first menstrual bleeding. When ovulation is not followed by fertilization, the body sheds out the uterine endometrial lining, which is termed menstruation. Menarche is a relatively late pubertal marker that usually occurs before the secondary sex characteristics have reached full maturation and after peak height velocity (1). Although menarche is considered a marker for reaching reproductive function, the first menstrual cycles often occur without ovulation. Menstrual cycles are often irregular during adolescence and particularly during the first years after menarche (9, 10). The timing of menarche is an easy pubertal marker to assess and requires only a yes/no answer from children and adolescents of different ages that

cover most of the pubertal years (status quo data) or alternatively, older girls and women can be asked to report when their first menstruation occurred (recall data).

Hormonal changes in puberty

The hypothalamic-pituitary-gonadal (HPG) axis is in control of central pubertal development. The HPG axis is activated in the first months of life (11) and is silenced during childhood due to the inhibition of gonadotropin-releasing hormone (GnRH) secretion. Puberty begins when the HPG-axis is re-activated in late childhood when GnRH is released from the hypothalamus with increased frequency and amplitude (2). Although the mechanism that initiates puberty is not fully understood, it is likely that there is both a suppression of the inhibitors of GnRH secretion along with a reinforcement of activators of GnRH (2). The release of GnRH drives the secretion of luteinizing hormones and follicle stimulating hormones from the pituitary that stimulate ovarian growth and the production of estradiol (E2). E2 also stimulates the maturation and growth of the breasts. Breast development is recognized as the first physical sign of estrogenic activity and thus a clinical sign of the re-activation of the HPG axis (12). E2 also stimulates the changing distribution of adipose tissue and bone mineral accretion. In puberty, increasing serum levels of E2 lead to the pubertal growth spurt. Estrogens stimulates skeletal growth by increasing growth hormone secretion and locally within the growth plate. Higher concentrations of E2 stimulate the fusion of the epiphyseal plates and stop linear growth; in other words, oestrogen has a biphasic effect on growth (13).

The activation of the hypothalamic-pituitary-adrenal-axis leads to a rise in dehydroepiandrosterone and androstenedione, two adrenal androgens produced in the zona reticularis of the adrenal gland. This event is termed adrenarche (2), and it occurs independently of gonadal (ovarian) activation (HPG axis) and is hormonally distinct from gonadarche (14). Clinical and visual signs of adrenarche are pubic hair growth (pubarche), axillary hair growth, acne and adult body odour; these hormones also contribute to advancements in skeletal maturation (2).

Genetic control of puberty

The timing of menarche in mothers is associated with the age at menarche in their female offspring (15-20), and similar associations have been shown for pubertal markers such as pubic hair and breast development (18-20). This indicates that pubertal timing is heritable. A large genome-wide association study identified almost 400 significant single nucleotide polymorphisms (SNPs) that were associated with the age at menarche (21). These SNP variants were found to account for 7.4% of the variance in age at menarche, which corresponded to ~25% of the estimated heritability (21).

1.2 Assessing breast development

The assessment of breast development is important when evaluating the timing of female puberty in clinical settings. In the 1960s, the British paediatrician James Tanner (1920–2010) described the longitudinal development of normal puberty using a system based on the development of secondary sex characteristics that is still in use today (22). Breast and pubic hair development were classified into one of five consecutive stages of maturation: breast stages B1–B5 (Figure 1) and pubic hair stages PH1–PH5, respectively (1). Stage 1 is the prepubertal stage, while Stage 2 marks the first appearance of the breast bud (including enlargement of the areola) and/or pubic hair. Stages 3 and 4 are the subsequent pubertal phases, while Stage 5 indicates the mature or adult stage (1). The Tanner scale is almost universally used in clinical and scientific settings and is therefore regarded as the standard when alternative assessment methods are tested (23-26). The widespread use of the same classification system makes it possible to compare data from different studies. The pubertal references currently in use in Norway, based on the Tanner staging system for breast and pubic hair development, are based on data originating from the Copenhagen Puberty Study conducted from 1991 to 1993 (27-29).

The Harpenden Growth study, from which the Tanner staging system evolved, also provided normal age references for breast and pubic hair development, and menarche. The inter-individual variation in timing is rather large, and the age limits

where normal puberty start are relatively wide (1). The mean age at Tanner B2, thelarche, was 11.2 years. The age at onset of pubic hair (Tanner PH2) was 11.7 years, and menarche occurred at a mean of 13.5 years (1). To determine a normal age range for pubertal onset, lower and upper age limits have been statistically defined as the 95% prediction interval of the age at onset. Based on this, the breast development should start between the age of 8 and 13 years. This lower limit of precocious puberty (thelarche before 8 years) is still maintained in the current Norwegian guidelines (29), and in several other countries (30-32).

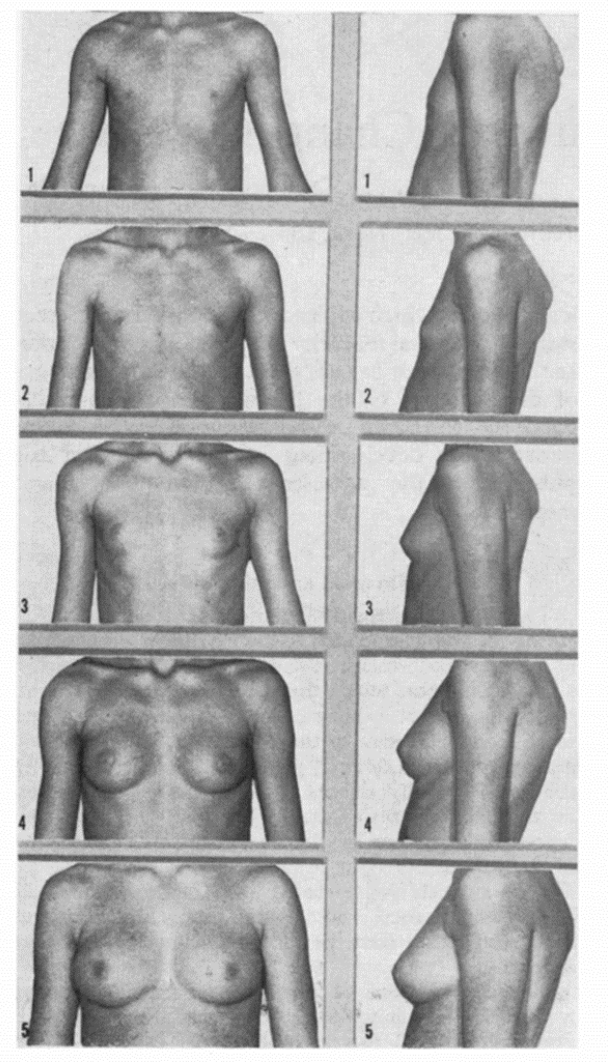


Figure 1 Tanner stages of breast development. *(The original photographs used in the paper by Marshall and Tanner (1))*

1: Tanner B1: Prepubertal, elevation of papilla only 2: Tanner B2: Breast bud stage; elevation of breast and papilla as a small mound, enlargement of areola diameter. 3: Tanner B3: Further enlargement of the breast and areola, with no separation of their contours. 4: Tanner B4: Projection of the areola and papilla to form a secondary mound above the level of the breast. 5: Tanner B5: The mature stage; the projection of the papilla only, due to recession of the areola to the general contour of the breast (1). Reproduced from [Variations in pattern of pubertal changes in girls, W. A. Marshall and J. M. Tanner, 44, 291–303, Copyright 1969] with permission from the BMJ Publishing Group, Ltd.

Palpation of the breast bud, a method in which the fingers are used to examine the 'texture' of the breast, was not conducted in the Harpenden Growth Study. Instead, the assessments were based on photographs, which do not allow for palpation by definition (1). However, several reviews and researchers have cautioned that bias may occur when assessing pubertal development using the Tanner B stages, particularly in overweight and obese girls (2, 25, 33-35). Especially worrisome is the misclassification of adipose breast tissue (lipomastia) as glandular breast tissue, which may result in over-rating of breast maturation. Thus, girls may be considered as pubertal when they are not. Both visual inspection and palpation are therefore advised when examining breast development according to Tanner B staging to decide if thelarche has occurred (2, 23, 25, 33, 35, 36).

The interobserver agreement, which is the ability of a single or different observers to obtain a similar Tanner B stage in the same patient, has shown varying results (37-40) and has not always been found reliable (41), meaning that Tanner B assessments of pubertal breast maturation can be challenging. The method has also been referred to as subjective (35). Therefore, although Tanner staging remains the standard method for staging puberty, its reliability is not necessarily promising. Ultrasound has been proposed as a method that can better determine breast maturation (42), and has the ability to discriminate glandular breast tissue from adipose tissue (26).

1.2.1 Breast ultrasound

Ultrasound is the use of sound waves above 20 kHz. Ultrasound images show how ultrasound waves are reflected by different tissues, and results in different grades of brightness. Thus, the normal appearance of the breast as seen on ultrasound varies depending on the proportion and composition of glandular, fatty and fibrous tissue (43). To date, paediatric breast ultrasound is occasionally performed if breast abnormalities, such as pain or a palpable mass, are found on clinical examination. Breast pathology in women is often associated with breast cancer. However, such pathology is very rare in children and adolescents (44). In most cases where breast disease is suspected based on the clinical examination, a breast ultrasound typically

reveals a normal appearance or benign findings (45, 46). Ultrasound is the preferred initial imaging modality in children because the developing breast is susceptible and vulnerable to biopsies and is also sensitive to ionizing radiation from mammography (47, 48).

Breast ultrasound to stage pubertal development

The breasts undergo changes during infantile growth, puberty, pregnancy, lactation and post-menopause (49). In the newborn, the breast is a rudimentary organ that consists mainly of a network of extending ducts (45). During childhood, the mammary gland development keeps pace with the general growth but does not appear to develop any further (49). At pubertal onset, there is an expansive proliferation of ducts, and the breast starts to develop by growing breast buds. The subcutaneous adipose and connective tissues increase in volume, and the ducts elongate and extend into the adipose tissue. The adult breast tissue mainly consists of two components: fibroglandular tissue and fat (50).

An ultrasound examination of a pubertal breast shows the skin, nipple, fat, Cooper's ligaments, ducts, breast parenchyma, pectoralis muscle, pleura and ribs. The skin and Cooper's ligaments appear hyperechoic (lightly shaded grey), while the mammilla is hypoechoic (dark grey to black) (51). Fatty tissue appears hypoechoic (43). The echogenicity of fibrous tissue can appear intermediate on ultrasound (43). The fibres called Cooper's ligaments can be characterized as the framework holding the glandular and fatty tissues together (51). The glandular tissue appears hypoechoic until the mature phase, when the ducts and lobules spread into the increasingly fatty breast and then take on a more heterogeneous appearance (43) (for illustrations, see Figure 2 of this thesis and Figure 1-6 in Paper I).

Because of certain identifiable morphological hallmarks that occur during puberty, an interpretation of the changing appearance of breast tissue facilitates a classification of breast development using ultrasound. These appearances were first observed and characterized by Bruni et al. in a study of 48 healthy girls from 8–14 years of age that represented Tanner B1 to B4 stages (26). They compared breast

ultrasound findings with the Tanner B stages, breast circumference and oestrogen levels and found that ultrasound grading of the pubertal breast into one of five levels (A–E) was closely related to increasing oestrogen levels (26). Similar descriptions of five ultrasound stages that occur during puberty were reported by Garcia et al. in 2000 (52) and indicated a close correlation to the Tanner B stages with their names of Tanner Stage 1, Tanner Stage 2, etc. The differences between these ultrasound descriptions were primarily regarding the first stages; Stage A was described as the absence of a glandular bud corresponding to both the complete absence of clinically evident mammary growth and adiposity in the breast region (26), while Tanner Stage 1 was defined as prepubertal with an ultrasound that shows ill-defined, hyperechoic, retro-areolar tissue in most cases (52). However, while the descriptions from Bruni et al. relied on shape and size, Garcia et al. included a more detailed description of the ultrasound images' appearance (descriptions are given in Table 1 of Paper I) (26, 52).

Ultrasound staging is used in clinical studies to determine the extent of breast maturation (42, 53-56), and rising gonadotropin and E2 levels have been shown to be correlated with the maturation of breast tissue as observed using ultrasound (morphological stage) (26, 42, 53, 55). There is some variation between the Tanner B staging system and ultrasound-based systems. For example, the Tanner B2 stage does not necessarily equal Stage B or Tanner Stage 2 (26, 42, 53, 55). However, these studies are small (48–80 participants) and vary regarding their recruitment and inclusion criteria as follows: the inclusion of girls that are expected to have already entered puberty (42, 55), obese girls only (55), girls with a limited age range, or inclusion based on probable pathology (e.g. central precocious puberty) (42, 54). Thus, there seems to be a lack of data on how breast ultrasound compares to Tanner B staging in a larger study population of healthy girls of different ages and developmental stages.

Direct measurements of glandular breast tissue

Ultrasound has also been used to measure the size of glandular tissue as either the diameter (54) or both the diameter and the depth to facilitate a volumetric estimate using an ellipsoid formula (V = D1 x D2 x D3 x 0.523) (42). Breast size is associated

with increasing ultrasound stages (54). None of the studies using ultrasound for staging the breast maturation and/or measurements of glandular tissue have made a systematic report of the reliability or repeatability of such assessments, and the observer agreement of the staging systems has not been tested.

In summary, ultrasound allows for phenotyping a hormone-sensitive and developing organ by characterizing its appearance and size without radiation and appears to be a promising alternative to Tanner B examinations of breast development. Accurate interpretation of the ultrasound images depends on in-depth knowledge of breast anatomy as well as on normative data and reference estimates based on healthy, normal-weight girls across all pubertal stages. However, the intra- and inter-observer agreement and agreement with the standard method (Tanner B) is currently unknown. The presence of slightly different definitions also limits the comparability of findings between studies.

1.3 Secular trends in the timing of puberty

1.3.1 Historical trends in the timing of menarche

The age at menarche declined between the 19th century and the first half of the 20th century in Western countries (57-60). In Norway, the menarcheal age fell from above 15.6 years for women born around 1850 to about 13.3 years for women born after 1940 (58) and has been stable at around 13 years since then (61-66). This finding is consistent with trends in other European countries (27, 30, 31, 67). The rapid decline in the age at menarche from the 19th to the 20th century has been attributed to improved living conditions and health, a better nutritional status and improved hygiene (59, 60).

1.3.2 Recent trends in pubertal timing

Two large recent cross-sectional studies from the United States have received particular attention in light of renewed secular changes towards earlier maturation in girls: The Pediatric Research in Office Settings (PROS) (40) and the Third National Health and Nutrition Examination Survey (NHANES III) (68, 69). The PROS study

included 17,077 girls aged 3–12 years who were seen in paediatric practices from 1992–1993. The NHANES III was a nationally representative study that included girls aged 8–17 years and took place from 1988–1994. The PROS study reported that participants reached the Tanner B2 stage at an average age of 10.0 years in white American girls and 8.9 years in African-American girls, and that girls demonstrated breast development at younger ages (0.6–1.2 years earlier) than previously reported (40). The age at onset of both breast and pubic hair development was lowest in the PROS study, but both the PROS and NHANES III reported an earlier onset when compared to previous studies. The American studies also emphasized racial differences in the timing of female puberty, with black girls maturing at younger ages than white girls (40, 68, 69). There was also a steeper decline in the age at pubertal timing among black girls (40, 70). Both of these studies have been the starting point for methodological discussions regarding puberty studies (35), with an emphasis on the ages of participants, definitions of when puberty is precocious (32), Tanner B assessments (with or without breast bud palpation) and also the relationship between anthropometry and puberty timing (71). A further decline in age at onset of breast development has been shown in a more recent study from the United States (34).

European studies from Denmark, Belgium and the Netherlands conducted around the same time as the PROS study (27, 30, 31) could not document the same trends in pubertal onset. However, a study published in 2009 showed a decline in the mean age at pubertal onset during a 15-year-period from 10.9 years in 1991–1993 to 9.8 years in 2006–2008 in Denmark (72). The mean age of menarche did not decrease to the same extent (0.3 years). A decrease in the age at pubertal onset has also been suggested in Greece (73) and in a British study based on self-reported data (15). Studies reporting ages at thelarche were recently reviewed by Eckert-Lind and colleagues (74). They concluded that there has been a worldwide decline in age at pubertal onset based on studies published from 1977–2013. Subsequently, the trends in pubertal timing seem somewhat ‘asymmetrical’, with earlier breast development and a stabilization of age at menarche during the same period. It has been discussed if the the observed decline in age at thelarche and not in age at menarche is attributable to the

challenges of assessing breast development reliably, with interpretations of fat tissue as glandular breast tissue (35).

1.4 Health outcomes associated with early pubertal onset

The early timing of menarche is associated with adverse health outcomes in adulthood (75). Increased risk of cardiovascular disease, type 2 diabetes (76), all-cause mortality (77, 78), increased risk of adult obesity (79-81), and different types of cancers (75), especially breast cancer (82, 83), have all been associated with early menarche.

Earlier pubertal timing is also related with psychosocial problems in girls, such as depression (84), mental disorders, lower self-esteem, a non-favourable body image (85) and higher rates of substance use and abuse (86). Earlier menarche has been related to an increased risk of depressive symptoms and antisocial behaviour in early-middle adulthood, and is suggestive of that the elevated risk of poorer mental health during adolescence may not attenuate in adulthood (87).

1.5 Factors that influence pubertal timing

1.5.1 Perinatal factors

In girls, rapid weight gain in infancy has been associated with a higher risk of earlier pubertal development (88-92), and so has a small birth size (91, 92) and a high birth weight (90). Other perinatal factors that have been linked to pubertal timing are maternal smoking during pregnancy (19, 93) and maternal pre-pregnancy body mass index (BMI) (94). Findings from a recent review on pubertal timing after preterm birth suggested that the existing evidence was not supportive of an earlier (or later) onset of puberty among premature girls (95).

1.5.2 Endocrine-disrupting chemicals

Endocrine-disrupting chemicals (EDCs) are single or mixtures of chemicals that interfere with or alter hormonal action by exhibiting endocrine properties, i.e. estrogenic, antiandrogenic or thyroid actions (96). There are many different groups of EDCs and although the majority of them are man-made, some are naturally occurring.

It has been suggested that pubertal timing and distribution are affected by exposure to EDCs (97-99). Because the presence of and exposure to man-made EDCs are widespread (100), they have been implicated as a possible contributor to the trend of earlier pubertal onset (40, 72, 100).

1.5.3 Stress and socioeconomic factors

Stressful conditions during childhood and adolescence have also been associated with pubertal timing, particularly a younger age at menarche. Stressful events may include family conflicts, divorce, and the absence of a parent (101, 102) or adoption (33, 103). The very poor socioeconomic conditions that typically occur during wartime have been related to a later occurring menarche (104). The effect of socioeconomic status on pubertal timing has shown varying results in Western countries (34, 105).

1.5.4 Body mass index, body composition and the timing of puberty

Studies have found that girls who mature early have higher BMI or BMI standard deviation scores (SDS) than their less mature peers (71, 105-108). Similarly, several studies have observed that girls with an early menarche have a higher BMI than girls with an average or late timing of menarche (31, 109, 110). The association between BMI and early pubertal timing has therefore been related to the epidemic of paediatric overweight and obesity (34, 40, 111-115). High prevalence of overweight and obesity in children is a worldwide phenomenon (116) that is also found in Norway, where the prevalence has increased from the 1970s to the 2000s (117).

Decline in ages at menarche have been attributed to the increased prevalence of overweight and obesity or differences in BMI when similar cohorts investigated at different time points have been compared (72, 109, 118). Nevertheless, causality is not established by these studies due to their cross-sectional design. Longitudinal studies have also found associations between a higher weight status in childhood before pubertal onset (down to three years in one study), and younger ages at pubertal timing (15, 34, 90, 119-123). These studies suggest that increased body weight, including the presence of overweight and obesity, may lead to the early onset of puberty.

However, there are some inconsistencies in these findings, and not all studies point towards overweight and obesity as the only reason for the trend towards an earlier timing of puberty. A Danish longitudinal study observed that the trend towards earlier age at puberty between 1930 and 1969 occurred irrespective of the children's prepubertal BMI. Thus, obesity could not be the only responsible factor for the decline (123). Other studies have found that the age at menarche has remained stable despite an increasing prevalence in overweight and obesity (17, 30, 67) and that the decline in age at menarche remains significant when controlling for the BMI SDS between studies (105).

In 1970, Frisch and Revelle published the critical weight hypothesis, which suggested that a body weight of approximately 48 kg (124) and then later also body fat (125) were required for the occurrence of menarche. Their hypothesis received criticism, as several studies could not replicate their findings (79, 126-128). However, this publication did lead several studies to focus on the impact of body composition on puberty and also the possible mechanisms behind this relationship. Leptin, an appetite-regulating hormone, has been identified as one potential link between nutrition and reproductive function because it acts on the hypothalamus as a signal of energy status to the reproductive system by modulating the Kiss1/Kiss1R system (129). In cases with a negative energy balance, the hypothalamic Kiss1 system may contribute to the suppression of reproductive function (129). Leptin is secreted by fat tissue, and leptin levels are correlated with the amount of fat that is present (e.g. BMI and circulating leptin levels) (130). A threshold level of leptin appears to be a permissive factor for puberty to begin, but although leptin stimulates puberty, it cannot initiate the pubertal onset alone (131). The interrelationship between overweight and obesity and the timing of puberty is still not fully understood (12, 132).

1.5.5 Weight-related anthropometric measures

BMI, which is calculated as kg/m^2, is the most frequently used measure of weight status in clinical settings and population studies. In adults, a BMI > 25 kg/m^2 corresponds to overweight, while > 30 kg/m^2 is defined as obesity. Standard international definitions for underweight, normal weight, overweight and obesity, as

well as for underweight in children and adolescents, have been developed by the International Obesity Task Force (IOTF) (133-135). Consistency with adult cut-offs for overweight and obesity was estimated by extrapolating the cut-offs for a BMI of 25 and 30, at age 18 to children and adolescents using centile BMI curves, and with similar exercice done for the definitions of underweight (134). Calculation of the SDS for BMI and other anthropometric measurements (e.g. waist circumference (WC) and skinfolds) is carried out using growth references. The SDS refer to the position of a child on the reference chart (e.g. the number of SDs above or below the median), while the SD refers to the standard deviation of the distribution.

BMI is a weight-for-height measure and does not offer information about body composition or fat distribution. However, BMI in the overweight or obesity range in children is usually explained by excess fat tissue. The amount of fat is known to vary between children with similar BMIs (136-138). Although BMI has a high specificity for diagnosing high amounts of body fat after being adjusted for age and sex, its sensitivity is lower. Thus, BMI has the ability to detect children who are (clearly) overweight and obese, but this measure may be less accurate at identifying those with a high amount of fat mass and a BMI within the normal range (139).

Skinfold thicknesses in children are commonly assessed at the subscapular (SSF) and/or triceps (TSF) skinfolds. Skinfold measurements are direct measures of subcutaneous fat and have been shown to correlate better with measurements of total body fat percentage than BMI or WC (5). In Norway, the increase in skinfold thickness exceeded the increase in weight-for-height from the 1970s to 2003 (117). Compared to BMI, skinfolds are interesting because they demonstrate the amount of subcutaneous fat, while the BMI represents changes or amounts of other tissues as well (140). Up-to-date Norwegian references for SSF and TSF currently exist (141).

WC reflects the amount of abdominal fat. Central adiposity has been shown to identify both a high (5, 142) and a low trunk fat mass in children and adolescents (142) and also has been found to explain the largest part of variance in BMI followed by the waist-to-height ratio and skinfold assessments (140). Increased abdominal fat (WC) in

childhood has been associated with an earlier pubertal timing (120). Norwegian references for WCs are available (143).

Within normal weight and moderately overweight children, defined by BMI, there are large variations in weight-related anthropometric measurements, and anthropometric measurements beyond BMI may have a potential to add information about fat accumulation in children that are not obese (140). A link between earlier pubertal timing and overweight and obesity is well documented, but few studies have investigated how this association could depend on body composition, more specifically body fat mass, rather than indices of total body weight like the BMI, which may be a method to explore the relation between adiposity and pubertal timing. Weight-related anthropometric measures beyond BMI might add valuable information into the literature on the association between body composition/weight status and timing of pubertal development.

Literature review completed June 2020.

2. Aims and objectives

The overall aims of this thesis were to develop a framework for ultrasound assessments of pubertal breast development by providing descriptive data on the use of ultrasound to assess breast development, to estimate the timing of puberty with this new method, to provide data on the current timing of puberty in healthy Norwegian girls and adolescents and to study how anthropometric measurements are associated with pubertal timing.

The specific aims of the papers included in this thesis are as follows:

I To provide a description of the ultrasound staging of pubertal breast development, to assess the intra- and inter-observer agreement of breast maturity staging with ultrasound, and to estimate the precision of direct measurements of the depth and diameter of glandular tissue in healthy peripubertal girls.

Hypothesis: Ultrasound is a reliable method that can be used to assess the breast developmental stage and size of the mammary gland, with acceptable intra- and inter-observer agreement.

II To estimate the current timing of pubertal development and descriptive reference ranges for ultrasound-based stages of breast development, Tanner B and PH stages, and menarche, to compare the ultrasound and Tanner methods of staging pubertal breast development, and to assess recent changes in age at menarche.

Hypothesis: Ultrasound stages require their own reference ranges because they are based on different characteristics. Girls are more often assessed as pubertal by the Tanner B method than by the ultrasound method due to the interpretation of adipose tissue as pubertal breast tissue. The age at menarche is similar as previously reported.

III To compare the association between menarche and an overall index of body weight (BMI) with the association between menarche and more direct measures of fat mass (skinfolds and WC).

Hypothesis: Central and subcutaneous fat have stronger impact on pubertal timing than the BMI.

3. Methods

3.1 Bergen Growth Studies: Setting and design

This thesis is based on data from the first and second Bergen Growth Studies (BGS1 and BGS 2), which were both cross-sectional studies. As this thesis is solely focused on the pubertal development in girls, data on boys, which were also included in both studies, are not discussed here.

3.1.1 The Bergen Growth Study 1

The BGS1 aimed to identify and describe normal patterns of growth and development. Data were collected from more than 8000 children from 0–19 years of age recruited from health clinics, kindergartens and public schools from November 2003 to December 2006. The study was population-based, and the current national growth references are based on data from the BGS1 (141, 143, 144). Other than recording menarche in girls, the issue of puberty was not investigated in the BGS1.

3.1.2 The Bergen Growth Study 2

The main aim of the BGS2 was to describe normal pubertal development in Norwegian girls. The study was conducted from January to June in 2016 and also in February 2017. The participants were recruited from a random selection of public schools stratified by their town area in the municipality of Bergen, Norway. All girls in the selected schools were invited to participate. The children were investigated during school hours. All participants, regardless of their ethnic origin, were included in the analyses unless otherwise stated in the results. A short summary of the study populations and characteristics in BGS1 and BGS2 is shown in Table 1.

Table 1 *Short summary of the study populations used in this study from the Bergen Growth Studies 1 and 2.*

Study characteristics	**Bergen Growth Study 1**		**Bergen Growth Study 2**
Year of study	2003–2006		2016 + 2017 (test-retest study)
Study setting	19 randomly selected schools (all pupils invited)		7 randomly selected schools (all pupils invited)
N	1481		703
Age, mean (SD)	11.8 (2.2)		11.1 (2.8)
BMI, mean (SD)	18.5 (2.9)		18.3 (3.2)
BMI z-score, mean (SD)	-0.01 (1.0)		0.04 (1.0)
Exclusion	Chronic illness (n=39)		Chronic illness (n=27)
Pubertal markers	Menarche (status quo and recall)		US B Tanner B Tanner PH Menarche (status quo and recall) Blood samples *
Participation rate	69% in primary school (6–12 years) 53% in middle school (13–14 years)		52.1% in primary school (6-12 years) 43.3% in middle school (13-16 years)
Ethnic composition (%)	Norwegian/Nordic: 1294 (87.1%) European: 67 (4.2%) Non-European: 120 (8.1%)		Girls with a known origin: 466 Norwegian/Nordic origin: 381 (81.2 %) European origin: 27 (5.8 %) Non-European: 51 (11.1%)
Parental education (highest degree)			
No secondary education	71 (7.6%)		16 (3.3%)
Secondary education	280 (29.9%)		82 (17.1%)
Higher education	583 (62.4%)		382 (79.6%)
Weight status (IOTF)			
Underweight	144 (9.7%)		47 (7.2%)
Normal	1128 (75.9%)		504 (78.1%)
Overweight	180 (12.1%)		80 (12.4%)
Obese	33 (2.2%)		14 (2.2%)

Abbreviations: BMI: Body mass index, IOTF: International Obesity Task Force, US B: Ultrasound-based breast stage: Tanner B: Tanner breast stage, Tanner PH: Tanner pubic hair stage. *Blood samples were collected in 2016, and are only briefly discussed in this thesis. *Modified from Supplementary Table 1 in Paper II, with permission.*

3.2 Participants

Paper I is based on data from a test-retest study that was conducted in February 2017 as part of the BGS2. It is a methodological paper that investigated the agreement both within and between observers when using ultrasound to assess breast development in puberty. For this study, a random sample of 116 girls aged 6–16 years were invited to participate; a total of 76 (65.5%) agreed, and 57 girls aged 6.1–15.9 years made up the final sample. Due to time constraints, not all girls who agreed could be included.

The descriptive references for pubertal development in Paper II are based on data collected between Januarty and June 2016 (n=667) and the 57 girls that participated in the test-retest study in February 2017. The seven schools were situated in mostly urban environments in the Bergen municipality. All girls (n=1349) attending the selected schools were invited to participate. Parental consent was obtained for 667 girls in the BGS2 in 2016 (participation rate: 49.4%). Children with any known chronic illnesses that were likely to affect their growth were excluded (e.g. coeliac disease, diabetes, heart disease, epilepsy, hypothyroidism, anorexia, cancer or a kidney disorder), which led to the exclusion of 27 girls from the BGS2.

Paper III is based on data from girls who participated in the BGS1. This study included a total of 4035 girls from 0-19 years of age. Data on menarche was available from 1481 girls who were between 8 and 15.5 years of age. A total of 39 girls within the target age group were excluded due to chronic illness following the same criteria as above.

3.3 Ultrasound technique and assessments

The ultrasound-based scoring system used to characterize the maturation of glandular breast tissue (US B) in Papers I and II was primarily based on a description published by Garcia et al. (52) but was adapted to reflect relevant details and characteristic features given by Bruni et al. (26) (further details are given in Table 1 of Paper I). For the purpose of observer training and quality control, the ultrasound method was piloted in healthy girls, and rater agreement was sought regarding standard views and

measurement points. Moreover, all ultrasound examinations that took place during the first three days of data collection in the BGS2 were done jointly by the study nurse (Ingvild Særvold Bruserud, ISB) and an experienced pediatric radiologist (Karen Rosendahl, KR). These training and calibration sessions were used to standardize the ultrasound procedure and lead to an adjustment in the ultrasound protocol by adding a distinct second prepubertal stage: US B0 (Figure 2).

3.3.1 Ultrasound procedure

All ultrasound examinations were performed with the girl in the supine position and her arms along her sides. The left breast was examined in all participants. For the purpose of the test-retest study, we included one breast only to avoid issues with dependent data. In the BGS2, the left breast was examined in all girls, and when the right breast appeared more mature, the right breast also. This was the case in three girls. The left breast was chosen over the right because examining the left breast allowed the observer's (right) arm to rest stable and calm, with the observer sitting on the girls' right side during all examinations.

The ultrasound device used for examination was a SonoSite Edge (Fujifilm SonoSite, USA) machine with a 15-6 MHz (5-cm) linear transducer. The probe was placed at 90º to the skin and centred on the nipple to produce a sagittal standard section that was used for all measurements and staging procedures. Based on this standard section, the depth and diameter of the breast were measured, followed by morphological staging on a scale from US B0 to US B5 (Figure 2).

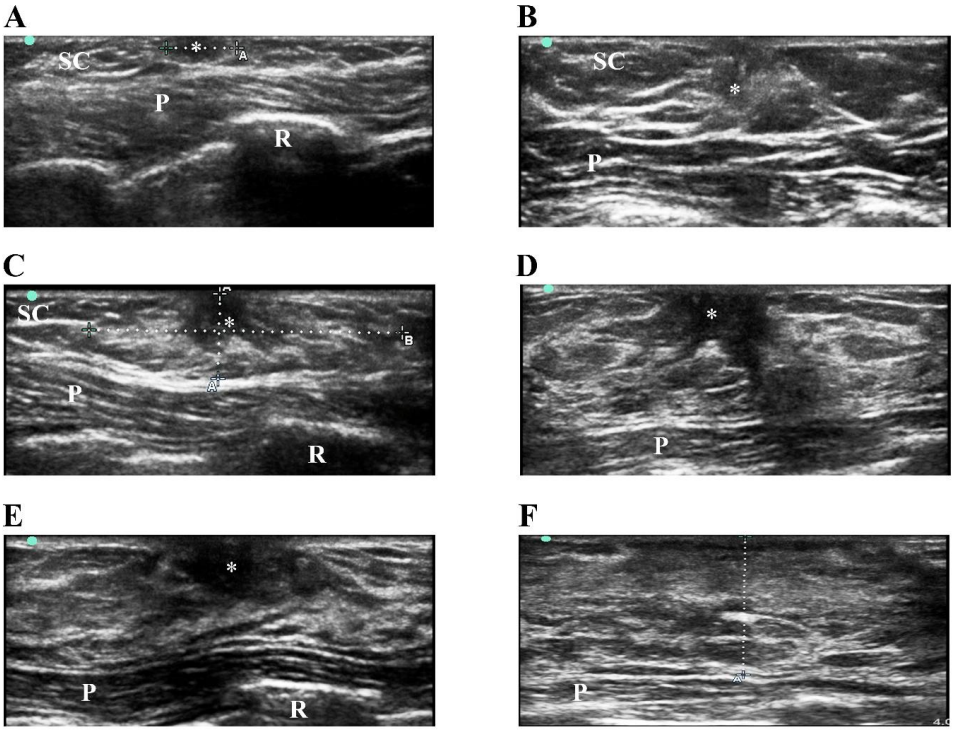


Figure 2 *Ultrasound-based stages of pubertal breast development (US B)*

A *A midsagittal view through the left breast in a 7.1-year-old girl conforming to stage US B0. The image shows a small elongated mass (asterisk) that is hypoechoic relative to the surrounding subcutaneous (SC) adipose tissue. The ribs (R) are located posterior to the pectoral muscle (P) (145).*

B *A midsagittal view through the left breast in a 10.8-year-old girl that corresponded to stage US B1. The breast tissue is hyperechoic compared to the surrounding subcutaneous tissue (SC), but not to the pectoral muscle (P) and has a triangular shape (asterisk) (145).*

C *A midsagittal view through the left breast in a 9.7-year-old girl conforming to stage US B2. The image shows well-defined breast tissue that appears hyperechoic in relation to the subcutaneous (SC) tissue, but not to the pectoral muscle (P), and has a small hypoechoic roundish center (asterisk). Rib (R) (145).*

D *A midsagittal view through the left breast in a 10.7-year-old girl that indicated stage stage US B3. The breast tissue appears hyperechoic with a hypoechoic 'spider-shaped' center (asterisk). Pectoral muscle (P) (145).*

E *A midsagittal view through the left breast in a 12.0 –old-girl conforming to stage US B4. The image shows hypoechoic glandular tissue in the*

subareolar area that has a rounder shape compared to the US B3 stage. The ribs (R) are located posterior to the pectoral muscle (P) (145).

***F** A midsagittal view through the left breast in a 13.9-old-girl conforming to stage US B5. Mature breast tissue is seen as a heterogeneous mass without any hypoechoic center. Pectoral muscle (P) (145).The figure and figure text are only slightly modified from the Supplemental Figure 1 in Paper II. Reprinted with permission.*

The longest diameter of the fibroglandular area was measured. In this study we used a 5-cm long linear transducer. For breast diameters larger than 5 cm but less than 10 cm, we combined the measurements from two scans in the same plane (Figure 3). The depth was measured from the nipple and then vertically down towards the pectoral muscle and/or the end of the glandular tissue. The degree of compression was held to a minimum, as discussed during the training and standardization sessions. Direct measurements of the depth and diameter were chosen for the purpose of calculating the glandular volumes using the formula for a conical shape (formula: volume = (π/3) * radius2 * depth), as described elsewhere (25, 42).

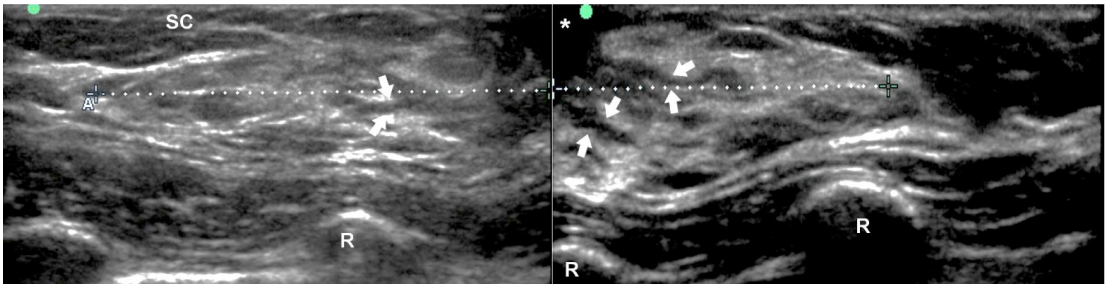


***Figure 3** Mid-sagittal view through the breast in an 11-year-old girl with ultrasound stage 3 shows well-defined tissue which appears hyperechoic in relation to the subcutaneous tissue (SC) with a hypoechoic "spider/octopus-shaped" center (asterisk). The hypoechoic "arms" extend into the glandular hyperechoic tissue (arrows). Two consecutive measurements of the diameter are shown. The upper part of the glandular tissue (left image) is 3.1 cm, and the lower part of glandular tissue (image to the right) is 4.1 cm**. The ultrasound image of the same breast showed a depth of 1.6 cm measured from the interface between glandular tissue and the pectoral muscle to the mammilla (not shown). Figure is identical with Figure 4 a in Paper I. Reprinted with permission*

A preliminary US B stage was recorded during the examinations of 166 girls in the BGS2, but the standardized ultrasound images were stored to allow a final decision on the stage at a later time. This procedure was carried out because a final decision could not be reached within a certain time, and we wanted to avoid prolonged exposure of the girls. To examine the agreement between live scoring and scoring by using the ultrasound images only, a total of 122 ultrasound images (> 10 from each age year) were re-scored after a period of two years by ISB, who was blinded to the previous results,. The agreement between the original and rescored stage had a Cohen's kappa with linear weights of 0.76 (95% confidence interval [CI]: 0.698–0.812), which is considered good.

3.3.2 Reliability of the ultrasound examination of the breast

To assess the intra- and inter-observer agreement of the ultrasound measurements of glandular depth and diameter and staging, two observers (ISB and KR) independently performed ultrasound examinations of the breast in 57 girls (Paper I). Each examination included the determination of the development stage and direct measurements of the glandular tissue. The first examination by the main observer was compared to her second examination performed after an interval of at least 20–35 minutes (intra-observer) and also to the examination carried out by an expert (second) observer (inter-observer).

3.4 Tanner assessments of breast and pubic hair development

The breast stages according to Tanner were assessed by visual inspection and palpation (1), as illustrated in Figure 1 of this thesis. The Tanner B stage was determined before the ultrasound assessment, with the girl in the supine position. Pubic hair development was also assessed according to the method described by Tanner and Marshall (1). The Tanner PH stage was not included in the initial protocol but was systematically recorded from February 22nd until the end of the data collection in BGS2 in 2016. Out of the 453 girls who were asked for the assessment of pubic hair, 372 (82%) consented and had their Tanner PH stage assessed. The original

photographs published by Marshall and Tanner (1) and the modified illustrations published by Rasmussen et al. (38) were used as guide for the assessment of both Tanner B and PH stage.

3.5 Menarche

All girls in the BGS2 were asked if they had experienced menarche, except for those who participated in the test-retest study. In the BGS1, all girls from the 1st to 10th grade of primary school (ranging from 6 to 15–16 years of age) with clinical signs of breast development were asked if they had experienced menarche. Girls younger than 8 years (in the BGS1) were not included in the analysis because none of them had reached menarche. Signs of breast development were not recorded. Girls older than 15.5 years of age from the BGS1 were also excluded from the analysis of menarche because menarcheal status was not systematically recorded above this age. In both studies (BGS1 and BGS2), girls who had experienced menarche were also asked if they could recall the age or date (year and month) of menarche.

3.6 Anthropometric measurements

All examinations of the girls in the BGS2 were conducted by the same female nurse (ISB), while children in the BGS1 were examined by a total of 13 study nurses and one paediatrician (Pétur B. Júlíusson). All observers were trained ahead of the data collection and used standardized techniques (146). Height was measured with the girls standing against a Harpenden Portable Stadiometer (Holtain Ltd., Crosswell, UK) and recorded to the nearest 0.1 cm. Weight was assessed using an electronic scale while the participants wore light clothing (Tanita MC-780MA, Tanita Corp. of America, Inc., Illinois, USA) in the BGS2 and with a Seca personal digital scale (Hamburg, Germany) in the BGS1. Both scales had a precision of 0.1 kg.

A Holtain Tanner/Whitehouse Skinfold Caliper (Cosswell, UK) was used for the skinfold measurements. Skinfolds were assessed on the participants' left sides. The SSF was measured by picking up a skinfold inferomedially with the thumb and

forefinger and setting the caliper about 2 cm below the inferior angle of the left scapula. The skinfold measure was read 1–3 seconds after releasing the skin from the caliper. The TSF was measured between the acromion and the caput radii of the posterior overarm with the arm hanging straight down. Further details are available in Brannsether et al. (141).

The WC was measured with a non-stretch tape on bare skin in a horizontal position at the mid-point between the lowest rib and the top of the iliac crest. The measurement was read from the side at the end of normal expiration and recorded at the nearest millimetre. All measurements were performed only once.

3.7 Questionnaire

A parental questionnaire was distributed to all participants in the BGS2 (n = 667), except those in the test-retest study, and to all participants in the BGS1 (n = 1524). A total of 510 (74%) girls from the BGS2 and 958 (64.5%) participants in the BGS1 returned the completed questionnaire. Data on chronic illnesses, parental educational level and parental ethnical origin were retrieved from similar questionnaires in both studies. The educational level was categorized as low (primary school), medium (secondary school, e.g. college) or higher (more than 12 years of education). The BGS2 questionnaire also contained a question about the timing of menarche, which was used when the girls could not recall the date themselves.

3.8 Quality control

The equipment was checked in the morning before every session. A test-re-test study was conducted every six months during the BGS1 to standardize the measurement technique and also to assess the observer reliability of the anthropometric measurements. The results from these sessions have been published for length, height (144), skinfolds (141) and WC (147). The measurement error was expressed as the technical error of measurement (TEM=$\sqrt{(\sum d2 /2N)}$ (where d represents duplicate

measurements across all observers, and N is the number of duplicates), which was 0.28 cm for height, 0.80 cm for TSF, 0.64 cm for SSF and 0.80 cm for WC.

3.9 Statistical methods

An overview of the samples, variables and statistical methods for each analysis is provided in Table 2. A more detailed description of the statistical methods is given below and is also provided in the respective papers. The statistical analysis was mainly carried out using the IBM Statistical Package for the Social Sciences (SPSS) Statistics, versions 24, 25 and 26 (IBM Corp, Armonk NY, 2016, 2017 and 2019, respectively) and R version 3.5 (R Foundation for Statistical Computing, Vienna, Austria, 2018).

The SDS for anthropometric measures (BMI, WC, SSF and TSF) were calculated according to the Norwegian growth references, which were derived from the BGS1 (141, 143, 144). Underweight (thinness), overweight and obesity were defined according to the International Obesity Task Force (134) and Cole et al. (133). Measurements with an SDS outside +/-3 SD were checked and excluded if clearly erroneous.

Table 2 *An overview of the material and statistical methods used in the three papers of this thesis.*

	Paper I: Observer agreement (test-retest study) of the ultrasound examination of the breast	**Paper II: Pubertal references and method comparisons of ultrasound and Tanner staging of breast development**	**Paper III: Timing of menarche in relation to weight-related anthropometric measures**
Study (year of data collection)	BGS2 (2017)	BGS2 (2016 and 2017) and BGS1 (2003–2006)	BGS1 (2003 to 2006)
Study population (n)	57 girls	703 (BGS2) 1481 girls (BGS1)	1481 girls
Median age (range)	10.9 (6.1–15.9)	10.7 (6.1–16.2)	11.8 (8–15.5)
Statistical methods	Cohen's kappa coefficient Bland-Altman plot One-sample t-test TEM	GLM (probit) GAM Logistic regression	Cox regression Kaplan Meier GLM (probit)
Main variables	US B Depth and diameter of the mammary gland	US B Tanner B Tanner PH Menarche (current status) BMI SDS IOTF-determined weight-classes	Menarche (current status) and age at menarche BMI SDS WC SDS SSF SDS TSF SDS IOTF-determined weight-class

Abbreviations: BGS1 and BGS2: Bergen Growth Study 1 and 2, BMI: body mass index, GAM: Generalized Additive model, GLM: Generalized Linear model (probit), IOTF: International Obesity Task Force, SDS: standard deviation scores, SSF: subscapular skinfold thickness, Tanner B: Tanner breast stage, Tanner PH: Tanner pubic hair stage, TEM: technical error of measurement, TSF: triceps skinfold, US B: ultrasound-based breast stage, WC: waist circumferences.

3.9.1 Observer agreement

Continuous and categorical data require different approaches to determine the degree of agreement either within or between observers. In this study, we used the Cohen's kappa for categorical variables and a Bland-Altman analysis for the continuous measurements.

Agreement: Cohen's kappa (categorical data)

Cohen's kappa is a widely used statistic for observer agreement of categorical outcomes like breast development (37-41, 148), as well as in the radiological field (149, 150). Kappa is a measure of the difference between the observed agreement (percentage agreement/concordance) and the agreement that could be obtained due to chance alone. The coefficient gives a measure between 0 (observed agreement does not exceed the agreement by chance) and 1 (perfect agreement) (151). According to Landis and Koch (151), the coefficient is interpreted as poor when ≤ 0.2, fair when 0.21–0.40, moderate when 0.41–0.60, good when 0.61–0.8 or very good when 0.81–1.0. The amount of disagreement between ordinal outcomes (e.g. a difference of 1, 2, … stages) can be accounted for by assigning weights. For the analysis of observer agreement of US staging and the method agreement with Tanner B staging (five ordinal levels each), we used the Cohen's kappa with the linear weights. The Kappa coefficient was reported with a 95% confidence interval (CI). Agreement of the dichotomized classification of breast development as pubertal (US B2–5 or Tanner B2–5) or pre-pubertal (US B0 and B1 or Tanner B1) was tested with a simple (unweighted) Cohen's kappa coefficient.

Bland-Altman Plot

Bland-Altman mean-difference plots were used to analyse the observer agreement of the (continuous) measurements of glandular depth and diameter by one (intra) and two (inter) observers (152). The Bland-Altman plots show the difference between the two repeated measurements (either by the same observer or by different observers) on the y-axis against the mean of both measurements on the x-axis ('difference against the mean'). The mean value was used because the true value is assumed to be unknown (152). One-sample t-tests were used to test if the pairwise differences were significantly different from zero, which would indicate bias. The 95% limits of agreement (mean difference ± 1.96 SD of difference) indicate where 95% of the measurement differences are located (152). The mean difference and upper and lower limits of agreement are typically marked with a horizontal lines in the mean difference plots. The same method can be used to examine differences between two measurements by the same observer, two different observers or two methods that

measure the same quantity (153). An acceptable level of variance should be pre-defined (152), and this threshold was set to 15% of the actual measurement in our study.

In Paper I, we also calculated the technical error of measurement (TEM) by dividing the SD of the differences by $\sqrt{2}$ (154). The TEM is the standard deviation (measurement error) of a single measurement. Both the Bland-Altman plot and the associated limits of agreement and the TEM provide an idea of how far the measurements may be from the true value, which is highly relevant information.

3.9.2 Logistic regression and Cox proportional hazards

Logistic regression was used to test for differences in the occurrence of the four pubertal markers (US B2, Tanner B2, Tanner PH2 and menarche) between Norwegian and non-Norwegian girls adjusted for age. Likewise, the occurrence of menarche in the BGS1 (2003–2006) and BGS2 (2016) studies was also tested with logistic regression, adjusting for age, BMI SDS and parental education level. The aim of this analysis was to investigate if the occurrence of menarche (by age) was significantly different between the studies. Only girls with a documented Norwegian or Nordic origin were included in the latter analysis to avoid interference from the diverse ethnic background of the girls. The outcomes of these analyses are given as (adjusted) odds ratios (ORs) with 95% confidence intervals.

In Paper III, the association between the weight-related anthropometric measures and the age at menarche was analysed using Cox proportional hazards models. The Cox regression uses the reported age at menarche and allows for censoring if menarche has not yet occurred at the time of study. The advantage of this approach is that more information is used, but it can only be applied when the age at transition is known (e.g. menarche but not B2, PH2). The anthropometric predictor variables (BMI, WC, SSF and TSF) were classified according to their SDS as low (< -1 SDS), normal (-1 ≤ SDS ≤ 1 SDS) or high (> + 1 SDS), and ‘normal’ was used as the reference category to which both ‘low’ and ‘high’ were compared in the analyses. We reported results from unadjusted simple Cox regression models for each marker

separately, as well as fully adjusted multiple regression models in which the effect of all anthropometric measures is jointly estimated and final models, which are the result of a stepwise (backwards) removal of statistically non-significant predictor variables. A p-value below 0.1 was used as a removal criterion, but only factors with a p-value below 0.05 are considered statistically significant. The results are expressed as hazard ratio's (HRs) with 95% confidence intervals, which indicate the probability of having menarche at a given age according to the grouping level of a covariate (e.g. 'high BMI') relative to the reference category ('average BMI'). The fully adjusted and final models allowed to estimate the effect of anthropometric variables independent of the other (somewhat correlated) variables. Because of the possibility that girls may have had a different weight status at the time of examination and at the time of menarche, all analyses were repeated in a subgroup of participants who had experienced menarche during the 12 months prior to investigation.

In Paper III, the median ages at menarche for all girls, and then also girls in different weight classes were estimated with Kaplan-Meier analysis. Similar to the data sampling for Paper II, the occurrence of menarche implies left censored variables, girls that have not yet experienced menarche were right censored.

3.9.3 References for the age at entry of pubertal markers

Reference ranges for pubertal development are based on the cumulative distribution of ages at entry into the various puberty markers/stages (US B, Tanner B, Tanner PH stages and menarche). The cumulative distribution was estimated from the age and current status of the girls with either probit regression (a generalized linear model with probit link function) or with a generalized additive model (GAM) for binary outcomes with probit link. These two statistical methods differ by the assumption of a normal distribution of ages (probit regression) or not (GAM). The age reference curves were estimated for each stage of each marker separately with both the probit and GAM analyses. The degree of smoothing of the GAM was determined by generalized cross-validation, and all models were checked with diagnostic plots using the mgcv package in R (155) (version 3.5, R Foundation for Statistical Computing, Vienna, Austria, 2018). The probit model has the advantage that the whole age distribution is

characterized by the mean and SD, but the validity depends on a normal (Gaussian) distribution of ages. In the tables and text, we primarily reported numbers from the nonparametric additive model, because few age distributions were found Gaussian. Probit analysis was also used to estimate the mean and SD of age at menarche in BGS1, both for the whole group, and for the sub-group of girls who had experienced menarche during the 12 months before the measurements were performed ('girls with recent menarche').

3.9.4 Power calculations

The initial aim of the BGS2 was to recruit 100 girls per age stratum from 6 – 16 years of age (i.e. a total of 1000 girls) for the construction of pubertal references. Due to the time required for the examination of individual girls, the difficult recruitment of children for a puberty study and the logistical challenges, we did not reach this target number of participants. Some dropouts or missing responses were, however, expected, and the current sample size allowed us to estimate the median age at the transition from a particular stage of pubertal development to the next (e.g. menarche, or from B2 to B3) with a standard error of about one month, and the outer limits of the 'normal range' (-2 to + 2 SD) with a standard error of 2–3 months, which was considered sufficiently precise.

The sample size of the BGS1 was estimated for the purpose of detecting secular changes in height and weight during a 30-year period (117). A post hoc power analysis for Paper III showed that the sample allowed us to detect statistically significant hazard ratios of approximately 1.3 for all girls and 1.5 for the 181 girls with recent menarche.

3.10 Ethics

Parental consent and participant assent were obtained before the children were enrolled in the study (consent letter in Appendix). Participants who withdrew their assent were omitted (approximately 5 girls in the BGS2). All children received age-appropriate information in writing and also verbally by the study nurse prior to their

participation. The BGS1 and BGS2 were approved by the Regional Committee for Medical and Health Research Ethics in Western Norway (BGS1, REK number 09740 and for BGS2, 2015/128/REK Vest). The studies complied with the Declaration of Helsinki (156).

4. Summary of results

4.1 Paper I

Paper I was a methodological paper that investigated the agreement both within and between observers when using ultrasound to assess breast development throughout puberty. Fifty-seven girls aged 6.1–15.9 years (mean age: 10.9 years) participated in this study. The intra-observer comparison (for one observer, ISB) of breast staging had a linear-weighted kappa coefficient of 0.84 (95% CI: 0.78–0.91), and 70.2% (40/70; 95% CI: 56.4–81.2%) concordance, respectively, and the inter-observer (nurse vs. specialist) comparison demonstrated a kappa coefficient of 0.71 (0.62–0.80) and a concordance of 51.8% (29/56; 95% CI: 38.1–65.2%) when using the six stages of breast development. When the two prepubertal stages (US B0 and US B1) and the pubertal stages (US B2 and above) were combined, perfect agreement (e.g. 100% concordance) was found for one observer, and 96.4% (95% CI: 86.6%–99.4%) concordance was calculated for the inter-observer assessments.

For the measurements of the depth and diameter of the mammary gland, the mean differences between the first and second measurement of one observer were not significantly different from zero (one sample t-test, $p = 0.86$ and $p = 0.070$ for the diameter and depth, respectively), which indicated no systematic bias. For the two observers, the mean difference was not significantly different from zero for the measurements of diameters ($p = 0.86$), but a mean difference of 0.1 cm was found for the depth ($p < 0.01$). However, the limits of agreement were relatively wide for both the depth (29% of the sample mean) and diameter (45.0%). The percentages were found by dividing the limits of agreement (range from lower to upper) (e.g. 0.4 cm for the depth and 2.4 cm for the diameter for one observer) by the average depth and diameter. A constant variance across the range of measurements was observed in the Bland-Altman plots (shown in the supplemental material of Paper I).

4.2 Paper II

Reference ranges for the timing of pubertal development are represented in Paper II (Table 3). The pubertal references were based on 696 girls for US B, 700 for Tanner B, 372 for Tanner PH, and 643 for menarche. The median (p3, p97) age at onset of breast development was 10.2 (7.8, 12.6) years according to ultrasound staging (US B2), and 10.4 (8.2, 12.6) years according to Tanner staging (B2). The median (p3, p97) age at Tanner PH2 was 10.9 (8.18, 12.58) years, while that of menarche was 12.7 (11.0, 15.9) years. Ultrasound breast staging (US B) indicated an earlier age at pubertal onset (0.2 years) compared to the Tanner method, while the opposite was found for the higher maturational stages (Tanner B4 and B5) where the age at transition with Tanner staging was ahead of the ultrasound assessment.

The ultrasound and Tanner methods had a good overall level of agreement (kappa = 0.87 (95% CI: 0.85–0.89)) and were concordant in 551 of 695 (79.3%) assessments. When dichotomising the breast developmental stage into thelarche (B2 or higher) or not thelarche (US B0/B1 or Tanner B1), the agreement was very good (kappa = 0.94 (95% CI; 0.91–0.96)). The kappa coefficients were comparable in girls with average weight (kappa = 0.88 (95% CI: 0.86–0.91)) and overweight/obesity (kappa = 0.85 (95% CI: 0.79–0.90).

The onset of all pubertal markers in the BGS2 occurred earlier in girls with a known non-Norwegian origin (n = 92) compared to girls of a known Norwegian origin (n = 374). A comparison with the BGS1 demonstrated that the occurrence of menarche had significantly decreased from 13.3 (SD 1.7) years in 2006 to 13.1 (SD 1.2) years in 2016 ($p < 0.05$) in Norwegian girls (Figure 4). This difference remained statistically significant (OR: 2.0; 95% CI: 1.1–3.6; $p = 0.016$), when additionally adjusted for the BMI SDS and parental educational level.

Table 3 *Age references for breast development, pubic hair and menarche: the median and distribution (-2, -1, 1, 2 SD) of the age (in years) at attainment of the specified stage.*

	Generalized Additive Model					Probit
	Age at attainment					
Pubertal marker	**-2SD**	**-1SD**	**Median**	**1SD**	**2SD**	**mean (SD) age**
Breast (ultrasound)						
US B1	NA	5.6*	8.3	10.7	12.4	8.3 (2.2)
US B2	7.7	8.9	10.2	11.5	12.8	10.2 (1.3)
US B3	9.1	10.3	11.5	12.7	13.9	11.5 (1.2)
US B4	10.4	11.3	12.3	13.9	NA	12.7 (1.5)
US B5	11.3	12.7	15.9	NA	NA	15.3 (2.3)
Breast (Tanner)						
B2	8.0	9.2	10.4	11.6	12.7	10.4 (1.2)
B3	9.5	10.4	11.4	12.5	13.8	11.5 (1.1)
B4	10.7	11.6	12.5	13.8	NA	12.8 (1.3)
B5	11.1	12.3	14.7	NA	NA	14.8 (2.2)
Pubic Hair (Tanner)						
PH2	8.5	9.7	10.9	12.1	13.3	10.9 (1.2)
PH3	9.7	10.9	11.9	12.9	13.7	11.9 (1.1)
PH4	10.5	11.8	13.1	14.4	15.7	13.1 (1.3)
PH5	11.4	13.0	15.1	NA	NA	15.1 (1.9)
Menarche	11.0	11.7	12.7	14.1	16.2	12.9 (1.2)

Abbreviations: NA: not available (outside the data range); SD: standard deviation; US B: ultrasound-based breast stage, B: breast; PH: pubic hair * extrapolated age that falls just below the data range. *The table is similar to Table 1 in Paper II and was reprinted with permission.*

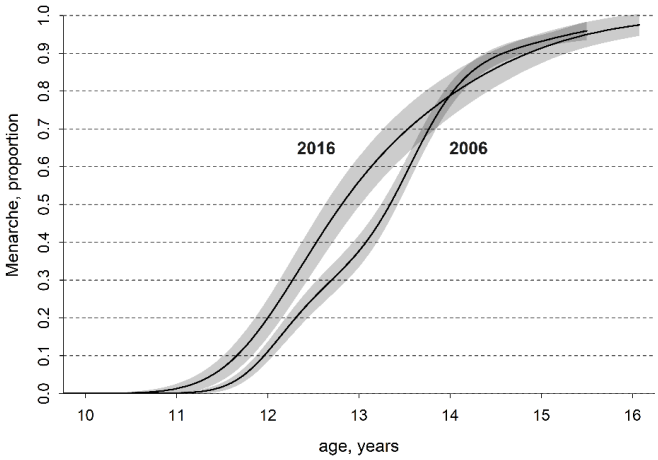


***Figure 4**. The cumulative proportion of girls who had reached menarche by age estimated with a generalized additive model for binary data with a probit link. Data from girls of Norwegian or Nordic origin in the studies from 2006 (BGS1) and 2016 (BGS2). Grey areas indicate the standard errors below/above the estimated age prevalence. The figure is similar to Supplemental figure 3 in Paper II and has been reprinted with permission.*

4.3 Paper III

In Paper III, we explored the association between the age at menarche and weight-related anthropometric measurements (BMI, WC, SSF and TSF) in the BGS1. A total of 477 of 1481 (32.2%) girls had experienced menarche, and 181 had done so during the last 12 months before the data collection. The overall mean age at menarche was 13.3 years (95% CI; 13.1-13.4, SD 0.9 years). We demonstrated that both high (> 1 SDS) and low levels (< -1 SDS) of all weight-related anthropometric variables were significantly associated with earlier and later menarche, respectively, when compared to the girls with normal levels (-1 ≤ SDS ≤ 1) in the unadjusted models. In a fully adjusted analysis of all the measurements together, only a high BMI was significantly related to earlier menarche (HR 1.41, 95% CI: 1.07–1.85), and both low BMI and low SSF with later menarche (HR 0.53 95% CI: 0.38–0.75 and HR 0.54, 95% CI 0.39–0.75, respectively). Similar results were found when using the IOTF classification of weight status. In girls with recent menarche, only a high BMI was significantly associated with an earlier age at menarche (HR 1.79, 95% CI: 1.23–2.62), but this analysis may have lacked sufficient statistical power because of the small sample size.

5. Discussion

In this thesis we have investigated a new method to assess puberty in girls based on utrasound examination of the breast. Utrasound staging was found a reliable alternative for traditional Tanner staging. Further we have developed a framework to evaluate utrasound staging of the breast by estimating age-references which show the typical distribution of maturity by age in healthy girls living in Norway. Further we could document that the age at menarche has again decreased in girls of Norwegian origin. Finally we have demonstrated that the age at menarche is related to weight status in girls.

5.1 Methodological considerations

When investigating pubertal development with the aim to estimate pubertal references, the sample of participating children and adolescents should be representative of the target population. This implies that a population-based sample is preferable. However, selection bias may arise when the study population differs from the general population in relation to factors either known or unknown to influence the outcome: in this case, normal pubertal onset and development. In this section, potential biases related to participation, generalizability and selection will be discussed.

5.1.1 Participants and study design

Participation and generalizability

The design, recruitment and target populations were similar in the BGS2 and BGS1. Non-participation is a challenge in most, if not all, population studies. The overall participation rates were 50.3% in the BGS 2 and 67% in the BGS1. Higher participation rates were observed among primary school children (from 6–12 years of age) than they were in upper secondary school in both studies (compared in Table 2).

The BGS2 was the first pubertal reference study conducted in Norway. Earlier Norwegian studies that reported data on pubertal development have either included data on menarche only, or documented Tanner stages from datasets with a limited age

range without estimating the ages at onset (63, 157-159). The published references in Paper II are the first to be based on a Norwegian sample, and also the very first references for ultrasound breast staging. The participation rate of 50.3% may seem rather low when representativeness of the target population is aimed at, but it is comparable to studies including physical assessments of pubertal development in other European countries with reported rates from < 40 to 80% in the Netherlands (160), and better than the rates reported in the Copenhagen Puberty Study in Denmark (~ 40% (27) and ~35 % (72)). There were only a few missing observations for the US B and Tanner B assessments (98.9 % and 99.6% consented, respectively), while the assessment of pubic hair included 82% of the girls who were asked.

National growth reference curves for height, weight, BMI, skinfolds and WC have previously been published based on data from the BGS1, and the study is regarded as nationally representative. Possible selection bias was examined among 149 non-participants in BGS1 and revealed no statistically significant difference in the prevalence of overweight and obesity between those who participated and those who did not (144). However, selection bias could not be ruled out based on this sub-study as the participation rate was 38%. The similar prevalence of overweight and obesity documented in other population-based samples from the same time period strengthens the case of generalizability in terms of weight-status for the BGS1 (161, 162). The prevalence of overweight (including obesity) was 14.6% in BGS2 participants and very similar to data from the girls aged 8 to 15.5 years in BGS1 (14.3%) (163). A population-based study of 8- to 9-year-old Norwegian girls, found a slightly higher prevalence of overweight and/or obesity of 17% in 2015 (164). Regarding the prevalence rate of overweight and obesity we believe that our sample is fairly representative of the current Norwegian female child population.

The difference in participant rates between the BGS1 and the BGS2 may be partly attributable to their different aims, (i.e. puberty vs. normal growth and development). One can speculate that girls with late or early puberty compared to their peers may not have consented to participate or that some girls were reluctant to

participate due to being unhappy with their physical appearance in general. These are potential sources of a self- selection bias that most pubertal studies will encounter.

Socioeconomic status

A recent survey of living conditions from 2016 found that the Bergen municipality is overall very similar to the rest of the Norwegian population in terms of its socioeconomic structure (165). Parental educational level has been shown to be a valid variable for socioeconomic status in Norway (166). The prevalence of parents with a higher level of education was 79.6% in the BGS2 versus 62.4% in the BGS2, 40.4% in the population in Bergen in 2016 (167) and 34.1% in the general Norwegian population (168). Overrepresentation of participants with a higher education level has been reported in other studies as well (169-171). In the BGS1, the prevalence of overweight (including obesity) was higher in children of parents with a lower educational level (22% vs. 14%) (172). A similar trend was found in the BGS2 (19% vs. 11%, unpublished data). However, we could not find evidence that that the parental education level had a statistically significant impact on the occurrence of menarche in either the BGS2 or BGS1, which is in line with the findings of other relevant studies (34, 67, 105, 173).

Ethnic origin

The study population from BGS2 reflects the overall composition of ethnicity/race in the population of Norwegian girls in 2016 (Table 4 (174)), and data from all girls were included in the estimation of references. Although this approach might increase the applicability of the references in the Norwegian population as a whole, it could also be challenging to compare with follow-up studies with other ethnical compositions. Differences in pubertal timing according to ethnicity are well-known (33) which is why we also carried out analysis stratified for parental origin to test the potential impact on our data. There was a tendency for earlier puberty in non-Norwegian girls, but exclusions based on ethnicity had only minor impact on the lower cut-offs (-2SD) for defining precocious puberty (breast development before the age of 8.0 years including all ethnicities, and 8.1 for girls with a known Norwegian origin (not published)). When studying changes in the timing of menarche from BGS1 to BGS2,

only girls with a Norwegian origin were included to avoid interference from changes in the ethnic background. Girls with a Nordic origin (1.5%) were also included in the sample from the BGS1 because the Nordic girls were included in all growth references from the BGS1.

Table 4 *The origin of non-Norwegian girls in the study sample and in the entire population in Norway in 2016 (174).*

	Europe (%)	**Africa (%)**	**Asia (%)**	**North-America (%)**	**South-America (%)**	**Oceania (%)**	**Other**
In Norway							
6-15 years	5.9 %	3.1%	5.8 %	0.1 %	0.3%	0.01%	9.4%
In the BGS2							
6-16 years	5.7 %	2.6%	6.4%	0.0%	1.5%	-	-

Chronic illness

Girls with known chronic illnesses likely to affect growth and development were excluded. However, any participant with a chronic illness who did not complete a questionnaire that disclosed such information may have been included. Girls with chronic illness are known to enter puberty later (97), and this may have influenced our estimates towards later ages. However, the exclusion/inclusion of children with known chronic illness showed non-significant and inconclusive results, but the numbers were very small.

To conclude, we believe that the participants in Paper II are fairly representative of the contemporary population of children and adolescents in Norway with respect to ethnicity and weight status. Although the socioeconomic status of BGS2 participants was higher, this potential difference did not seem to impact on our results. Thus, we believe the data can reliably indicate the pubertal timing and may be generalized to the whole of Norway.

5.1.2 Ultrasound procedure

In Paper I, the ultrasound procedure (staging and direct measurements of glandular breast tissue) was performed two times by the main observer (ISB). It may be argued that the 20 to 35 minutes interval between examinations was not sufficient to exclude recall bias for the intra-observer agreement. However, all 57 girls in the test-retest study were examined over three busy days, which implies a large amount of similar measurements were obtained within a limited time frame. For this reason, we believe that the risk of recall bias was small, although it cannot be entirely excluded. In terms of external validity, it is of note that the results for agreement in this study may be different for other observers in other settings, due to factors such as the experience and training of the observers (175). Thus, a limitation of our test-retest study was that the assessments were performed by two observers only. However, a study including more than two observers would have posed an added burden on the participating girls and also introduced some logistical challenges.

Tanner B assessments by ISB were included in the examinations of each girl in Paper I and Paper II. The Tanner B assessments were done prior to the investigation with ultrasound which may have introduced a systemic bias toward better agreement between ratings. Ideally, the assessments of breast development by ultrasound should have been done by someone who was blinded for the Tanner B examination. However, the Tanner B assessments were done blinded for the examinations with ultrasound, and the recorded Tanner B stages were not changed 'retrospectively'. A second scoring session where ISB re-scored 122 ultrasound images blinded for previous scorings, found a good agreement of 0.76 between the initial and new ultrasound stages. In a clinical setting, an observer performing ultrasound assessment is not likely to be blinded for the visual appearance of the breasts either. We believe that the preceding Tanner B assessment and postponed staging procedure did not have a large impact on our results regarding ultrasound breast staging, although this cannot be ruled out.

Characterizing pubertal breast development with ultrasound

Six distinct ultrasound-based stages were described in Paper I. Systems created to characterize biological features may be challenging due to inter-individual variations

in appearances, i.e. that the observed morphological appearance does not necessarily 'fit' into a predefined scheme/protocol. In our study we found it necessary to establish two stages that were assumed to be prepubertal to accommodate the different appearances. Since ultrasound can only assess the appearance and not the nature of tissues, we were not able to determine potential physiological differences between US B0 and US B1. It is, however, reasonable to believe that the white areas on the image in US B1 contain more/different amounts of fat as compared to the small, dark area that defines US B0.

Direct measurements

We used a 5-cm long linear transducer for all ultrasound assessments in Papers I and II. For all diameters larger than 5 cm but less than 10 cm we combined the results of two scans. Extending this procedure for diameters of more than 10 cm to capture the whole breast diameter was considered too difficult to do in a reliable and standardized way without introducing a high degree of measurement error. While the truncation at 10 cm would be an important aspect to consider when presenting descriptive data on diameter (or volume) of glandular tissue in more mature breasts, we do not believe that it is a major problem for the purpose of Paper I because the Bland-Altman plots showed that measurement errors were largely independent of size (at least up to 10 cm).

5.1.3 Tanner staging

The pubertal assessments of breast and pubic hair development in Paper II was based on Tanner staging. According to recommendations from reviewers and expert opinions, we included breast bud palpation into the Tanner B staging (2, 35). Ideally we should have included an assessment of observer agreement of Tanner B and PH staging as well, but the focus of our study was on ultrasound assessment. The skill of the observer is important when assessing breast development, and observer agreement for Tanner B staging has shown to be highly variable (39-41). However, not all studies specified whether breast bud palpation was used, and the intra-observer agreement was rarely tested. Due to the cross-sectional design of our study, we do not know if girls

had transient thelarche, which has been shown to be a relatively frequent phenomenon (12%) in one Danish study (176).

Tanner B staging including palpation was validated against MRI in 100 healthy girls, and the study concluded that the method could reliably detect glandular breast tissue development (25). The study nurse in our study (ISB) received training by an experienced paediatrician before the examinations. In addition, the drawings and descriptions of the Tanner B stages were always available for the observer. We believe that this step improved the validity of the assessments. The issue of intimate shaving practices came up frequently, and may have biased the assessments of the Tanner PH stage, in particular the higher stages (and thus ages). However, the visual inspection gives a fair idea of how widespread the pubic hair development is and where the 'hairline' stops.

5.1.4 Menarche

Data on the occurrence (yes/no) of menarche from BGS2 and BGS1 was used in Papers II and III, while recall age (year/month) from BGS1 was also used in Paper III. In the BGS1, girls with visual signs of breast development were asked for occurrence and recalled age at menarche, while all girls in the BGS2 were asked. The accuracy of the recalled age at menarche is probably dependent on the recall interval, with a higher accuracy for shorter intervals (177). The time since menarche varied from 1 day to 4.7 years in the BGS1. Possible recall bias was tested by comparing the probit mean age found in all girls and in the group of girls with recent menarche (during the last 12 months) in Paper III, and we found only a small difference (0.1 year). Thus, we expect our data on menarche to be valid. Descriptive age references for menarche were based on status quo data in Paper II.

5.1.5 Anthropometric measurements

In the BGS2, the same nurse performed all anthropometric measurements (height and weight used in this study). The observer agreement in performing the anthropometric measures was tested in test-retest studies in BGS1 and found to be satisfactory. Both the WC and skinfold (178) are regarded as technically difficult assessments, and the

limited number of trained observers is therefore considered a strength of these procedures.

5.1.6 Questionnaire data

The response rate for the questionnaire data in the BGS1 and BGS2 were relatively good, with rates of 67% in the BGS1 and 75.7% in the BGS2. In Paper II, questionnaire data were used to obtain information about parental ethnic origins, chronic illness, parental educational level, and the girl's occurrence of (and age at) menarche.

5.1.7 Age range

The inclusion of girls from 6 years and onwards was sufficient to cover the age where no girls had yet started breast or pubic hair development. The inclusion of girls above 16 years would probably have improved the precision (with a smaller statistical uncertainty) of our estimates of the ages when girls attained stages 4 or 5. Including girls above this age would also have required data sampling in high schools, and our priority was instead to cover the stages nearer to the pubertal onset.

5.1.8 Statistical considerations

Cross-tables were included to show agreement of repeated measures of ultrasound staging in Paper I and agreement of different methods in Paper II. These tables were included because the kappa coefficient does not provide information about the nature of the disagreements, such as whether they appear systematic. This is visualised in cross-tables.

Generalized additive models (GAM) are extensions of generalized linear models (GLM). These nonparametric models are very flexible and do not depend on a predefined shape (e.g. logistic, probit, linear) of the curve, but are also more susceptible to overfitting, and extrapolation beyond the range of actual observations should be avoided. To avoid overfitting, we relied on cross-validation to determine the degree of smoothing, and visually inspected smoothness of the GAM curves and fit of the data (observations) together. Extrapolated age estimates outside the data range

were omitted from the tables when they occurred before the age of 6 or above the age of 16 years (Table 1 in Paper II).

Data on anthropometric measurements and the occurrence of menarche were collected at the same time in BGS1, and menarche could have occurred years before the measurements were taken. This introduced methodological challenges in paper III because the measurement were not obtained at the same time point as menarche. An increase or decrease in anthropometric measurements may have happened after menarche. Additional analyses showed that the same pattern of associations was found for girls who had experienced menarche during the last year, i.e. close to the actual measurement. A similar approach has been used in other studies, with also similar results in both groups (67).

5.1.9 Ethical considerations

Ethical issues arise when inviting children to participate in research. The Declaration of Helsinki states that 'no individual capable of giving informed consent may be enrolled in a research study unless he or she freely agrees' (156). In this study, we performed examinations that may be perceived as intimate. The children were informed about their individual right and also their ability to withdraw at any time from the examinations independently of their parents' consent for participation. All children were rewarded with a cinema ticket for their participation, but were not further encouraged or persuaded to participate if there was doubt about their consent.

5.2 Discussion of results

5.2.1 Observer agreement

To establish a relatively new method to determine breast development and to estimate the associated reference values, we had to document observer error. We did this by evaluating the intra- and inter-observer agreement. We have used the term *agreement* when referring to the testing and quantification of how close two assessments were when performing ultrasound-determined morphological staging and/or when performing direct measurements of the same breast by one or two observers (175). We

also used agreement when referring to the comparison of the two methods (US B and Tanner B) used to assess pubertal breast development (Paper II). However, repeatability of the ultrasound method was only tested in Paper I.

According to the common interpretation of kappa values (151), our ultrasound scoring system performed 'very good' for one (kappa 0.81; 95% CI 0.78–0.91), and 'good' (kappa 0.71; 95% CI: 0.62–0.80) for two observers when using the ultrasound scoring system. Comparisons of the kappa coefficients *only* between studies are of limited interest, because they are dependent on the distribution of stages (179), which is rarely reported, and different inclusion criteria are applied in the studies (age range, level of maturation, clinical assessment or photos, etc.). With these limitations in mind, we compared our results with earlier studies that tested the inter-observer agreement in Tanner B staging and found kappa values that varied between 0.50 (n = 25) (41), 0.67 (n = 127) (37), 0.78 (n = 22 and n = 20) (38, 39), and 0.86 (n = 56) (40). The kappa values kappa values achieved in our study were similar or higher compared to most of these previous studies that tested observer agreement. The main discrepancies between observers were between the two prepubertal stages (US B0 and US B1), and merging these two stages (and therefore making the scale more similar to the Tanner B system with five stages) did improve the kappa coefficient for intra- (kappa from 0.81 to 0.85) and inter- (kappa from 0.71 to 0.78) observer agreement slightly. Because merging these stages had a limited impact on the kappa coefficient, we chose to use both when comparing the staging with ultrasound with the staging according to Tanner B in Paper II.

Further, we demonstrated that one and two observers had very good concordance, (i.e. corresponding scores), when determining if breast development had started or not with ultrasound (US B0/US B1 vs. US B≥2). According to existing literature that questioned the reliability of Tanner B staging to determine whether breast development had started or not, detection of this transition may be the most important advantage of new methods. Nevertheless, Paper I provides novel and previously unavailable data on observer agreement when using ultrasound to stage breast development, so no other studies are available for a direct comparison. Because

the kappa values were good and the difference between assessments was rarely more than one stage within (n = 1) or between observers (n = 4) in Paper I, we reasoned that all six stages could be used reliably to characterize pubertal breast development with ultrasound.

The measurements of glandular size obtained by one and two observers were found to have no or small systematic differences but were imprecise. The variance for both depth and diameters seemed homogeneous across all measurements, and implied that the difference between measures was relatively smaller for larger depths and diameters than for the smaller depths and diameters. We concluded that we could not use the direct measurements in our study.

5.2.2 Method comparison

In Paper II, ultrasound staging of glandular tissue was further investigated by comparing it to the standard Tanner B method used to assess breast development. We found a high level of agreement between the two methods in determining pubertal onset (kappa 0.94; 96% CI: 0.9-0.96). We suggested, in Paper I that using two prepubertal ultrasound stages could potentially allow the detection of early signs of breast development during the prepubertal period. When investigating the different US B stages within the prepubertal Tanner B1 (n = 313), we found that 166 (53%) girls had US B0, 129 (41.2%) girls had US B1, and that 18 (5.8%) girls were classified as having US B2, indicating a pubertal appearance of glandular tissue with ultrasound. In contrast, 3 (3.8%) girls had signs of pubertal breast development according to Tanner B staging (Tanner B≥2) with no glandular breast tissue visible on ultrasound (US B0 or US B1). When estimating the age references, there was a difference in median age between the methods, with ultrasound-determined pubertal onset (US B≥2) being 0.2 years earlier than that identified with Tanner B≥2. This result was opposite of what we expected based on the literature, which suggested that Tanner B staging, particularly without palpation, may overestimate the appearance of breast development (180, 181). On the other hand, ultrasound allows to directly observe the glandular tissue and may (theoretically) allow to detect small psysiological changes before they become apparent by visual inspection. In a recently submitted paper, we found that the girls

with Tanner B1 and US B≥2 (of who 14 girls with blood samples), was associated with significantly higher levels of E2 compared to girls with (prepubertal) US B0 or US B1 (182).

We were, however, expecting to find prepubertal breast tissue with ultrasound in a higher number of girls that were assessed as pubertal with Tanner B (Tanner B≥2). This was not the case, and only happened in three cases. Although our hypothesis was related to all girls, we thought that particularly girls with overweight would be classified as pubertal with Tanner B and prepubertal with ultrasound. The overall prevalence of girls with overweight and obesity was, however, low, and none of these (three) girls had overweight or obesity.

In the studies that have applied ultrasound to stage breast development, inclusion has mostly been based on the visual appearance of breast development (by parents) or appearence of breast development according to Tanner B staging (by clinicians) (42, 53-55), i.e. included girls that had entered puberty. Therefore, these previous findings comparing the ultrasound and Tanner B staging around the time of pubertal onset is of limited value with respect to our unexpected findings. However, Bruni et al. (26) performed both Tanner B and ultrasound staging in healthy girls ranging from Tanner B1-B4 stages and found that 6 out of 48 girls presented with Tanner B1 and a 'Stage 2' (i.e. "first appearance of the glandular bud"). Of note, no report was made about palpation when assessing the Tanner B stage. In our study, all girls were palpated to detect the presence of glandular tissue, and we do not believe this unexpected discrepancy was due to assessment error.

Previous studies that compared ultrasound-determined breast maturation with Tanner B stages did not find that the staging by each of the methods were necessarily appearing simultaneously (26, 42, 53). The two methods (US B and Tanner B) had an overall good agreement when determining breast maturation throughout puberty as shown in Paper II. However, there were several discrepancies, also for the higher maturational stages. For the Tanner B2 to B4 stages, both higher and lower US B stages were observed, indicating different grades of maturation. The methodological

differences were also evident when comparing the estimated median ages at attaining corresponding breast stages from each method. The difference varied from 0.1 years (for US B3 and Tanner B3) to -1.2 years (for US B5 and Tanner B5). In Paper I, we found that agreement within and between observers improved when dichotomizing the stages as prepubertal or pubertal. The morphological appearances of pubertal breast tissue that correspond to US B2, US B3 and US B4 may be perceived as similar, recognised by a subareolar hypoechoic area (Figure 2 of this thesis). Consequently, stages 3 and 4 have been described as coinciding in the literature (47, 183, 184). Merging Tanner B stages into less narrow categories such as prepubertal, in- puberty and (near) adult phases, has also been suggested for self-assessments of pubertal development with the intention to increase accuracy (185). Further, Tanner B4 and B5 are regularly combined when reporting findings (72). This might be because both stages are regarded as mature, that not all girls experience Tanner B4 before B5 (1) and because the distinction between these two stages may be of limited clinical importance. However, in the aforementioned (unpublished) paper from this research group, we did find that the good agreement that was established by the kappa coefficient between the two methods was also confirmed when comparing hormonal values for the pubertal girls (comparing Tanner B2-5 and US B2-5), which demonstrated that these two methods seem to be endocrinologically comparable, across the whole range of stages (182).

The observed discrepancies between the methods may not be surprising, as one approach is based on the outer characteristics while the other is based on the morphological appearances observed with ultrasound. Neither method was regarded as a gold standard. Therefore, we have been reluctant to propose or hypothesise that one of both methods misinterprets the glandular maturation but instead suggest that they may provide different information.

5.2.3 Ultrasound and Tanner staging and adipose tissue

Adipose tissue in breasts without actual breast maturation (lipomastia) is probably the main reason of misclassifications with Tanner B staging of breast development. This concern seem to be based more on ‘clinical empirical’ experiences than on

scientifically observed data, since only few studies have actually compared the observer agreement of Tanner B staging in girls with overweight and/or obesity to girls of a normal weight (37, 38). These studies did not find any significant differences.

We observed a small difference in agreement between girls with a normal weight and girls who were overweight when comparing the two methods (kappa coefficients were 0.88 (95% CI: 0.86–0.91) and 0.85 (95% CI: 0.79–0.90), respectively). Full concordance between US B and Tanner B staging was found in 81.6% and 73.4% of girls with a normal weight and overweight/obesity, respectively. In total, 94 out of 644 girls with a known weight status had overweight or obesity (47 had underweight), and these girls were spread over a 10-year age span and across all five stages of pubertal development. We were therefore reluctant to use this limited group to draw conclusions on how ultrasound staging performed on girls with overweight vs. normal weight. However, we did find that the Tanner B staging tended to overestimate the breast development compared to the corresponding US B stage for the same breast in girls with overweight (including obesity). Increased inconsistency with increasing BMI has also been found in other studies comparing the ultrasound staging and Tanner B staging (55). Interestingly, Fugl et al. (25) observed that girls with an increased BMI had *less* glandular breast tissue (as estimated using MRI) compared to girls with a normal weight in the mature stages. The study suggested that the clinical examination could not distinguish between glandular tissue and fat in more mature girls and indicated that in girls with a higher BMI, the Tanner B stage may be overestimated. Longitudinal studies in samples emphasizing girls with normal weight and overweight would be ideal to find out more about how the ultrasound-determined breast development appears in these groups, and whether there are differences in how they relate to Tanner B staging.

5.2.4 Pubertal references compared to current references

Compared to the current pubertal references used in Norway based on Danish data from 1991–1993, our estimated ages at onset for breast (B2) and pubic hair development (PH2) and menarche were advanced by 0.5, 0.4 and 0.7 years (27),

respectively (Table 5). The Tanner B3 and B4 stages occurred 0.9 and 0.8 years earlier (27), indicating that the entire distribution has shifted towards younger ages when compared with current references. Danish data from the same group in 2006 also found that all stages of breast development occurred earlier; however, they observed a steeper decrease in pubertal onset (the average age at onset of the Tanner B2 stage was 9.8 years) during a shorter time (72). Evidence suggests a larger variance for the onset of puberty than in the age at menarche. When comparing ages at pubertal onset with relevant studies from Europe, we could not find that Norwegian girls matured at particularly young ages (e.g. Tanner B2 at 10.7 years in Belgium (30), 10.0 years in Greece (73), 10.7 years in the Netherlands (31) and 10.1 years in British girls (15).

Table 5 *Mean ages at reaching breast and pubic hair development and menarche in 1100 healthy Danish girls and 703 healthy Norwegian girls investigated from 1991–1993 and in 2016, respectively. The results are presented as the mean and lower and upper bounds for the probit analyses for Tanner B and PH, and Generalized additive model for menarche from the BGS2.*

	Copenhagen Puberty Study, Denmark (27) (N = 1100)		**Bergen Growth Study 2, Norway (N = 703)**	
	Mean	**± 2 SD**	**Mean**	**± 2 SD**
Tanner B2	10.9	8.7–13.1	10.4	8.0–12.7
Tanner PH 2	11.3	9.3–13.3	10.9	8.5–13.3
Menarche	13.4	11.2–15.7	12.7*	11.0–16.2

Abbreviations: Tanner B2: Tanner breast stage 2; Tanner PH 2: Tanner pubic hair stage 2; SD: standard deviation. *median age at menarche

5.2.5 Secular trend in the age at menarche

Considering the increased prevalence of paediatric overweight and obesity in Norway during the 30 years preceding the BGS1 (117), a lower age at menarche would not be unexpected due to the suggested association between overweight and earlier pubertal timing. However, the age at menarche had remained relatively stable just above 13 years in the BGS1 (Paper III). A trend with a stable age at menarche despite an increase in the prevalence of overweight has also been found in e.g. Germany and

Belgium (30, 67). However, a statistical significant decline in age at menarche of 0.1 year from 1997 to 2009 and 0.3 years from 1991 to 2006 was observed in the Netherlands and Denmark, respectively (72, 105). In the Danish study, however, the significance did not hold after adjusting for differences in BMI. They, like us, could not find large differences in the BMI between the two studies (72). However, the Dutch study found, similar to us, that the difference in occurrence of menarche remained significant, also when adjusting for BMI SDS.

Until 1985 (from 1920) data on growth, development and menarche in Oslo schoolgirls were systematically recorded and processed (65, 186). No Norwegian studies have compared ages at menarche in a similar way after this. Therefore, a comparison with previously published ages at menarche had to be based on various studies in Paper III. The comparison we were able to perform in Paper II was therefore particularly interesting since it was the first to study the age at menarche in comparable cohorts in almost 35 years in Norway (65). Although the decline was more pronounced when all girls from both studies (the BGS1 and the BGS2) were compared, the decline of 2.8 months was still significant when limited to Norwegian girls and adjusted for the BMI SDS in Paper II. It is also noteworthy that the BMI in the BGS2 was slightly lower compared to that in the BGS1, which suggest that the change is not explained by increased overweight.

Is there a renewed trend towards earlier maturation among Norwegian girls? Our estimated age of 12.7 years seems to be lower than the age reported in most studies conducted within Norway in recent years. The North-European countries have shown a limited age range from 12.8–13.13 years during the last two decades (15, 30, 67, 72, 105). A recently published population-based study including 312,656 Norwegian women born between 1936 and 1964, recalled ages at menarche that ranged from 13.18–13.42 years with a u-shaped pattern between the oldest and youngest women suggesting that the age at menarche fluctuated within a limited range, but had not declined (66). When reviewing the published ages at menarche in Norway, our estimate for girls with a known Norwegian origin (mean age: 13.1 years) was still in the lower range compared to previous estimates (Figure 5). However, larger

population-based studies following similar designs and conducted in different geographical areas within Norway should be carried out to explore how the trend of pubertal timing develops in Norway.

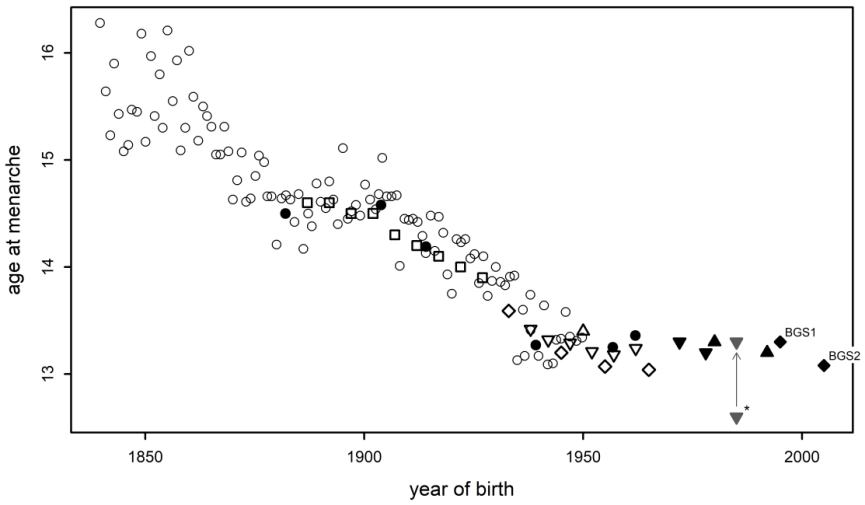


Figure 5 *An overview based on Norwegian studies including estimates on age at menarche for children, adolescents or women born from 1840–2005 (58, 65, 66, 78, 187-194). Open markers are retrospective surveys in adult women. Filled markers are studies with schoolchildren. *Arithmetic mean of reported age at menarche according to Lien et al. (188), and mean age at menarche calculated with probit regression from the prevalence by year of age (marked with an arrow).*

Since the observed change in our study did not appear to be attributable to changes in BMI, further studies should also investigate the potential role of EDCs on pubertal timing in Norway. Interestingly, the ages at menarche were positively skewed in the BGS2 but not in the BGS1. EDCs have frequently been implicated when trying to find explanations for changes in pubertal timing. Parent et al. recently suggested that changes in the age distribution for both earlier (Tanner B2) and later (menarche) markers may be due to the influence of EDCs (97).

5.2.6 Tempo of maturation versus adiposity

In Paper III, we speculated that if menarche is triggered or associated with an increase in fat mass, menarche would be more strongly associated with skinfolds and waist circumferences because these are more direct measures of fat tissues, while BMI also measures lean mass and fat mass. However, the risk of an earlier age at menarche among girls with an elevated BMI SDS could not be explained by increased measures of WC, SSF or TSF in our study. Thus, our results were not suggestive of a stronger association between central fat or subcutaneous fat mass and the timing of menarche than for the BMI SDS and the timing of menarche. Although the timing of menarche was related to all weight-related anthropometric measurements, the age at menarche had remained stable before the BGS1, despite an overall increase in overweight and obesity. Therefore, we believe that the strong association between timing of menarche and these measurements to be related to the tempo of maturation more than adiposity.

The maturational effect on BMI was also recently shown by Sorensen and Juul (195), by estimating adiposity with BMI and body fat percent (from bioelectrical impedance analyses). They found that adiposity was not significantly increased in early maturing girls compared to normal and late maturing girls within the same pubertal stage, although the BMI SDS were higher among the early maturers (195).

Interestingly, low levels of SSF and not TSF tissue remained significantly associated with a later timing of menarche in the fully adjusted model. Low adiposity has been related to later pubertal timing and is regularly seen in girls with anorexia nervosa or frequent and/or intense physical activity (196). However, the aforementioned Danish study warned that the BMI SDS may underestimate adiposity in pubertal children with late maturation (195). Our finding that a low level of BMI SDS was significantly associated with a late timing of menarche also when adjusting for the level of SSF SDS suggests that there is an association between truncal subcutaneous thinness and later maturation. Interestingly, the more ‘peripheral’ measure of subcutaneous tissue, TSF, lacked significance.

6. Conclusions

The following conclusions are drawn in accordance with the specific aims and hypothesis of the thesis:

- Ultrasound can reliably be used to assess pubertal breast development on a scale from 0 to 5 in healthy girls. The ability of the ultrasound method to determine if girls had signs of pubertal breast maturation or not, was very good both within and between observers.

- Direct measurements of the size and volume of glandular breast tissue lacked the necessary precision to be of use in clinical practice.

- The stage of breast development according to ultrasound was in overall good agreement and often in accordance with the Tanner B staging. Contrary to what we hypothesised, the onset of breast development was detected earlier when assessed with ultrasound compared to the Tanner B method.

- Pubertal development started earlier than what the current pubertal references imply, and this finding was expected with regards to the current trends in pubertal timing that have been observed in other Western countries. Norwegian girls do not seem to enter puberty significantly earlier than their peers in neighbouring countries.

- A decline in the age at menarche between the BGS1 and the BGS2 was observed. The decline was significant, also when adjusting for body mass index and ethnic background. This finding warrants further investigation in population-based studies to monitor ongoing changes in pubertal timing and to find potential causes for this trend.

- Measures of BMI, subcutaneous fat (skinfolds) and central fat (WC) were all associated with earlier and later timing of menarche. In a fully adjusted analysis

of all the measurements together, only a high BMI was related to earlier menarche, and a low BMI and a low SSF were associated with later menarche. This outcome was not what we hypothesised, and we concluded that the association between BMI and the timing of menarche may reflect the degree of maturation more than the adiposity.

7. Future perspectives

Ultrasound has emerged as a feasible method to determine the breast developmental stage. It may be used as an additional assessment when in doubt about whether breast development has started in a girl after a clinical examination. In case of longitudinal follow-up of girls to monitor their pubertal progression, ultrasound facilitates a retrospective analysis of previous ultrasound images and a comparison with the current appearance. For these applications, an ultrasound-based staging system and references will help guide radiologists or paediatric endocrinologists in interpreting the findings. Another advantage of ultrasound is that it is readily available. However, the method may be perceived as technically more challenging to perform than Tanner B staging, and practice is required for proficiency. Our findings apply to healthy girls and need to be replicated in other populations, such as in girls with overweight and obesity.

The study suggests that ultrasound is more sensitive to the early changes in breast tissue before pubertal onset. However, longitudinal studies of breast development from before pubertal onset will also be needed to clarify how these different appearances relate to the pubertal onset and if all girls experience all stages. Such studies would also have the potential to shed light on the discrepancies between ultrasound staging and Tanner B staging around the time of pubertal onset.

This research also has clinical implications for the evaluation of puberty in paediatric and adolescent female patients in Norway by providing the first Norwegian pubertal references. Very few studies have included the assessments of pubertal development other than menarche in Norway, and our results provide a necessary scientific base that can be used for comparison in future studies. Our study support the finding that the pubertal timing (including menarche) may still be changing. Changes in growth and maturation may reflect conditions that also affect the health and development of the overall society, and should be of interest and concern for health policy makers. Routine measurements of school-aged children are recommended in Norway. The third and last one takes place at the age of 13–14 years. No assessments on pubertal maturation are included in these measurements; however, asking the girls

if they have experienced menarche may be an easy way to monitor the pubertal timing, although this is a late sign.

It is important to identify possible causes and determinants of pubertal timing, and early puberty in particular, because many diseases and conditions have been linked to the timing of puberty. The decline in age at menarche did not seem to be explained by overweight. Another contributor to the trend may be EDCs, an issue that will soon be explored in blood samples collected in the BGS2. Further, most children (including approximately 500 girls) who participated in the BGS2 agreed that we could use their previously registered data on growth in infancy and childhood (0–6 years), so that we may link these data to the pubertal status in our study.

References

1. Marshall WA, Tanner JM. Variations in pattern of pubertal changes in girls. Archives of Disease in Childhood. 1969;44(235):291-303.
2. Abreu AP, Kaiser UB. Pubertal development and regulation. The lancet Diabetes & endocrinology. 2016;4(3):254-64.
3. Abbassi V. Growth and normal puberty. Pediatrics. 1998;102(2 Pt 3):507-11.
4. de Ridder CM, Thijssen JH, Bruning PF, Van den Brande JL, Zonderland ML, Erich WB. Body fat mass, body fat distribution, and pubertal development: a longitudinal study of physical and hormonal sexual maturation of girls. The Journal of Clinical Endocrinology & Metabolism. 1992;75(2):442-6.
5. Wohlfahrt-Veje C, Tinggaard J, Winther K, Mouritsen A, Hagen CP, Mieritz MG, et al. Body fat throughout childhood in 2647 healthy Danish children: agreement of BMI, waist circumference, skinfolds with dual X-ray absorptiometry. European journal of clinical nutrition. 2014;68(6):664-70.
6. Taylor RW, Grant AM, Williams SM, Goulding A. Sex differences in regional body fat distribution from pre- to postpuberty. Obesity. 2010;18(7):1410-6.
7. Boot AM, de Ridder MAJ, Pols HAP, Krenning EP, de Muinck Keizer-Schrama SMPF. Bone mineral density in children and adolescents: relation to puberty, calcium intake, and physical activity. The Journal of Clinical Endocrinology & Metabolism. 1997;82(1):57-62.
8. Mølgaard C, Thomsen BL, Michaelsen KF. Whole body bone mineral accretion in healthy children and adolescents. Archives of disease in childhood. 1999;81(1):10-5.
9. Committee Opinion No. 651 Summary: Menstruation in Girls and Adolescents: Using the Menstrual Cycle as a Vital Sign. Obstetrics and gynecology. 2015;126(6):1328.
10. Flug D, Largo RH, Prader A. Menstrual patterns in adolescent Swiss girls: a longitudinal study. Annals of human biology. 1984;11(6):495-508.
11. Chellakooty M, Schmidt IM, Haavisto AM, Boisen KA, Damgaard IN, Mau C, et al. Inhibin A, inhibin B, follicle-stimulating hormone, luteinizing hormone, estradiol, and sex hormone-binding globulin levels in 473 healthy infant girls. The Journal of clinical endocrinology and metabolism. 2003;88(8):3515-20.
12. Reinehr T, Roth CL. Is there a causal relationship between obesity and puberty? The Lancet Child & adolescent health. 2019;3(1):44-54.
13. Cutler GB. The role of estrogen in bone growth and maturation during childhood and adolescence. The Journal of Steroid Biochemistry and Molecular Biology. 1997;61(3):141-4.
14. Sklar CA, Kaplan SL, Grumbach MM. Evidence for dissociation between adrenarche and gonadarche: studies in patients with idiopathic precocious puberty, gonadal dysgenesis, isolated gonadotropin deficiency, and constitutionally delayed growth and adolescence. The Journal of Clinical Endocrinology & Metabolism. 1980;51(3):548-56.
15. Rubin C, Maisonet M, Kieszak S, Monteilh C, Holmes A, Flanders D, et al. Timing of maturation and predictors of menarche in girls enrolled in a contemporary British cohort. Paediatric and perinatal epidemiology. 2009;23(5):492-504.

16. Ersoy B, Balkan C, Gunay T, Egemen A. The factors affecting the relation between the menarcheal age of mother and daughter. Child: Care, Health and Development. 2005;31(3):303-8.
17. Biro FM, Pajak A, Wolff MS, Pinney SM, Windham GC, Galvez MP, et al. Age of menarche in a longitudinal US cohort. Journal of pediatric and adolescent gynecology. 2018;31(4):339-45.
18. Wohlfahrt-Veje C, Mouritsen A, Hagen CP, Tinggaard J, Mieritz MG, Boas M, et al. Pubertal onset in boys and girls is influenced by pubertal timing of both parents. The Journal of Clinical Endocrinology & Metabolism. 2016;101(7):2667-74.
19. Maisonet M, Christensen KY, Rubin C, Holmes A, Flanders WD, Heron J, et al. Role of prenatal characteristics and early growth on pubertal attainment of British girls. Pediatrics. 2010;126(3):e591-600.
20. Sorensen S, Brix N, Ernst A, Lauridsen LLB, Ramlau-Hansen CH. Maternal age at menarche and pubertal development in sons and daughters: a Nationwide Cohort Study. Human reproduction (Oxford, England). 2018;33(11):2043-50.
21. Day FR, Thompson DJ, Helgason H, Chasman DI, Finucane H, Sulem P, et al. Genomic analyses identify hundreds of variants associated with age at menarche and support a role for puberty timing in cancer risk. Nature genetics. 2017;49(6):834-41.
22. Tanner JM. Growth at adolescence. 1962. 2nd ed. Blackwell, Oxford.
23. Dorn LD, Dahl RE, Woodward HR, Biro F. Defining the boundaries of early adolescence: a user's guide to assessing pubertal status and pubertal timing in research with adolescents. Applied Developmental Science. 2006;10(1):30-56.
24. Walker IV, Smith CR, Davies JH, Inskip HM, Baird J. Methods for determining pubertal status in research studies: literature review and opinions of experts and adolescents. Journal of developmental origins of health and disease. 2019:1-20.
25. Fugl L, Hagen CP, Mieritz MG, Tinggaard J, Fallentin E, Main KM, et al. Glandular breast tissue volume by magnetic resonance imaging in 100 healthy peripubertal girls: evaluation of clinical Tanner staging. Pediatric Research. 2016;80(4):526-30.
26. Bruni V, Dei M, Deligeoroglou E, Innocenti P, Pandimiglio AM, Magini A, et al. Breast development in adolescent girls. Adolescent and Pediatric Gynecology. 1990;3(4):201-5.
27. Juul A, Teilmann G, Scheike T, Hertel NT, Holm K, Laursen EM, et al. Pubertal development in Danish children: comparison of recent European and US data. International journal of andrology. 2006;29(1):247-55; discussion 86-90.
28. Juliusson PB, Roelants M, Eide GE, Moster D, Juul A, Hauspie R, et al. [Growth references for Norwegian children]. Tidsskrift for den Norske laegeforening : tidsskrift for praktisk medicin, ny raekke. 2009;129(4):281-6.
29. Barnelegeforening N. 2.4 Normal pubertet helsebiblioteket.no: Norsk barnelegeforening. Den norske legeforening; [Available from: https://www.helsebiblioteket.no/pediatriveiledere?menuitemkeylev1=5962&menuitemkeylev2=5964&key=144406].
30. Roelants M, Hauspie R, Hoppenbrouwers K. References for growth and pubertal development from birth to 21 years in Flanders, Belgium. Annals of human biology. 2009;36(6):680-94.

31. Mul D, Fredriks AM, van Buuren S, Oostdijk W, Verloove-Vanhorick SP, Wit JM. Pubertal development in The Netherlands 1965-1997. Pediatric Research. 2001;50(4):479-86.
32. Midyett LK, Moore WV, Jacobson JD. Are pubertal changes in girls before age 8 benign? Pediatrics. 2003;111(1):47-51.
33. Parent AS, Teilmann G, Juul A, Skakkebaek NE, Toppari J, Bourguignon JP. The timing of normal puberty and the age limits of sexual precocity: variations around the world, secular trends, and changes after migration. Endocrine reviews. 2003;24(5):668-93.
34. Biro FM, Greenspan LC, Galvez MP, Pinney SM, Teitelbaum S, Windham GC, et al. Onset of breast development in a longitudinal cohort. Pediatrics. 2013;132(6):1019-27.
35. Euling SY, Herman-Giddens ME, Lee PA, Selevan SG, Juul A, Sorensen TI, et al. Examination of US puberty-timing data from 1940 to 1994 for secular trends: panel findings. Pediatrics. 2008;121 Suppl 3:S172-91.
36. Ong KK, Ahmed ML, Dunger DB. Lessons from large population studies on timing and tempo of puberty (secular trends and relation to body size): the European trend. Molecular and cellular endocrinology. 2006;254-255:8-12.
37. Biro FM, Galvez MP, Greenspan LC, Succop PA, Vangeepuram N, Pinney SM, et al. Pubertal assessment method and baseline characteristics in a mixed longitudinal study of girls. Pediatrics. 2010;126(3):e583-e90.
38. Rasmussen AR, Wohlfahrt-Veje C, Tefre de Renzy-Martin K, Hagen CP, Tinggaard J, Mouritsen A, et al. Validity of self-assessment of pubertal maturation. Pediatrics. 2015;135(1):86-93.
39. Britton JA, Wolff MS, Lapinski R, Forman J, Hochman S, Kabat GC, et al. Characteristics of pubertal development in a multi-ethnic population of nine-year-old girls. Annals of Epidemiology. 2004;14(3):179-87.
40. Herman-Giddens ME, Slora EJ, Wasserman RC, Bourdony CJ, Bhapkar MV, Koch GG, et al. Secondary sexual characteristics and menses in young girls seen in office practice: a study from the Pediatric Research in Office Settings network. Pediatrics. 1997;99(4):505-12.
41. Hergenroeder AC, Hill RB, Wong WW, Sangi-Haghpeykar H, Taylor W. Validity of self-assessment of pubertal maturation in African American and European American adolescents. Journal of Adolescent Health. 1999;24(3):201-5.
42. Calcaterra V, Sampaolo P, Klersy C, Larizza D, Alfei A, Brizzi V, et al. Utility of breast ultrasonography in the diagnostic work-up of precocious puberty and proposal of a prognostic index for identifying girls with rapidly progressive central precocious puberty. Ultrasound in Obstetrics and Gynecology. 2009;33(1):85-91.
43. Venta LA, Dudiak CM, Salomon CG, Flisak ME. Sonographic evaluation of the breast. RadioGraphics. 1994;14(1):29-50.
44. Kreftregisteret. Kreftregisteret: Norwegian Cancer Registry.; 2020 [Available from: https://sb.kreftregisteret.no/insidens/].
45. Bock K, Duda VF, Hadji P, Ramaswamy A, Schulz-Wendtland R, Klose K-J, et al. Pathologic Breast Conditions in Childhood and Adolescence. Journal of Ultrasound in Medicine. 2005;24(10):1347-54.

46. Durmaz E, Oztek MA, Arioz Habibi H, Kesimal U, Sindel HT. Breast diseases in children: the spectrum of radiologic findings in a cohort study. Diagnostic and interventional radiology (Ankara, Turkey). 2017;23(6):407-13.
47. Gao Y, Saksena MA, Brachtel EF, terMeulen DC, Rafferty EA. How to approach breast lesions in children and adolescents. European Journal of Radiology. 2015;84(7):1350-64.
48. Kaneda HJ, Mack J, Kasales CJ, Schetter S. Pediatric and Adolescent Breast Masses: A Review of Pathophysiology, Imaging, Diagnosis, and Treatment. American Journal of Roentgenology. 2013;200(2):W204-W12.
49. Russo J, Russo IH. Development of the human breast. Maturitas. 2004;49(1):2-15.
50. Chen J-H, Gulsen G, Su M-Y. Imaging Breast Density: Established and Emerging Modalities. Translational Oncology. 2015;8(6):435-45.
51. Rapelyea J, Marks C. Breast Ultrasound Past, Present, and Future. In: Kuzmiak CM, ed. Breast Imaging. Croatia: Intech open; 2018: 21- 48.
52. García CJ, Espinoza A, Dinamarca V, Navarro O, Daneman A, García H, et al. Breast US in children and adolescents. RadioGraphics. 2000;20(6):1605-12.
53. Carlson L, Flores Poccia V, Sun BZ, Mosley B, Kirste I, Rice A, et al. Early breast development in overweight girls: does estrogen made by adipose tissue play a role? International Journal of Obesity. 2019;43(10):1978-1987.
54. Youn I, Park SH, Lim IS, Kim SJ. Ultrasound assessment of breast development: Distinction between premature thelarche and precocious puberty. American Journal of Roentgenology. 2015;204(3):620-4.
55. Yuce O, Sevinc D. Ultrasonographic assessment of pubertal breast development in obese children: compliance with the clinic. Journal of pediatric endocrinology & metabolism : JPEM. 2018;31(2):137-41.
56. Ramadan SU, Gokharman D, Kacar M, Kosar P, Kosar U. Assessment of vascularity with color Doppler ultrasound in gynecomastia. Diagnostic and interventional radiology (Ankara, Turkey). 2010;16(1):38-44.
57. Tanner JM. Trend towards Earlier Menarche in London, Oslo, Copenhagen, the Netherlands and Hungary. Nature. 1973;243(5402):95-6.
58. Brudevoll JE, Liestol K, Walloe L. Menarcheal age in Oslo during the last 140 years. Annals of human biology. 1979;6(5):407-16.
59. Wyshak G, Frisch RE. Evidence for a secular trend in age of menarche. The New England journal of medicine. 1982;306(17):1033-5.
60. Zacharias L, Wurtman RJ. Age at menarche. Genetic and environmental influences. The New England journal of medicine. 1969;280(16):868-75.
61. Rosenberg M. Menarcheal age for Norwegian women born 1830-1960. Annals of human biology. 1991;18(3):207-19.
62. Brundtland GH, Walloe L. Menarcheal age in Norway in the 19th century: a re-evaluation of the historical sources. Annals of human biology. 1976;3(4):363-74.
63. Tell GS. Cardiovascular disease risk factors related to sexual maturation: the Oslo Youth Study. Journal of chronic diseases. 1985;38(8):633-42.
64. Bratberg GH, Nilsen TI, Holmen TL, Vatten LJ. Early sexual maturation, central adiposity and subsequent overweight in late adolescence. a four-year follow-up

of 1605 adolescent Norwegian boys and girls: the Young HUNT study. BMC public health. 2007;7:54.
65. Liestol K, Rosenberg M. Height, weight and menarcheal age of schoolgirls in Oslo--an update. Annals of human biology. 1995;22(3):199-205.
66. Gottschalk MS, Eskild A, Hofvind S, Gran JM, Bjelland EK. Temporal trends in age at menarche and age at menopause: a population study of 312 656 women in Norway. Human reproduction (Oxford, England). 2020;35(2):464-71.
67. Bau AM, Ernert A, Schenk L, Wiegand S, Martus P, Gruters A, et al. Is there a further acceleration in the age at onset of menarche? A cross-sectional study in 1840 school children focusing on age and bodyweight at the onset of menarche. European journal of endocrinology. 2009;160(1):107-13.
68. Wu T, Mendola P, Buck GM. Ethnic differences in the presence of secondary sex characteristics and menarche among US girls: the Third National Health and Nutrition Examination Survey, 1988-1994. Pediatrics. 2002;110(4):752-7.
69. Sun SS, Schubert CM, Chumlea WC, Roche AF, Kulin HE, Lee PA, et al. National estimates of the timing of sexual maturation and racial differences among US children. Pediatrics. 2002;110(5):911-9.
70. Freedman DS, Khan LK, Serdula MK, Dietz WH, Srinivasan SR, Berenson GS. Relation of age at menarche to race, time period, and anthropometric dimensions: the Bogalusa Heart Study. Pediatrics. 2002;110(4):e43.
71. Kaplowitz PB, Slora EJ, Wasserman RC, Pedlow SE, Herman-Giddens ME. Earlier onset of puberty in girls: relation to increased body mass index and race. Pediatrics. 2001;108(2):347-53.
72. Aksglaede L, Sorensen K, Petersen JH, Skakkebaek NE, Juul A. Recent decline in age at breast development: the Copenhagen Puberty Study. Pediatrics. 2009;123(5):e932-9.
73. Papadimitriou A, Pantsiotou S, Douros K, Papadimitriou DT, Nicolaidou P, Fretzayas A. Timing of pubertal onset in girls: evidence for non-Gaussian distribution. The Journal of clinical endocrinology and metabolism. 2008;93(11):4422-5.
74. Eckert-Lind C, Busch AS, Petersen JH, Biro FM, Butler G, Bräuner EV, et al. Worldwide secular trends in age at pubertal onset assessed by breast development among girls: A systematic review and meta-analysis. JAMA Pediatrics. 2020:e195881-e.
75. Day FR, Elks CE, Murray A, Ong KK, Perry JR. Puberty timing associated with diabetes, cardiovascular disease and also diverse health outcomes in men and women: the UK Biobank study. Scientific reports. 2015;5:11208.
76. Cheng TS, Day FR, Lakshman R, Ong KK. Association of puberty timing with type 2 diabetes: A systematic review and meta-analysis. PLoS medicine. 2020;17(1):e1003017.
77. Charalampopoulos D, McLoughlin A, Elks CE, Ong KK. Age at menarche and risks of all-cause and cardiovascular death: a systematic review and meta-analysis. American journal of epidemiology. 2014;180(1):29-40.
78. Jacobsen BK, Heuch I, Kvale G. Association of low age at menarche with increased all-cause mortality: a 37-year follow-up of 61,319 Norwegian women. American journal of epidemiology. 2007;166(12):1431-7.

79. Garn SM, LaVelle M, Rosenberg KR, Hawthorne VM. Maturational timing as a factor in female fatness and obesity. The American journal of clinical nutrition. 1986;43(6):879-83.
80. van Lenthe FJ, Kemper CG, van Mechelen W. Rapid maturation in adolescence results in greater obesity in adulthood: the Amsterdam Growth and Health Study. The American journal of clinical nutrition. 1996;64(1):18-24.
81. Helm P, Munster K, Schmidt L. Recalled menarche in relation to infertility and adult weight and height. Acta obstetricia et gynecologica Scandinavica. 1995;74(9):718-22.
82. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. The Lancet Oncology. 2012;13(11):1141-51.
83. Clavel-Chapelon F. Cumulative number of menstrual cycles and breast cancer risk: results from the E3N cohort study of French women. Cancer Causes & Control. 2002;13(9):831-8.
84. Ge X, Conger RD, Elder GH, Jr. Pubertal transition, stressful life events, and the emergence of gender differences in adolescent depressive symptoms. Developmental psychology. 2001;37(3):404-17.
85. Striegel-Moore RH, McMahon RP, Biro FM, Schreiber G, Crawford PB, Voorhees C. Exploring the relationship between timing of menarche and eating disorder symptoms in Black and White adolescent girls. The International journal of eating disorders. 2001;30(4):421-33.
86. Patton GC, McMorris BJ, Toumbourou JW, Hemphill SA, Donath S, Catalano RF. Puberty and the onset of substance use and abuse. Pediatrics. 2004;114(3):e300-e306.
87. Mendle J, Ryan RM, McKone KMP. Age at menarche, depression, and antisocial behavior in adulthood. Pediatrics. 2018;141(1).
88. Salgin B, Norris SA, Prentice P, Pettifor JM, Richter LM, Ong KK, et al. Even transient rapid infancy weight gain is associated with higher BMI in young adults and earlier menarche. International journal of obesity (2005). 2015;39(6):939-44.
89. Ong KK, Emmett P, Northstone K, Golding J, Rogers I, Ness AR, et al. Infancy weight gain predicts childhood body fat and age at menarche in girls. The Journal of clinical endocrinology and metabolism. 2009;94(5):1527-32.
90. Wang Y, Dinse GE, Rogan WJ. Birth weight, early weight gain and pubertal maturation: a longitudinal study. Pediatric obesity. 2012;7(2):101-9.
91. Hvidt JJ, Brix N, Ernst A, Lauridsen LLB, Ramlau-Hansen CH. Size at birth, infant growth, and age at pubertal development in boys and girls. Clinical epidemiology. 2019;11:873-83.
92. Adair LS. Size at birth predicts age at menarche. Pediatrics. 2001;107(4):e59-e.
93. Brix N, Ernst A, Lauridsen LLB, Parner ET, Olsen J, Henriksen TB, et al. Maternal smoking during pregnancy and timing of puberty in sons and daughters: A population-based cohort study. American journal of epidemiology. 2019;188(1):47-56.
94. Brix N, Ernst A, Lauridsen LLB, Parner ET, Arah OA, Olsen J, et al. Maternal pre-pregnancy body mass index, smoking in pregnancy, and alcohol intake in pregnancy in relation to pubertal timing in the children. BMC pediatrics. 2019;19(1):338.

95. James E, Wood CL, Nair H, Williams TC. Preterm birth and the timing of puberty: a systematic review. BMC pediatrics. 2018;18(1):3.
96. Schell LM, Gallo MV, Ravenscroft J. Environmental influences on human growth and development: historical review and case study of contemporary influences. Annals of human biology. 2009;36(5):459-77.
97. Parent A-S, Franssen D, Fudvoye J, Gérard A, Bourguignon J-P. Developmental variations in environmental influences including endocrine disruptors on pubertal timing and neuroendocrine control: Revision of human observations and mechanistic insight from rodents. Frontiers in Neuroendocrinology. 2015;38:12-36.
98. Biro FM, Wolff MS. Puberty as a window of susceptibility. In: Russo J, editor. Environment and Breast Cancer. New York, NY: Springer New York; 2011. p. 29-41.
99. Buck Louis GM, Gray LE, Jr., Marcus M, Ojeda SR, Pescovitz OH, Witchel SF, et al. Environmental factors and puberty timing: expert panel research needs. Pediatrics. 2008;121 Suppl 3:S192-207.
100. Mouritsen A, Aksglaede L, Sorensen K, Mogensen SS, Leffers H, Main KM, et al. Hypothesis: exposure to endocrine-disrupting chemicals may interfere with timing of puberty. International journal of andrology. 2010;33(2):346-59.
101. Moffitt TE, Caspi A, Belsky J, Silva PA. Childhood experience and the onset of menarche: a test of a sociobiological model. Child Development. 1992;63(1):47-58.
102. Wierson M, Long PJ, Forehand RL. Toward a new understanding of early menarche: the role of environmental stress in pubertal timing. Adolescence. 1993;28(112):913-24.
103. Teilmann G, Pedersen CB, Skakkebæk NE, Jensen TK. Increased risk of precocious puberty in internationally adopted children in denmark. Pediatrics. 2006;118(2):e391-e9.
104. Tahirović HF. Menarchal age and the stress of war: an example from Bosnia. European journal of pediatrics. 1998;157(12):978-80.
105. Talma H, Schonbeck Y, van Dommelen P, Bakker B, van Buuren S, Hirasing RA. Trends in menarcheal age between 1955 and 2009 in the Netherlands. PloS one. 2013;8(4):e60056.
106. Biro FM, McMahon RP, Striegel-Moore R, Crawford PB, Obarzanek E, Morrison JA, et al. Impact of timing of pubertal maturation on growth in black and white female adolescents: The National Heart, Lung, and Blood Institute Growth and Health Study. The Journal of Pediatrics. 2001;138(5):636-43.
107. Rosenfield RL, Lipton RB, Drum ML. Thelarche, pubarche, and menarche attainment in children with normal and elevated body mass index. Pediatrics. 2009;123(1):84-8.
108. Wang Y. Is obesity associated with early sexual maturation? A comparison of the association in American boys versus girls. Pediatrics. 2002;110(5):903-10.
109. Anderson SE, Dallal GE, Must A. Relative weight and race influence average age at menarche: results from two nationally representative surveys of US girls studied 25 years apart. Pediatrics. 2003;111(4 Pt 1):844-50.
110. Biro FM, Lucky AW, Simbartl LA, Barton BA, Daniels SR, Striegel-Moore R, et al. Pubertal maturation in girls and the relationship to anthropometric changes: pathways through puberty. The Journal of Pediatrics. 2003;142(6):643-6.

111. Wagner IV, Sabin MA, Pfäffle RW, Hiemisch A, Sergeyev E, Körner A, et al. Effects of obesity on human sexual development. Nature Reviews Endocrinology. 2012;8(4):246-54.
112. Ahmed ML, Ong KK, Dunger DB. Childhood obesity and the timing of puberty. Trends in Endocrinology & Metabolism. 2009;20(5):237-42.
113. Adair LS, Gordon-Larsen P. Maturational timing and overweight prevalence in US adolescent girls. American journal of public health. 2001;91(4):642-4.
114. Biro FM, Khoury PR, Morrison JA. Influence of obesity on timing of puberty. International Journal of Andrology. 2006;29(1):272-7.
115. Kaplowitz PB. Link between body fat and the timing of puberty. Pediatrics. 2008;121 Suppl 3:S208-17.
116. WHO. The World Bank Group joint child malnutrition estimates: levels and trends in child malnutrition: key findings of the 2020 edition. 2020. Contract No.: ISBN: 9789240003576.
117. Juliusson PB, Roelants M, Eide GE, Hauspie R, Waaler PE, Bjerknes R. Overweight and obesity in Norwegian children: secular trends in weight-for-height and skinfolds. Acta paediatrica (Oslo, Norway : 1992). 2007;96(9):1333-7.
118. Wattigney WA, Srinivasan SR, Chen W, Greenlund KJ, Berenson GS. Secular trend of earlier onset of menarche with increasing obesity in black and white girls: the Bogalusa Heart Study. Ethnicity and Disease. 1999;9(2):181-9.
119. Lee JM, Appugliese D, Kaciroti N, Corwyn RF, Bradley RH, Lumeng JC. Weight status in young girls and the onset of puberty. Pediatrics. 2007;119(3):e624-30.
120. Davison KK, Susman EJ, Birch LL. Percent body fat at age 5 predicts earlier pubertal development among girls at age 9. Pediatrics. 2003;111(4):815-21.
121. He Q, Karlberg J. Bmi in childhood and its association with height gain, timing of puberty, and final height. Pediatric Research. 2001;49(2):244-51.
122. Brix N, Ernst A, Lauridsen LLB, Parner ET, Arah OA, Olsen J, et al. Childhood overweight and obesity and timing of puberty in boys and girls: cohort and sibling-matched analyses. International Journal of Epidemiology. 2020;49(3):834-844.
123. Aksglaede L, Juul A, Olsen LW, Sorensen TI. Age at puberty and the emerging obesity epidemic. PloS one. 2009;4(12):e8450.
124. Frisch RE, Revelle R. Height and weight at menarche and a hypothesis of critical body weights and adolescent events. Science (New York, NY). 1970;169(3943):397-9.
125. Frisch RE, Revelle R, Cook S. Components of weight at menarche and the initiation of the adolescent growth spurt in girls: estimated total water, llean body weight and fat. Human biology. 1973;45(3):469-83.
126. Cameron N. Weight and skinfold variation at menarche and the critical body weight hypothesis. Annals of human biology. 1976;3(3):279-82.
127. Johnston FE, Roche AF, Schell LM, Norman H, Wettenhall B. Critical weight at menarche. Critique of a hypothesis. American journal of diseases of children (1960). 1975;129(1):19-23.
128. de Ridder CM, de Boer RW, Seidell JC, Nieuwenhoff CM, Jeneson JA, Bakker CJ, et al. Body fat distribution in pubertal girls quantified by magnetic resonance imaging. International journal of obesity and related metabolic disorders : journal of the International Association for the Study of Obesity. 1992;16(6):443-9.

129. Fernandez-Fernandez R, Martini AC, Navarro VM, Castellano JM, Dieguez C, Aguilar E, et al. Novel signals for the integration of energy balance and reproduction. Molecular and cellular endocrinology. 2006;254-255:127-32.
130. Ahmed ML, Ong KKL, Morrell DJ, Cox L, Drayer N, Perry L, et al. Longitudinal Study of Leptin Concentrations during Puberty: Sex Differences and Relationship to Changes in Body Composition1. The Journal of Clinical Endocrinology & Metabolism. 1999;84(3):899-905.
131. Grumbach MM. The neuroendocrinology of human puberty revisited. Hormone research. 2002;57 Suppl 2:2-14.
132. Demerath EW, Li J, Sun SS, Chumlea WC, Remsberg KE, Czerwinski SA, et al. Fifty-year trends in serial body mass index during adolescence in girls: the Fels Longitudinal Study. The American journal of clinical nutrition. 2004;80(2):441-6.
133. Cole TJ, Flegal KM, Nicholls D, Jackson AA. Body mass index cut offs to define thinness in children and adolescents: international survey. BMJ (Clinical research ed). 2007;335(7612):194.
134. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. BMJ (Clinical research ed). 2000;320(7244):1240-3.
135. Cole TJ, Lobstein T. Extended international (IOTF) body mass index cut-offs for thinness, overweight and obesity. Pediatric Obesity. 2012;7(4):284-94.
136. Wells JC. Toward body composition reference data for infants, children, and adolescents. Advances in nutrition (Bethesda, Md). 2014;5(3):320s-9s.
137. Maynard LM, Wisemandle W, Roche AF, Chumlea WC, Guo SS, Siervogel RM. Childhood body composition in relation to body mass index. Pediatrics. 2001;107(2):344-50.
138. McCarthy HD, Cole TJ, Fry T, Jebb SA, Prentice AM. Body fat reference curves for children. International journal of obesity (2005). 2006;30(4):598-602.
139. Javed A, Jumean M, Murad MH, Okorodudu D, Kumar S, Somers VK, et al. Diagnostic performance of body mass index to identify obesity as defined by body adiposity in children and adolescents: a systematic review and meta-analysis. Pediatric obesity. 2015;10(3):234-44.
140. Brannsether B, Eide GE, Roelants M, Bjerknes R, Juliusson PB. Interrelationships between anthropometric variables and overweight in childhood and adolescence. American journal of human biology : the official journal of the Human Biology Council. 2014;26(4):502-10.
141. Brannsether B, Roelants M, Bjerknes R, Juliusson PB. References and cutoffs for triceps and subscapular skinfolds in Norwegian children 4-16 years of age. European journal of clinical nutrition. 2013;67(9):928-33.
142. Taylor RW, Jones IE, Williams SM, Goulding A. Evaluation of waist circumference, waist-to-hip ratio, and the conicity index as screening tools for high trunk fat mass, as measured by dual-energy X-ray absorptiometry, in children aged 3-19 y. The American journal of clinical nutrition. 2000;72(2):490-5.
143. Brannsether B, Roelants M, Bjerknes R, Juliusson PB. Waist circumference and waist-to-height ratio in Norwegian children 4-18 years of age: reference values and cut-off levels. Acta paediatrica (Oslo, Norway : 1992). 2011;100(12):1576-82.

144. Juliusson PB, Roelants M, Nordal E, Furevik L, Eide GE, Moster D, et al. Growth references for 0-19 year-old Norwegian children for length/height, weight, body mass index and head circumference. Annals of human biology. 2013;40(3):220-7.
145. Bruserud IS, Roelants M, Oehme NHB, Eide GE, Bjerknes R, Rosendahl K, et al. Ultrasound assessment of pubertal breast development in girls: intra- and interobserver agreement. Pediatric Radiology. 2018;48(11):1576-83.
146. Juliusson PBV, S., Nilsen, B. Måling av vekst og vekt: en oversikt over anbefalte teknikker. Pediatrisk Endokrinologi. 2005;19.
147. Juliusson PB. Overweight and obesity in Norwegian children. Trends, current precalence, effect of socio-demographic factors and parental perception.: University of Bergen, Bergen, Norway; 2010.
148. Ernst A, Lauridsen LLB, Brix N, Kjersgaard C, Olsen J, Parner ET, et al. Self-assessment of pubertal development in a puberty cohort. Journal of pediatric endocrinology & metabolism : JPEM. 2018;31(7):763-72.
149. Shelmerdine SC, Di Paolo PL, Rieter JFMM, Malattia C, Tanturri de Horatio L, Rosendahl K. A novel radiographic scoring system for growth abnormalities and structural change in children with juvenile idiopathic arthritis of the hip. Pediatric radiology. 2018;48(8):1086-95.
150. Lambot K, Boavida P, Damasio MB, Tanturri de Horatio L, Desgranges M, Malattia C, et al. MRI assessment of tenosynovitis in children with juvenile idiopathic arthritis: inter- and intra-observer variability. Pediatr Radiol. 2013;43(7):796-802.
151. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics. 1977;33(1):159-74.
152. Martin Bland J, Altman D. Statistical methods for assessing agreement between two methods of clinical measurement. The Lancet. 1986;327(8476):307-10.
153. Altman DG, Bland JM. Commentary on Quantifying Agreement between Two Methods of Measurement. Clinical Chemistry. 2002;48(5):801-2.
154. World Health Organization (WHO). Measuring change in nutritional status. Guidelines for assessing the nutritional impact of supplementary feeding programmes for vulnerable groups. WHO, Geneva,1983. ISBN: 9241541660
155. Wood S. Generalized Additive Models : An Introduction with R, Second Edition. Philadelphia, PA, United States: Chapman and Hall/CRC; 2017.
156. Association WM. World Medical Association Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects. JAMA. 2013;310(20):2191-4.
157. Ogland B, Nilsen ST, Forman MR, Vatten LJ. Pubertal development in daughters of women with pre-eclampsia. Archives of Disease in Childhood. 2011;96:740-3.
158. Steene-Johannessen J, Kolle E, Andersen LB, Anderssen SA. Adiposity, aerobic fitness, muscle fitness, and markers of inflammation in children. Medicine and science in sports and exercise. 2013;45(4):714-21.
159. Grydeland M, Bergh IH, Bjelland M, Lien N, Andersen LF, Ommundsen Y, et al. Intervention effects on physical activity: the HEIA study - a cluster randomized controlled trial. The international journal of behavioral nutrition and physical activity. 2013;10:17.

160. Fredriks AM, van Buuren S, Burgmeijer RJ, Meulmeester JF, Beuker RJ, Brugman E, et al. Continuing positive secular growth change in The Netherlands 1955-1997. Pediatric Research. 2000;47(3):316-23.
161. Bjørnelv S, Lydersen S, Mykletun A, Holmen TL. Changes in BMI-distribution from 1966–69 to 1995–97 in adolescents. The Young-HUNT study, Norway. BMC public health. 2007;7(1):279.
162. Andersen LF, Lillegaard IT, Overby N, Lytle L, Klepp KI, Johansson L. Overweight and obesity among Norwegian schoolchildren: changes from 1993 to 2000. Scandinavian journal of public health. 2005;33(2):99-106.
163. Juliusson PB, Eide GE, Roelants M, Waaler PE, Hauspie R, Bjerknes R. Overweight and obesity in Norwegian children: prevalence and socio-demographic risk factors. Acta paediatrica (Oslo, Norway : 1992). 2010;99(6):900-5.
164. Norwegian Institute of Public Health. Barn og overvekt: Norwegian Institute of Public Health (NIPH); 2017 [cited 2020 21.01.2020]. Available from: https://www.fhi.no/nettpub/hin/levevaner/overvekt-og-fedme/.
165. Bergen Kommune. Levekår og helse i Bergen 2016. Available from: https://www.bergen.kommune.no/hvaskjer/tema/folkehelse/statistikk-og-rapporter/folkehelseoversikt-for-bergen-rapporter
166. Lien N, Kumar BN, Holmboe-Ottesen G, Klepp KI, Wandel M. Assessing social differences in overweight among 15- to 16-year-old ethnic Norwegians from Oslo by register data and adolescent self-reported measures of socio-economic status. International journal of obesity (2005). 2007;31(1):30-8.
167. SSB. 08921: Personer 16 år og over, etter region, kjønn, alder, utdanningsnivå, statistikkvariabel og år: Statistics Norway; 2019 [cited 2020. Available from: https://www.ssb.no/statbank/table/08921].
168. SSB. Utdanningsnivå for menn og kvinner 16 år og eldre: Statistics Norway; 2020 [cited 2020. Available from: https://www.ssb.no/utniv.
169. Elinder LS, Heinemans N, Zeebari Z, Patterson E. Longitudinal changes in health behaviours and body weight among Swedish school children - associations with age, gender and parental education – the SCIP school cohort. BMC public health. 2014;14(1):640.
170. Reinikainen J, Tolonen H, Borodulin K, Harkanen T, Jousilahti P, Karvanen J, et al. Participation rates by educational levels have diverged during 25 years in Finnish health examination surveys. European journal of public health. 2018;28(2):237-43.
171. Langhammer A, Krokstad S, Romundstad P, Heggland J, Holmen J. The HUNT study: participation is associated with survival and depends on socioeconomic status, diseases and symptoms. BMC Med Res Methodol. 2012;12:143-.
172. Kristiansen H, Juliusson PB, Eide GE, Roelants M, Bjerknes R. TV viewing and obesity among Norwegian children: the importance of parental education. Acta paediatrica (Oslo, Norway : 1992). 2013;102(2):199-205.
173. Fassler CS, Gutmark-Little I, Xie C, Giannini CM, Chandler DW, Biro FM, et al. Sex hormone phenotypes in young girls and the age at pubertal milestones. The Journal of Clinical Endocrinology & Metabolism. 2019.
174. Statistics Norway. Befolkning 05196: Statsborgerskap, alders- og kjønnsfordeling for hele befolkningen 1977 - 2019: Statistics Norway; 2019 [cited 2019. Available from: https://www.ssb.no/statbank/table/05196/.

175. Bartlett JW, Frost C. Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. Ultrasound in Obstetrics and Gynecology. 2008;31(4):466-75.
176. Lindhardt Johansen M, Hagen CP, Mieritz MG, Wolthers OD, Heuck C, Petersen JH, et al. Pubertal progression and reproductive hormones in healthy girls with transient thelarche. The Journal of Clinical Endocrinology & Metabolism. 2017;102(3):1001-8.
177. Koo MM, Rohan TE. Accuracy of short-term recall of age at menarche. Annals of human biology. 1997;24(1):61-4.
178. Mei Z, Grummer-Strawn LM, Wang J, Thornton JC, Freedman DS, Pierson RN, Jr., et al. Do skinfold measurements provide additional information to body mass index in the assessment of body fatness among children and adolescents? Pediatrics. 2007;119(6):e1306-13.
179. Altman DG. Practical statistics for medical research. London: Chapman and Hall; 1991.
180. Biro FM FF, Khoury P, Morrison JA, Lucky AW. Areolar and breast staging in adolescent girls. Adolescent and Pediatric Gynecology. 1992;5:271-2.
181. Kaplowitz PB, Oberfield SE. Reexamination of the age limit for defining when puberty is precocious in girls in the United States: implications for evaluation and treatment. Drug and Therapeutics and Executive Committees of the Lawson Wilkins Pediatric Endocrine Society. Pediatrics. 1999;104(4 Pt 1):936-41.
182. Madsen A, Bruserud, IS and Bertelsen, BF, Roelants M, Oehme NHB, Viste K, Bjerknes R, Almås B, Rosendahl K, Mellgren G, Sagen JV and Júlíusson, PB. Hormone references for ultrasound breast staging and endocrine profiling to predict female onset of puberty. submitted june 2020. 2020.
183. Chung EM, Cube R, Hall GJ, González C, Stocker JT, Glassman LM. Breast Masses in Children and Adolescents: Radiologic-Pathologic Correlation. RadioGraphics. 2009;29(3):907-31.
184. Valeur NS, Rahbar H, Chapman T. Ultrasound of pediatric breast masses: what to do with lumps and bumps. Pediatric Radiology. 2015;45(11):1584-99.
185. Campisi SC, Marchand J, Siddiqui FJ, Islam M, Bhutta ZA, Palmert MR. Can we rely on adolescents to self-assess puberty stage? The Journal of clinical endocrinology and metabolism. 2020.
186. Brundtland GH, Liestol K, Walloe L. Height, weight and menarcheal age of Oslo schoolchildren during the last 60 years. Annals of human biology. 1980;7(4):307-22.
187. Lundblad MW, Jacobsen BK. The reproducibility of self-reported age at menarche: The Tromso Study. BMC women's health. 2017;17(1):62.
188. Lien L, Dalgard F, Heyerdahl S, Thoresen M, Bjertness E. The relationship between age of menarche and mental distress in Norwegian adolescent girls and girls from different immigrant groups in Norway: Results from an urban city cross-sectional survey. Social Science & Medicine. 2006;63(2):285-95.
189. Kloven B, Hoftun GB, Romundstad PR, Rygg M. Relationship between pubertal timing and chronic nonspecific pain in adolescent girls: the Young-HUNT3 study (2006-2008). Pain. 2017;158(8):1554-60.

190. Vatten LJ, Romundstad PR, Holmen TL, Hsieh CC, Trichopoulos D, Stuver SO. Intrauterine exposure to preeclampsia and adolescent blood pressure, body size, and age at menarche in female offspring. Obstetrics and gynecology. 2003;101(3):529-33.
191. Hellevik AI, Nordsletten L, Johnsen MB, Fenstad AM, Furnes O, Storheim K, et al. Age of menarche is associated with knee joint replacement due to primary osteoarthritis (The HUNT Study and the Norwegian Arthroplasty Register). Osteoarthritis and cartilage. 2017;25(10):1654-62.
192. Brundtland GH, Walloe L. Menarchal age in Norway: halt in the trend towards earlier maturation. Nature. 1973;241(5390):478-9.
193. Bjelland EK, Hofvind S, Byberg L, Eskild A. The relation of age at menarche with age at natural menopause: a population study of 336 788 women in Norway. Human reproduction (Oxford, England). 2018;33(6):1149-57.
194. Nafstad P, Stray-Pedersen B, Solvberg M, Tangen T. [Menarche and menstruation problems among teenagers in Oslo 1993]. Tidsskrift for den Norske laegeforening : tidsskrift for praktisk medicin, ny raekke. 1995;115(5):604-6.
195. Sorensen K, Juul A. BMI percentile-for-age overestimates adiposity in early compared with late maturing pubertal children. European journal of endocrinology. 2015;173(2):227-35.
196. Brooks-Gunn J, Warren MP. Mother-daughter differences in menarcheal age in adolescent girls attending national dance company schools and non-dancers. Annals of human biology. 1988;15(1):35-43.

Appendix

II

References for Ultrasound Staging of Breast Maturation, Tanner Breast Staging, Pubic Hair, and Menarche in Norwegian Girls

Ingvild Særvold Bruserud,[1,2] Mathieu Roelants,[3] Ninnie Helén Bakken Oehme,[1,2] Andre Madsen,[1,4] Geir Egil Eide,[5,6] Robert Bjerknes,[1,2] Karen Rosendahl,[7,8] and Petur B. Juliusson[1,2,9]

[1]Department of Clinical Science, University of Bergen, Bergen, Norway; [2]Department of Paediatrics, Haukeland University Hospital, Bergen, Norway; [3]Environment and Health, Department of Public Health and Primary Care, KU Leuven–University of Leuven, Leuven, Belgium; [4]Hormone Laboratory, Haukeland University Hospital, Bergen, Norway; [5]Centre for Clinical Research, Haukeland University Hospital, Bergen, Norway; [6]Department of Global Public Health and Primary Care, University of Bergen, Bergen, Norway; [7]Department of Radiology, University Hospital of North Norway, Tromsø, Norway; [8]Department of Clinical Medicine, University in Tromsø, The Artic University of Norway, Norway; and [9]Department of Health Registries, Norwegian institute of Public Health, Bergen, Norway

ORCiD numbers: 0000-0001-8307-2293 (I. S. Bruserud); 0000-0002-3749-0475 (M. Roelants).

Context: Discriminating adipose and glandular tissue is challenging when clinically assessing breast development. Ultrasound facilitates staging of pubertal breast maturation (US B), but has not been systematically compared to Tanner breast (Tanner B) staging, and no normative data have been reported.

Objective: To present normative references for US B along with references for Tanner B, pubic hair (PH), and menarche.

Design, Setting, and Participants: A cross-sectional sample of 703 healthy girls aged 6 to 16 years were examined.

Main Outcome Measures: Breast development was determined with US B and Tanner B staging. Tanner PH and menarcheal status were recorded. The age distributions of entry in US B, Tanner B, and PH stages and menarche were estimated with generalized linear and generalized additive models with a probit link. Method agreement was tested with weighted Cohen's kappa.

Results: The median (±2SD) ages for thelarche, US B2 and Tanner B2, were 10.2 (7.7, 12.8) and 10.4 (8.0, 12.7) years. The median (±2SD) ages at Tanner PH2 and menarche were 10.9 (8.5, 13.3) and 12.7 (11.0, 16.2) years. Cohen's kappa of agreement (95% confidence interval) between US B and Tanner B was 0.87 (0.85–0.88). When the methods disagreed, US B was usually more advanced.

Conclusion: Thelarche occurred at a slightly younger age when assessed with ultrasound compared to clinical Tanner staging, although the 2 methods had a very good agreement when determining pubertal breast maturation. A significant decrease of 2.8 months in age at menarche was observed during the past decade in Norwegian girls. (***J Clin Endocrinol Metab*** **105: 1–9, 2020)**

Key Words: puberty, breast development, menarche, female

ISSN Print 0021-972X ISSN Online 1945-7197
Printed in USA

Received 15 October 2019. Accepted 2 March 2020.
First Published Online 6 March 2020.
Corrected and Typeset 25 March 2020.

Abbreviations: BGS 1, Bergen Growth Study 1; BGS 2, Bergen Growth Study 2; US B, ultrasound breast stage; Tanner B, Tanner breast stage; Tanner PH, Tanner pubic hair stage.

Breast development in girls is usually defined clinically by the 5-stage scale as described by Marshall and Tanner in 1969 (1,2). The Tanner breast (Tanner B) scale ranges from B1 (prepubertal) to B5 (mature), in which stage B2 marks the first appearance of glandular breast tissue (thelarche) and thus the start of central puberty. Clinical breast staging relies on visual inspection and palpation to distinguish the glandular breast tissue from adipose tissue, but this approach is often perceived as challenging. With ultrasound (US), we, and others, have been able to identify distinctive breast developmental stages based on changes in the appearance and relative amount of glandular, adipose and fibrous tissues (3–5). Despite being a promising method, US staging of the breast has not yet been systematically compared with clinical Tanner B staging in a large developmentally diverse sample, and no normative references have yet been published.

In Norway and other northern European countries, only marginal changes in age at menarche have been reported during the last decades (6–10). However, data from Denmark and the United States have shown that the onset of breast development might be advancing toward younger ages (7,11,12). A worldwide secular trend toward younger ages at thelarche according to race/ethnicity and geography was also shown in a recent systematic review (13). Accordingly, these studies suggest an asymmetrical secular trend of earlier breast development, while age at menarche remains more stable. Early puberty, and early menarche in particular, have been associated with increased all-cause mortality (14), a higher risk of breast cancer (15), cardiovascular disease (16), and mental health problems (17). It has been suggested that this secular trend may be caused by overweight or endocrine disrupting chemicals (18,19).

The present study is a part of the Bergen Growth Study 2 (BGS 2), which is the first large-scale pubertal reference study conducted in Norway. In addition to data on menarche and traditional Tanner B and pubic hair (PH) stages, we collected data on breast developmental stage assessed with US (US B) as a novel methodological approach. The aim of the current work was to provide descriptive reference ranges for these phenotypic traits in a large sample of contemporary girls living in Norway, to compare both methods of staging breast development, and to assess recent changes in age at menarche.

Materials and Methods

Study design and participants

A total of 1349 girls between 6 and 16 years of age from 6 randomly selected public schools in Bergen municipality, Norway, were invited to participate. Bergen is the second largest city in Norway and the demographic profile and living conditions are very similar to those in the country at large (20). Parental consent was received prior to examination from 678 girls (50.3%) from January to June 2016. The participation rates varied with age, from 52.1% before 12 years to 43.4% above 13 years, and between schools, from 29.1% to 59.4%. In addition, data on breast development from 57 girls examined in February 2017 for a reliability study were included in the analysis. Details from this study have been published previously (5). A parental questionnaire with information on chronic illness, parental origin, socioeconomic status, and menarche was obtained from 482 (68.6%) girls. Five girls were absent on the day of examination, and 27 girls were excluded due to a disease likely to affect growth or maturation, leaving 703 girls eligible for the analysis. Of 466 girls with known origin of both parents, 374 (80.2%) were Norwegian, 408 (88.9%) European, and 51 girls (11.1%) had 1 or 2 non-European parents, mostly from Asia (n = 30), Africa (n = 12), or South America (n = 7) (Supplementary Table 1, all supplementary materials and figures are located in a digital data repository (21)). The analyses are based on all girls, regardless of their origin. The highest educational level of both parents was used as an indicator of socioeconomic status. This level was classified as primary education (9 years of school), secondary education (12 years of school), or higher education (college degree). Demographics of the study populations from Bergen Growth Study 1 (BGS 1) and BGS 2 are summarized in Supplementary Table 1 (21).

The median (range) age in the final sample was 10.7 (6.1–16.2) years. Due to missing data (menarche was not recorded in 3 girls, 3 girls declined examination of the breast, and 81 declined examination of PH, and the US image could not be retrieved for off-site scoring in 4 girls), the pubertal references are based on 696 observations of US B, 700 of Tanner B, and 643 of menarche. Data on the Tanner PH stage were available for 372 girls as further explained in the following discussion.

Examinations

The girls were examined during school hours. All examinations were performed by a single female nurse (ISB), who was trained in US staging of breast development by an experienced pediatric radiologist (KR) before and at the start of the actual data collection. Pubertal development was assessed with the girl in the supine position. The US examinations were performed using a Sonosite Edge machine (Fujifilm SonoSite, US) with a 15–6 MHz (5 cm) linear probe. Based on a standard mid-sagittal section, breast development was scored on a scale from 0 to 5, consistent with our previously published descriptions (5). US breast staging was conducted based on radiologic characteristics (Supplementary Figure 1, (21)). In brief, a small, hypoechoic retro-areolar mass is scored as US B0, while a hyperechoic area (with or without a small hypoechoic mass) is US B1. Both US B0 and US B1 are considered prepubertal. US B2 is characterized by the presence of hyperechoic retro-areolar glandular breast tissue with a round or star-shaped hypoechoic center. This stage is assumed to mark the start of puberty (3,5). US B3 is characterized by a larger, spider-shaped hypoechoic center, becoming increasingly roundish in US B4 (5). The mature US B5 is characterized by heterogeneous glandular breast tissue without the hypoechoic center observed in

US B2 to B4. The left breast was examined according to the protocol. Due to logistic reasons, the US stage was determined at a later time in 166 girls, by the same examiner, based on the standardized US images. This was done to avoid the girls becoming uncomfortable due to prolonged exposure. Based on a reassessment of 120 saved images, we determined that this had little impact on the results. Intra- and inter-observer agreement of US B staging was determined in an additional study of 57 girls (5). Cohen's kappa was estimated at 0.84 and 0.71, respectively, which is commonly considered good to very good agreement (22).

The development of the breasts and PH was clinically scored on a scale from 1 to 5 (B1–B5 and PH1–PH5) according to the method described by Tanner (1), which is based on a visual inspection, although palpation of the breast was included in this procedure. Pubertal onset is defined as Tanner B2, which is characterized by the visual and/or palpable appearance of glandular tissue. The left breast was always examined, however, in 3 cases the right breast was found more advanced according to Tanner B staging and therefore used in the references (Tanner B staging recorded additionally on the right side in 453 girls). A subsample of 453 girls were asked if the examiner could assess the Tanner PH by visual inspection, to which 372 (82.1%) assented. Because of assumed time constrains, Tanner B staging of the right breast and Tanner PH staging were not included in the protocol at the start of the study but were only added and systematically recorded from February 22 until the end of the data collection. Tanner PH2 marks the first appearance of PH (1,2). All participants were asked if they had experienced their first menstruation (menarche) and the month/year, if applicable. Missing data on menarche were completed with data from the parental questionnaire for 16 girls. Two girls above 15 years of age were encouraged to see their general practitioner due to primary amenorrhea.

Height was measured in the standing position with a Harpenden Portable Stadiometer (Holtain Ltd Crosswell, UK) and recorded to the nearest 0.1 cm. Weight was measured in light clothing with an electronic scale (Tanita MC-780MA, Tanita Corp. of America, Inc. Illinois, US) with a precision of 0.1 kg. Body mass index (BMI) was calculated by dividing weight (in kg) by the square of height (in meters). The BMI was available for 645 (91.7%) girls of whom 14.6% were overweight (including obesity) and 7.3% were underweight according to the definitions from the International Obesity Task Force (23,24).

Statistical analysis

Age at menarche or entry into a certain puberty stage was estimated from age and pubertal status of the girls with a generalized linear model (probit regression) and with a nonparametric generalized additive model (GAM) for binary outcomes with probit link. The methods differ by the assumption of a normal age distribution (probit) or not (GAM). The probit and GAM curves were estimated for each consecutive stage separately. The degree of smoothing of the GAM was determined by generalized cross-validation and models were checked with diagnostic plots using the mgcv package (25) in R version 3.5 (R Foundation for Statistical Computing, Vienna, Austria, 2018). Primarily, the median, percentiles, and equivalent z-scores of age at entry from the nonparametric additive model are reported in the tables. In addition, we report the mean and standard deviation (SD) from the parametric probit models to allow a comparison with data from other studies. In addition to reference data based on all eligible girls, we estimated age at menarche and transition to US B2, Tanner B2, and Tanner PH2 in the subgroup of girls with a documented Norwegian or non-Norwegian origin. The statistical significance of differences between Norwegian and non-Norwegian girls was tested with logistic regression using age as a covariate and origin (Norwegian, non-Norwegian) as the variable of interest. This analysis was limited to girls with a documented origin only.

Agreement between US B and Tanner B staging was analyzed with simple Cohen's kappa and Cohen's kappa test with linear weights. The kappa statistic gives a coefficient between 0 (no agreement) and 1 (perfect agreement) (22). For these comparisons, the US B0 and B1 stages were collapsed into a single prepubertal stage. Separate analysis of agreement and percentage concordance between the methods were done for girls with normal weight and overweight (including obesity).

To investigate if age at menarche had advanced in the Bergen municipality since the first Bergen Growth Study (BGS 1), conducted in 2003 to 2006, we compared girls with a documented Norwegian or Nordic (1.5%) background in both studies to avoid interference from demographic changes in the population. Details from BGS 1 have been published elsewhere (10). In brief, girls from 8 to 15.5 years (n = 1481) were asked if they had experienced their menarche, and this was the only pubertal marker recorded in BGS 1. The occurrence of menarche was compared between both studies with logistic regression using age, BMI z-score, and parental educational level as covariates. Logistic regression, and agreement analyses were done in Statistical Package for Social Sciences (SPSS version 24) and R (version 3.5).

Ethical considerations

The project was approved by the Regional Committee for Medical and Health Research Ethics West Norway (number 2015/128/REK Vest). Written informed consent was obtained from a parent or legal guardian of each participant and also from the participants 12 years and older. All girls received age-appropriate information in writing and orally by the study nurse ahead of participation. Assent from the participant was an additional requirement for inclusion. All participants received a cinema ticket for their collaboration.

Results

The medians and equivalent z-scores, means, and SD for age at entry into consecutive US B, Tanner B and PH stages, and menarche are presented in Table 1. Traditional age percentiles of all pubertal markers are provided in Supplemental Table 2 (21). Reference quantiles were calculated nonparametrically from the cumulative distribution estimated with a GAM, which in some cases coincided with the normal distribution as shown in Figs. 1A–D.

Table 1. Age References for Breast Development, Pubic Hair and Menarche: Median and Distribution (–2, –1, 1, 2 SD) of the Age (in Years) at Attainment of the Specified Stage

	Generalized Additive Model Age at Attainment					Probit
Pubertal marker	**–2 SD**	**–1 SD**	**Median**	**1 SD**	**2 SD**	**Mean (SD) age**
Breast (ultrasound)						
US B1	NA	5.6[a]	8.3	10.7	12.4	8.3 (2.2)
S B2	7.7	8.9	10.2	11.5	12.8	10.2 (1.3)
US B3	9.1	10.3	11.5	12.7	13.9	11.5 (1.2)
US B4	10.4	11.3	12.3	13.9	NA	12.7 (1.5)
US B5	11.3	12.7	15.9	NA	NA	15.3 (2.3)
Breast (Tanner)						
B2	8.0	9.2	10.4	11.6	12.7	10.4 (1.2)
B3	9.5	10.4	11.4	12.5	13.8	11.5 (1.1)
B4	10.7	11.6	12.5	13.8	NA	12.8 (1.3)
B5	11.1	12.3	14.7	NA	NA	14.8 (2.2)
Pubic Hair (Tanner)						
PH2	8.5	9.7	10.9	12.1	13.3	10.9 (1.2)
PH3	9.7	10.9	11.9	12.9	13.7	11.9 (1.1)
PH4	10.5	11.8	13.1	14.4	15.7	13.1 (1.3)
PH5	11.4	13.0	15.1	NA	NA	15.1 (1.9)
Menarche	11.0	11.7	12.7	14.1	16.2	12.9 (1.2)

Distribution of age at entry in consecutive stages of breast development measured with ultrasound (US B; $n = 696$), breast development according to Tanner B ($n = 700$), Pubic hair (PH) according to Tanner ($n = 372$) and menarche ($n = 643$) in 6- to 16-year-old healthy girls in Bergen, Norway. Median age and ±1 or 2 reference limits were estimated with nonparametric generalized additive models. Mean (SD) ages were estimated with probit regression.

Abbreviations: B, breast; NA, not applicable (outside the data range); PH, pubic hair; SD, standard deviation; US B, ultrasound-based breast stage.

[a]Extrapolated age that falls just below the data range.

The distribution of age at onset of breast development measured with US was slightly ahead of the Tanner method. However, the opposite was true for the higher maturational stages, where the age at transition with the Tanner method was ahead of the assessment with US. The mean age (SD) of thelarche was 10.2 (1.3) years with US (US B2), and 10.4 (1.2) years with the Tanner method (Tanner B2). The lower end of the normal range defined by a *z*-score of –2 SD was 7.7 and 8.0 years, respectively, while the corresponding upper limits were 12.8 and 12.7 years.

The breast developmental stage assessed with US and the Tanner method was concordant in 551 of 695 (79.3%) girls for whom both assessments were available (kappa = 0.87; 95% confidence interval [CI]; 0.85–0.89) (Table 2). When dichotomizing the breast developmental stage into thelarche or not thelarche, the agreement was very good (kappa = 0.94; 95% CI; 0.91–0.96). US identified breast development (US B ≥2) in 18 girls in whom the mammary gland was not palpable (corresponding to Tanner B1). Opposite, three girls had clinical signs of pubertal breast development (Tanner ≥B2) with no glandular breast tissue visible on US (US B0 or US B1). The kappa coefficients when comparing the 2 methods were 0.88 (95% CI 0.86–0.91) and 0.85 (95% CI: 0.79–0.90) for girls with normal and overweight/obesity, respectively. Full concordance between US B and Tanner B staging was found in 81.6% and 73.4% of girls with normal weight and overweight/obesity, respectively.

Pubic hair development (Tanner ≥PH2) was observed in 182 of 372 (48.9%) girls. The mean (SD) age at Tanner PH2 was estimated as 10.9 (1.2) years. The age distribution of each Tanner PH stage was close to normal (Fig. 1C).

In total, 193 of 643 (30.0%) girls had experienced menarche. The median (±2 SD) age at menarche was 12.7 (11.0–16.2) years. This is 2.5 years after thelarche according to the US method (US B2) and 2.3 years after thelarche according to the Tanner method (B2). A comparison of the probit model and the GAM indicated a positive skewness in the distribution of age at menarche (Fig. 1D).

The onset of all pubertal markers occurred earlier in girls with a non-Norwegian origin ($n = 92$) compared to girls with two parents of Norwegian origin ($n = 374$), in the subset of girls with known origin (74%) (Supplemental Table 3 (21)). The difference was only statistically significant for the start of breast development according to US ($P < .05$). Compared to girls from BGS 1, mean (SD) age at menarche in girls of Norwegian, including Nordic, origin had significantly declined from 13.3 (1.7) years (not previously published) in 2006 to 13.1 (1.2) years in 2016 ($P < 0.05$).

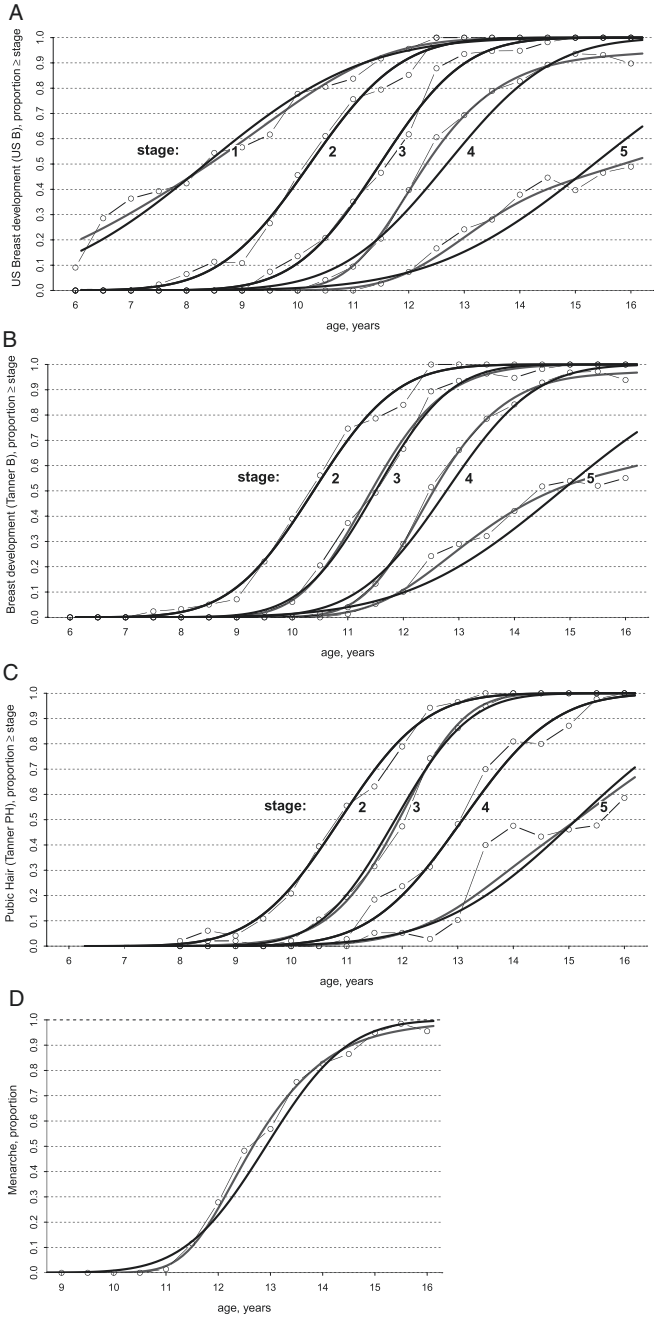


Figure 1. Age reference curves. Age distribution of pubertal development in 6- to 16-year-old girls living in Norway. **A:** Stages of breast development with ultrasound (n = 696); **B**: Breast development stage by Tanner stage (n = 700); **C:** Pubic hair (PH) according to Tanner (n = 372); **D:** Menarche (n = 643). Black lines show the parametric probit model, assuming a Gaussian distribution of the age at transition; grey lines show the corresponding curves of the nonparametric generalized additive model with a probit link; connected dots show the cumulative proportion of girls who have reached a particular pubertal stage in overlapping one-year intervals.

Table 2. Distribution of Breast Development with Ultrasound and Tanner Staging

	Breast Development With Ultrasound						
Tanner stage	**US B0**	**US B1**	**US B2**	**US B3**	**US B4**	**US B5**	**Total**
B1	**166**	**129**	18	0	0	0	313
B2	1	2	**63**	12	2	0	80
B3	0	0	13	**43**	25	2	83
B4	0	0	0	15	**77**	16	108
B5	0	0	0	5	33	**73**	111
Total	167	131	94	75	137	91	695

Agreement of the breast developmental stage according to Tanner (B1 to B5) and ultrasound examination (US B0 to B5) in 695 6- to 16-year-old girls in Bergen, Norway. US B0 and US B1 and Tanner B1 are prepubertal stages; stages US B0 and US B1 are merged for the analysis of method agreement.
Abbreviations: US B, ultrasound-based breast stage.

The difference in occurrence of menarche between both studies remained statistically significant (odds ratio 2.0; 95% CI 1.1–3.6; P = .016), when accounting for the BMI z-score and parental educational level. Parental education was not retained in the final models because it was not significantly associated with menarche in either study. The age distribution in both studies is shown in Supplementary Figure 2 (21) for all girls in the study and in Supplementary Figure 3 (21) for girls of Norwegian/Nordic origin.

Discussion

In this study, we describe female pubertal development using a large contemporary sample of healthy girls living in Norway. The results are summarized as age percentiles, which can serve as normative reference intervals for pubertal milestones in the Norwegian and other comparable populations. Up-to-date references are important to support clinical decision-making. In addition to the traditional Tanner stages, we have also documented breast development based on US examination of glandular breast tissue. Our results show a high degree of concordance between the 2 methods, but the start of breast development is detected almost 2 months earlier with US staging. Finally, our results indicate that there might be an ongoing decrease in the age at menarche in the Norwegian population.

The principle of assessing breast development with US examination of the mammary gland has been proposed by several authors (4,5,26–28), but US staging of breast development was not yet compared with clinical staging in a large sample of healthy pubescent girls, and no reference data were available. The 4 studies that could be identified were small (<80 subjects) or based on a heterogeneous clinical sample (4,26–28). US staging is a promising technique with a good intra- and inter-observer reliability (5) that may overcome several known limitations of the visual and palpatory method described by Tanner. In the present study, we have shown that both methods agree well with each other in a large sample of girls at various stages of their development, also when stratifying by weight category. The kappa statistic showed good agreement over all five stages and very good agreement when girls were classified as having thelarche or not. The biggest discrepancy was observed between Tanner B5 and US staging, but in most cases, this difference was 1 stage only. This may be due to the statistical model because only about half of the girls have reached stage 5 by the end of the age range of girls included in the study (differences are smaller at –1 or –2 SD) but could also reflect a discrepancy between the size of the breast and maturation of the gland as visualized with US or reflect an overlap of the characteristics of stages 4 and 5 with the Tanner and US methods.

However, the latter comparison also indicated that the US method is probably more sensitive to pick up changes that are consistent with the start of pubertal breast development. In addition, US examination identifies two distinct developmental stages (denoted as US B0 and US B1) in girls who are prepubertal (B1) according to the Tanner method. This has not been acknowledged in previous studies by other authors (5). Both these stages are considered by us as prepubertal and future research will indicate to what extent they correlate with physiological and hormonal changes. For the comparison of the US and Tanner methods both prepubertal stages were merged.

The mean age at thelarche according to the Tanner method (B2) was 10.4 years with an SD of 1.2 years. While no historical data are available for the Norwegian population, other studies reported a substantial decline in onset of breast development (Tanner B2) from 10.9 years in 1993 to 9.9 years in 2008 in Danish girls living in the Copenhagen area (7). However, a more recent population-based longitudinal study of self-reported onset of breast development found a mean age at thelarche of 10.4 years, which is more comparable to our findings (29). A declining age at

thelarche has also been reported in the United States (11,30) and in the United Kingdom (9), but studies from Belgium (6) and the Netherlands(31) reported a mean age at thelarche of 10.7 years, without an apparent trend toward younger ages. The lower reference limit (–2 SD) for age at attaining Tanner B2 was 8.0 years in the current study. Overall, these findings suggest that the onset of breast development in Norwegian girls is neither earlier nor later when compared to other northern European countries and that the current age limit of precocious puberty in girls (ie, start of breast development before 8 years) can be maintained, as is the case in other countries (6,32). Likewise, we reported a mean (SD) age of the start of PH development (Tanner PH2) of 10.9 (1.2) years, which corresponds well with the mean ages of 11.0 to 11.1 years reported in northern European populations (6,7,31).

Mean age at menarche fluctuates within a narrow range from 12.9 to 13.1 years in most studies from northern Europe (6,7,9,29,33). Although age at menarche has been remarkably stable for more than half a century, a small but statistically significant decline has been reported in Denmark and the Netherlands (0.3 and 0.1 years, respectively) (7,33). A recent study from the United States reported an overall median age at menarche of 12.3 years and in white girls 12.7 years (12), which is only a modest change from a previous U.S. study where the age at menarche in white girls was 12.9 years (11). In the current study, we found a median age at menarche of 12.7 years, which is in line with the white girls from the United States, but earlier than the cited northern European studies. However, we previously reported an overall median age at menarche of 13.1 years in BGS 1 from 2006 (10). The age at menarche in Norway has been stable just above 13 years during the last 50 years (10,34). To rule out possible interference from recent changes in the composition of the target population such as origin, BMI, and socioeconomic background, we compared a subsample of girls with a documented Norwegian origin from both studies and included BMI and parental educational level in the analysis. Parental education was not significantly associated with age at menarche, but there was a significant effect of the BMI as was documented before (10). The statistically significant decrease with 2.8 months in age at menarche since 2006 was however not affected by the inclusion of the BMI *z*-score in the models because the BMI distribution was highly comparable in both studies. An increased exposure to endocrine disruptive chemicals remains a possible hypothesis, to which we cannot offer a conclusive answer in the current analysis (19). Given the potential importance of this finding, we believe it should be confirmed in a larger population-based study.

For the descriptive reference ranges of menarche and secondary sexual characteristics we have included all girls regardless their origin in the reference sample because their impact on the reference limits is relatively small from a clinical perspective, and implementation of a single reference might be more feasible. Establishing and adapting reference curves to fit an increasingly demographically heterogenic composition where growth and puberty timing of the native population is different to that of immigrant populations remain an epidemiological challenge.

Although the BGS 2 included a large number of children at or near puberty and covers many aspects of pubertal development, some limitations should be noted. The schools were randomly selected and cover various socioeconomic conditions, but they were all situated in mostly urban areas of the municipality of Bergen. The prevalence of overweight (including obesity) of 14.6% in our study is slightly lower than, but still comparable to, data from the first Bergen growth study (17.7%) (35) and to data from a nationally representative sample of 9-year-old girls (17%) (36). Including girls from different Norwegian regions would have increased generalizability but would also have introduced logistic challenges and could have negatively impacted the standardization of measurement procedures. The recruitment strategy is also a potential source of selection bias, that is, when girls with a relatively early or late puberty are more reluctant to participate. Our study sample is representative of the Norwegian population with respect to origin (37). Girls of non-Norwegian origin were included in the analysis since this reflects the natural variation in the target population. Although this increases the applicability of the references in the Norwegian population at large, it may also hamper the comparability of the data with other nations or future studies. For this reason, we also provided references based on a subsample of girls with a Norwegian background as supplementary material (21). Our knowledge about the nonparticipants is limited to school and grade. However, there was a slightly lesser participation rate in schools with a larger number of children with a non-Norwegian background. Increased nonparticipation in these schools may have been caused by language barriers or cultural preferences of parents.

One of the potential advantages of US staging of breast development relates to difficulties with Tanner staging in girls with obesity. We observed that the agreement and concordance between the two breast staging methods were slightly lower among the girls with overweight and obesity and that the clinical assessment, Tanner B, showed a more mature stage compared to the US B in more cases among the girls with overweight than in the normal weight group (data not shown). A higher prevalence of girls with Tanner

B2 would have been preferable to further examine the relation between US B and Tanner B around the clinical stage that defines pubertal onset. Another limitation is the low response, even though we consider the participation rate of 50% as relatively high for this particular type of study. Information on health and ethnicity were only known when the questionnaire was answered and returned by parents (69%), which may have resulted in the inclusion of some girls with relevant pathology unknown to us. However, the prevalence of such conditions was low in girls with a questionnaire available, and including or excluding them had little impact on the resulting estimates (data not shown).

In conclusion, we found that US and Tanner B staging are in close agreement when determining the degree of pubertal breast maturation. US allows for a direct examination of the mammary gland and detects pubertal breast development slightly ahead of clinical Tanner B staging. Moreover, this method allows to discriminate 2 distinct prepubertal stages, the clinical relevance of which will be the subject of further research. Pubertal breast and PH development in Norwegian girls occur at comparable ages as in their northern European peers. In girls of Norwegian origin, age at menarche occurred at a significantly younger age than 10 years before. This trend toward earlier menarche merits further investigation.

Acknowledgments

We would like to thank all the children who participated in Bergen Growth Study 2. We also acknowledge Magnus Rise Sveen for his substantial help during data collection.

Financial Support: This study was supported by a grant (number 912221) from the Western Norway Regional Health Authority.

Additional Information

Correspondence and Reprint Requests: Ingvild Særvold Bruserud, Department of Clinical Science, University of Bergen, Haukelandsbakken 15, 5021 Bergen, Norway. E-mail: ingvild.servold.bruserud@helse-bergen.no.

Disclosure Summary: The authors have nothing to disclose.

Data Availability: Restrictions apply to the availability of data generated or analyzed during this study to preserve patient confidentiality or because they were used under license. The corresponding author will on request detail the restrictions and any conditions under which access to some data may be provided.

References

1. Tanner JM. *Growth at Adolescence*. 2nd ed. Springfield, IL: Thomas; 1962.
2. Marshall WA, Tanner JM. Variations in pattern of pubertal changes in girls. *Arch Dis Child*. 1969;**44**(235):291–303.
3. García CJ, Espinoza A, Dinamarca V, et al. Breast US in children and adolescents. *Radiographics*. 2000;**20**(6):1605–1612.
4. Bruni V, Dei M, Deligeoroglou E, et al. Breast development in adolescent girls. *Adolesc Pediatr Gynecol*. 1990; **3**:201–205
5. Bruserud IS, Roelants M, Oehme NHB, et al. Ultrasound assessment of pubertal breast development in girls: intra- and interobserver agreement. *Pediatr Radiol*. 2018;**48**(11):1576–1583.
6. Roelants M, Hauspie R, Hoppenbrouwers K. References for growth and pubertal development from birth to 21 years in Flanders, Belgium. *Ann Hum Biol*. 2009;**36**(6):680–694.
7. Aksglaede L, Sørensen K, Petersen JH, Skakkebaek NE, Juul A. Recent decline in age at breast development: the Copenhagen Puberty Study. *Pediatrics*. 2009;**123**(5):e932–e939.
8. Juul A, Teilmann G, Scheike T, et al. Pubertal development in Danish children: comparison of recent European and US data. *Int J Androl*. 2006; **29**:247–255; discussion 286–290
9. Rubin C, Maisonet M, Kieszak S, et al. Timing of maturation and predictors of menarche in girls enrolled in a contemporary British cohort. *Paediatr Perinat Epidemiol*. 2009;**23**(5):492–504.
10. Bratke H, Bruserud IS, Brannsether B, et al. Timing of menarche in Norwegian girls: associations with body mass index, waist circumference and skinfold thickness. *BMC Pediatr*. 2017;**17**(1):138.
11. Herman-Giddens ME, Slora EJ, Wasserman RC, et al. Secondary sexual characteristics and menses in young girls seen in office practice: a study from the Pediatric Research in Office Settings network. *Pediatrics*. 1997;**99**(4):505–512.
12. Biro FM, Pajak A, Wolff MS, et al. Age of menarche in a longitudinal US cohort. *J Pediatr Adolesc Gynecol*. 2018;**31**(4): 339–345.
13. Eckert-Lind C, Busch AS, Petersen JH, et al. Worldwide secular trends in age at pubertal onset assessed by breast development among girls: a systematic review and meta-analysis. *JAMA Pediatrics* 2020. doi:10.1001/jamapediatrics.2019.5881. [Epub ahead of print]
14. Charalampopoulos D, McLoughlin A, Elks CE, Ong KK. Age at menarche and risks of all-cause and cardiovascular death: a systematic review and meta-analysis. *Am J Epidemiol*. 2014;**180**(1):29–40.
15. Collaborative Group on Hormonal Factors in Breast Cancer. Menarche, menopause, and breast cancer risk: individual participant meta-analysis, including 118 964 women with breast cancer from 117 epidemiological studies. *Lancet Oncol*. 2012;**13**:1141–1151.
16. Lakshman R, Forouhi NG, Sharp SJ, et al. Early age at menarche associated with cardiovascular disease and mortality. *J Clin Endocrinol Metab*. 2009;**94**(12):4953–4960.
17. Mendle J, Ryan RM, McKone KMP. Age at menarche, depression, and antisocial behavior in adulthood. *Pediatrics* 2018;**141**(1):e20171703.
18. Abreu AP, Kaiser UB. Pubertal development and regulation. *Lancet Diabetes Endocrinol*. 2016;**4**(3):254–264.
19. Parent AS, Teilmann G, Juul A, Skakkebaek NE, Toppari J, Bourguignon JP. The timing of normal puberty and the age limits of sexual precocity: variations around the world, secular trends, and changes after migration. *Endocr Rev*. 2003;**24**(5):668–693.
20. The Municipality of Bergen. Levekår og helse i Bergen 2016. https://issuu.com/hg-9/docs/bg-kommune-levekarsrapport-2016?e=19530043/41371289. Accessed March 18, 2020.
21. Bruserud IS, Roelants M, Oehme NHB, et al. Supplemental tables, figures and figure legends. Figshare Digital Repository 2020. Deposited 15 February 2020. doi:10.6084/m9.figshare.11858355.v1.
22. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;**33**(1):159–174.
23. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. *BMJ*. 2000;**320**(7244):1240–1243.
24. Cole TJ, Flegal KM, Nicholls D, Jackson AA. Body mass index cut offs to define thinness in children and adolescents: international survey. *BMJ*. 2007;**335**(7612):194.
25. Wood S. *Generalized Additive Models: An Introduction with R*. 2nd ed. Philadelphia, PA: Chapman and Hall/CRC; 2017.

26. Yüce Ö, Sevinç D. Ultrasonographic assessment of pubertal breast development in obese children: compliance with the clinic. *J Pediatr Endocrinol Metab*. 2018;**31**(2):137–141.
27. Calcaterra V, Sampaolo P, Klersy C, et al. Utility of breast ultrasonography in the diagnostic work-up of precocious puberty and proposal of a prognostic index for identifying girls with rapidly progressive central precocious puberty. *Ultrasound Obstet Gynecol*. 2009;**33**(1):85–91.
28. Carlson L, Flores Poccia V, Sun BZ, et al. Early breast development in overweight girls: does estrogen made by adipose tissue play a role? *Int J Obes (Lond)*. 2019;**43**(10):1978–1987.
29. Brix N, Ernst A, Lauridsen LLB, et al. Timing of puberty in boys and girls: A population-based study. *Paediatr Perinat Epidemiol*. 2019;**33**(1):70–78.
30. Biro FM, Greenspan LC, Galvez MP, et al. Onset of breast development in a longitudinal cohort. *Pediatrics*. 2013;**132**(6):1019–1027.
31. Fredriks AM, van Buuren S, Burgmeijer RJ, et al. Continuing positive secular growth change in The Netherlands 1955–1997. *Pediatr Res*. 2000;**47**:316–323.
32. Mul D, Fredriks AM, van Buuren S, Oostdijk W, Verloove-Vanhorick SP, Wit JM. Pubertal development in The Netherlands 1965–1997. *Pediatr Res*. 2001;**50**:479–486.
33. Talma H, Schonbeck Y, van Dommelen P, Bakker B, van Buuren S, Hirasing RA. Trends in menarcheal age between 1955 and 2009 in the Netherlands. *PloS One* 2013;**8**:e60056
34. Brudevoll JE, Liestøl K, Walløe L. Menarcheal age in Oslo during the last 140 years. *Ann Hum Biol*. 1979;**6**(5):407–416.
35. Juliusson PB, Roelants M, Eide GE, Hauspie R, Waaler PE, Bjerknes R. Overweight and obesity in Norwegian children: secular trends in weight-for-height and skinfolds. *Acta Paediatrica*. 2007;**96**:1333–1337.
36. Norwegian Institute of Public Health. Barn og overvekt. Folkehelserapporten 2017. https://www.fhi.no/nettpub/hin/levevaner/overvekt-og-fedme/. Accessed January 21, 2020.
37. Statistics Norway. Befolkning 05196: Statsborgerskap, alders- og kjønnsfordeling for hele befolkningen 1977–2019. Statistikkbanken 2019. https://www.ssb.no/statbank/table/05196/. Accessed October 1, 2019.

Supplemental tables for the paper entitled: "References for ultrasound staging of breast maturation, Tanner breast staging, pubic hair and menarche in Norwegian girls"

Bruserud IS, Roelants M, Oehme N, Madsen A, Eide GE, Bjerknes R, Rosendah K, Juliusson PB

Supplementary Table 1 Comparison of the Bergen Growth Studies 1 and 2

Demographic characteristics of the study samples in the Bergen Growth studies 1 and 2.

Study characteristics	Bergen Growth Study 1		Bergen Growth Study 2
Year of study	2003-2006		2016
Study setting	randomly selection of 20 schools and 34 kindergardens (all children 0-19 years of age)		6 randomly selected schools (all pupils 6-16 years invited)
Included in comparison of occurrence of menarche	1485 (8 to 15.5 years)		643
Age, mean (SD)	11.8 (2.2)		11.1 (2.8)
BMI, mean (SD)	18.5 (2.9)		18.3 (3.2)
BMI z-score, mean (SD)	-0.01 (1.0)		0.04 (1.0)
Exclusion	chronic illness (n=39)		chronic illness (n=27)
Participation rate	69% in primary schools (6-12 years) 53% in middle schools (13-14 years)		52.1% in primary schools(6-12 years) 43.3% in middle schools (13-16 years)
Ethnic composition (%)	Norwegian/Nordic: 1294 (87.1%) European: 67 (4.2%) Non-European: 120 (8.1%)		Girls with known origin: 466 Norwegian/Nordic origin: 381 (81.2 %) European origin: 27 (5.8 %) Non-European: 51 (11.1%)
Parental education (highest degree)			
No secondary education	71 (7.6%)		16 (3.3%)
Secondary education	280 (29.9%)		82 (17.1%)
Higher education	583 (62.4%)		382 (79.6%)
Weight status (IOTF)			
Underweight	144 (9.7%)		47 (7.2%)
Normal	1128 (75.9%)		504 (78.1%)
Overweight	180 (12.1%)		80 (12.4%)
Obese	33 (2.2%)		14 (2.2%)

Abbreviations: IOTF: International Obesity Task Force, BMI: Body Mass Index

Supplemental Table 2. Age percentiles of ultrasound (US B) and Tanner (B) based stages of breast development, Tanner stages of pubic hair (PH) and menarche (M) in healthy 6 – 16 years old girls in Norway. The median (P50) and percentiles 3 to 97 of age at entry were derived from nonparametric generalized additive models (GAM) with a probit link function.

Pubertal marker	P_3	P_{10}	P_{25}	P_{50}	P_{75}	P_{90}	P_{97}
Breast (Ultrasound; n=969)							
US B1	NA	NA	6.5	8.4	10.0	11.2	12.2
US B2	7.8	8.6	9.4	10.2	11.1	11.9	12.6
US B3	9.2	10.0	10.7	11.5	12.3	13.0	13.7
US B4	10.5	11.0	11.6	12.3	13.3	14.7	NA
US B5	11.5	12.2	13.3	15.9	NA	NA	NA
Breast (Tanner; n=700)							
Tanner B2	8.2	8.9	9.6	10.4	11.2	11.9	12.6
Tanner B3	9.7	10.2	10.8	11.4	12.1	12.9	13.7
Tanner B4	10.8	11.3	11.9	12.5	13.3	14.3	16.4
Tanner B5	11.2	11.9	12.9	14.7	NA	NA	NA
Pubic Hair (Tanner; n=372)							
Tanner PH2	8.6	9.3	10.1	10.9	11.7	12.4	13.1
Tanner PH3	9.8	10.63	11.3	11.9	12.6	13.1	13.6
Tanner PH4	10.6	11.4	12.2	13.1	13.8	14.8	15.5
Tanner PH5	11.6	12.6	13.6	15.1	NA	NA	NA
Menarche (n = 643)							
	11.0	11.5	12.0	12.7	13.5	14.5	15.9

Abbreviations: SD: standard deviation; US B: ultrasound-based breast stage Tanner B: Tanner breast stage; PH: pubic hair, NA: not available (outside the data range)

Supplemental Table 3. Estimated mean ages at reaching ultrasound-based breast stage 2 (US B2), Tanner breast stage 2 (B2), Tanner pubic hair stage 2 (PH2) and menarche, in the girls with known non-Norwegian (n = 92) and in girls that are known to be of Norwegian origin (n = 374).

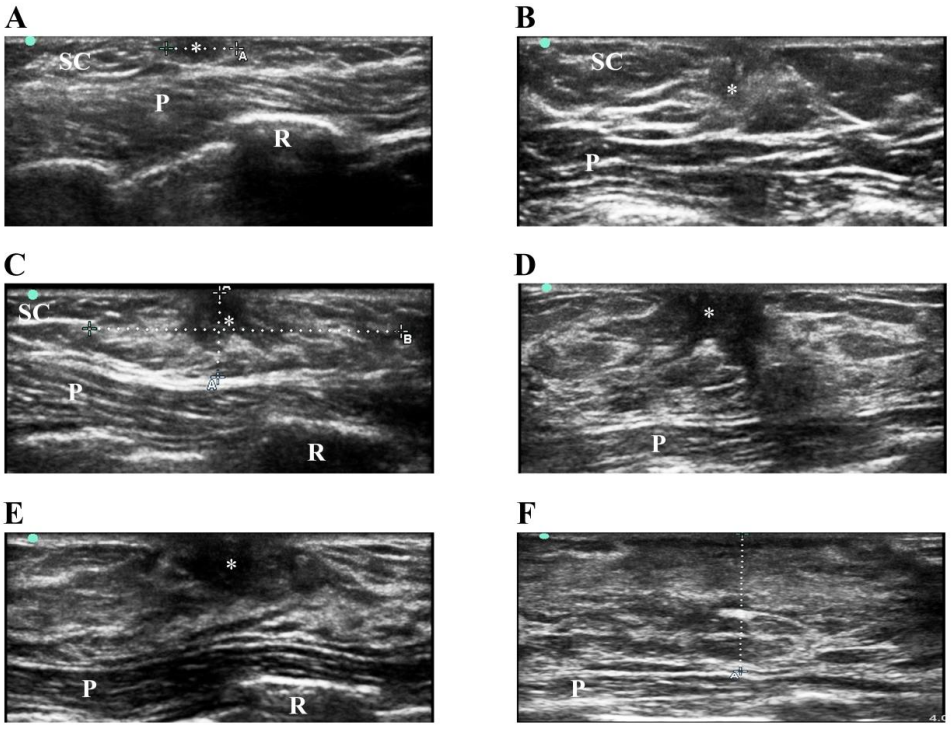
	Non Norwegian				Norwegian origin				
Puberty measure	**n**	**Mean**	**SD**	**95% CI**	**n**	**Mean**	**SD**	**95%CI**	**P-value**
US B2	92	9.84	1.16	(9.37 - 10.3)	371	10.46	1.29	(10.21 - 10.71)	0.04
Tanner B2	92	10.13	1.00	(9.70 - 10.56)	372	10.56	1.24	(10.32 - 10.80)	0.10
Tanner PH2	59	10.44	1.11	(9.86 - 11.01)	205	11.08	1.16	(10.75 - 11.40)	0.08
Menarche	92	12.64	1.33	(12.01 - 13.27)	374	13.08	1.19	(12.79 - 13.37)	0.10

Abbreviations: SD: standard deviation; CI: confidence interval; P-value from logistic regression using age as a covariate and origin (Norwegian, Non-Norwegian) as the variable of interest

Supplemental material: "References for ultrasound staging of breast maturation, Tanner breast staging, pubic hair and menarche in Norwegian girls"

Bruserud IS, Roelants M, Oehme NHB, Madsen A, Eide GE, Bjerknes R, Rosendahl K, Juliusson P

Figures and figure legends to Supplemental Figures 1, 2 and 3



Supplemental Figure 1. Ultrasound-based stages of pubertal breast development (US B)

A Midsagittal view through the left breast in a 7.1-year-old girl conforming to stage US B0. Image shows a small elongated mass (asterisk) hypoechoic relative to the surrounding subcutaneous (SC) adipose tissue. The ribs (R) are located posterior to the pectoral muscle (P) (1).

B Midsagittal view through the left breast in a 10.8 –year-old girl conforming to stage US B1. The breast tissue is hyperechoic compared to the surrounding subcutaneous tissue (SC), but not to the pectoral muscle (P) and has a triangular shape (asterisk) (1).

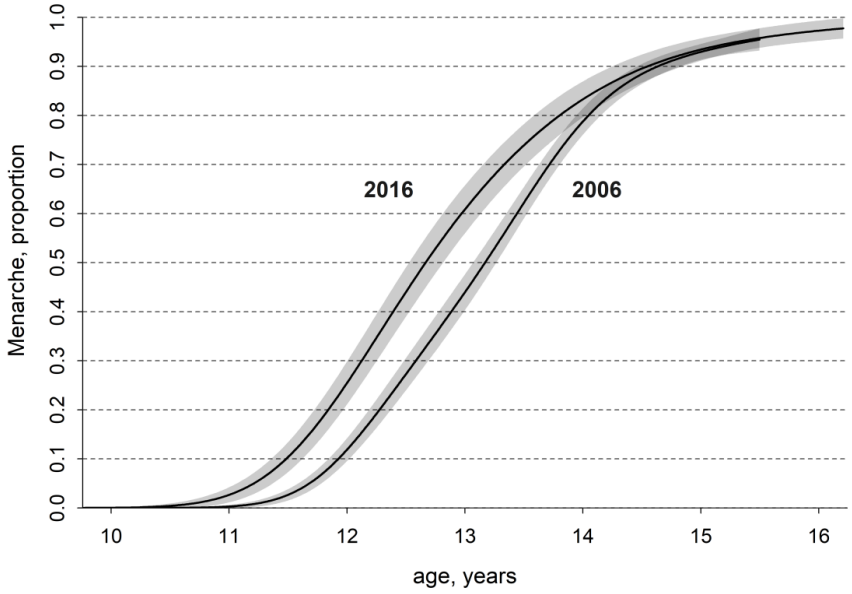
C Midsagittal view through the left breast in a 9.7-year-old girl conforming to stage US B2. The image shows well-defined breast tissue that appears hyperechoic in relation to the

subcutaneous (SC) tissue, but not to the pectoral muscle (P), and has a small hypoechoic roundish centre (asterisk). Rib (R) (1).

D Midsagittal view through the left breast in a 10.7 –year-old girl conforming to stage US B3. The breast tissue appears hyperechoic with a hypoechoic "spider-shaped" centre (asterisk). Pectoral muscle (P) (1).

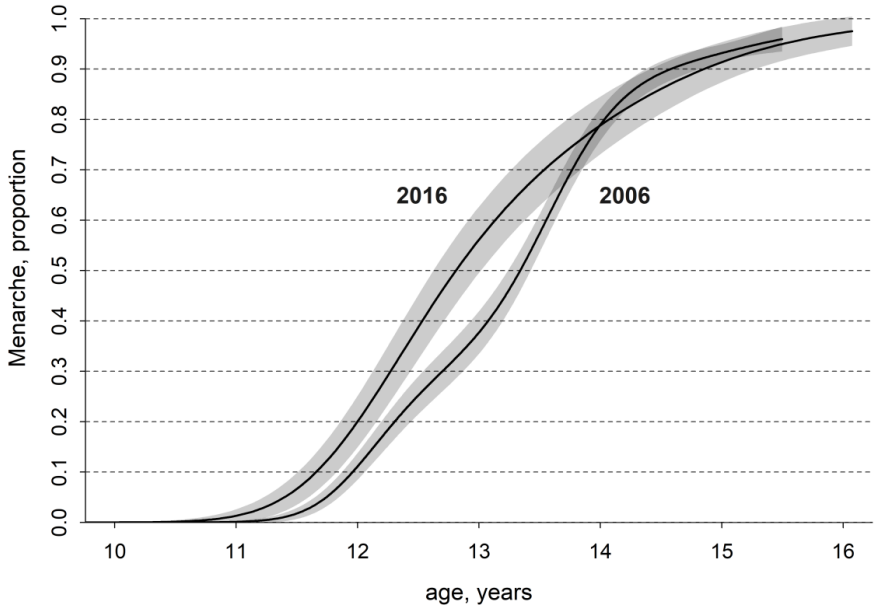
E Midsagittal view through the left breast in a 12.0 –old-girl conforming to stage US B4. The image shows hypoechoic glandular tissue in the subareolar area that has a rounder shape compared to US B3. The ribs (R) are located posterior to the pectoral muscle (P) (1).

F Midsagittal view through the left breast in a 13.9-old-girl conforming to stage US B5. Mature breast tissue is seen as a heterogeneous mass without any hypoechoic centre. Pectoral muscle (P) (1).



Supplemental Figure 2

Cumulative proportion of girls with menarche by age estimated with a generalized additive model for binary data with probit link. Data from all girls included in the studies from 2006 (BGS1) and 2016 (current study, BGS2). Grey areas indicate the standard errors below/above the estimated age prevalence.



Supplemental Figure 3

Cumulative proportion of girls with menarche by age estimated with a generalized additive model for binary data with probit link. Data from girls of Norwegian or Nordic origin in the studies from 2006 (BGS1) and 2016 (current study, BGS2). Grey areas indicate the standard errors below/above the estimated age prevalence.

Reference

1. Bruserud IS, Roelants M, Oehme NHB, Madsen A, Eide GE, Bjerknes R, Rosendahl K, Juliusson P. Ultrasound assessment of pubertal breast development in girls: intra- and interobserver agreement. Pediatr Radiol 2018; 48(11):1576-1583.

III

Bratke et al. BMC Pediatrics (2017) 17:138
DOI 10.1186/s12887-017-0893-x

BMC Pediatrics

RESEARCH ARTICLE Open Access

Timing of menarche in Norwegian girls: associations with body mass index, waist circumference and skinfold thickness

Heiko Bratke[1], Ingvild Særvold Bruserud[2], Bente Brannsether[3,4], Jörg Aßmus[5], Robert Bjerknes[3], Mathieu Roelants[6] and Pétur B. Júlíusson[2,3*]

Abstract

Background: Research studies show conflicting results regarding the association between menarche and body weight. The purpose of the present study was to investigate if anthropometric indicators of body composition, body mass index (BMI), waist circumference (WC), triceps (TSF) and subscapular skinfold (SSF) thicknesses, were differentially associated with age at menarche in Norwegian girls.

Methods: The association between menarche and BMI, WC, TSF and SSF was investigated in 1481 girls aged 8–15. 5 years, and in a subgroup of 181 girls with menarche during the 12 months prior to examination. Anthropometric measures were categorized as low (< −1SDS), average (−1 ≤ SDS ≤ +1) or high (> 1SDS), and menarche according to this classification was analysed with Kaplan-Meier curves and unadjusted and adjusted Cox regression.

Results: The median age at menarche in the total sample was 13.1 years. In the unadjusted models, low categories of all traits were associated with later menarche, and high categories with earlier menarche. When adjusted for other covariates, earlier menarche was only related with a high BMI (Hazard Ratio 1.41, 95% confidence interval (CI) 1.07, 1.85), and later menarche with a low BMI (HR 0.53, 95%CI 0.38, 0.75) and low SSF (HR 0.54, 95%CI 0.39, 0.75). In girls with recent menarche, early menarche was significantly associated with a high BMI in the final model (HR 1.79, 95%CI 1.23, 2.62).

Conclusions: The timing of menarche was associated with the BMI, WC, TSF and SSF, but more strongly so with the BMI. These associations may be related to a common tempo of growth, as the mean age at menarche has remained stable during the last decades during a time period while the prevalence of overweight and obesity has increased significantly.

Keywords: Menarche age, Puberty, Overweight, BMI, Skinfold thickness, Waist circumference

Background

In Western populations, the age at menarche has declined about 0.3 years per decade until the 1950's [1, 2]. In Norwegian girls, the mean age at menarche was reported to be above 16 years in 1830, whereas it has been stable around 13 years since the 1940s [3–6]. In contrast, studies from Denmark and the US showed a further decline in the age of pubertal onset in girls, in particular the age of breast development [7, 8]. Changes in age at menarche were however modest, suggesting asymmetrical secular trends in the timing of pubertal development [7, 9].

The observed variation in the activation of the hypothalamic-pituitary-gonadal axis and menarche, reflects interactions between genetic background and various endogenous and exogenous factors [1]. Prenatal conditions, nutrition and body composition, light exposure, endocrine disrupting chemicals (EDC) and psychosocial stress have been suggested as possible regulators of this development [1, 10].

* Correspondence: petur.juliusson@uib.no
[2]Department of Paediatrics, Haukeland University Hospital, Bergen, Norway
[3]Department of Clinical Science, Section of Paediatrics, University of Bergen, 5021 Bergen, Norway
Full list of author information is available at the end of the article

In the early 1970s, Frisch and Revelle suggested a "critical weight" theory, pointing out the relationship between weight and pubertal timing [11]. Later, the leptin hormone secreted by fat tissue, has been shown to function as a signal for energy status and as a permissive factor for pubertal onset by modulating the Kiss1/Kiss1R system [12]. The Kiss1/Kiss1R system seems to represent a link between the reproductive and the energy system, influenced not only by leptin, but also ghrelin, insulin, pro-opiomelanocortin and neuropeptide Y [13]. Therefore, one can speculate that the increase in the prevalence of overweight and obesity should result in a downward trend in pubertal timing, and several studies have provided evidence for this association [14, 15]. However, a Danish study showed that, although a relatively high BMI was consistently associated with an earlier onset of the pubertal growth spurt in children born between 1930 and 1969, such a trend was also seen in girls with a low BMI. Therefore, the BMI alone could not explain the observed downward trend [16].

In the present study, age at menarche was compared to anthropometric indicators of body composition in Norwegian girls, notably body mass index (BMI), waist circumference (WC), and triceps (TSF)- and subscapular (SSF) skinfold thickness. Our hypothesis was that measures of subcutaneous and/or central fat tissue would correlate more strongly with age at menarche compared to an overall index of body weight.

Methods

The present analysis is based on a sample of 8–15.5 year old girls who participated in the Bergen Growth Study (BGS). This cross-sectional study on growth from birth to 19 years of age included 4035 girls measured in 2003–2006 and representative for Bergen County. All participating girls in the included primary schools (grades 1–10) were visually assessed by a study nurse for signs of breast development, and subsequently asked about menarche and, if applicable, age at menarche (month/year) [17]. From this study population, 68 girls were excluded because of diseases known to affect growth, and 4 because of incomplete information. In total, 1481 girls were between 8 and 15.5 years and thus eligible for inclusion in the present study ("total study sample"). Girls who had their menarche during the 12 months preceding the measurement represented a subgroup of "girls with recent menarche". Age at menarche was recorded and height, weight, WC, TSF and SSF were measured according to standardized procedures [17–19]. Briefly, height was measured to the nearest 0.1 cm with a portable Holtain stadiometer (Crosswell, UK) and weight to nearest 0.1 kg with a Seca personal digital scale (Hamburg, Germany). Waist circumference was measured at the end of normal expiration with a Lufkin W606 PM metal measurement tape placed at the midpoint between the lowest rib and the top of the iliac crest. Skinfolds were measured on the left side with a Holtain Skinfold Caliper (Crosswell, UK). The TSF was measured on the back side of the upper arm midway between the acromion and the radial head. The SSF was measured 2 cm below the inferior angle of the scapula. The rationale for choosing these were threefold: [1] the triceps and subscapulae are common sites for taking skinfolds; [2] at these sites, the measurements are easier to perform and standardize compared to other measurement sites; and [3] we wanted to include both a "peripheral" (TSF) and more central or "truncal" (SSF) site. All anthropometric measures were converted to SD scores (SDS) using national growth references, and categorized as low (< −1SDS), average (−1 ≤ SDS ≤ +1) or high (> 1SDS).

Statistical analysis

Median age at menarche was estimated with Kaplan-Meier analysis, taking into account that girls who had not yet reached menarche are right censored. The mean and variance were estimated with probit analysis of status presens data assuming a Gaussian distribution of age at menarche in the total sample, and calculated as the arithmetic mean and SD in the subgroup with recent menarche. The association between the menarche and the anthropometric measures (BMI, WC, SSF and TSF) was analysed with Cox proportional hazards models. Predictor variables were grouped as low, average or high as described above, and the average group was used as the reference category. In addition, BMI was also classified as an ordinal variable with four levels according to IOTF (International Obesity Task Force) criteria for overweight (the equivalent of a BMI ≥ 25 kg/m^2 in adults) and obesity (equivalent of BMI ≥ 30 kg/m^2), and the analogous criteria for underweight (the equivalent of a BMI ≤ 18.5 kg/m^2) [20, 21]. The IOTF-classification was included because its common use in clinical and research settings. Results are presented for unadjusted simple regression models of each marker separately, fully adjusted multiple regression models including all anthropometric measures, and final models which are the result of a (backwards) stepwise removal of statistically not significant covariates (using $p > 0.1$ as a conservative criterion for removal). The sample size of the Bergen Growth Study was estimated with the aim to detect secular changes in height and weight since the 1970s. A post hoc power analysis of the present study shows that the sample allows to detect a statistically significant hazard ratio of approximately 1.3 in the analysis of all girls, and a hazard ratio of approximately 1.5 in the subgroup of girls with recent menarche.

Test results with a p-value less than 0.05 were considered as statistically significant. Data analysis was done

using IBM SPSS release 21.0 and in R version 3.2 (R Foundation for Statistical Computing, Vienna, Austria, 2015).

Results

Of the total study sample of 1481 girls, 477 reported menarche, of whom 181 had experienced their first menstruation during the 12 months preceding measurement ("girls with recent menarche"). In the total study sample, the Kaplan-Meier median age at menarche was 13.1 years (95%CI 13.0–13.3), while the probit mean was 13.3 years (95%CI 13.1–13.4) with a corresponding SD of 0.9 years. In girls with recent menarche, the median age at menarche was 13.2 years (95%CI 13.0–13.4), and the probit mean 13.2 (95%CI 13.0–13.3) with a corresponding SD of 0.9 years.

In the total study sample, the IOTF-defined prevalence of overweight including obesity was 14.4%, and that of underweight 9.7%, and in girls with recent menarche, 11.1% and 13.8%, respectively.

In the unadjusted analysis of the total study sample, low BMI SDS was associated with later menarche and a high BMI SDS with earlier menarche (Table 1, Fig. 1). Comparable results were obtained for WC, TSF and SSF. In the fully adjusted model, a high BMI was significantly associated with menarche at an earlier age, and both a low BMI and low SSF with menarche at a later age. Backward elimination of non-significant variables preserved the level of significance and did not alter the corresponding Hazard ratios by more than 0.04. The timing of menarche was associated with the IOTF weight classes in a similar way: median age (95%CI) was 12.5 (12.1–13.0) years in girls with overweight including obesity (n = 213), 13.1 (13.0–13.2) years in girls with normal weight (n = 1124), and 14.1 (13.7–14.6) years in girls with underweight (n = 144).

The same trends were observed in the group of girls with recent menarche, but with a lower significance level, probably because of the smaller sample size (Table 1). Also, the effect sizes were usually larger for the high categories of BMI, WC, TSF and SSF, and smaller for the low categories, when compared to the total cohort.

Discussion

In the present study, we found a median age at menarche in Norwegian girls of 13.1 years. Although the timing of menarche was strongly related to all weight related anthropometric variables, early menarche was only significantly associated with a high BMI, and late menarche with a low SSF or low BMI in multiple Cox regression analysis.

The mean age of menarche in Norwegian girls has been stable for more than half a century, while the prevalence of overweight and obesity has continued to rise during the past decades [22, 23]. Rosenberg et al. reported the age at menarche between 1830 and 1960 using recall data from Norwegian maternity hospitals, and showed a steady decline that reached a nadir just above 13 years of age in the 1940s [3]. More recently, comparable estimates of age at menarche in Norwegian girls have been published by Tell et al. [5], using youth data from Oslo, and by Bratberg et al., using data collected from 1995 to 2001 for the Young-HUNT study [6]. The estimates from our own study conclude this list and confirm that age at menarche in Norwegian girls has not changed since the 1940s. The current median age at menarche of 13.1 years in Norwegian girls is similar to that observed in other North-European countries [1]. Recent findings from the USA, Denmark and the Netherlands showed a slight decrease in menarcheal age during the last decade [7, 9, 24]. We could not document such a trend, in Norwegian girls.

All weight-related anthropometric variables were strongly associated with the timing of menarche. A relatively low value of the BMI, WC, TSF and SSF was associated with menarche at a later age, and a relatively high value with an earlier age at menarche. Earlier studies have linked early menarche to higher weight status in childhood [25, 26], and Adair et al. found a higher prevalence of overweight in girls with menarche before 11 years compared to those with a menarche after 14 years of age [27]. This difference remains up to young adulthood at least, as follow-up of young women with early menarche showed higher BMI and larger skinfold-thickness at the age of 21 and 27 years, when compared to women with late menarche [28]. Finally, using data from 34 European countries and North America, Currie et al. could show that at on an individual level, age at menarche was 1 month earlier for each unit increase in BMI [14]. Further, mean age at menarche was approximately 1 week earlier with each percentage point increase in prevalence of overweight/obesity at country level [14]. About 14.4% of the girls in our study were overweight according to the IOTF. Although this estimate roughly corresponds to a threefold increase in the prevalence of overweight during the past decades [22], it is surprising that the age at menarche has not changed during this period. This indicates that the mechanism behind the association between weight status and tempo of maturation, at least for menarche as an endpoint, may be more complex than a direct causal relation.

In the multiple Cox regression model, a relatively high BMI was related with early menarche and low BMI or SSF was with late menarche. Our hypothesis was that measures of subcutaneous fat tissue like TSF and SSF, would show a stronger relation with menarche than the BMI, which measures both fat mass and lean mass. However, the present findings show the opposite, as we

Table 1 Kaplan-Meier estimates and Cox regression of menarche according to the BMI, WC, TSF and SSF[a] in the total sample and in girls with recent menarche

			Unadjusted model[b]		Fully adjusted model[b]		Final model[b]	
	N	Median	HR (95% CI)	*P*	HR (95% CI)	*P*	HR (95% CI)	*P*
All girls (*N* = 1481)								
BMI								
Low	262	14.1	0.40 (0.30–0.55)	<0.001	0.56 (0.39–0.82)	0.001	0.53 (0.38–0.75)	<0.01
Normal	960	13.1	1.00		1.00		1.00	
High	259	12.6	1.46 (1.16–1.84)	0.001	1.45 (1.03–2.05)	0.01	1.41 (1.07–1.85)	0.01
WC								
Low	236	13.7	0.54 (0.40–0.71)	<0.001	0.93 (0.67–1.30)			
Normal	1000	13.1	1.00		1.00			
High	244	12.8	1.33 (1.04–1.71)	0.01	0.86 (0.59–1.24)			
TSF								
Low	235	13.6	0.59 (0.45–0.78)	<0.001	0.97 (0.72–1.30)			
Normal	988	13.1	1.00		1.00			
High	248	12.7	1.33 (1.04–1.71)	0.01	1.01 (0.73–1.41)			
SSF								
Low	237	14	0.40 (0.29–0.54)	<0.001	0.53 (0.37–0.75)	<0.001	0.54 (0.39–0.75)	<0.01
Normal	965	13.1	1.00		1.00		1.00	
High	273	12.8	1.21 (0.95–1.54)		1.03 (0.73–1.45)		0.96 (0.72–1.27)	
Girls with recent menarche (*N* = 181)								
BMI								
Low	26	13.60	0.76 (0.50–1.17)		0.92 (0.51–1.66)		0.76 (0.50–1.17)	
Normal	118	13.10	1.00		1.00		1.00	
High	37	12.70	1.79 (1.23–2.62)	0.001	1.19 (0.67–2.12)		1.79 (1.23–2.62)	0.003
WC								
Low	19	13.70	0.82 (0.50–1.33)	<0.001	0.99 (0.52–1.87)			
Normal	131	13.20	1.00		1.00			
High	31	12.60	1.96 (1.32–2.94)	0.01	1.40 (0.79–2.48)			
TSF								
Low	36	13.50	0.80 (0.55–1.17)		1.08 (0.69–1.70)			
Normal	119	13.10	1.00		1.00			
High	23	12.70	1.74 (1.11–2.74)		1.11 (0.61–2.04)			
SSF								
Low	25	13.50	0.65 (0.42–1.00)	0.050	0.66 (0.39–1.14)			
Normal	124	13.10	1.00		1.00			
High	32	12.70	1.71 (1.15–2.53)	0.01	1.25 (0.73–2.13)			

[a]Grouped as low (< −1 SDS), normal (> −1 SDS – > +1 SDS) or high (> +1 SDS) BMI, WC, TSF and SSF. Estimates and models based on the total sample take into account that some girls have not yet reached menarche (censored)

[b] Unadjusted models are simple Cox regression models for each marker separately; fully adjusted models are multiple Cox regression models including all anthropometric markers; and final models are multiple Cox regression models after backward elimination of statistically not significant covariates ($p > 0.1$)

found the BMI to be a consistent predictor of early and late menarche. This suggests a strong association between the increase in BMI and tempo of maturation, and could explain why menarche is more closely related to the BMI than to measures of subcutaneous (skinfolds) or central (WC) fat tissue. However, the later menarche in girls with a low SSF is equally interesting as one can speculate that SSF, as a measure of truncal subcutaneous fat tissue, has more impact on maturational processes than TSF, which is a measure of peripheral fat tissue.

A limitation in the present work is that girls might have changed weight status between menarche and the

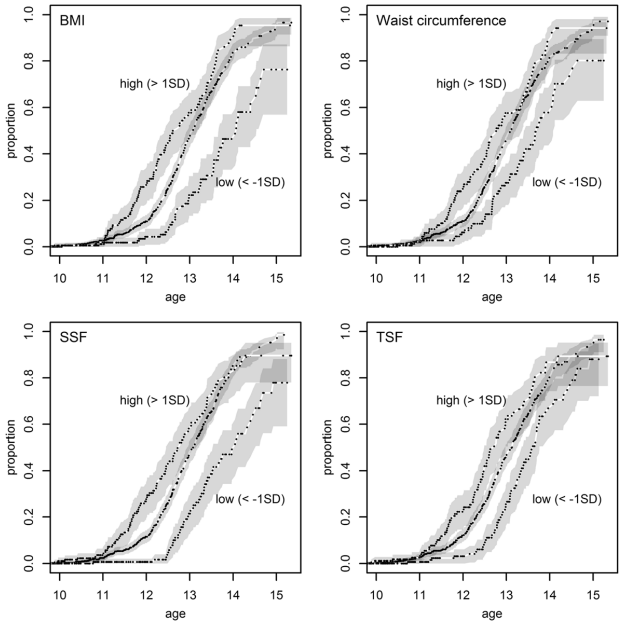


Fig. 1 Kaplan-Meier curves of age at menarche showing that menarche occurs at an earlier age in girls with a high (> +1 SDS) BMI, Waist circumference, Subscapular (SSF) or Triceps (TSF) skinfold, and at a later age in girls with a low (< −1 SDS) value for these anthropometric indicators. The line in between shows the Kaplan-Meier estimate in girls with average (−1 to +1 SDS) measurement values. Dots are censored observations, and shaded areas indicate the 95% confidence interval

time of examination, in particular during puberty. Because of this possibility, we repeated our analysis in a subsample of girls who had their menarche within 12 months (mean time since menarche was 6.3 months) prior to the date of measurement [29]. The analysis of this subgroup showed the same trends as in the total study group, although not always statistically significant, probably due to the loss of statistical power because of the much smaller sample size. A further limitation is the cross-sectional nature of the data. Because of this, no causal conclusions can be made.

Conclusion

Although timing of menarche was clearly associated with weight-related anthropometric measures in Norwegian girls, the increasing prevalence of overweight and obesity seems to have had little impact on the mean age at menarche. Therefore, the observed association between weight-related anthropometric measures and menarche might be reflecting maturation to a larger extent than the degree of adiposity. This is also supported by the findings in the current study that anthropometric measures of subcutaneous fat tissue (TSF, SSF) and central fat (WC) did not show stronger relationship to menarche than BMI.

Abbreviations

BMI: Body mass index; EDC: Endocrine disrupting chemicals; IOTF: International obesity task force; SSF: Subscapularis skinfolds; TSF: Triceps skinfolds; WC: Waist circumference

Acknowledgments

Not applicable.

Funding

No external funding was secured for this study.

Availability of data and materials

The data generated or analysed during this study are available on reasonable request to the corresponding author.

Authors' contributions

HB analyzed the data and drafted the initial manuscript. MR and JA analyzed the data. PBJ, MR and RB conceptualized and designed the study. ISB and BB interpreted the data. All authors were involved in writing of the paper and had final approval of the submitted and published versions.

Competing interests

The authors declare that they have no competing interest.

Consent for publication

Not applicable.

Ethics approval and consent to participate
The BGS was approved by the Regional Committee for Medical and Health Research Ethics Region West. According to the Norwegian legislation, written informed consent was obtained from a parent of each participating child. In addition, written informed assent was obtained from all children above 12 years of age.

Publisher's Note

Author details
[1]Department of Internal Medicine, Section of Paediatrics, Haugesund District Hospital, Haugesund, Norway. [2]Department of Paediatrics, Haukeland University Hospital, Bergen, Norway. [3]Department of Clinical Science, Section of Paediatrics, University of Bergen, 5021 Bergen, Norway. [4]Department of Paediatrics, Stavanger University Hospital, Stavanger, Norway. [5]Haukeland University Hospital, Centre for Clinical Research, Bergen, Norway. [6]KU Leuven –Department of Public Health and Primary Care, KU Leuven - University of Leuven, Leuven, Belgium.

Received: 24 November 2016 Accepted: 31 May 2017
Published online: 06 June 2017

References

1. Parent AS, Teilmann G, Juul A, et al. The timing of normal puberty and the age limits of sexual precocity: variations around the world, secular trends, and changes after migration. Endocr Rev. 2003;24:668–93.
2. Ong KK, Ahmed ML, Dunger DB. Lessons from large population studies on timing and tempo of puberty (secular trends and relation to body size): the European trend. Mol Cell Endocrinol. 2006;254-255:8–12.
3. Rosenberg M. Menarcheal age for Norwegian women born 1830-1960. Ann Hum Biol. 1991;18:207–19.
4. Brundtland GH, Walloe L. Menarcheal age in Norway in the 19th century: a re-evaluation of the historical sources. Ann Hum Biol. 1976;3:363–74.
5. Tell GS. Cardiovascular disease risk factors related to sexual maturation: the Oslo Youth Study. J Chronic Dis. 1985;38:633–42.
6. Bratberg GH, Nilsen TI, Holmen TL, Vatten LJ. Early sexual maturation, central adiposity and subsequent overweight in late adolescence. a 4-year follow-up of 1605 adolescent Norwegian boys and girls: the Young HUNT study. BMC Public Health. 2007;7:54.
7. Aksglaede L, Sorensen K, Petersen JH, et al. Recent decline in age at breast development: the Copenhagen Puberty Study. Pediatrics. 2009;123:e932–9.
8. Herman-Giddens ME, Slora EJ, Wasserman RC, et al. Secondary sexual characteristics and menses in young girls seen in office practice: a study from the Pediatric Research in Office Settings network. Pediatrics. 1997;99:505–12.
9. Talma H, Schonbeck Y, van Dommelen P, et al. Trends in menarcheal age between 1955 and 2009 in the Netherlands. PLoS One. 2013;8:e60056.
10. Bourguignon JP, Juul A, Franssen D, et al. Contribution of the endocrine perspective in the evaluation of endocrine disrupting chemical effects: the case study of pubertal timing. Horm Res Paediatr. 2016;86:221-32
11. Frisch RE, Revelle R. Height and weight at menarche and a hypothesis of critical body weights and adolescent events. Science. 1970;169:397–9.
12. Fernandez-Fernandez R, Martini AC, Navarro VM, et al. Novel signals for the integration of energy balance and reproduction. Mol Cell Endocrinol. 2006; 254-255:127–32.
13. De Bond JA, Smith JT. Kisspeptin and energy balance in reproduction. Reproduction. 2014;147:R53–63.
14. Currie C, Ahluwalia N, Godeau E, Nic Gabhainn S, Due P, Currie DB. Is obesity at individual and national level associated with lower age at menarche? Evidence from 34 countries in the Health Behaviour in School-aged Children Study. J Adolesc Health. 2012;50:621–6.
15. Lee JM, Appugliese D, Kaciroti N, Corwyn RF, Bradley RH, Lumeng JC. Weight status in young girls and the onset of puberty. Pediatrics. 2007;119: e624–30.
16. Aksglaede L, Juul A, Olsen LW, Sorensen TI. Age at puberty and the emerging obesity epidemic. PLoS One. 2009;4:e8450.
17. Juliusson PB, Roelants M, Nordal E, et al. Growth references for 0–19 year-old Norwegian children for length/height, weight, body mass index and head circumference. Ann Hum Biol. 2013;40:220–7.
18. Brannsether B, Roelants M, Bjerknes R, Juliusson P. Waist circumference and waist-to-height ratio in Norwegian children 4–18 years of age: reference values and cut-off levels. Acta Paediatr. 2011;100:1576–82.
19. Brannsether B, Roelants M, Bjerknes R, Juliusson PB. References and cutoffs for triceps and subscapular skinfolds in Norwegian children 4–16 years of age. Eur J Clin Nutr. 2013;67:928–33.
20. Cole TJ, Bellizzi MC, Flegal KM, Dietz WH. Establishing a standard definition for child overweight and obesity worldwide: international survey. BMJ. 2000;320:1240–3.
21. Cole TJ, Flegal KM, Nicholls D, Jackson AA. Body mass index cut offs to define thinness in children and adolescents: international survey. BMJ. 2007;335:194.
22. Juliusson P, Eide G, Roelants M, et al. Overweight and obesity in Norwegian children: Secular trends in weight-for-height and skinfolds. Acta Paediatr. 2007;96:1333–7.
23. Juliusson PB, Eide GE, Roelants M, et al. Overweight and obesity in Norwegian children: prevalence and socio-demographic risk factors. Acta Paediatr. 2010;99:900–5.
24. McDowell MA, Brody DJ, Hughes JP. Has age at menarche changed? Results from the National Health and Nutrition Examination Survey (NHANES) 1999-2004. J Adolesc Health. 2007;40:227–31.
25. Salgin B, Norris SA, Prentice P, et al. Even transient rapid infancy weight gain is associated with higher BMI in young adults and earlier menarche. Int J Obes. 2015;39:939–44.
26. Stark O, Peckham CS, Moynihan C. Weight and age at menarche. Arch Dis Child. 1989;64:383–7.
27. Adair LS, Gordon-Larsen P. Maturational timing and overweight prevalence in US adolescent girls. Am J Public Health. 2001;91:642–4.
28. van Lenthe FJ, Kemper CG, van Mechelen W. Rapid maturation in adolescence results in greater obesity in adulthood: the Amsterdam Growth and Health Study. Am J Clin Nutr. 1996;64:18–24.
29. Bau AM, Ernert A, Schenk L, et al. Is there a further acceleration in the age at onset of menarche? A cross-sectional study in 1840 school children focusing on age and bodyweight at the onset of menarche. Eur J Endocrinol. 2009;160:107–13.

IV

Forespørsel om deltakelse i forskningsprosjektet
Barn under 12 år
"Vekststudien i Bergen 2"

Kjære foreldre/foresatte!

I 2003-6 gjennomførte vi «Vekststudien i Bergen 1» som gjorde det mulig å tegne vekstkurver som er brukt i hele landet. Med dette brevet ber vi om deres samtykke til å inkludere deres barn i en studie for å utforske pubertetsutviklingen til norske barn. Vi vet svært lite om denne utviklingen, noe som er problematisk når en skal vurdere om et barn har for tidlig eller har for sen utvikling (da trenger vi informasjon om hva som er normalt). I andre land, for eksempel Danmark og USA, har en de siste årene funnet tegn på at pubertetsutviklingen kommer tidligere nå enn før. Det er uklart hva dette skyldes. Det er mulig at barn går tidligere i pubertet pga endret kroppssammensetning (økt fettvev) eller stoffer i omgivelsene (såkalte hormonhermere) som kan påvirke utviklingen. Tidlig pubertetsutvikling kan være uheldig og øke risikoen for overvekt og høyt kolesterol senere i livet. I tillegg har noen studier vist assosiasjon mellom for tidlig pubertet og enkelte kreftformer. Hensikten med studien er å kartlegge normal pubertetsutvikling hos norske barn ved å vurdere vekst og utvikling, måling av hormoner i blod og spytt og kartlegge de genene som er involvert. I tillegg ønsker vi å undersøke faktorer som kan påvirke utviklingen med måling av kroppssammensetning og forekomsten av hormonhermere i kroppen.

Hva innebærer studien?
Studien innebærer målinger av høyde, vekt og mageomfang. Alle undersøkelsene vil skje på skolen til barnet. En kvinnelig sykepleier vil måle størrelsen av brystkjertlene med ultralyd hos jentene og en mannlig radiograf vil måle størrelsen av testiklene hos guttene, også med ultralyd, og utvikling av kjønnshår registreres. En blir godt dekket til og barna skal ikke føle seg eksponert. Vanligvis vil en ikke trenge å ta av seg undertøy/singlet. En og en blir undersøkt om gangen og adskilt fra andre. I tillegg blir kroppssammensetningen målt med bioimpetanse, et utstyr som måler hvor mye muskler, knokler og fett det er i kroppen. I etterkant av dette blir det tatt blodprøve og spyttprøve. Alle undersøkelsene tar kort tid og er smertefrie (unntatt blodprøvestikk, men hvis barnet ditt ønsker det, kan en få bedøvelsessprey på huden først). Vi ber deg/dere også om å fyllet ut et spørreskjema.

Mulige fordeler og ulemper
Hvis en oppdager avvikende høyde, vekt eller pubertetsutvikling hos ditt barn vil vi informere om dette. Barnet vil få en skriftlig tilbakemelding hvor sin nåværende høyde er oppgitt. Ellers vil en ikke ha noen spesiell fordel av deltakelse, men barnet er med på å lage en «normal» for pubertetsutviklingen i Norge. Slik informasjon vil hjelpe andre barn som har avvikende utvikling fordi det er lettere å identifisere dem når en vet hvordan normal pubertetsutvikling er. Studien vil også gi verdifull informasjon om faktorer som påvirker pubertetsutviklingen. Ulempe med studien vil være litt ubehag i forbindelse med blodprøvetaking.

Hva skjer med prøvene og informasjonen om barnet?
Vi vil oppbevare prøver og informasjon om barnet slik at det kan brukes i pågående og fremtidige forskningsprosjekter for å finne mekanismer som ligger bak pubertetsutviklingen. Prøvene og informasjonen som registreres vil bli behandlet konfidensielt og vil bare bli brukt til forskning innen pubertetsutvikling.

Det er kun autorisert personell knyttet til prosjektet som har adgang til navnelisten og som kan finne tilbake til barnet. Det vil ikke være mulig å identifisere barnet i resultatene av studien når disse publiseres.

For noen forskningsprosjekter kan det være aktuelt å sammenstille data fra våre studier om pubertet med offentlige registre, som for eksempel Reseptregisteret, Medisinsk fødselsregister, Kreftregisteret, Dødsårsaksregisteret og andre registre som er basert på opplysninger fra sykehus og primærhelsetjeneste. Alle slike sammenstillinger krever godkjenning av de offentlige instanser som er fastsatt i norsk lov.

Du/dere kan få jevnlig oppdatert informasjon om forskningsprosjekter som er omfattet av dette samtykket på www.vekststudien.no.

Frivillig deltakelse
Det er frivillig å delta i studien. Dere kan når som helst og uten å oppgi noen grunn trekke samtykket til å delta i studien. Dette vil ikke få konsekvenser. Dersom dere ønsker å delta, undertegner dere samtykkeerklæringen på siste side. Om dere nå sier ja til å delta, kan dere senere trekke tilbake deres samtykke. Dersom dere senere ønsker å trekke dere eller har spørsmål til studien, kan du/dere kontakte postdoktor/overlege Pétur B. Júlíusson (mobil tlf. 97777079 eller epost petur.juliusson@helse-bergen.no).

Alle som deltar i studien vil få en kinobillett og et diplom med informasjon om høyden sin!

Ytterligere informasjon om studien finnes i kapittel A – utdypende forklaring av hva studien innebærer.
Ytterligere informasjon om biobank, personvern og forsikring finnes i kapittel B – Personvern, biobank, økonomi og forsikring.

Samtykkeerklæring følger etter kapittel B.

Med vennlig hilsen og på vegne av prosjektgruppen,

Pétur B. Júlíusson
Overlege/Postdoktor
Barneklinikken
HELSE BERGEN HF

Kapittel A- utdypende forklaring av hva studien innebærer

- Alle barn på skolen blir invitert til å delta.
- I Norge har vi ingen informasjon om normal pubertetsutvikling hor norske barn. I Danmark og USA har en observert utvikling mot tidligere pubertet de senere år. Det er ukjent hva som forårsaker dette. I vår studie ønsker vi å få informasjon om normal pubertet hos barn i Norge og å finne faktorer som kan påvirke utviklingen.
- Vi vil måle barnets høyde og vekt samt mageomfang, og også bruke ultralyd til å måle utviklingen av bryst og testikler. Hos jentene blir «lydhode» satt på brystet mens området rundt blir godt dekket til. Hos guttene blir en slik «lydhode» lagt på en testikkel, mens området rundt er godt dekket til. Utviklingen av kjønnshår registreres. Barnet skal ikke på noe tidspunkt oppleve at det blir eksponert. Vanligvis vil en ikke trenge å ta av seg undertøy/singlet. Kun ett og ett barn blir målt om gangen. I etterkant blir kroppssammensetningen målt med såkalt bioimpetanse (utstyr som måler hvor mye muskel, skjelett og fett er i kroppen din) og avslutningsvis blir det tatt blodprøve og spyttprøve. En undergruppe blir bedt om å levere inn urinprøve.
- Dere foreldre/foresatte blir bedt om å fylle ut et spørreskjema.
- Undersøkelsen vil foregå på skolen, og ett og ett barn blir undersøkt adskilt fra hverandre. Det hele vil ta ca 15 minutter.
- Undersøkelsen skal ikke gi noe ubehag utenom blodprøvetaking. Hvis ønskelig kan barnet få bedøvelsessprey på huden før det tas blodprøve.
- Studiens ansatte vil gi grundig informasjon om forskningsprosjektet i klasserommet. Barna tar med seg hjem dette informasjonsskrivet/samtykkeerklæringen/spørreskjema. Kun de barna som leverer tilbake undertegnet samtykke blir undersøkt.

Kapittel B - Personvern, biobank, økonomi og forsikring

Personvern
Så snart opplysningene og prøvene er samlet inn vil de bli lagret uten navn og fødselsnummer eller andre direkte gjenkjennende opplysninger. En kode knytter barnets opplysninger og prøver gjennom en navneliste. Det er kun autorisert personell knyttet til prosjektet som har adgang til navnelisten og som kan finne tilbake til barnet. Alle medarbeidere i prosjektene har taushetsplikt. Dere kan føle dere trygg på at informasjonen om barnet og prøvene vil bli behandlet med respekt for personvern og privatliv, og i samsvar med lover og forskrifter. Det vil ikke være mulig å identifisere barnet i resultatene av studien når disse publiseres. Data som inngår i studien barnet deltar i vil anonymiseres ved studienes slutt.

Biobank
Blod- og spyttprøvene som blir tatt vil bli lagret i en forskningsbiobank ved Haukeland Universitetssykehus. Hvis du/dere sier ja til å delta i studien gir du/dere også samtykke til at det biologiske materialet og analyseresultater inngår i biobanken. Pétur B. Júlíusson er ansvarshavende for forskningsbiobanken. Det biologiske materialet kan bare brukes etter godkjenning fra Regional komité for medisinsk og helsefaglig forskningsetikk (REK). Hvis det i fremtiden blir aktuelt å bruke prøver eller opplysninger om barnet til noe annet enn pubertetsstudier, kan det bli nødvendig å be om et nytt samtykke. Vi vil da sende dere et brev. Du/dere kan også bli spurt om et nytt samtykke hvis det blir aktuelt å samarbeide med private aktører om genetisk forskning. Slikt samarbeid vil være

underlagt offentlig regulering og kontroll. Det vil ikke i noen tilfeller være aktuelt å selge blodprøver eller annet biologisk materiale.

Biobanken vil inneholde følgende:
De opplysningene som inngår er barnets navn og fødselsnummer samt prøvematerialet. I Biobanken lagres nedfryst biologisk materiale fra barna som følges ved Barneklinikken ved Helse Bergen HF. Prøvene vil analyseres i forbindelse med spesifiserte forskningsprosjekter.

Hvem kan få tilgang til prøvematerialet?
Prøvene lagres permanent. Alle analysene foregår på avidentifisert materiale, men det kan bli aktuelt å koble materialet med opplysninger fra f.eks. registrere (om diagnose, behandling o.s.v.) og da brukes fødselsnummeret for å sammenstille prøve med opplysninger fra registeret. Når dette er gjort avidentifiseres prøver og datafiler. Alle som arbeider med analysene har taushetsplikt om forhold de får kjennskap til gjennom dette arbeidet.

Innsynsrett, endring og sletting av opplysninger
Du/dere kan til enhver tid få innsyn i hvilke opplysninger/prøver som er registrert om barnet. Du/dere har videre rett til å få korrigert eventuelle feil i de opplysningene vi har registrert. I tillegg kan du/dere når som helst kreve at innsamlede prøver og opplysninger om barnet blir slettet/destruert, uten at du må oppgi noen grunn. Sletting av data og biobankmateriale vil ikke innebære sletting fra anonymiserte forskningsfiler som allerede er benyttet i forskning. Det vil ikke ha noe betydning for barnet dersom du/dere velger å ikke avgi prøve, eller dersom du/dere senere ønsker å trekke deg.

Utlevering av materiale og opplysninger til andre
Utlevering av prøvemateriale vil bare skje til forskere i form av avidentifiserte prøver. Det vil si at alle prøver vil bli behandlet uten navn og fødselsnummer eller andre direkte gjenkjennende opplysninger. En kode knytter barnet til barnets prøver gjennom en navneliste. Det er kun autorisert personell knyttet til biobanken som har adgang til navnelisten og kan finne tilbake til barnet. Det kan bli aktuelt å sende prøver til analyser i utlandet, fortrinnsvis land i EU/EØS, men også andre land kan være aktuelle, f.eks. USA. Navnelisten vil uansett forbli i Norge. Skulle dette bli aktuelt vil nærmere opplysninger vil bli lagt ut på biobankens hjemmesider.

For forskningsprosjekter kan det være aktuelt å sammenstille informasjon fra biobanken med opplysninger fra sykejournal, helseundersøkelser, helseregistre og andre offentlige administrative registre. Alle forskningsprosjekter må imidlertid forhåndsgodkjennes av Den regionale komité for medisinsk forskningsetikk og andre offentlige instanser som loven krever.

Du kan følge med på aktuelle forskningsprosjekter på vår internettside: www.vekststudien.no
Registeransvarlig/kontaktperson: Overlege/postdoktor Pétur B. Júlíusson, Barneklinikken, Haukeland Universitetssykehus. Tlf: 55975200.

Rett til innsyn og sletting av opplysninger om deg og sletting av prøver
Hvis du/dere sier ja til å delta i studien, har dere rett til å få innsyn i hvilke opplysninger som er registrert om barnet. Dere har videre rett til å få korrigert eventuelle feil i de opplysningene vi har registrert. Dersom barnet trekker seg fra studien, kan du/dere kreve å få slettet innsamlede prøver og opplysninger, med mindre opplysningene allerede er inngått i analyser eller brukt i vitenskapelige publikasjoner.

Samtykke til deltakelse i studien og biobanken

Vi gir herved tillatelse til at vårt barn

(Barnets navn (blokkskrift) og fødselsdato)

kan delta i «Vekststudien i Bergen 2» og avgi prøver til biobanken i henhold til informasjonsskrivet.

Sted, dato (Underskrift foreldre/foresatte)

Errata

Page 13 1st paragraph: Incorrect statement without precise definition of what is meant by reproduction: "(…) occurrence of the first menstruation (menarche, which results in the attainment of reproductive function" corrected to "(…)occurrence of the first menstruation (menarche)".

Page 53 2nd paragraph: Missing number: "The prevalence of parents with a higher level of education was 79.6% in the BGS2 versus 62.4% in the BGS", - corrected to "The prevalence of parents with a higher level of education was 79.6% in the BGS2 versus 62.4% in the BGS1"

Page 68 1st paragraph: Misspelling: "(…) remained stable before BGS2" – corrected to "(…) remained stable before BGS1".

Graphic design: Communication Division, UiB / Print: Skipnes Kommunikasjon AS

uib.no

ISBN: 9788230864029 (print)
9788230847985 (PDF)