Oestrogen versus androgen in hormone-replacement therapy for complete androgen insensitivity syndrome: a multicentre, randomised, double-dummy, double-blind crossover trial

Wiebke Birnbaum, Louise Marshall, Ralf Werner, Alexandra Kulle, Paul-Martin Holterhus, Katharina Rall, Birgit Köhler, Annette Richter-Unruh, Michaela F Hartmann, Stefan A Wudy, Matthias K Auer, Anke Lux, Siegfried Kropf, Olaf Hiort

Summary

Background Women with complete androgen insensitivity syndrome (CAIS) after gonadectomy have complained about reduced psychological wellbeing and sexual satisfaction. The aim of this study was to compare the effectiveness of hormone-replacement therapy with either androgen or oestrogen in women with 46,XY karyotype and CAIS after gonadectomy.

Methods This national, multicentre, double-blind, randomised crossover trial was performed at three university medical centres and three specialised treatment institutions in Germany. Eligible participants were women aged 18–54 years with 46,XY karyotype, genetically diagnosed CAIS, and removed gonads. Participants were randomly assigned (14:12) by a central computer-based minimisation method to either oestradiol 1·5 mg/day for 6 months followed by crossover to testosterone 50 mg/day for 6 months (sequence A) or to testosterone 50 mg/day for 6 months followed by crossover to oestradiol 1·5 mg/day for 6 months (sequence B). Participants also received oestradiol or testosterone dummy to avoid identification of the active substance. All participants received oestradiol 1·5 mg/day during a 2 months' run-in phase. The primary outcome was mental health-related quality of life, as measured with the standardised German version of the SF-36 questionnaire. Secondary outcomes were psychological wellbeing, as measured with the Brief Symptom Inventory (BSI), sexual function, as measured with the Female Sexual Function Index (FSFI), and somatic effects, such as signs of virilisation and effects on metabolic blood values. The primary analysis included all patients who were available at least until visit 5, even if protocol violations occurred. The safety analysis included all patients who received at least oestradiol during the run-in phase. This trial is registered with the German Clinical Trials Register, number DRKS00003136, and with the European Clinical Trials Database, number 2010-021790-37.

Findings We enrolled 26 patients into the study, with the first patient enrolled on Nov 7, 2011, and the last patient leaving the study on Jan 23, 2016. 14 patients were assigned to sequence A and 12 were assigned to sequence B. Ten participants were withdrawn from the study, two of whom attended at least five visits and so could be included in the primary analysis. Mental health-related quality of life did not differ between treatment groups (linear mixed model, p=0·794), nor did BSI scores for psychological wellbeing (global severity index, p=0·638; positive symptom distress index, p=0·378; positive symptom total, p=0·570). For the FSFI, testosterone was superior to oestradiol only in improving sexual desire (linear mixed model, p=0·018). No virilisation was observed, and gonadotrophin concentrations remained stable in both treatment groups. Oestradiol and testosterone concentrations changed substantially during the study in both treatment groups. 28 adverse events were reported for patients receiving oestradiol (23 grade 1 and five grade 2), and 38 adverse events were reported for patients receiving testosterone (34 grade 1, three grade 2, and one grade 3). One serious adverse event (fibrous mastopathy) and 20 adverse events (16 grade 1 and four grade 2) were reported during the run-in phase, and 12 adverse events during follow-up (nine grade 1 and three grade 2).

Interpretation Testosterone was well tolerated and as safe as oestrogen for hormone-replacement therapy. Testosterone can be an alternative hormone substitution in CAIS, especially for woment with reduced sexual functioning.

Funding German Federal Ministry of Education and Research.

Introduction

Disorders of sex development describe a group of rare congenital conditions leading to a discrepancy between chromosomal, gonadal, and phenotypic sex. Patients with disorders of sex development have claimed that management of their condition has been a major burden in the past,[1,2] leading to psychological comorbidities. So far, treatment concepts in disorders of sex development

Lancet Diabetes Endocrinol 2018

Published **Online** July 31, 2018 http://dx.doi.org/10.1016/S2213-8587(18)30197-9

See **Online/Comment** http://dx.doi.org/10.1016/S2213-8587(18)30206-7

Division of Paediatric Endocrinology and Diabetes, Department of Paediatric and Adolescent Medicine, University of Lübeck, Germany (W Birnbaum MD, L Marshall MS, R Werner PhD, Prof O Hiort MD); **Department of Paediatrics, Christian-Albrechts-University, Kiel, Germany** (A Kulle PhD, Prof P-M Holterhus MD); **Department of Women's Health, Centre for Rare Female Genital Malformations, Women's University Hospital, Tübingen University Hospital, Tübingen, Germany** (K Rall MD); **Department of Pediatric Endocrinology and Diabetology, Charité—Universitätsmedizin Berlin, Corporate Member of Freie Universität Berlin, Humboldt-Universität zu Berlin, Berlin, Germany** (B Köhler MD); **Paediatric Endocrinology, Department of Paediatrics, Universitätsklinikum Münster, Westfälische Wilhelms-Universität Münster, Germany** (Prof A Richter-Unruh MD); **Steroid Research & Mass Spectrometry Unit, Laboratory for Translational Hormone Analytics in Paediatric Endocrinology, Division of Paediatric Endocrinology & Diabetology, Centre of Child and Adolescent Medicine, Justus Liebig University, Giessen, Germany** (M F Hartmann PhD, Prof S A Wudy MD); **Research Group Clinical Neuroendocrinology, Max**

Planck Institute of Psychiatry, Munich, Germany (M K Auer MD)**; and Institute for Biometrics and Medical Informatics, Otto-von-Guericke University, Magdeburg, Germany** (A Lux PhD, Prof S Kropf PhD)

Correspondence to: Prof Olaf Hiort, University of Lübeck, Division of Paediatric Endocrinology and Diabetes, Department of Paediatric and Adolescent Medicine, 23538 Lübeck, Germany
olaf.hiort@uksh.de

Research in context

Evidence before this study

Patients with complete androgen insensitivity syndrome (CAIS) after gonadectomy report significantly reduced psychological wellbeing and sexual satisfaction.

Added value of this study

We showed that testosterone is a well tolerated hormone treatment for patients with CAIS. Serum hormone profiles correspond to typically male reference ranges, which is similar to those of patients with CAIS before gonadectomy. We found that patients with CAIS can benefit from continuous hormone substitution. However, compared with oestradiol, testosterone only improved sexual functioning. Yet sexual satisfaction is a relevant topic in CAIS, as reflected by very low scores on sexual functioning measured with the Female Sexual Function Index. Intra-individual differences were detected, with some participants gaining greater benefit from testosterone treatment.

Implications of all the available evidence

Patients with CAIS must be supplied with continuous hormone treatment. However, a relevant percentage of patients with CAIS do not receive adequate hormone substitution. Poor adherence to life-long medication because of lack of evidence of benefit and the inadequate medical expertise of caregivers contribute to this phenomenon. Testosterone treatment can be an alternative to oestradiol, especially to improve sexual functioning. Individual trends during both treatments must be analysed and correlated to endocrine profiles in blood and urine in further analyses.

associated with a 46,XY karyotype are based on personal experiences rather than medical evidence.

Complete androgen insensitivity syndrome (CAIS) is the largest entity within the 46,XY spectrum of disorders of sex development.[3] According to recent estimates from Denmark, the prevalence of androgen insensitivity syndrome with predominantly female phenotype is 4·1:100 000 girls.[4] X-linked recessive mutations within the *AR* gene cause resistance towards androgens in all tissues that express the androgen receptor.[5,6]

Individuals with CAIS are of special interest for clinical management strategies because of the major implications on sex and gender due to a single gene mutation in *AR*. In individuals with 46,XY karyotype who have CAIS the external phenotype is unremarkably female, including breast development at the time of puberty and lack of any androgenisation such as sexual hair growth. A very distinct endocrine profile is present after puberty, with testosterone concentrations in the normal to upper male reference range. The degree of aromatisation of testosterone can be estimated by the ratio of oestradiol to testosterone concentrations. This ratio is within the upper male reference range in women with CAIS and intact gonads. Oestradiol concentrations are normal to slightly increased relative to normal male references. However, despite aromatisation, oestradiol concentrations are below the normal female reference range.[7]

Individuals with CAIS can have high degrees of psychological distress and reduced psychological wellbeing.[8,9] Patients have reported increased wellbeing during testosterone replacement. Additionally, recent study findings indicate a profound dissatisfaction with sexual wellbeing in patients with CAIS,[10] even in patients who do not have cosmetic surgery on the external genitalia, which might affect sexual sensation. Until recently, surgical procedures for patients with CAIS comprised creation of a neo-vagina and gonadectomy because of an assumed risk for malignant transformation.[11] Women with CAIS have considerably reduced sexual satisfaction compared with women with vaginal aplasia due to Mayer-Rokitansky-Küster-Hauser syndrome.[12] These findings lead to the assumption that removal of the gonads and the consequential loss of physiological hormonal balance could have a key role in sexual function and quality of life in these patients.

So far, the treatment of patients with CAIS after gonadectomy has followed the usual concepts of therapy for female hypogonadism. However, previously high androgen concentrations are replaced by oestrogens to mimic an endocrine profile matching the female pattern.

The question of an optimal hormone treatment after gonadectomy seems to be most urgent. To gain reliable insight into the effects of testosterone treatment in CAIS, conventional replacement with oestrogens has to be reviewed and compared with testosterone treatment.

Methods

Study design and participants

This national, multicentre, randomised, double-dummy, double-blind crossover trial was done at three university medical centres and three specialised treatment institutions in Germany (Lübeck, Berlin, Regensburg, Tübingen, Bochum [Dortmund], and Munich), and was coordinated from the centre of referral for disorders of sex development in Lübeck, Germany. Ethical approval was obtained from all participating study sites, with Lübeck's Ethics Committee being the leading institution (reference 11-066). The protocol is available online.

This study was done in accordance with the Declaration of Helsinki for Biomedical Research Involving Human Participants, Good Clinical and Epidemiological Practice, and the German Medicines Act.

Eligible participants (aged 18–54 years) with genetically proven CAIS gave written informed consent. Gonadectomy had to date back more than 1 year. Exclusion criteria are listed in the appendix. Patients

For the **protocol** see http://www.uksh.de/Kinderhormonzentrum_Luebeck/Forschung_CAIS_Studie_Studienprotokoll.html

See Online for appendix

were contacted through study sites. Further recruitment methods encompassed an informative brochure for medical professionals, contact to the German 46,XY-Frauen support group, scientific meetings, and a trial website. Advanced search via diagnostic codes completed the recruitment strategy. Two prolongation periods of the trial for patient recruitment were approved in 2013, and in 2015.

Randomisation and masking

Using the minimisation method described by Pocock and Simon,[13] and stratifying by study centre only, participants were randomly assigned (14:12) by the Institute of Biometry at the Otto-von-Guericke-University in Magdeburg to receive oestradiol 1·5 mg/day and testosterone dummy for 6 months followed by crossover to testosterone 50 mg/day and oestradiol dummy for 6 months (sequence A) or to receive testosterone 50 mg/day and oestradiol dummy for 6 months followed by crossover to oestradiol 1·5 mg/day and testosterone dummy for 6 months (sequence B; figure 1). The crossover of active component after 6 months was done in a double-blinded manner. A double-dummy design was chosen to avoid any identification of study medication. Study drugs and their dummies were controlled and labelled by the study pharmacy (Einhorn Apotheke, Hamburg, Germany) after randomisation, and study participants and investigators were masked to treatment allocation. A sealed opaque envelope with the randomisation result was kept in a trial master file at the coordinating centre in Lübeck.

Procedures

Diagnosis of CAIS was confirmed by molecular genetic analysis of the *AR* gene in the paediatric endocrine laboratory at the University of Lübeck. The timeline for the clinical visits for clinical assessments and drug crossover during the run-in and treatment phases is shown in figure 1. Baseline variables were measured at visit 1 and included detailed information about medical and psychosocial history. We used a questionnaire designed for adult patients with disorders of sex development that was originally applied in a large-scale German clinical evaluation study by the German Network of Disorders of Sex Development/Intersexuality (variables are listed in the appendix).[14] Efficacy and safety data were collected at all study visits. Clinical and laboratory measurements are listed in the appendix. Each study visit was completed with a physical examination. A follow-up visit was scheduled for 3 months after completion (close-out).

All participants received standard oestradiol 1·5 mg/day during a 2 month run-in phase. Assuming that some patients with CAIS were not taking continuous oestrogen replacement, this run-in phase was introduced to accomplish a homogeneous hormonal milieu.

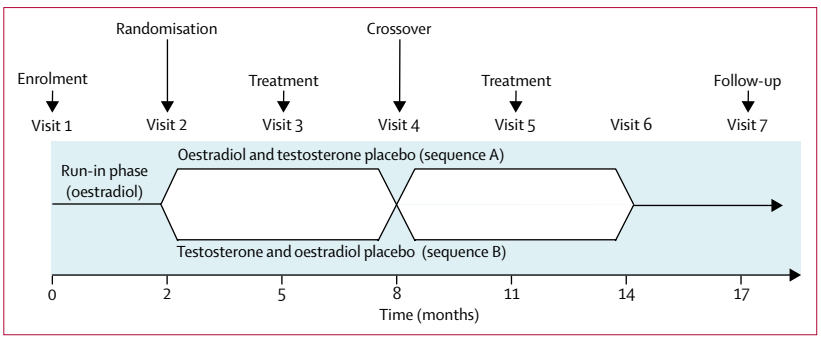

Figure 1: **Study design**

After 12 months of treatment with study drug, patients could choose their preferred type of hormone-replacement therapy to continue until the close-out visit 3 months later.

Testosterone 50 mg (Testogel; BESINS Healthcare SA, Brussels, Belgium) and oestradiol 1·5 mg (Gynokadin; Dr Kade Pharmaceuticals, Berlin, Germany) were provided as gel preparations for transdermal application in daily doses. The dummy for each study drug was provided by BESINS and Dr Kade Pharmaceuticals.

Mental health-related quality of life (MHRQoL) and health-related quality of life (HRQoL) were measured with the standardised German version of the SF-36.[15] The SF-36 is a generic, multipurpose, and validated short-form health survey with 36 questions and eight scaled scores that yields a profile of functional health and wellbeing (physical functioning, role-physical, role-emotional, bodily pain, vitality, mental health, social functioning, and general health) and psychometrical physical and mental health summary measures (appendix). High scores are indicative of improved health status, with a score of 50 being the mean for the general population. The SF-36 was used in the German clinical evaluation study by the German network of disorders of sex development/intersexuality.[14]

Psychological wellbeing was measured with the standardised German version of the Brief Symptom Inventory (BSI). The BSI is a short screening instrument for self-reported and clinically relevant psychological symptoms and includes nine symptom dimensions (somatisation, obsession-compulsion, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychoticism; appendix). The BSI includes three global indices of distress: the Global Severity Index (GSI; the total score of psychological distress); the Positive Symptom Distress Index (PSDI; indicating the intensity of the distress), and the Positive Symptom Total (PST; indicating the total number of symptoms). High scores are indicative of reduced psychological wellbeing. Data were transformed into standardised T values using the gender-matched standard values for women provided in the manual,[16] which allows a direct comparison to be made with the

non-clinical reference samples. T values of 63 or more (cutoff) in the overall score or in at least two subscales are indicative of considerable psychological distress (clinical cases). The BSI was used by Schützmann and colleagues[8] in 2009, and by Krupp and colleagues[17] in 2012 to measure self-reported psychological distress in the context of disorders of sex development.

To assess sexual functioning, including six domains of desire, arousal, lubrication, orgasm, satisfaction, and pain, we used the German version of the Female Sexual Function Index (FSFI) designed for women in clinical trials (appendix). The FSFI is a brief validated questionnaire[18] sensitive to alterations in sexual functioning. Total scores range from 2·0 to 36·0, with high values indicative of good sexual functioning. Classification of total sores range from poor (<23), satisfactory or good (24–29), to very good (>30) sexual functioning.[19] Wiegel and colleagues[20] defined a cutoff score of 26·55; any value below this cutoff is indicative of a risk of sexual dysfunction. In 2014, Fliegner and colleagues[12] used the FSFI to investigate women with CAIS, and Rall and colleagues[21] used the FSFI to study patients with vaginal agenesis.

Laboratory investigations included molecular genetic analysis of the *AR* gene, endocrine profiles in serum,[22] targeted metabolome analysis of 36 urinary steroids,[23] baseline laboratory findings, and safety parameters. Further details are provided in the appendix.

A safety desk managed the reporting of adverse events and serious adverse events and issued annual safety reports to the authorities.

Outcomes

The primary efficacy endpoint was the mental health summary score of the standardised German version of the SF-36 health survey.[15] Secondary endpoints were psychological wellbeing, as measured with the standardised German version of the BSI,[16] sexual functioning, as measured with the Female Sexual Function Index. Assessment of somatic effects included evaluation of possible virilisation, using the Ferriman-Gallwey Score for hirsutism as well as effects on metabolic blood values (appendix).

Statistical analysis

The sample size calculation was based on the primary endpoint, on data provided by the developers of the SF-36, and on results from the German clinical evaluation study by the German network of disorders of sex development/intersexuality,[14] in which a subgroup of nine patients matching the inclusion and exclusion criteria of the study described here had a mean score of 40·7 (SD 9·5) on the psychological sum scale of SF-36. According to the providers of SF-36, the normal female population has a mean score of 50·7 (SD 8·4), yielding a mean deficit of about 10 in the patient group. Assuming a difference in score of 5 (half of the expected deficit *vs* reference value) between both treatments as relevant, and further assuming an SD of 8·4 and a correlation coefficient of 0·5 (or more) between both treatment phases, about 25 patients were necessary to achieve 80% power (simplified calculation as two-sided *t* test for paired samples; α=5%). Anticipating a 15% drop-out, we planned to recruit 30 patients.

A p value less than 0·05 was used to indicate statistical significance. We chose a 2×2 crossover design to use the few patients available most effectively and to compensate for the remaining heterogeneity in this patient group. The patients were characterised by their demographic data and by baseline values of all primary and secondary endpoints. Means and SEs of patient-reported outcome variables, hormone concentrations, and other laboratory parameters were compared between the two treatment sequences and between the two treatments (with treatment periods correspondently matched).

To compensate for unbalanced dropout between the treatment groups, differences between the two treatments for the primary and secondary endpoints were tested in a

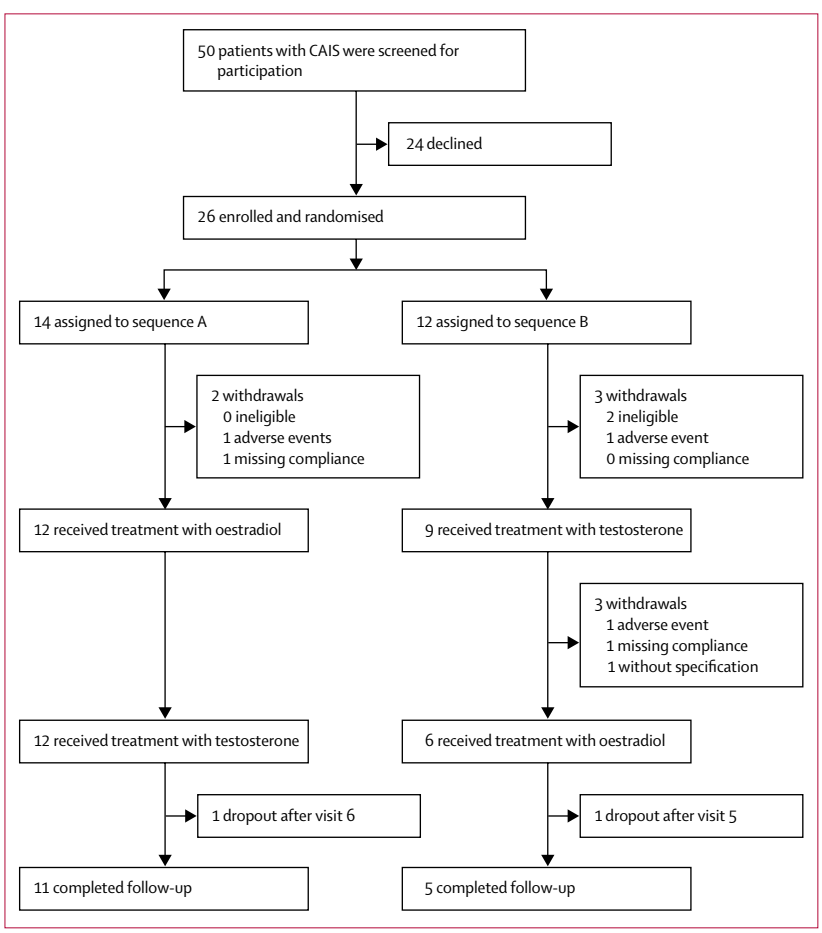

Figure 2: **Trial profile**
Sequence A involved initial treatment with oestradiol for 6 months and cross-over to testosterone for 6 months. Sequence B involved initial treatment with testosterone for 6 months and cross-over to oestradiol for 6 months. CAIS=complete androgen insensitivity syndrome.

mixed linear model analysis for crossover designs, with fixed effects for treatment (oestradiol *vs* testosterone), period (first *vs* second treatment phase), and sequence (oestradiol to testosterone *vs* testosterone to oestradiol) and with a random patient effect (in addition to an ANOVA model for a 2×2 crossover design). This approach enables analyses in unbalanced or incomplete designs, including test decisions and estimation of CIs for the treatment effect. The analysis was based on data at the end of the two treatment phases (visits 4 and 6, respectively), as originally planned, and on data from previous visits (visits 3 and 5, respectively).

The primary analysis included all patients who were available at least until visit 5, even if protocol violations occurred (modified intention-to-treat). Secondary analyses included the per-protocol population only. Serious adverse events and adverse events were recorded for all patients who received at least oestradiol in the run-in phase.

The secondary outcome variables were analysed analogously, whereas oestradiol and testosterone concentrations (and other laboratory parameters, if necessary) were log-transformed to approximate a Gaussian distribution. In the further exploratory secondary analyses, the Wilcoxon paired difference test was used to compare data from oestradiol or testosterone treatments (independent of sequence or period) against the respective baseline values and to compare data from oestradiol or testosterone treatments with reference values in healthy controls and with cutoff values, if available.

We used SAS Power and Sample Size 3·1 for sample size calculations, and IBM SPSS Statistics version 24 for statistical analyses.

An independent data monitoring and safety committee oversaw the study. This trial is registered with the German Clinical Trials Register, number DRKS00003136, and with the European Clinical Trials Database, number 2010-021790-37.

Role of the funding source

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Results

Starting May 24, 2011, we invited 50 patients to participate in the study. 24 patients declined the invitation. The first patient was enrolled on Nov 7, 2011, and recruitment was stopped at about 90% of the planned sample size because of lack of recruitment progress. 14 patients were randomly assigned to receive oestradiol with crossover to testosterone (sequence A), and 12 patients were randomly assigned to receive testosterone with crossover to oestradiol (sequence B; figure 2). Ten patients left the study before completion of the whole trial. One patient did not attend visit 6 and one the follow-up visit (visit 7), but treatment with study medication was complete and data for both could be included. Two additional patients were not compliant. The last patient left the study on Jan 23, 2016. 12 patients in sequence A and six patients in sequence B were included in the modified intention-to-treat population. 12 patients in sequence A and five patients in sequence B finished the essential parts of the study without protocol violations (the per-protocol population) and were included in the secondary analyses.

Baseline characteristics of the modified intention-to-treat population are listed in table 1. Results of the molecular genetic analysis revealed a disorder of sex development other than CAIS in two patients. Genetic investigation of the *AR* gene was completed for all participants in the run-in-phase. All patients who received study drug had mutations within the corresponding *AR* gene associated with CAIS. Thus, the diagnosis of CAIS was proven. Patients with clinical signs of residual activity were not enrolled. Genetic findings are shown in the appendix.

26 patients were included in the safety analysis (table 2). Three patients dropped out because of adverse events. Fibrous mastopathy in one patient receiving oestradiol during the run-in phase was categorised as a serious adverse event. The other two patients dropped out (one during the run-in phase and the other during testosterone treatment) because of episodes of depression and hot flush symptoms. Adverse events were grouped by intensity (grade 1 for mild, grade 2 for moderate, and grade 3 for intense). 23 grade 1 adverse events occurred while patients were receiving oestradiol and 34 grade 1

	Sequence A (n=14)	Sequence B (n=12)
Age (years)	36 (23–53)	29 (19–41)
Age at gonadectomy (years)	20 (14–50)	14 (1–23)
BMI (kg/m²)	25·3 (20·1–34·1)	24·8 (19·6–33·1)
Education		
Junior high school	1 (7%)	1 (8%)
University-entrance diploma	8 (57%)	4 (33%)
College of higher education	1 (7%)	2 (17%)
University degree	4 (29%)	5 (42%)
Partnership		
Heterosexual	5 (36%)	5 (42%)
Homosexual	0	0
None	9 (64%)	7 (58%)
Hormone replacement in the preceding month		
Oestrogens	9 (64%)	11 (92%)
Testosterone	0	0
None	4 (29%)	1 (8%)
Unknown	1 (7%)	0

Data are n (%) or mean (range).

Table 1: **Baseline characteristics of the modified intention-to-treat population**

	Oestradiol		Testosterone			Run-in phase		Follow-up	
	Grade 1	Grade 2	Grade 1	Grade 2	Grade 3	Grade 1	Grade 2	Grade 1	Grade 2
Adverse events	23 (6)	5 (2)	34 (9)	3 (3)	1 (1)	17 (6)	4 (2)	9 (6)	3 (3)
Serious adverse events*	..	..	..	..	..	1	..	..	..
Fatigue	2 (1)	1	6 (2)	0	0	4 (2)	0	0	0
Infection of upper airways	6 (5)	2 (2)	5 (4)	0	0	1	0	2 (1)	1
Weight gain	1	0	0	1	0	0	0	2 (1)	0
Mastodynia	4 (2)	1	1	0	0	5 (3)	0	1	0
Headache	7 (3)	0	15 (4)	1	1	4 (1)	3 (1)	1	0
Epistaxis	..	..	..	..	..	..	..	0	1
Poor concentration	..	..	..	..	..	1	0	1	0
Depression	..	..	1	1	0	..	..	..	..
Increased sexual desire	..	..	2 (1)	0	0	..	..	..	..
Hypercholesterolaemia	1	0	1	0	0	..	..	1	0
Elevated GPT (alamine transaminase)	1	0	..	..	..	..	..	1	0
Hot flush	..	..	1	0	..	1	0	..	..
Dejectedness	..	..	2 (2)	0	0	..	..	..	..
Fracture digitus type IV	..	..	..	..	..	..	..	0	1
Migraine	1	1	..	..	..	..	..	..	..
Bladder infection	..	..	..	..	..	0	1	..	..

Numbers in parentheses are numbers of affected patients in the intention-to-treat population. Grade 1 is slight, with no impact on daily activities. Grade 2 is moderate, with daily activities impaired. Grade 3 is intense, not allowing daily activities. *Fibrous mastopathy leading to study discontinuation in one patient.

Table 2: **Adverse events stratified by treatment group and grade**

	Mean	95% CI lower bound	95% CI upper bound	p value
Mental health summary score				
Descriptive data				
Baseline	42·14	35·79	48·48	..
Baseline (*Z* transformation)	–0·75	–1·47	–0·04	..
After run-in	44·72	39·35	50·09	..
During oestradiol treatment	44·54	38·73	50·35	..
During testosterone treatment	45·31	39·89	50·74	..
Linear mixed model				
During oestradiol treatment	46·80	38·73	50·35	..
During testosterone treatment	46·36	39·89	50·74	..
Difference	0·45	–2·98	3·88	0·794
Physical summary score				
Descriptive data				
Baseline	54·71	51·39	58·02	..
Baseline (*Z* transformation)	0·55	0·1	0·99	..
After run-in	54·42	51·04	57·82	..
During oestradiol treatment	53·79	49·42	58·1	..
During testosterone treatment	55·12	52·96	57·27	..
Linear mixed model				
During oestradiol treatment	55·97	53·82	58·02	..
During testosterone treatment	55·65	53·65	57·66	..
Difference	0·32	–2·08	2·72	0·791

Table 3: **Short-form-36 health survey (SF-36) mental health and physical summary scores**

adverse events were recorded during testosterone treatment (affecting >10% of the intention-to-treat population). Grade 2 adverse events were recorded in two patients in the oestradiol group and in three in the testosterone group. One grade 3 event was reported (severe headache), and this was in the testosterone group. 21 adverse events occurred (17 grade 1 and four grade 2) during the run-in phase and 12 during the follow-up phase after treatment (nine grade 1 and three grade 2). No relevant signs of virilisation were detected in any case during the whole study. The Ferriman-Gallwey-Score did not reach the clinically relevant value of 7.

At baseline (visit 1 and 2), MHRQoL was significantly reduced compared with reference data (z-transformation mean z –0·75, CI –1·47 to –0·04).[24] By contrast, mean values for the physical health summary score were higher than in the reference data (0·55, 0·1 to 0·99), indicating significantly improved physical health compared with the normal population (table 3). No significant difference was found in the effect of oestradiol and testosterone on mental health scores (p=0·794) or physical summary scores (p=0·791) in the linear mixed model.

MHRQoL scores and physical HRQoL scores for patients receiving treatment did not differ significantly from baseline scores (MHRQoL: p=0·065 for patients receiving oestradiol *vs* p=0·207 for patients receiving testosterone; physical HRQoL: p=0·782 *vs* p=0·854).

BSI scores for psychological wellbeing did not differ significantly between the two treatment groups (p=0·638 for GSI; p=0·570 for PST; p=0·378 for PSDI; appendix). At baseline (visit 1) all t-scores for indices were higher than 50, which is the reference value for healthy controls (mean 58·28 for GSI; mean 59·06 for PST; mean 56·67

for PSDI; appendix). Eight of 17 participants scored higher than 63 in the GSI or two subscales. These findings show that patients had psychological distress at the beginning of the trial. Relative to baseline values at visit 1, all global indices showed significant improvement in psychological wellbeing and mental health in patients receiving oestradiol treatment. Patients receiving testosterone had reduced scores, but the difference relative to baseline values was only significant for the PSDI (appendix).

We found a relevant risk for sexual dysfunction for most of the study participants at baseline. Mean values during visit 1 (15·78) were markedly lower than the 26·55 cutoff value (appendix). Only two of 18 women scored more than 26·55, which is consistent with the low prevalence of satisfactory sexual functioning at baseline. Relative to baseline values at visit 1, patients in the testosterone group had significant improvement for the FSFI total score and for the desire, arousal, lubrication, and orgasm domains of sexual functioning (appendix). Oestradiol treatment had no significant efficts on sexual functioning (appendix). The mean FSFI total scores for sexual functioning and all subscales except for the satisfaction domain were higher during testosterone treatment than during oestradiol treatment (table 4). For the desire domain, the difference between testosterone and oestradiol was significant (p=0·018), but the difference between the effects of oestradiol and testosterone on the total FSFI score was not significant (p=0·141).

Seven participants scored more than zero in all domains and were identified as sexually active. Three women scoring zero in the pain or lubrication domains (but in no other domain) were included in the group of sexually active participants. Initial scores at visit 1 (23·8) were markedly higher than in the total study population but nevertheless below the cutoff-value of 26·55 (appendix). The descriptive analysis within this group did not reveal any significant changes in sexual activity in response to testosterone or oestradiol treatment.

The FSFI total scores and scores for all domains were significantly lower for patients with CAIS at baseline and after the run-in phase than in healthy controls described by Rosen and colleagues.[18] The FSFI score remained less than the cutoff value during both treatments and was significantly lower than reference values (appendix).

Hormone concentrations were measured before treatment and during oestradiol and testosterone treatment (figure 3; figure 4; appendix). The concentrations of luteinising hormone and follicle-stimulating hormone were high before treatment (in the reference range of postmenopausal women and above the reference range of women with CAIS and intact gonads[7,11]) and remained high after treatment. No significant difference was found in gonadotrophin concentrations between treatment sequences. This correlates with the findings of similar effects of oestradiol treatment in a case of CAIS reported by Taes and colleagues.[25]

	Mean score	95% CI lower bound	95% CI upper bound	p value
FSFI total score	..	..	..	0·141
Oestradiol	18·36	13·97	22·75	..
Testosterone	20·95	16·75	25·15	..
Difference	–2·59	–6·07	0·89	..
Desire	..	..	..	0·018
Oestradiol	3·22	2·70	3·74	..
Testosterone	3·74	3·25	4·24	..
Difference	–0·52	–0·95	–0·09	..
Arousal	..	..	..	0·071
Oestradiol	3·21	2·37	4·06	..
Testosterone	3·86	3·06	4·67	..
Difference	–0·65	–1·36	0·0	..
Lubrication	..	..	..	0·225
Oestradiol	3·74	2·75	4·73	..
Testosterone	4·29	3·35	5·23	..
Difference	–0·55	–1·46	0·35	..
Orgasm	..	..	..	0·052
Oestradiol	2·86	2·04	3·68	..
Testosterone	3·60	2·82	4·37	..
Difference	–0·73	–1·47	0·01	..
Satisfaction	..	..	..	0·396
Oestradiol	3·39	2·45	4·34	..
Testosterone	3·15	2·23	4·06	..
Difference	0·25	–0·33	0·83	..
Pain	..	..	..	0·250
Oestradiol	2·11	0·95	3·28	..
Testosterone	2·60	1·46	3·74	..
Difference	–0·487	–1·34	0·35	..

Data are derived from linear mixed model analysis. FSFI=Female Sexual Function Index.

Table 4: **Effect of treatment on sexual functioning**

After run-in treatment with oestrogens, oestradiol concentrations were within the lower reference range for women. The median concentration did not change during oestradiol treatment. The median oestradiol concentration during testosterone treatment was less than the adult male reference range.

Testosterone concentrations during the run-in phase and during oestradiol treatment were in the lower range for adult women. The median testosterone concentration during testosterone treatment was within the range for young adult men.

After the run-in phase with oestradiol treatment the median concentrations of oestrogens in urine (sum of oestrone, oestradiol, and oestriol) were in the upper reference range for women (appendix). The median concentrations of urinary oestrogens were similar during oestradiol treatment and during the run-in phase, showing good compliance. During testosterone treatment the median concentration of all urinary oestrogens was below the normal range. Testosterone concentrations in urine after the run-in phase and during oestradiol treatment were in the lower range for adult

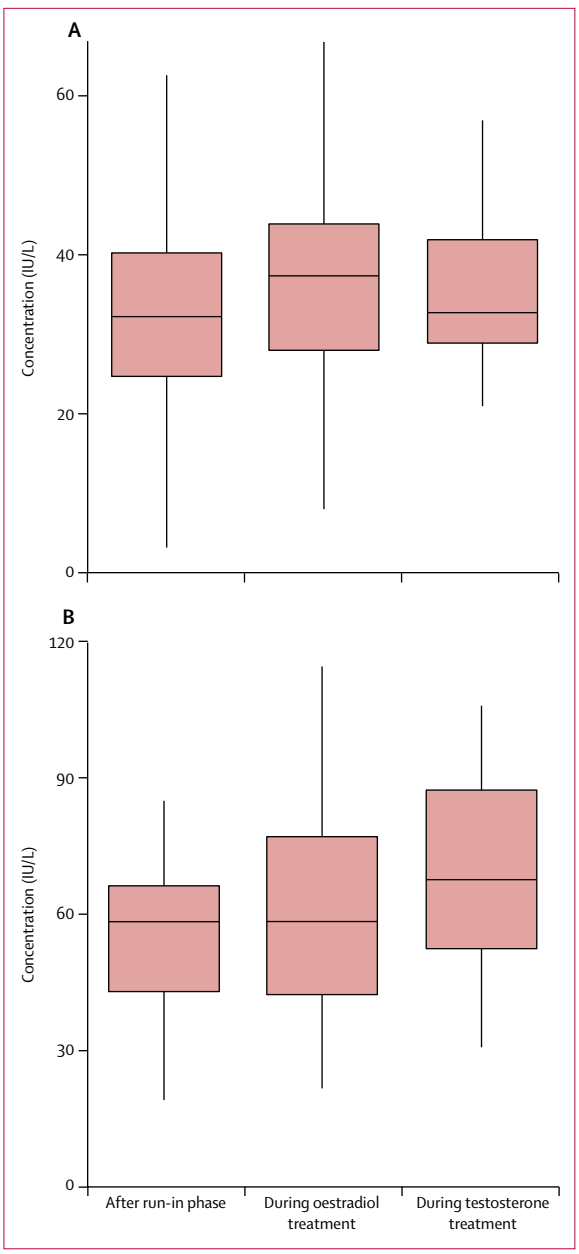

Figure 3: **Gonadotropin concentrations after the run-in phase and during treatment phases**
(A) Luteinising hormone. (B) Follicle-stimulating hormone. IU=international units.

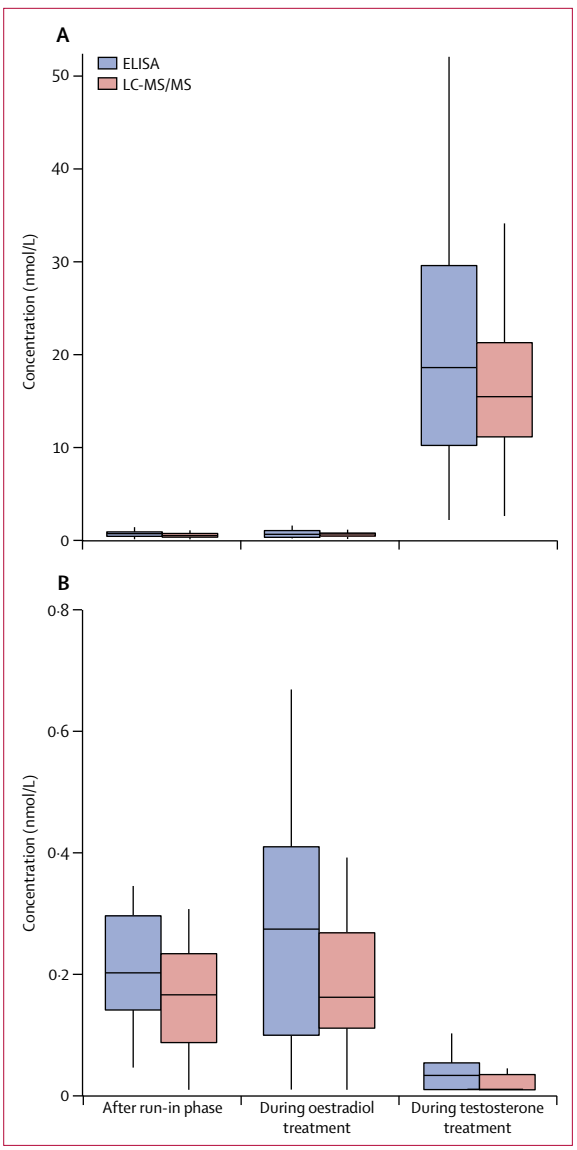

Figure 4: **Sex hormone concentrations after the run-in phase and during treatment phases**
(A) Testosterone. (B) Oestradiol. LC-MS/MS=liquid chromatography with mass spectrometry.

women. During testosterone treatment, the median concentrations of urinary testosterone were in the upper or even above the normal range for young adult men.

Discussion

Clinical trials in rare diseases pose a great challenge, especially in medical treatment. Small numbers of patients and heterogeneous groups of diagnoses impede the generation of statistically significant evidence of efficacy. In this study we gained prospective data of a relatively homogeneous group that was well defined by genetic diagnosis.

This trial shows the difficulty of recruiting patient cohorts that are large enough to reach high statistical power in rare diseases. Considering the rareness of CAIS, the study cohort was adequate in size and provides valuable information. With slightly reduced sample sizes and enlarged drop-out rates, the power of efficacy analyses was somewhat reduced. Because of the small sample sizes, the reported exploratory pairwise comparisons of visits are unadjusted for multiple testing. Nevertheless, for the first time, reliable data from a prospective randomised study in the field of disorders of sex development were obtained and trends were visible.

Sexual satisfaction was markedly reduced in the study cohort. Higher sexual dysfunction in patients with CAIS seems plausible when compared with healthy controls.[18] Hormone replacement has an effect on sexual functioning in CAIS, but other factors contribute to poor outcomes. The approach to disclosure, social support, and contact to the community for disorders of sex development affects the coping process.

Testosterone treatment seems to have a stronger effect on sexual functioning in patients with CAIS than oestrogen treatment does, but the difference was significant only for the desire domain of sexual functioning. Oestradiol treatment did not have any effect on sexual functioning. Although applying an advanced statistical model to rule out confounding variables, testosterone treatment could not be proven to be superior in all aspects of sexual functioning. Yet scores in the FSFI were higher during testosterone treatment than during oestradiol treatment.

Central regulation of sexual functioning is complex and involves various distinct brain sites. In both sexes, oestradiol and testosterone affect responsiveness to sexual stimuli and sexual behaviour.[26] However, assuming that the *AR* gene is non-functioning in patients with CAIS, how might testosterone improve sexual functioning? One possibility, as shown in rodent models, is that site-specific conversion of testosterone in the brain by steroidogenic enzymes such as aromatase or 5-α-reductase into oestradiol and 3α-androstanediol, respectively, are involved in activating sexual behavior,[27] hereby allowing for preserved central effects of testosterone in women with CAIS, despite the absence of a functioning *AR* gene.

Desire is a relevant domain of sexual functioning in patients with CAIS, especially when sexual activity is practiced without a partner. This was the case in some of the study participants. Lubrication, pain, and satisfaction with emotional aspects of partnership are rather unsuitable scales to analyse in these circumstances. Unfortunately, no questionnaires are specific for disorders of sex development. The FSFI is a validated and widely used tool but provides an informative basis mostly for patients in a relationship.

Baseline scores in sexually active patients (appendix) were higher than in the whole cohort. Yet mean values were below the cutoff value, indicating a risk of sexual dysfunction for many of these participants. This finding underlines the importance of endocrine assessment and psychosexual counselling for patients with CAIS to initiate hormone treatment and improve coping strategies.

The SF-36 also is a common patient-reported outcome instrument in clinical studies. The mental health summary score encompasses aspects of vitality, social functioning, role limitations, and psychological wellbeing. MHRQoL was significantly reduced in the study cohort compared with healthy controls, whereas physical HRQoL was significantly increased. Similar results were found in the German clinical evaluation study[28] and in a national registry-based study from Denmark.[29] Additionally, participants had higher psychological distress, as measured with the BSI. 47% of the participants had to be classified as clinical cases before beginning the trial, corroborated by results reported by Bennecke and colleagues in 2017.[28]

Some participants were not taking any medication before the study (table 1). In some cases, uncertainty about the best hormone replacement led to refusal to take any medication. The lack of lifelong concepts of care and of experienced caregivers contributes to poor compliance with hormone replacement and difficulty in coping with the diagnosis. Positive effects on mental health, psychological wellbeing, and sexual functioning, to varying extents, could be detected when continuous hormone treatment was implemented with run-in medication. Reliable hormone substitution for patients with CAIS seems important for an improved quality of life. Continuous therapy could also lead to positive long-term effects, which were not measured in this study, such as preservation of bone mineral density. Bone mineral density impairment after removal of the gonads has been reported in patients with CAIS, but the underlying mechanisms are not understood.[30]

Taking into account that testosterone was well tolerated and just as safe as oestrogen treatment, we conclude that testosterone can be an alternative hormone substitution for patients with CAIS, especially when sexual satisfaction is reduced. Complementary treatment with oestrogens (in case of low aromatisation) for potential beneficial effects on bone metabolism should be discussed. Long-term follow-up will be crucial to assess effects on physical and psychological wellbeing.

Our findings also emphasise the need for life-long concepts of care for patients with CAIS to ensure they are given the best tailored hormone replacement and to reduce relevant psychological distress, improve sexual functioning, and improve mental health-related quality of life.

Contributors

WB, LM and OH planned the study and coordinated the study centers. WB, LM, KR, BK, AR-U, MKA, and OH were in charge of individual study centers, recruited patients, and did the study. RW did the molecular genetic studies. AK, P-MH, MFH, and SAW did the metabolic studies. AL and SK were in charge of statistical analysis. All authors contributed to the writing of the report.

Declaration of interests

We declare no competing interests.

Acknowledgments

This study was funded by the German Federal Ministry of Education and Research (BMBF), project no 01KG1003.

References

1 Hiort O, Birnbaum W, Marshall L, et al. Management of disorders of sex development. *Nat Rev Endocrinol* 2014; **10:** 520–29.

2 Diamond M, Sigmundson HK. Management of intersexuality. Guidelines for dealing with persons with ambiguous genitalia. *Arch Pediatr Adolesc Med* 1997; **151:** 1046–50.

3 Ahmed SF, Rodie M, Jiang J, Sinnott RO. The European disorder of sex development registry: a virtual research environment. *Sex Dev* 2010; **4:** 192–98.

4 Berglund A, Johannsen TH, Stochholm K, et al. Incidence, prevalence, diagnostic delay, and clinical presentation of female 46,XY disorders of sex development. *J Clin Endocrinol Metab* 2016; **101:** 4532–40.

5 Quigley CA, De Bellis A, Marschke KB, el-Awady MK, Wilson EM, French FS. Androgen receptor defects: historical, clinical, and molecular perspectives. *Endocr Rev* 1995; **16:** 271–321.

6 Werner R, Holterhus PM. Androgen action. *Endocr Dev* 2014; **27:** 28–40.

7 Doehnert U, Bertelloni S, Werner R, Dati E, Hiort O. Characteristic features of reproductive hormone profiles in late adolescent and adult females with complete androgen insensitivity syndrome. *Sex Dev* 2015; **9:** 69–74.

8 Schützmann K, Brinkmann L, Schacht M, Richter-Appelt H. Psychological distress, self-harming behavior, and suicidal tendencies in adults with disorders of sex development. *Arch Sex Behav* 2009; **38:** 16–33.

9 D'Alberton F, Assante MT, Foresti M, et al. Quality of life and psychological adjustment of women living with 46,XY differences of sex development. *J Sex Med* 2015; **12:** 1440–49.

10 Köhler B, Kleinemeier E, Lux A, et al. Satisfaction with genital surgery and sexual life of adults with XY disorders of sex development: results from the German clinical evaluation study. *J Clin Endocrinol Metab* 2012; **97:** 577–88.

11 Döhnert U, Wünsch L, Hiort O. Gonadectomy in complete androgen insensitivity syndrome: why and when? *Sex Dev* 2017; **11:** 171–74.

12 Fliegner M, Krupp K, Brunner F, et al. Sexual life and sexual wellness in individuals with complete androgen insensitivity syndrome (CAIS) and Mayer-Rokitansky-Kuster-Hauser Syndrome (MRKHS). *J Sex Med* 2014; **11:** 729–42.

13 Pocock SJ, Simon R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial. *Biometrics* 1975; **31:** 103–15.

14 Lux A, Kropf S, Kleinemeier E, Jürgensen M, Thyen U, Group DSDNW. Clinical evaluation study of the German network of disorders of sex development (DSD)/intersexuality: study design, description of the study population, and data quality. *BMC Public Health* 2009; **9:** 110.

15 Bullinger M. German translation and psychometric testing of the SF-36 health survey: preliminary results from the IQOLA Project. International quality of life assessment. *Soc Sci Med* 1995; **41:** 1359–66.

16 Franke G. BSI. Brief Symptom Inventory von L R Derogartis—Deutsche Version. Göttingen: Beltz Test, 2000.

17 Krupp K, Fliegner M, Brunner F, Brucker S, Rall K, Richter-Appelt H. How do individuals with complete androgen insensitivity syndrome, Mayer-Rokitansky-Kuster-Hauser syndrome or polycystic ovary syndrome experience contact to other affected persons? *Geburtshilfe Frauenheilkd* 2012; **72:** 1009–17.

18 Rosen R, Brown C, Heiman J, et al. The Female Sexual Function Index (FSFI): a multidimensional self-report instrument for the assessment of female sexual function. *J Sex Marital Ther* 2000; **26:** 191–208.

19 Allen LM, Lucco KL, Brown CM, Spitzer RF, Kives S. Psychosexual and functional outcomes after creation of a neovagina with laparoscopic Davydov in patients with vaginal agenesis. *Fertil Steril* 2010; **94**: 2272–76.

20 Wiegel M, Meston C, Rosen R. The female sexual function index (FSFI): cross-validation and development of clinical cutoff scores. *J Sex Marital Ther* 2005; **31:** 1–20.

21 Rall K, Schickner MC, Barresi G, et al. Laparoscopically assisted neovaginoplasty in vaginal agenesis: a long-term outcome study in 240 patients. *J Pediatr Adolesc Gynecol* 2014; **27:** 379–85.

22 Kulle AE, Welzel M, Holterhus PM, Riepe FG. Principles and clinical applications of liquid chromatography—tandem mass spectrometry for the determination of adrenal and gonadal steroid hormones. *J Endocrinol Invest* 2011; **34:** 702–08.

23 Wudy SA, Schuler G, Sanchez-Guijo A, Hartmann MF. The art of measuring steroids: principles and practice of current hormonal steroid analysis. *J Steroid Biochem Mol Biol* 2018; **179:** 88–103.

24 Bullinger M, Kirchberger I. Der SF-36 Fragebogen zum Gesundheitszustand. Handbuch für die deutschprachige Fragebogen-Version. Göttingen: Hogrefe, 1998.

25 Taes Y, Lapauw B, Vandewalle S, et al. Estrogen-specific action on bone geometry and volumetric bone density: longitudinal observations in an adult with complete androgen insensitivity. *Bone* 2009; **45:** 392–97.

26 Pfaus JG. Pathways of sexual desire. *J Sex Med* 2009; **6:** 1506–33.

27 Christensen LW, Clemens LG. Intrahypothalamic implants of testosterone or estradiol and resumption of masculine sexual behavior in long-term castrated male rats. *Endocrinology* 1974; **95:** 984–90.

28 Bennecke E, Thyen U, Grüters A, Lux A, Köhler B. Health-related quality of life and psychological well-being in adults with differences/disorders of sex development. *Clin Endocrinol (Oxf)* 2017; **86:** 634–43.

29 Berglund A, Johannsen TH, Stochholm K, et al. Morbidity, mortality, and socioeconomics in females with 46,XY disorders of sex development: a nationwide study. *J Clin Endocrinol Metab* 2018; **103:** 1418–28.

30 Bertelloni S, Baroncelli GI, Mora S. Bone health in disorders of sex differentiation. *Sex Dev* 2010; **4:** 270–84.