Behav Genet (2015) 45:35–50
DOI 10.1007/s10519-014-9670-x

ORIGINAL RESEARCH

Contrast Effects and Sex Influence Maternal and Self-Report Dimensional Measures of Attention-Deficit Hyperactivity Disorder

J. L. Ebejer · S. E. Medland · J. van der Werf · M. J Wright · A. K. Henders · N. A. Gillespie · I. B. Hickie · N. G. Martin · D. L. Duffy

Received: 17 January 2014 / Accepted: 30 July 2014 / Published online: 24 August 2014

Abstract The heritability of attention-deficit/hyperactivity disorder (ADHD) is higher for children than adults. This may be due to increasing importance of environment in symptom variation, measurement inaccuracy when two raters report behavior of a twin-pair, a contrast effect resulting from parental comparison of siblings and/or dimensionality of measures. We examine rater contrast and sex effects in ADHD subtypes using a dimensional scale and compare the aetiology of self, versus maternal-report. Data were collected using the Strengths and Weaknesses of ADHD and Normal Behaviour Scale (SWAN): maternal-report for 3,223 twins and siblings (mean age 21.2, $SD = 6.3$) and self-report for 1,617 twins and siblings (mean age 25.5, $SD = 3.2$). Contrast effects and magnitude of genetic and environmental contributions to variance of ADHD phenotypes (inattention, hyperactivity-impulsivity, combined behaviours) were examined using structural equation modeling. Contrast effects were evident for maternal-report hyperactivity-impulsivity ($b = -0.04$) and self-report inattention (-0.09) and combined ADHD (-0.08). Dominant genetic effects were shared by raters for inattention, hyperactivity-impulsivity and combined ADHD. Broad-sense heritability was equal across sex for maternal-report inattention, hyperactivity-impulsivity and combined ADHD (0.72, 0.83, 0.80). Heritability for corresponding subtypes in self-reported data were best represented by sex (0.46, 0.30, 0.39 for males; 0.69, 0.41, 0.65 for females). Heritability difference between maternal and self-report ADHD was due to greater variance of male specific environment in self-report data. Self-reported ADHD differed across sex by magnitude of specific environment and genetic effects.

Keywords Twin studies · Contrast effects · Rater effects · ADHD · Inattention · Hyperactivity-impulsivity · Sex-limitation

Edited by Yoon-Mi Hur.

J. L. Ebejer (✉) · J. van der Werf
Environmental and Rural Sciences, University of New England, Armidale, NSW, Australia
e-mail: ebejer.j@gmail.com

J. L. Ebejer · A. K. Henders · N. A. Gillespie · N. G. Martin · D. L. Duffy
Genetic Epidemiology, QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia

S. E. Medland
Quantitative Genetics, QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia

M. J Wright · N. A. Gillespie
Neuroimaging Genetics, QIMR Berghofer Medical Research Institute, Brisbane, QLD, Australia

M. J Wright · N. A. Gillespie
Virginia Institute for Psychiatric and Behavioral Genetics, Virginia Commonwealth University, Richmond, VA, USA

I. B. Hickie
Brain & Mind Research Institute, University of Sydney, Sydney, NSW, Australia

Introduction

In childhood, behavioural measures of attention-deficit hyperactivity disorder (ADHD) symptoms are collected from parents, teachers or trained interviewers. Adult ADHD data are most often self-reported. The change in rater with time corresponds to a change in heritability estimates of ADHD. A heritability estimate of ~80 % is

consistently found for children (Derks et al. 2008; Knopik et al. 2005; Martin et al. 2002) and adult heritability is estimated at ~45 % (Boomsma et al. 2010; Larsson et al. 2012; Kan et al. 2012). Sex effects are suggested, but not clarified in these studies. The drop in heritability is confounded with the move from parental to self-report of symptoms (Kan et al. 2012) and could be due to the increased influence of environmental factors in symptom moderation for adults. For example adults are more able to select environments in which their symptoms are not problematic. This decline could also be due to rater effects and a change in model parameters when two rather than one person reports on the behaviour of paired twins.

Classical twin studies examining parental report of ADHD, show sex effects. Eaves and colleagues (1997) show fathers' but not mothers' report of childhood ADHD differed for males and females. They also show evidence for a sibling contrast effect (Simonoff et al. 1998; Eaves et al. 1997). Simonoff and colleagues (1998) speculate this is due to parental bias rather than true sibling interaction because the same effect is not evident in teacher report of children's behavior. It may also be that parents' see aspects of their children's behavior not evident during school. A contrast effect results when for example, high levels of ADHD in one twin lead to lower levels of ADHD in a cotwin. The effect reduces the within twin-pair correlation with the decrease in genetic relationship between siblings. Therefore differences between DZ twins are exaggerated in comparison to MZ twins, an effect also indicative of the action of dominant genes or epistasis (Eaves 1976). It is important to clarify the nature of contrast and sex effects to increase measurement accuracy of ADHD symptoms.

Two recent studies have sought to explain the variation in ADHD measurement that occurs with sex and rater. Merwood and colleagues (2013) found a contrast effect for parent ratings of ADHD and a dominant genetic effect in self-report symptoms within an 11–12 year old twin sample. The contrast effect was equal for males and females, but there were scalar sex effects in parent, teacher and self-reported ADHD in this sample. The heritability estimates they found when two raters reported the behaviour of each twin within a pair (teachers = 49 % and self = 48 %), were of similar magnitude. They were also lower than heritability estimates calculated with parental (82 %) and same teacher (60 %) behavioural report.

Chang et al. (2013) examined sex effects and change in genetic effects on attention problems that occur with age. Their sample consisted of 8–20 year old twins with parental and self-reported data collected across time. Their results indicated variation across sex and a dominant genetic effect accounting for variation in attention problems. In contrast to the drop in heritability estimates of ADHD with age, there was no decline in genetic effects influencing attention problems over time (range = 77–82 %) when a latent measure including maternal and self-reported symptoms was used. This study did not test for contrast effects and both parent and self-report data were collected using a severity scale (The Child Behavior Checklist).

The scale used to collect data has been suggested to exert an additional influence on research findings. Pinto and colleagues (2012) found consistent contrast effects and a sex effect influencing ADHD for children aged 7 and 12 when parents' reported ADHD using severity scales (The Strengths and Difficulties Questionnaire and The Revised Rutter Scale). These authors suggest rater contrast effects are due to the limitations of a severity scale. A study by Hay and colleagues (2007) found contrast effects for ADHD when a severity scale was used (Australian Twin Behavioural Rating Scale; Levy et al. 1997), but not when ADHD data were collected using the Strengths and Difficulties of ADHD and Normal behaviour Scale (SWAN).

The SWAN scale (Swanson et al. 2005) measures ADHD-related behaviours along a continuum, ranging from high levels of attention and highly appropriate activity, to the clinically relevant inattention and hyperactivity-impulsivity characteristic of ADHD. The use of this dimensional scale could provide a more accurate description of behavior because respondents are not required to say whether or not the participant has a symptom. The requirement is to rate the way in which a symptom is expressed, this can be in a negative or positive direction. The increased specificity of behavioural description possible with the SWAN could remove a reporting bias imposed by the base category of a severity scale.

This study explores the presence of rater contrast effects in maternal and self-reported ADHD related behaviours using the SWAN. We estimate the similarity and difference in genetic and environmental factors contributing to maternal and self-reported inattentive, hyperactive-impulsive and combined ADHD related behaviours. We also test for variation in the genetic and environmental factors influencing these behaviours across sex in a sample of adolescents and young adults, extending the age range of previous studies.

Methods

Samples

Data were from two ongoing sub-studies of the Brisbane Longitudinal Twin Study (BLTS; Wright & Martin 2004) in which the inattentive and hyperactive-impulsive behaviours comprising ADHD were collected using the SWAN. The first sample came from a study of melanocytic naevi (Zhu et al. 1999) and included 3,236 twins and family members, 51.0 % female with a mean age of 21.2

Table 1 Mean, standard deviation and within twin-pair correlations for sample 1: maternal-report SWAN by data collection method and zygosity group

	Twin pairs	Unpaired twins	1 Sibling *n*	2 Siblings *n*	*N*	Inattention *M (SD)*	Hyperactivity impulsivity *M (SD)*	Combined ADHD *M (SD)*	Pearson's product moment correlations		
									Inattention	Hyperactivity impulsivity	Combined ADHD
Clinic current report N = 504, age range = 10–17											
MZ female	32	0	10	0	74	0.12 (0.77)	0.14 (0.81)	0.15 (0.78)	0.67 (0.43, 0.82)	0.71 (0.48, 0.85)	0.71 (0.48, 0.85)
MZ male	36	0	4	3	82	0.37 (0.98)	0.42 (0.87)	0.42 (0.92)	0.79 (0.63, 0.89)	0.78 (0.61, 0.88)	0.81 (0.66, 0.90)
DZ female	60	1	13	3	140	−0.02 (0.81)	−0.15 (0.84)	−0.09 (0.79)	0.18 (−0.07, 0.41)	0.51 (0.29, 0.67)	0.34 (0.10, 0.55)
DZ male	39	0	13	1	93	0.33 (0.92)	0.36 (0.91)	0.37 (0.92)	0.37 (0.07, 0.62)	0.56 (0.30, 0.74)	0.55 (0.29, 0.74)
DZ female/male	23	0	3	0	49	0.07 (0.85)	−0.14 (0.93)	−0.36 (0.86)	0.19 (−0.24, 0.56)	0.63 (0.29, 0.83)	0.48 (0.08, 0.74)
DZ male/female	29	0	6	1	66	0.21 (1.00)	0.15 (0.94)	0.19 (0.90)	0.16 (−0.10, 0.42)	0.72 (0.52, 0.84)	0.57 (0.32, 0.75)
Siblings	–	–	–	–	–	0.11 (0.91)	−0.90 (0.95)	0.01 (0.93)	0.21 (0.05, 0.35)	0.60 (0.49, 0.69)	0.46 (0.33, 0.58)
Online current report N = 918, age range = 5–20											
MZ female	71	1	25	5	178	−0.18 (0.96)	−0.27 (0.89)	−0.24 (0.92)	0.72 (0.58, 0.81)	0.87 (0.80, 0.92)	0.84 (0.76, 0.90)
MZ male	59	1	26	5	155	0.28 (0.94	0.18 (0.99)	0.24 (0.93)	0.55 (0.34, 0.71)	0.47 (0.23, 0.65)	0.46 (0.21, 0.64)
DZ female	68	0	28	5	174	−0.23 (0.94)	−0.24 (1.01)	−0.25 (0.99)	0.09 (−0.14, 0.33)	0.22 (−0.02, 0.43)	0.17 (−0.06, 0.40)
DZ male	57	0	19	6	145	0.29 (1.31)	0.23 (1.20)	0.28 (1.28)	0.42 (0.17, 0.62)	0.44 (0.19, 0.64)	0.49 (0.24, 0.68)
DZ female/male	54	3	14	4	133	0.00 (1.10)	−0.14 (1.04)	−0.07 (1.08)	0.18 (−0.11, 0.44)	0.37 (011, 0.58)	0.32 (0.04, 0.55)
DZ male/female	52	1	18	5	133	0.06 (1.06)	−0.01 (1.06)	0.03 (1.07)	0.33 (0.06, 0.55)	0.19 (−0.08, 0.44)	0.24 (−0.03, 0.49)
Siblings	–	–	–	–		0.10 (1.01)	−0.01 (0.97)	0.05 (0.99)	0.20 (0.14, 0.27)	0.35 (0.29, 0.41)	0.30 (0.24, 0.36)
Online retrospective report N = 1,477, age range = 21–44											
MZ female	114	3	42	27	327	−0.24 (0.86)	−0.09 (0.92)	−0.17 (0.90)	0.68 (0.57, 0.77)	0.82 (0.75, 0.87)	0.79 (0.71, 0.85)
MZ male	97	2	34	29	288	0.15 (0.93)	0.05 (0.91)	0.11 (0.91)	0.80 (0.71, 0.86)	0.90 (0.85, 0.93)	0.89 (0.84, 0.93)
DZ female	72	2	23	18	205	−0.33 (0.99)	−0.24 (0.97)	−0.30 (0.98)	0.30 (0.08, 0.50)	0.54 (0.35, 0.69)	0.50 (0.30, 0.66)
DZ male	86	1	33	25	256	0.13 (1.01)	0.23 (1.03)	0.20 (1.04)	0.15 (−0.06, 0.36)	0.27 (0.05, 0.46)	0.23 (0.00, 0.43)
DZ female/male	75	1	36	10	207	0.06 (0.99)	0.16 (0.89)	0.12 (0.92)	0.03 (−0.20, 0.26)	0.35 (0.13, 0.54)	0.22 (−0.01, 0.44)
DZ male/female	73	1	21	13	194	0.01 (0.98)	0.07 (1.02)	0.04 (1.00)	0.04 (−0.19, 0.27)	0.19 (−0.05, 0.41)	0.09 (−0.15, 0.33)
Siblings	–	–	–	–	–	−0.13 (1.03)	−0.06 (1.06)	−0.10 (1.05)	0.20 (0.14, 0.27)	0.35 (0.29, 0.41)	0.30 (0.24, 0.36)
Data combined N = 3,141, age range = 5–44											
MZ female	217	4	86	51	626	−0.16 (0.89)	−0.11 (0.91)	−0.15 (0.90)	0.69 (0.62, 0.75)	0.84 (0.79, 0.87)	0.80 (0.75, 0.84)
MZ male	192	3	69	56	568	0.23 (0.95)	0.16 (0.94)	0.21 (0.93)	0.71 (0.64, 0.77)	0.74 (0.67, 0.79)	0.75 (0.68, 0.80)
DZ female	200	3	68	43	557	−0.20 (0.93)	−0.21 (0.94)	−0.22 (0.93)	0.23 (0.10, 0.35)	0.43 (0.32, 0.53)	0.38 (0.26, 0.48)
DZ male	182	1	68	49	531	0.23 (1.10)	0.26 (1.07)	0.26 (1.10)	0.30 (0.16, 0.42)	0.39 (0.26, 0.50)	0.37 (0.24, 0.48)
DZ female/male	152	4	61	28	425	0.04 (1.01)	0.01 (0.96)	0.03 (0.97)	0.13 (0.02, 0.23)	0.33 (0.22, 0.44)	0.24 (0.13, 0.34)
DZ male/female	154	2	49	31	421	0.07 (1.01)	0.06 (1.02)	0.07 (1.00)	0.13 (0.02, 0.25)	0.30 (0.21, 0.39)	0.24 (0.14, 0.34)
Siblings	–	–	–	–	–	−0.05 (1.02)	−0.05 (1.03)	−0.51 (1.03)	0.25 (0.19, 0.31)	0.41 (0.35, 0.47)	0.37 (0.24, 0.41)

SWAN 7-point scale. Higher values indicate the presence of ADHD symptoms. Twin pairs and first two additional siblings and only siblings with twins were included in analyses, excluding $n = 8$, $n = 98$, $n = 218$ and $n = 95$ participants from these descriptive statistics and subsequent analyses for clinic, current, retrospective and full samples respectively

($SD = 6.3$). ADHD data in sample 1 was provided by mothers. The second sample was drawn from a study of mental health in young adults (Gillespie et al. 2012). Sample 2 included 1,617 twins and family members aged on average 25.0 ($SD = 3.7$), and 58.1 % of these participants were female. These participants reported their own ADHD related behaviours. Participants with three or more missing inattentive or hyperactive-impulsive items were removed from the samples, reducing the size of sample 1 by 13.

Nine-hundred and twenty-four participants provided valid data to both projects, this allowed us to compare maternal and self-reported SWAN scores. We call this group sample 3. The mean age of sample 3 was 24.30 (3.94), and 41.2 % of these participants were female.

Twin zygosity was determined by typing nine independent polymorphic DNA markers with the AmpFLSTR Profile PCR Amplification Kit. Cross-checks compared ABO, MN and Rh blood groups and phenotypic information. The probability of error using this method is less than 10^{-4}. Subsequent genotyping with the Illumina 610 k array for the majority of the sample confirmed zygosity determination.

Participants and their parents were fully informed of study procedures and gave consent to participate.

Measurement

ADHD Data collection for samples 1 and 2 began in 2008 and 2009 respectively. The 18 SWAN items (Swanson et al. 2005) address the 18 ADHD criterion *A* symptoms listed in the Diagnostic and Statistical Manual of Mental Disorders, fourth edition-text-revision (DSM-IV-TR) but are worded to reflect normal behaviour. The mean of the first nine items represented participants' inattention score and the mean of items 10–18 represented their level of hyperactivity-impulsivity. The mean of all 18 items represented participants' level of combined ADHD related behaviours. Descriptive statistics for the SWAN scale are presented in Tables 1 and 2 by zygosity group, for samples 1 and 2 respectively.

ADHD measurement in sample 1

During the adolescent twins' first or second clinic visit, ADHD data were collected from mothers for twin-pairs and their siblings ($N = 512$). Data for twins completing clinic visits prior to introduction of the SWAN were collected using an online questionnaire completed at home, taking approximately 60–90 min ($N = 2,711$). If participants were aged twenty or younger, mothers were asked about childrens' *current symptoms* ($n = 1,016$). If participants were older than twenty, mothers' were asked about their childrens' ADHD symptoms when they were in primary school, scale items were *retrospective* ($n = 1,695$). Both clinic and online versions of the questionnaire recorded responses on a 7-point scale ranging from −3–3 describing participants' *ability to maintain attention* for example, as: *far above average, above average, slightly above average, average, slightly below average, below average, far below average.*

Mothers of 103 participants completed the SWAN during the clinic visit and approximately 2-years later ($SD = 6$-months) using the online questionnaire. Retest correlations followed by 95 % confidence intervals (CI) for the reported inattention, hyperactive-impulsivity and combined ADHD were 0.81 (0.73–0.87), 0.76 (0.67–0.83) and 0.82 (0.74–0.87) respectively.

ADHD measurement in sample 2

The SWAN scale used to collect self-reported ADHD symptoms in sample 2 was 5-point ranging from −2–2: *far above average, above average, average, below average* and *far below average*. Data were collected across three waves; the first wave included 373 participants and the SWAN scale was *retrospective*. Participants were asked to rate their behaviour in relation to peers when they were in primary school and aged around 7–10. The second wave included 711 participants and the third wave included 905, questionnaires in these two waves asked participants about *current* symptoms. Some participants contributed data to multiple waves of data collection allowing us to estimate retest correlations, these were listed in Table 3 along with the period of time between testing.

Less than 2 % of ADHD data points were missing within samples 1 and 2.

ADHD measurement in sample 3

Scales in samples 1 and 2 were standardized (mean = 0, variance = 1) this allowed us to compare ADHD scores across studies using sample 3. Maternal-report ADHD data for these participants was both current ($n = 248$) and retrospective ($n = 676$). Self-report ADHD came from each wave of sample 2 data collection (1 = 254, 2 = 293, 3 = 377), also including current and retrospective report of symptoms.

Analyses

Data from the first and second sibling within each twin's family was included in analyses whenever available, to increase the power to detect dominant genetic and common environmental effects (Posthuma and Boomsma 2000).

Table 2 Mean (SD) within-pair correlations for sample 2 data collection: self-report SWAN scores by data collection method and zygosity group

	Twin pairs	Unpaired twins	1 sibling *n*	2 siblings *n*	*N*	Inattention *M (SD)*	Hyperactivity impulsivity *M (SD)*	Combined ADHD *M (SD)*	Pearson's product moment correlations		
									Inattention	Hyperactivity impulsivity	Combined ADHD
Wave 1 retrospective report N = 373, age range = 18–30											
MZ female	44	10	–	–	98	−0.33 (0.88)	−0.16 (1.06)	−0.27 (0.96)	0.42 (0.14, 0.64)	0.41 (0.13, 0.63)	0.39 (0.11, 0.62)
MZ male	14	17	–	–	45	−0.20 (0.81)	−0.09 (0.88)	−0.16 (0.77)	−0.14 (−0.63, 0.41)	−0.34 (−0.73, 0.23)	−0.29 (−0.71, 0.29)
DZ female	27	22	–	–	76	−0.22 (0.93)	0.04 (1.05)	−0.10 (0.97)	0.34 (−0.04, 0.64)	0.17 (−0.22, 0.52)	0.34 (−0.05, 0.64)
DZ male	10	21	–	–	41	−0.35 (0.87)	−0.25 (1.01)	−0.33 (0.90)	−0.15 (−0.71, 0.53)	0.03 (−0.61, 0.65)	−0.12 (−0.70, 0.55)
DZ female/male	24	15	–	–	63	−0.15 (0.98)	0.08 (1.11)	−0.04 (1.02)	0.26 (−0.16, 0.60)	0.34 (−0.07, 0.66)	0.39 (−0.01, 0.69)
DZ male/female	19	12	–	–	50	−0.17 (0.97)	0.14 (1.02)	−0.02 (0.96)	0.03 (−0.43, 0.47)	−0.02 (−0.46, 0.44)	−0.02 (−0.47, 0.44)
Siblings	–	–	–	–	–	–	–	–	–	–	–
Wave 2 current report N = 427, age range = 19–38											
MZ female	34	41	6	2	119	−0.21 (1.02)	−0.19 (1.05)	−0.22 (1.06)	0.65 (0.39, 0.81)	0.44 (0.12, 0.68)	0.60 (0.33, 0.78)
MZ male	25	36	6	–	92	0.44 (1.01)	0.18 (1.12)	0.34 (1.04)	0.42 (0.03, 0.70)	0.54 (0.19, 0.77)	0.47 (0.09, 0.73)
DZ female	12	28	4	–	56	0.13 (1.05)	−0.08 (1.06)	0.03 (1.08)	−0.40 (−0.79, 0.23)	−0.43 (−0.81, 0.19)	−0.54 (−0.85, 0.05)
DZ male	10	27	6	1	55	0.42 (1.16)	0.26 (0.97)	0.38 (1.07)	−0.22 (−0.75, 0.48)	−0.45 (−0.84, 0.24)	−0.49 (−0.86, 0.19)
DZ female/male	12	32	2	1	60	−0.05 (1.17)	−0.08 (1.11)	−0.07 (1.14)	−0.33 (−0.76, 0.29)	−0.16 (−0.67, 0.45)	−0.21 (−0.70, 0.41)
DZ male/female	6	32	1	0	45	0.02 (0.95)	−0.23 (0.92)	−0.12 (0.92)	−0.89 (−0.99, 0.30)	−0.36 (−0.91, 0.63)	−0.79 (−0.97, 0.07)
Siblings	–	–	–	–	–	−0.03 (1.12)	−0.12 (1.01)	−0.08 (1.08)	0.23 (−0.20, 0.59)	0.02 (−0.40, 0.43)	0.14 (−0.29, 0.52)
Wave 3 current report N = 613, age range = 18–33											
MZ female	46	44	5	0	141	−0.03 (0.95)	−0.10 (0.93)	−0.07 (0.97)	0.54 (0.30, 0.72)	0.48 (0.23, 0.68)	0.55 (0.31, 0.72)
MZ male	24	37	5	0	90	0.22 (0.89)	0.20 (0.78)	0.23 (0.82)	0.34 (−0.08, 0.65)	0.30 (−0.11, 0.63)	0.33 (−0.09, 0.65)
DZ female	43	37	4	0	127	0.07 (1.02)	0.02 (0.99)	0.05 (1.03)	−0.23 (−0.50, 0.07)	−0.11 (−0.40, 0.19)	−0.19 (−0.47, 0.11)
DZ male	28	37	3	0	96	0.24 (0.81)	0.32 (0.81)	0.31 (0.80)	0.13 (−0.25, 0.48)	0.00 (−0.37, 0.37)	0.01 (−0.36, 0.39)
DZ female/male	13	37	2	0	65	−0.10 (0.92)	−0.10 (0.97)	−0.11 (0.95)	−0.39 (−0.77, 0.21)	−0.07 (−0.59, 0.50)	−0.18 (−0.66,0.41)
DZ male/female	28	37	1	0	94	0.09 (0.92)	0.17 (0.84)	0.15 (0.85)	−0.04 (−0.41, 0.33)	−0.26 (−0.57, −.13)	−0.14 (−0.49, 0.24)
Siblings	–	–	–	–		0.13 (1.04)	0.03 (1.12)	0.09 (1.12)	−0.12 (−0.29, 0.07)	−0.07 (−0.25, 0.11)	−0.10 (−0.28, 0.08)
Data combined N = 1507, age range = 18–38											
MZ female	139	65	27	4	378	−0.17 (0.96)	−0.14 (1.00)	−0.17 (1.00)	0.58 (0.46, 0.68)	0.47 (0.33, 0.59)	0.55 (0.43, 0.66)
MZ male	79	58	24	0	240	0.22 (0.95)	0.13 (0.95)	0.19 (0.92)	0.29 (0.08, 0.49)	0.27 (0.05, 0.46)	0.25 (0.04, 0.45)
DZ female	105	41	26	1	279	−0.01 (1.00)	0.01 (1.02)	0.00 (1.02)	−0.05 (−.23, 0.14)	−0.03 (−0.22, 0.17)	−0.05 (−0.24, 0.14)
DZ male	62	57	17	2	204	0.15 (0.96)	0.17 (0.93)	0.18 (0.94)	0.02 (−0.22, 0.27)	−0.13 (−0.36, 0.12)	−0.13 (−0.37, 0.12)
DZ female/male	60	62	15	5	207	−0.10 (1.02)	−0.03 (1.06)	−0.07 (1.03)	−0.05 (−0.30, 0.20)	0.20 (−0.05, 0.44)	0.16 (−0.10, 0.40)
DZ male/female	66	55	14	0	201	0.00 (0.94)	0.67 (0.92)	0.04 (0.90)	−0.03 (−0.27, 0.21)	−0.15 (−0.38, 0.10)	−0.10 (−0.33, 0.14)
Siblings	–	–	–	–	–	0.02 (1.10)	−0.08 (1.05)	−0.03 (1.09)	0.02 (−0.09, 0.13)	0.01 (−0.09, 0.12)	0.02 (−0.09, 0.13)

SWAN 5-point scale. Higher values indicate the presence of ADHD symptoms. Only twin pairs, the first two additional siblings within any family and siblings with twins were included in these descriptive statistics and subsequent analyses, excluding 0 participants in wave 1, 145 in wave 2 and 109 in the full sample

Twin methodology

Following classical twin methodology, monozygotic twins were considered genetically identical and any difference between twins within a pair was due to environmental experiences. Dizygotic twins shared on average 50 % additive genetic effects and 25 % dominant genetic effects, while both MZ and DZ twins could be assumed to share the family environment to the same extent. The genetic relatedness between family members was used to dissect the variance of ADHD ratings into genetic and environmental components: the additive effect of markers across multiple loci (A), variants acting in a dominant manner (D), environmental sharing between family members (C), and specific environmental experiences that made individuals within families differ (E). The E variance component also included measurement error. D and C were negatively confounded as both estimates were based on the difference in the MZ and DZ within pair correlation so could not be included in the same model for twins reared together.

Contrast effects act in the same way as dominance and could not be estimated in the same model as C. Therefore separate models were run for C and D parameters to determine which provided the best fit to the data and contrast parameters were excluded from the models including C. The contrast parameters were modeled as a direct pathway (b) from one twin's phenotype to that of their cotwin (or sibling) taking a negative value. The covariance for each ADHD phenotype was therefore represented as $(I-b)^{-1}\ (A + D)\ (I-b)^{-1}$ for MZ twins, and $(I-b)^{-1}\ (.5A + .25D)\ (I-b)^{-1}$ for DZ twins. The difference between MZ and DZ twin-pair covariance provided an estimate of the magnitude of genetic effects influencing ADHD symptoms. These analyses were run in classic Mx (Neale et al. 1999).

Univariate twin models testing contrast and sex effects on ADHD

Three nested models were used to estimate sex difference in the aetiology of ADHD: *(1)* the general model illustrated in Fig. 1, allowed qualitative and quantitative (scalar) A, C or D and E paths for males (am_{ij}, dm_{ij}, em_{ij}) and females (af_{ij}, df_{ij}, ef_{ij}), including a male specific genetic effect (mg_{ij}), *(2)* a scalar sex-limitation model allowed separate A, C or D and E paths across sex assuming the genetic effects contributing to symptoms were the same for males and females, but genetic and environmental paths differed by magnitude, and *(3)* a progressively constrained model in which path coefficients were equated between males and females ($am_{ij} = af_{ij}$, $dm_{ij} = df_{ij}$, $em_{ij} = ef_{ij}$). Contrast effects were included in each of these models for twins (b_{21}) and siblings (b_{32}, b_{43}) and initially allowed to differ between the sexes.

Bivariate modelling comparing maternal and self-report ADHD

The bivariate Cholesky decomposition illustrated in Fig. 1 allowed an estimation of genetic (a_{21}, d_{21}) and environmental factors (e_{21}) common to maternal and self-reported inattention, hyperactivity-impulsivity and combined ADHD related behaviours. This model also provided an estimate of genetic and environmental effects specific to either maternal (a_{11}, d_{11}, e_{11}) or self-reported (a_{22}, d_{22}, e_{22}) ADHD subtypes. Contrast effects for twins (b_{31}) and siblings (b_{53}, b_{75}) were also included in this model.

Common factors model

A common factors model was used to separate specific environmental factors (E) from error (R) for each rater, using sample 3. The common factor model can only be identified when four or more variables are included in the model. We included only two variables, so the loadings of the latent factors (l) onto ADHD subtypes collected at time 1 and 2 were equated and the latent variable was scaled to 1 to identify the model (Loehlin 1996) for our purposes. Siblings could not be included in this model due to the relatively low number completing data at two time points and the effect this had on parameter estimation.

Results

Mean current inattentive symptoms reported by mothers (sample 1) were higher than when these symptoms were reported retrospectively [$t\ (1{,}994) = 3.04,\ p = 0.002$]. This was not the case for hyperactivity-impulsivity [$t(2{,}091) = -0.93,\ p = 0.35$]. The average ages of participants with current and retrospectively reported data were 17.14 (2.31) and 26.26 (3.66) respectively. Self-reported inattention was also higher when participants reported on their current, in relation to retrospective symptoms [$t\ (674) = 5.77,\ p < 0.001$]. This was not the case for self-reported symptoms of hyperactivity-impulsivity [$t\ (589) = 0.87,\ p = 0.39$]. The ages of these groups with current and retrospective self-report data were 24.20 (3.30) and 25.18 (3.72) respectively.

Heritability estimates were calculated within each sample, for each method of data collection. Clinic data collection, retrospective and current report of symptoms in sample 1, and for each wave of data collection in sample 2. These results are presented in Table 4. Inspection of the

Table 3 Retest correlations (CI) for waves of data collection in sample 2

	Inattention		Hyperactivity-impulsivity		Combined ADHD	
	Wave 2	Wave 3	Wave 2	Wave 3	Wave 2	Wave 3
Wave 1	0.44	0.46	0.43	0.52	0.46	0.55
	(0.30, 0.57)	(0.34, 0.57)	(0.28, 0.56)	(0.41, 0.62)	(0.31, 0.59)	(0.45, 0.64)
Wave 2		0.66		0.52		0.62
		(0.43, 0.81)		(0.23, 0.73)		(0.37, 0.79)

Mean time lapse between waves: 1 and 2 = 1.6 years ($SD = 0.5$), 1 and 3 = 3.7 years ($SD = 0.3$), 2 and 3 = 2.1 ($SD = 2.1$). Size of re-test sample waves: 1 and 2 $n = 132$, 1 and 3 $n = 194$, 2 and 3 $n = 36$

estimates justified pooling data within each sample and including covariates to account for variation in method of data collection and age.

Likelihood-ratio tests (LRT) were used to estimate the homogeneity of means, variances across zygosity group within each sample. In sample 1, mean maternal-report inattention scores were approximately equal within sex, across zygosity groups. Females had lower mean scores than males ($\Delta -2LL = 77.1$, $\Delta df = 3$, $p < 0.001$). Variance of these scores was approximately equal for MZ and DZ zygosity groups, providing no evidence for a contrast effect. However variance was greater for DZ male than DZ female twins ($\Delta -2LL = 9.20$, $\Delta df = 3$, $p = 0.03$). Mean hyperactivity-impulsivity scores in sample 1 were also lower for females ($\Delta -2LL = 43.7$, $\Delta df = 3$, $p < 0.001$), but could be equated within sex across zygosity groups. The variance of hyperactivity-impulsivity was approximately equal for MZ and DZ zygosity groups and for males and females.

In sample 2 self-report data, mean inattention scores were approximately equal across DZ same and opposite sex twins, but MZ males scores were higher than MZ females ($\Delta -2LL = 13.67$, $\Delta df = 1$, $p < 0.001$). All variances were approximately equal. Mean hyperactivity-impulsivity scores were higher for males ($\Delta -2LL = 11.4$, $\Delta df = 3$, $p = 0.009$), but all variances and means within sex and across zygosity group were equal.

Univariate sex-limited analyses and contrast effects for ADHD subtypes

The presence of rater contrast effects (*b*) *in* SWAN measured ADHD was tested within each sex-limitation model. Parameters were consecutively constrained to determine their relevance to model fit. The best fitting models for inattention, hyperactivity-impulsivity and combined ADHD related behaviour within each dataset were presented in Table 5 respectively, and discussed below. Results of all model-fitting was provided in supplementary Tables 7 and 8 in Appendix Section for maternal and self-report ADHD subtypes respectively.

Inattention

ADE and *ADE* + *b* models were the best fit for maternal and self-report inattention. The *ADE* parameters were equal for males and females in maternal-report data but the magnitude of *D* and *E* differed across sex when inattention was self-reported. There was a contrast effect evident in self- but not maternal-report inattention (−0.09).

Hyperactivity-impulsivity

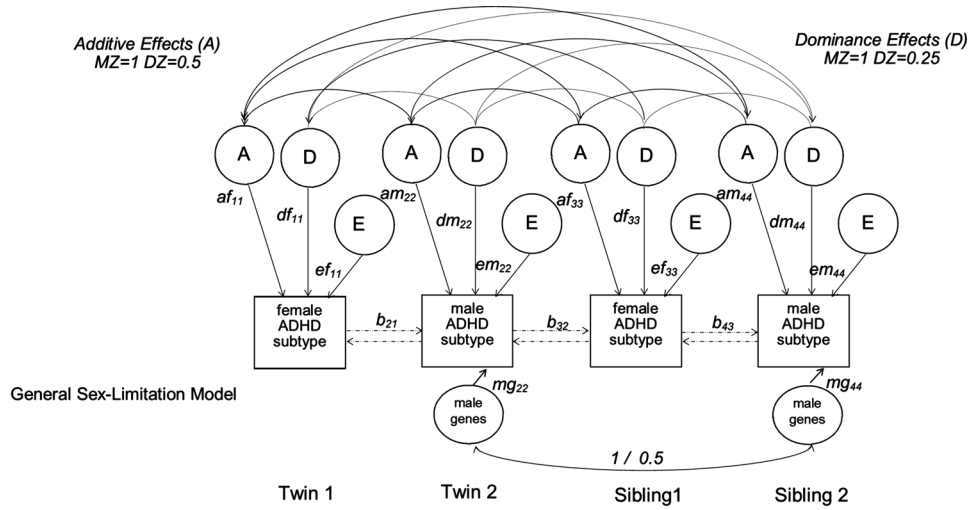
AE + *b* and *ADE* models provided the best fit for maternal and self-reported hyperactivity-impulsivity. The *A* and *E* factors could be equated for males and females in both samples, but the magnitude of *D* was greater for females than males in sample 2. The magnitude of *E* was greater for males than females in sample 1.

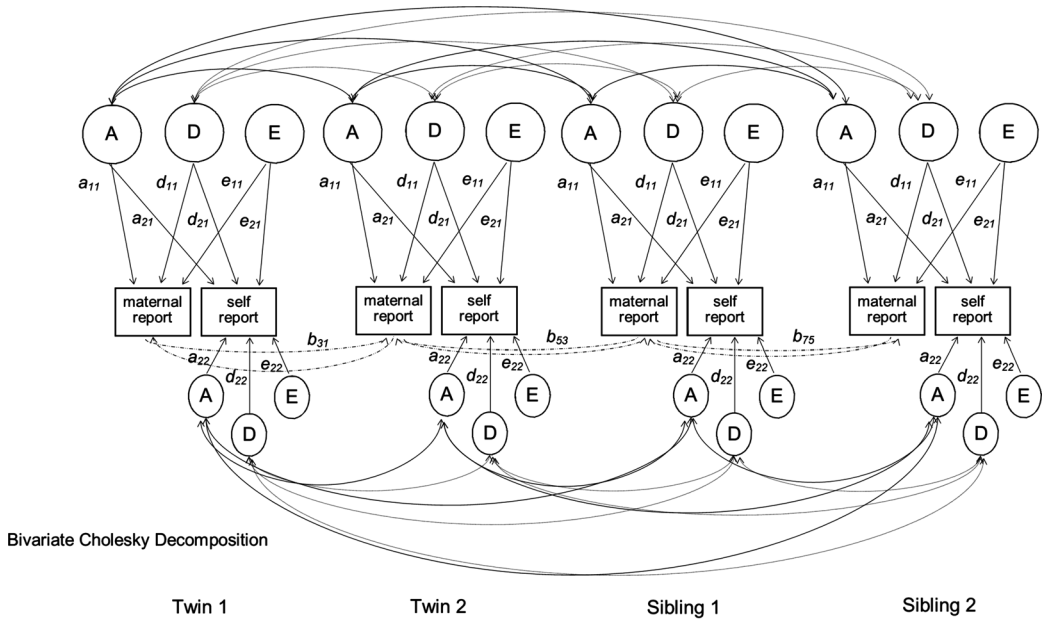
Combined ADHD

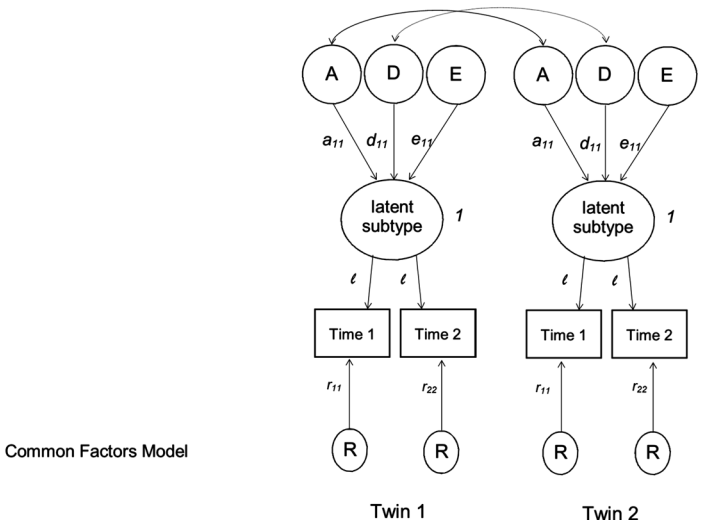
ADE and *ADE* + *b* models best described sample 1 and 2 combined ADHD behaviours. The magnitude of *A* was equal for males and females within both samples. Similarly *D* was equal across sex in sample 1 but differed for males and females when combined ADHD was self-reported. The magnitude of *E* was greater for males than females in both samples. There was a contrast effect evident in self-report, but not maternal-report of symptoms. There was however no difference between the MZ and DZ twin variance for combined ADHD related behaviours in sample 2 as would be expected in the presence of a contrast effect.

Bivariate Cholesky rater comparison

The bivariate Cholesky decomposition allowed us to estimate genetic and environmental effects that were shared by maternal and self-reported inattention, hyperactivity-impulsivity and combined ADHD ratings for the 924 participants contributing data to both samples. We were also able to identify *A*, *D* and *E* factors specific to each rater (see Table 6). Rater contrast parameters were included for


Additive Effects (A)
MZ=1 DZ=0.5
Dominance Effects (D)
MZ=1 DZ=0.25
A
D
E
af11
df11
ef11
am22
dm22
em22
af33
df33
ef33
am44
dm44
em44
female ADHD subtype
male ADHD subtype
female ADHD subtype
male ADHD subtype
b21
b32
b43
mg22
mg44
male genes
male genes
1 / 0.5
General Sex-Limitation Model
Twin 1
Twin 2
Sibling1
Sibling 2


A
D
E
a11
d11
e11
a21
d21
e21
maternal report
self report
b31
b53
b75
a22
d22
e22
Bivariate Cholesky Decomposition
Twin 1
Twin 2
Sibling 1
Sibling 2


A
D
E
a11
d11
e11
latent subtype
1
ℓ
Time 1
Time 2
r11
r22
R
Common Factors Model
Twin 1
Twin 2

◀**Fig. 1** Models used to decompose genetic and environmental variance in ADHD related behaviours. In all cases dominant genetic effects (*D*) provided a better model fit than common environment (*C*), *E* represents specific environmental factors, *b* represents contrast between twins (b_{21}, b_{31}) and siblings (b_{32}, b_{43}, b_{53}, b_{75}). Lower case *a, d,* and *e* represent additive genetic, dominant genetic and specific environmental loadings onto measured phenotypes. Lower case *f* and *m* represent the female and male specific effects in the test for sex-limitation and *mg* represents the male specific genetic effect. Subscripted numbers show the placement of effects in model matrices and 1 represents the loadings of latent ADHD subtypes onto data collected at times 1 and 2 for sample 3, in the common factors model

maternal and self-report data due to their presence in the univariate models. However this parameter could be dropped from each model without a significant deterioration in model fit (see Table 9 in Appendix Section of the supplementary, rows 4, 11 and 18).

The *A* factor loading onto each ADHD subtype was not shared by raters. The genetic factor common to maternal and self-reported behaviours showed a pattern of dominance (see Table 6 column 4, rows 3, 6 and 9). This genetic correlation was 0.64 (95 % CI 0.51, 0.89) for inattention, 1.00 (95 % CI 0.73, 1.00) for hyperactivity-impulsivity and 0.94 (95 % CI 0.69, 1.00) for combined ADHD.

The *E* factor for specific environment common to raters could not be dropped without a deterioration in model fit. These correlations were 0.17 (95 % CI = 0.03, 0.30) for inattention, 0.22 (95 % CI = 0.07, 0.37) for hyperactivity-impulsivity and 0.16 (95 % CI = 0.00, 0.30) and may represent a case of correlated error terms.

Common factors model examining the aetiology of specific environment (*E*)

Measurement error (*R*) and the specific environmental factor (*E*) were separable in this model allowing us to determine the magnitude of difference in *E* not shared by raters. The squared loading (l^2) of the latent factor in supplementary Table 10 (column 9) of Appendix, onto ADHD measures at times 1 and 2 provided as estimate of retest-correlations for each ADHD subtype and rater. These estimates should be similar to the estimates provided in the methods (ADHD Measurement in Sample 1) and Table 3. The magnitude of *R* was greater in

Table 4 Heritability Estimates for Samples 1 and 2 by Subtype

Model	AIC	−2LL	df	A	E
Maternal-report					
Clinic					
Inattention	241.17	1,239.17	499	0.63 (0.47, 0.75)	0.37 (0.25, 0.53)
Hyp-imp	157.53	1,155.53	499	0.79 (0.71, 0.84)	0.21 (0.16, 0.29)
Combined	156.74	1,154.74	499	0.75 (0.66, 0.82)	0.25 (0.18, 0.34)
Current					
Inattention	691.30	2,479.30	894	0.65 (0.54, 0.73)	0.35 (0.27, 0.46)
Hyp-imp	610.20	2,392.20	891	0.71 (0.63, 0.79)	0.29 (0.22, 0.37)
Combined	623.35	2,373.35	875	0.71 (0.63, 0.80)	0.29 (0.22, 0.37)
Retrospective					
Inattention	820.61	3,716.61	1,448	0.72 (0.65, 0.78)	0.28 (0.22, 0.35)
Hyp-imp	735.13	3,589.13	1,427	0.86 (0.83, 0.89)	0.14 (0.11, 0.17)
Combined	725.17	3,535.17	1,405	0.84 (0.80, 0.87)	0.16 (0.13, 0.20)
Self-Report					
Wave 1					
Inattention	239.32	975.32	368	0.31 (0.10, 0.49)	0.69 (0.51, 0.90)
Hyp-imp	340.05	1,076.05	368	0.31 (0.10, 0.49)	0.69 (0.51, 0.91)
Combined	268.93	1,005.93	368	0.30 (0.09, 0.47)	0.70 (0.52, 0.91)
Wave 2					
Inattention	394.88	1,238.88	422	0.59 (0.37, 0.73)	0.41 (0.27, 0.63)
Hyp-imp	373.03	1,217.03	422	0.46 (0.25, 0.62)	0.54 (0.38, 0.75)
Combined	384.85	1,228.85	422	0.56 (0.35, 0.70)	0.44 (0.30, 0.65)
Wave 3					
Inattention	412.48	1,628.48	608	0.33 (0.13, 0.51)	0.67 (0.49, 0.87)
Hyp-imp	388.96	1,604.96	608	0.26 (0.08, 0.43)	0.74 (0.57, 0.92)
Combined	402.07	1,618.07	608	0.32 (0.13, 0.48)	0.68 (0.52, 0.87)

AIC Akaike Information Criterion, *−2LL* −2 × log-likelihood, *df* degrees-of-freedom, *A* additive genetic effect, *E* unique environmental effect

Table 5 Standardised estimates for best fitting sex-limitation models with contrast parameters for maternal and self-reported SWAN

Best fitting model	Female variance					Male variance				Twin contrast
	A	C	D	E	Male specific	A	C	D	E	
Inattention										
Mother	0.17	–	0.55	0.28	–	0.17	–	0.55	0.28	–
Equal ADE	(−0.35, 0.35)		(0.37, 0.74)	(0.24, 0.32)		(−0.35, 0.35)		(0.37, 0.74)	(0.24, 0.32)	
Self	0.04	–	0.65	0.31	–	0.05	–	0.41	0.54	−0.09
Equal A, D and E differ	−0.52, 0.52)		(0.21, 0.82)	(0.23, 0.42)		(−0.58, 0.58)		(0.22, 0.65)	(0.39, 0.77)	(−0.17, −0.02)
Hyperactivity-impulsivity										
Mother	0.86	–	–	0.14	–	0.79	–	–	0.21	−0.04
Equal ADE + contrast	(0.79, 0.94)			(0.11, 0.17)		(0.72, 0.85)			(0.17, 0.26)	(−0.08, 0.00)
Self	0.00	–	0.41	0.59	–	0.00	–	0.30	0.70	–
Equal A and E	(−0.26, .0.26)		(0.11, 0.53)	(0.50, 0.70)		(−0.31, 0.31)		(−0.33, 0.45)	(0.59, 0.82)	
Combined symptoms										
Mother	0.62	–	0.21	0.18	–	0.57	–	0.19	0.24	–
Equal AD, E differs, no contrast	(0.43, 0.81)		(0.03, 0.38)	(0.15, 0.21)		(0.40, 0.75)		(0.03, 0.36)	(0.20, 0.29)	
Self	0.10	–	0.55	0.35	–	0.12	–	0.27	0.60	−0.08
Equal A, D and E differ	(−0.48, 0.48)		(0.50, 0.77)	(0.26, 0.47)		(−0.58, 0.58)		(0.14, 0.58)	(0.43, 0.84)	(−0.16, −0.02)

When paths were equated for females and males standardised estimates may differ due to greater variance in male scores. All model fitting for these analyses are provided in supplementary Table 7 and Table 8

self-report than maternal-report data (see column 10, Table 10 of Appendix), confidence intervals did not overlap. The magnitude of *E* (column 8, Table 10 of Appendix) was also greater for hyperactivity-impulsivity and combined ADHD when self-reported.

Discussion

We have addressed three questions raised in ADHD research: (1) the presence of rater contrast effects in maternal-report ADHD measured using a dimensional scale (Pinto et al. 2012), (2) the aetiology of difference between heritability estimates derived using maternal and self-reported data (Merwood et al. 2013; Chang et al. 2013), and (3) sex differences in genetic and environmental influences on ADHD-related behaviours.

We found there were rater contrast effects evident for hyperactivity-impulsivity in data collected from mothers of twins using a dimensional scale (SWAN). The size of the effect was consistent with previous findings (−0.04; Merwood et al. 2013) using maternal-report ADHD. Surprisingly there was a contrast effect evident in self-report inattention (−0.08) and combined ADHD related behaviours (−0.09). These effects could be due to greater contact and perhaps similarity between MZ twins in adulthood or sibling interaction—this was not examined. There was no evidence of contrast in the variances of MZ and DZ twins, leaving uncertainty about the origin of the contrast effect in our data.

The bivariate analyses we conducted show differences in heritability between maternal and self-reported data within our samples were due to the greater magnitude of rater specific *E* loading onto self-reported hyperactivity-impulsivity and combined behaviours. Other groups have found the *E* factor in rater specific self-report ADHD includes additive genetic effects (manuscript submitted), we did not test for this.

The retest-reliability of self-reported ratings within our data was moderate ($r = 0.46$), and lower than the retest-reliability of maternal-report ADHD ($r = 0.81$) over approximately 2 years. The phenotypic correlation between maternal and self-reported data was 0.34 (0.27, 0.41) for ADHD total score 0.38 (0.31, 0.44) for inattention and 0.27 (0.19, 0.34) for hyperactivity-impulsivity. These findings were consistent with reduced reliability of self-report data.

Researchers have previously suggested symptomatic adults do not provide an accurate report of their behaviour (Knouse et al. 2005) but others indicate adults with ADHD provide the more accurate assessment of their symptoms but minimize symptom severity (Kooij et al. 2008). We did not address this question but did find greater variation in

Table 6 Standardised path estimates for best fitting bivariate cholesky examining the aetiology of maternal and self-reported ADHD subtypes

Reporter	A		D		E		Twin contrast
	Mother	Self	Mother	Self	Mother	Self	
Inattention							
Mother	0.08		0.65		0.26		–
	(0.00, 0.25)		(0.48, 0.79)		(0.23, 0.30)		
Self	–	0.00	0.36	0.49	0.06	0.52	
		(0.00, 0.00)	(0.29, 0.44)	(0.27, 0.59)	(0.01, 0.12)	(0.44, 0.62)	
Hyperactivity-impulsivity							
Mother	0.65		0.16		0.18		–
	(0.52, 0.74)		(0.09, 0.28)		(0.16, 0.21)		
Self	–	0.00	0.25	0.38	0.07	0.62	
		(0.00, 0.14)	(0.18, 0.32)	(0.21, 0.48)	(0.02, 0.13)	(0.53, 0.73)	
Combined symptoms							
Mother	0.50		0.31		0.18		–
	(0.34, 0.62)		(0.20, 0.47)		(0.16, 0.44)		
Self	–	0.00	0.35	0.44	0.05	0.56	
		(0.00, 0.16)	(0.27, 0.42)	(0.26, 0.55)	(0.00, 0.10)	(0.47, 0.66)	

All model fitting results for these analyses are shown in supplementary Table 10

the specific environmental/error factor influencing self-reported data. A general review of the literature examining self-report of ADHD did however show, self-reported symptoms were reliable and valid in European (Adler et al. 2008; Magnússon et al. 2006) and Korean (Kim et al. 2013) samples. However, these results may not be comparable to ours due to the fact that they concentrate on more severely affected samples.

There were additive genetic factors specific to mothers' report of each ADHD subtype not evident when symptoms were self-reported. These factors could partly account for the severity of symptoms that has previously been reported as a cause of variation between mother and self-ratings (Kooij et al. 2008). There were also sex differences in self-reported data. The magnitude of dominant genetic effects influencing inattention was greater for females than males. The magnitude of unique environmental effects was greater for males within each self-reported subtype. Males generally have greater variance in ADHD scores (Chang et al. 2013; Merwood et al. 2013; Ebejer et al. 2013) and unique environmental experience may account for this. The magnitude of *D* was greater for females in self-reported data. We found scalar differences in the genetic and environmental effects across sex for self-report, but not maternal-report subtypes. Our findings differ from the work done by Merwood and colleagues (2013) in this regard.

Broad-sense heritability estimates of maternal report inattention, hyperactivity-impulsivity and combined behaviours were approximately equal for men and women (0.72, 0.79, 0.76 and 0.72, 0.86, 0.83 respectively). The broad-sense heritability of self-reported symptoms was lower for men than women for inattention (0.46 vs 0.69), hyperactivity-impulsivity (0.30 vs 0.41) and combined behaviours (0.39 vs 0.65).

The SWAN scale measures behaviours across the full spectrum, including high and low levels of attention and activity. Merwood and colleagues (2013) used a severity scale to measure ADHD within their sample of 11 and 12 year olds and the rater contrast effect they found for ADHD total score was approximately equal to the contrast effect we found influencing symptoms of hyperactivity-impulsivity in maternal-report data. Additionally a contrast effect unexpectedly appeared in self-report data, possibly due to the greater difference between MZ and DZ twin correlations for ADHD subtypes. The negative twin correlations evident in self-report data could account for this effect.

It is important to consider our findings in relation to the limitations of the study. There was a difference in the period of time to which scale items were addressed; approximately half of sample 1 and one-third of the sample 2 were retrospectively reporting on childhood symptoms. Symptoms of inattention were lower on average when they were reported retrospectively, suggesting participants were more attentive as children than they were as teenagers and young adults. Additionally the possible correlation of error terms in our comparison of maternal and self-reported variance components could show a model bias. This effect could also be due to the similarity of wording on the scale used to collect data from each group of informants (Bollen and Lennox 1991).

Despite these limitations our results show a contrast parameter for inattentive and hyperactive-impulsive subtypes when using a dimensional measure of symptoms. We also found ~14 % of the variation in self-reported ADHD was accounted for by variation in specific environmental experience and ~50 % of the variation was due to other factors falling into *E* when two raters rather than one report on twins within a pair.

Acknowledgments We thank Marlene Grace and Natalie Garden for conducting initial interviews with twins, Richard Parker for ongoing data collection and both David Smyth and Anthony Conciatore for IT support. But most importantly, we thank the twins and their families for continued involvement and interest in our studies. JLE was supported by an Australian Postgraduate Award and a Strategic Doctoral Scholarship. This work was also supported in part by the Australian National Health and Medical Research Council Grant 1009839. SEM was supported by an Australian Research Council (full) Future Fellowship 110100548. DLD was support by an NHMRC Senior Research Fellowship. Statistical analyses were carried out on the GenEpi Cluster which is financially supported by contributions from Grants from the NHMRC (389892; 496682; 496688; 496739; 613672) and ARC (FT0991022; FT0991360).

Conflict of Interest There are no conflicting interests to declare.

Human and Animal Rights and Informed Consent QIMR Human Research and Ethics Committee (HREC) approved samples 1 and 2 data collection. Data collection for sample 2 was also approved by the Virginia Commonwealth University Institutional Review Board (IRB).

Appendix

See Tables 7, 8, 9 and 10.

Table 7 Univariate sex-limited twin models with contrast effects for maternal-report ADHD subtypes

Model	AIC	−2LL	df	Δ−2LL	Δdf	*p* value	Best fitting model
Inattention							
ACE	2,012.23	8,258.23	3,123	–	–	–	
ADE + b	1,999.86	8,233.86	3,117	24.37	–	–	
Equate b sibs	1,996.12	8,238.12	3,121	4.26	4	0.37	
Drop b sibs	1,994.93	8,238.93	3,122	0.81	1	0.37	No contrast effect for twins or sibs
Drop b twins	1,996.53	8,242.53	3,123	3.60	1	0.06	No male specific genetic effects
Drop M	1,995.34	8,243.34	3,124	0.81	1	0.37	Equal A, D and E for males and females
Equate A	1,995.21	8,245.21	3,125	1.87	1	0.17	
Drop D	2,026.64	8,280.64	3,127	35.43	2	<0.001	
Equate AD	1,996.82	8,248.82	3,126	3.61	1	0.06	
Equate ADE	**1,997.74**	**8,251.74**	**3,127**	**2.92**	**1**	**0.09**	
Hyperactivity-impulsivity							
ACE	1,717.55	7,963.55	3,123	–	–	–	
ADE + b	1,720.56	7,954.56	3,117	8.99	–	–	
Equate b sibs	1,716.09	7,958.09	3,121	3.53	4	0.47	Contrast effect for twins
Drop b sibs	1,716.85	7,960.85	3,122	2.76	1	0.10	No male specific genetic effects
Drop b twins	1,718.62	7,964.62	3,123	3.77	1	0.05	Equal A and no D effects
Drop M	1,714.85	7,960.85	3,123	0.00	1	1.00	E differs for males and females
Equate A	1,712.86	7,960.86	3,124	0.01	1	0.92	
Drop D	**1,710.14**	**7,962.14**	**3,126**	**1.28**	**2**	**0.53**	
Equate E	1,727.77	7,981.77	3,127	19.63	1	<0.001	
Combined ADHD							
ACE	1,774.44	8,020.44	3,123	–	–	–	
ADE + b	1,776.55	8,010.55	3,117	9.89	–	–	
Equate b sibs	1,772.04	8,014.04	3,121	3.49	4	0.48	No contrast effect for twins or sibs
Drop b sibs	1,770.05	8,014.05	3,122	0.01	1	0.92	No male specific genetic effects
Drop b twins	1,771.61	8,017.61	3,123	3.56	1	0.06	Equal A and D parameters
Drop M	1,769.66	8,017.66	3,124	0.05	1	0.82	E differs for males and females
Equate A	1,768.38	8,018.38	3,125	0.72	1	0.40	
Drop D	1,772.73	8,026.73	3,127	8.25	1	0.004	
Equate AD	**1,769.54**	**8,021.54**	**3,126**	**3.16**	**1**	**0.08**	
Equate E	1,776.64	8,030.64	3,127	9.10	1	0.003	

AIC Akaike Information Criterion, *−2LL* −2 × log-likelihood, *df* degrees-of-freedom, *M* male specific genetic effects. Best fitting models are bolded and selected according to the least degree of change in −2LL when parameters are dropped from the model

Table 8 Sex-Limited Twin Models with Contrast Effects for Self-Reported ADHD Subtypes

Model	AIC	−2LL	df	Δ−2LL	Δdf	*p* value	Best fitting model
Inattention							
ACE	1,232.56	4,382.56	1,575	–	–	–	
ADE + b	1,217.83	4,355.83	1,569	26.73	6	<0.001	
Equate b sibs	1,216.97	4,362.97	1,573	7.14	4	0.13	
Drop b sibs	1,215.02	4,363.02	1,574	0.05	1	0.82	Contrast effect for twins
Drop b twins	1,220.56	4,370.56	1,575	7.54	1	0.006	No male specific genetic effects
Drop M	1,213.05	4,363.05	1,575	0.03	1	0.86	Equal A for males and females
Equate A	**1,211.05**	**4,363.05**	**1,576**	**0.00**	**1**	**1.00**	E and D differ for male and females
Drop D	1,222.33	4,378.33	1,578	15.28	2	<0.001	
Equate AD	1,216.68	4,370.68	1,577	7.63	1	0.006	
Equate AE	1,214.84	4,368.84	1,577	5.79	1	0.02	
Hyperactivity-Impulsivity							
ACE	1,263.27	4,413.26	1,575	–	–	–	
ADE + b	1,257.69	4,395.69	1,569	17.57	6	0.007	No contrast effect for twins or sibs
Equate b sibs	1,259.42	4,405.42	1,573	9.73	4	0.05	No male specific genetic effects
Drop b sibs	1,257.61	4,405.61	1,574	9.92	5	0.08	Equal A and E for males and females
Drop b twins	1,259.09	4,409.09	1,575	3.48	1	0.06	D differs for males and females
Drop M	1,257.09	4,409.09	1,576	0.00	1	1.00	
Equate A	1,255.19	4,409.19	1,577	0.10	1	0.75	
Drop D	1,262.12	4,420.12	1,579	10.93	2	0.004	
Equate AD	1,258.50	4,414.50	1,578	5.31	1	0.02	
Equate AE	**1,254.66**	**4,410.66**	**1,578**	**1.47**	**1**	**0.23**	
Combined ADHD							
ACE	1,244.66	4,374.66	1,575	–	–	–	
ADE + b	1,212.98	4,350.98	1,569	23.68	6	<0.001	
Equate b sibs	1,214.98	4,360.98	1,573	10.00	4	0.04	Contrast effect for twins
Drop b sibs	1,213.18	4,361.18	1,574	10.19	5	0.07	No male specific genetic effects
Drop b twins	1,217.34	4,367.34	1,575	6.16	1	0.01	Equal A for males and females
Drop M	1,211.18	4,361.18	1,575	0.00	1	1.00	D and E differ for males and females
Equate A	**1,209.31**	**4,361.31**	**1,576**	**0.13**	**1**	**0.72**	
Drop D	1,221.15	4,377.15	1,578	15.86	2	<0.001	
Equate AD	1,218.74	4,372.74	1,577	11.56	1	<0.001	
Equate AE	1,211.00	4,365.00	1,577	3.69	1	0.05	

AIC Akaike Information Criterion, *−2LL* −2 × log-likelihood, *df* degrees-of-freedom, *M* male specific genetic effects. Best fitting models are bolded and selected according to the least degree of change in −2LL when parameters are dropped from the model

Table 9 Model Fitting for Bivariate Cholesky Comparing Rater A, D and E Parameters

Model	AIC	−2LL	Df	Δ−2LL	Δdf	*p* value	Best fitting model
Inattention							
AE + b	3,717.74	13,901.74	5,092	–	–	–	
ADE + b	3,656.80	13,834.80	5,089	66.94	3	–	No contrast effects
Drop b twins	3,658.39	13,838.39	5,090	3.59	1	0.06	No common A
Drop A common	**3,657.43**	**13,839.43**	**5,091**	**1.04**	**1**	**0.31**	Common E
Drop AE common	3,660.79	13,844.79	5,092	5.36	1	0.02	Common D
Drop AD common	3,736.58	13,920.58	5,092	81.15	1	<0.001	
Hyperactivity-Impulsivity							
AE + b	3,436.74	13,618.74	5,091	–	–	–	
ADE + b	3,400.74	13,578.74	5,089	40.00	2	–	
Drop b twins	3,401.76	13,581.76	5,090	3.02	1	0.08	No contrast effects
Drop A common	**3,399.80**	**13,581.80**	**5,091**	**0.04**	**1**	**0.84**	No common A
Drop AE common	3,404.81	13,588.81	5,092	7.01	1	0.008	Common E
Drop D common	3,436.80	13,620.80	5,092	31.99	1	<0.001	Common D
Combined ADHD							
AE + b	3,450.66	13,634.66	5,092	–	–	–	
ADE + b	3,427.65	13,605.65	5,089	28.99	3	–	
Drop b twins	3,428.65	13,608.65	5,090	3.00	1	0.08	No contrast effects
Drop A common	**3,427.47**	**13,609.47**	**5,091**	**0.82**	**1**	**0.37**	No common A
Drop AE common	3,429.39	13,613.38	5,092	3.91	1	0.05	Common E
Drop AD common	3,502.98	13,686.98	5,092	77.51	1	<0.001	Common D

AIC Akaike Information Criterion, *−2LL* −2 × log-likelihood, *df* degrees-of-freedom, *M* male specific genetic effects. Best fitting models are indicated by non-significant change in −2LL when parameters are dropped from the model

Table 10 Path Estimates and Model Fit for Common Factors Model Examining Specific Environment and Error Components of ADHD Subtypes

Model	AIC	−2LL	df	A	D	C	E	l	R
Maternal-report									
Inattention									
ADE	**1,430.43**	**5,994.43**	**2,282**	**0.00 (.00, .00)**	**0.70 (.51, .77)**	–	**0.06 (.00, .12)**	**0.87 (.82, .92)**	**0.20 (.15, .27)**
ACE	1,459.31	6,029.87	2,282	0.67 (.59, .76)	–	0.00 (.00, .02)	0.11 (.03, .18)	0.88 (.83, .93)	0.20 (.15, .27)
AE	1,459.31	6,025.31	2,283	0.67 (.59, .76)	–	–	0.11 (.03, .18)	0.88 (.83, .93)	0.20 (.15, .27)
Hyperactivity-impulsivity									
ADE	1,247.36	5,811.36	2,282	0.61 (.36, .81)	0.14 (.00, .38)	–	0.00 (.00, .02)	0.86 (.82, .90)	0.20 (.18, .23)
ACE	1,248.68	5,812.69	2,282	0.75 (.67, .82)	–	0.00 (.00, .07)	0.00 (.00, .02)	0.87 (.83, .91)	0.21 (.18, .24)
AE	**1,246.20**	**5,812.20**	**2,283**	**0.75 (.68, .82)**	–	–	**0.00 (.00, .02)**	**0.87 (.83, .91)**	**0.21 (.18, .24)**
Combined symptoms									
ADE	**1,263.71**	**5,827.71**	**2,282**	**0.34 (.08, .59)**	**0.41 (.17, .67)**	–	**0.00 (.00, .04)**	**0.87 (.83, .91)**	**0.20 (.16, .22)**
ACE	1,275.34	5,839.34	2,282	0.76 (.69, .84)	–	0.00 (.00, .00)	0.00 (.00, .06)	0.87 (.83, .92)	0.21 (.16, .24)
AE	1,270.82	5,836.82	2,283	0.76 (.69, .84)	–	–	0.00 (.00, .05)	0.88 (.83, .92)	0.20 (.16, .24)
Self-report									
Inattention									
ADE	1,222.84	4,614.84	1,696	0.00 (.00, .00)	0.42 (.28, .51)	–	0.02 (.00, .11)	0.66 (60, .73)	0.51 (.45, .59)
ACE	1,238.32	4,630.32	1,696	0.36 (.26, .46)	–	0.00 (.00, .04)	0.07 (.00, .18)	0.66 (.58, .72)	0.52 (.45, .60)
AE	**1,235.08**	**4,631.08**	**1,698**	**0.36 (.26, .46)**	–	–	**0.07 (.00, .18)**	**0.66 (.59, .73)**	**0.52 (.45, .60)**

Table 10 continued

Model	AIC	−2LL	df	A	D	C	E	l	R
Hyperactivity-impulsivity									
ADE	**1,279.13**	**4,671.13**	**1,696**	**0.00 (.00, .00)**	**0.33 (.16, .43)**	–	**0.12 (.02, .23)**	**0.67 (.60, .73)**	**0.52 (.45, .60)**
ACE	1,288.37	4,680.37	1,696	0.27 (.17, .37)	–	0.00 (.00, .00)	0.18 (.07, .29)	0.67 (.60, .73)	0.52 (.46, .60)
AE	1,285.94	4,681.94	1,698	0.27 (.17, .37)	–	–	0.18(.08, .29)	0.67 (.60, .73)	0.52 (.45, .60)
Combined symptoms									
ADE	1,216.89	4,608.89	1,696	0.00 (.00, .00)	0.38 (.24, .48)	–	0.10 (.01, .20)	0.69 (.63, .76)	0.47 (.41, .54)
ACE	1,230.02	4,622.02	1,696	0.32 (.22, .42)	–	0.00 (.00, .00)	0.16 (.06, .27)	0.69 (.63, .75)	0.48 (.41, .55)
AE	**1,235.08**	**4,623.43**	**1,698**	**0.32 (.22, .42)**	–	–	**0.16 (.06, .27)**	**0.69 (.62, .76)**	**0.47 (.41, .55)**

AIC Akaike Information Criterion, *−2LL* −2 × log-likelihood, *df* degrees-of-freedom. Best fitting models are indicated by non-significant change in −2LL when parameters are dropped from the model using Chi square distribution with 1 df

References

Adler LA, Faraone SV, Spencer TJ, Michelson D, Reimherr FW, Glatt SJ, Marchant BK, Biederman J (2008) The reliability and validity of self- and investigator ratings of ADHD in adults. J Atten Disord 11:711–719

Bollen K, Lennox R (1991) Conventional wisdom on measurement: a structural equation perspective. Psychol Bull 110:305

Boomsma DI, Saviouk V, Hottenga J-J, Distel MA, De Moor MHM, Vink JM, Geels LM, Van Beek JHDA, Bartels M, De Geus EJC, Willemsen G (2010) Genetic epidemiology of attention deficit hyperactivity disorder (ADHD index) in adults. PLoS ONE 5:e10621

Chang Z, Lichtenstein P, Asherson PJ, Larsson H (2013) Developmental twin study of attention problems: high heritabilities throughout development. JAMA psychiatry 70:311–318

Derks EM, Hudziak JJ, Dolan CV, Van Beijsterveldt TCEM, Verhulst FC, Boomsma DI (2008) Genetic and environmental influences on the relation between attention problems and attention deficit hyperactivity disorder. Behav Genet 38:11–23

Eaves L (1976) A model for sibling effects in man. Heredity 36:205–214

Eaves LJ, Silberg JL, Maes HH, Simonoff E, Pickles A, Rutter M, Neale MC, Reynolds CA, Erikson MT, Heath AC, Loeber R, Truett KR, Hewitt JK (1997) Genetics and developmental psychopathology: 2. The main effects of genes and environment on behavioral problems in the Virginia Twin Study of Adolescent Behavioral Development. J Child Psychol Psychiatry 38:965–980

Ebejer JL, Medland SE, Van Der Werf J, Lynskey M, Martin NG, Duffy DL (2013) Variation in latent classes of adult attention–deficit hyperactivity disorder by sex and environmental adversity. J Atten Disord. doi:10.1177/1087054713506261

Gillespie NA, Henders AK, Davenport TA, Hermens DF, Wright MJ, Martin NG, Hickie IB (2012) The Brisbane Longitudinal Twin Study: pathways to Cannabis Use, Abuse, and Dependence project—current status, preliminary results, and future directions. Twin Res Hum Genet 1:1–13

Hay DA, Bennett KS, Levy F, Sergeant J, Swanson J (2007) A twin study of attention-deficit/hyperactivity disorder dimensions rated by the strengths and weaknesses of ADHD-symptoms and normal-behavior (SWAN) scale. Biol Psychiatry 61:700–705

Kan KJ, Dolan CV, Nivard MG, Middeldorp CM, Van Beijsterveldt CEM, Willemsen G, Boomsma DI (2012) Genetic and environmental stability in attention problems across the lifespan: evidence from the Netherlands twin register. J Am Acad Child Adolesc Psychiatry 52(1):12–25

Kim JH, Lee EH, Joung YS (2013) The WHO Adult ADHD Self-Report Scale: Reliability and Validity of the Korean Version. Psychiatry Investig 10:41–46

Knopik VS, Sparrow EP, Madden Paf, Bucholz KK, Hudziak JJ, Reich W, Slutske WS, Grant JD, Mclaughlin T, Todorov A (2005) Contributions of parental alcoholism, prenatal substance exposure, and genetic transmission to child ADHD risk: a female twin study. Psychol Med 35:625–635

Knouse LE, Bagwell CL, Barkley RA, Murphy KR (2005) Accuracy of self-evaluation in adults with ADHD. J Atten Disord 8:221–234

Kooij SJJ, Boonstra MA, Swinkels SHN, Bekker EM, De Noord I, Buitelaar JK (2008) Reliability, validity, and utility of instruments for self-report and informant report concerning symptoms of ADHD in adult patients. J Atten Disord 11:445

Larsson H, Asherson P, Chang Z, Ljung T, Friedrichs B, Larsson JO, Lichtenstein P (2012) Genetic and environmental influences on adult attention deficit hyperactivity disorder symptoms: a large Swedish population-based study of twins. Psychol Med 43(1):197–207

Levy F, Hay DA, Mcstephen M, Wood C, Waldman I (1997) Attention-deficit hyperactivity disorder: a category or a continuum? Genetic analysis of a large-scale twin study. J Am Acad Child Adolesc Psychiatry 36:737–744

Loehlin JC (1996) The Cholesky approach: a cautionary note. Behav Genet 26:65–69

Magnússon P, Smári J, Sigurðardóttir D, Baldursson G, Sigmundsson J, Kristjánsson K, Sigurðardóttir S, Hreiðarsson S, Sigurbjörnsdóttir S, Guðmundsson ÓÓ (2006) Validity of self-report and informant rating scales of adult ADHD symptoms in comparison with a semistructured diagnostic interview. J Atten Disord 9:494–503

Martin N, Scourfield J, Mcguffin P (2002) Observer effects and heritability of childhood attention-deficit hyperactivity disorder symptoms. Br J Psychiatry 180:260–265

Merwood A, Greven C, Price T, Kuntsi J, Mcloughlin G, Larsson H, Asherson P (2013) Different heritabilities but shared aetiological influences for parent, teacher and self-ratings of ADHD symptoms: an adolescent twin study. Psychol Med 1:12

Neale MC, Boker SM, Xie G, Maes HM (1999) Statistical modeling. Department of Psychiatry, Richmond

Pinto R, Rijsdijk F, Frazier-Wood A, Asherson P, Kuntsi J (2012) Bigger families fare better: a novel method to estimate rater contrast effects in parental ratings on ADHD symptoms. Behav Genet 42(6):875–885

Posthuma D, Boomsma DI (2000) A note on the statistical power in extended twin designs. Behav Genet 30:147–158

Simonoff E, Pickles A, Hervas A, Silberg JL, Rutter M, Eaves L (1998) Genetic influences on childhood hyperactivity: contrast effects imply parental rating bias, not sibling interaction. Psychol Med 28:825–837

Swanson J, Schuck S, Mann M, Carlson C, Hartman K, Sergeant J, Clevenger W, Wasdell M, Mccleary R (2005). Categorical and dimensional definitions and evaluations of symptoms of ADHD: The SNAP and the SWAN Ratings Scales [Draft] [Online]. Available: http://www.adhd.net/SNAP_SWAN.pdf. Accessed 24th February, 2012

Wright MJ, Martin NG (2004) Brisbane adolescent twin study: outline of study methods and research projects. Aust J Psychol 56:65–78

Zhu G, Duffy DL, Eldridge A, Grace M, Mayne C, O'gorman L, Aitken JF, Neale MC, Hayward NK, Green AC (1999) A Major Quantitative-Trait Locus for Mole Density Is Linked to the Familial Melanoma Gene CDKN2A: a Maximum-Likelihood Combined Linkage and Association Analysis in Twins and Their Sibs. Am J Hum Genet 65:483–492